INTEGRATING DECISION HEURISTICS AND BEHAVIOURAL REFINEMENTS INTO TRAVEL CHOICE MODELS

CAMILA BALBONTIN

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the Business School, University of Sydney, Australia

The Institute of Transport and Logistics Studies

The University of Sydney Business School

The University of Sydney

February 2018

Abstract

Discrete choice modelling has become the preferred empirical context to study individuals' preferences and willingness to pay. Although the *outcome* is important in decision making, so is the *process* that individuals adopt to assist them in reaching a decision. Both should be considered when analysing individual behaviour as they represent jointly the endogeneity of choice. Traditional choice studies assume, in the main, a linear in the parameters additive in the attributes (LPAA) approach, where individuals are rational, take into account all the attributes and alternatives presented to them when reaching a decision, and value the attribute levels exactly as were presented in the popular choice experiment paradigm. This has not always been shown to be a behaviourally valid representation of choice response, and there is a growing literature on the role of a number of alternative decision process strategies that individuals use when facing a decision, which are often referred to as heuristics, or simply as process rules.

The majority of choice studies also assume that respondents have a risk attitude that is risk neutral (i.e., a risky alternative is indifferent to a sure alternative of equal expected value) and that they perceive the levels of attributes in choice experiments in a way that suggests the absence of perceptual conditioning. Considering each in turn, there are people who are risk adverse, risk taking or risk neutral, and this heterogeneity in risk attitude does influence individuals' decisions when faced with different choice scenarios. Heterogeneity is also present for perceptual conditioning in cases where there is variability in the outcomes of an attribute(s), which allows for differences between the stated probability of occurrence (in a choice experiment) and the perceived probability used when evaluating the prospect. Finally, the (accumulated) experience that individuals' have with each alternative might also influence their decisions.

The objective of this research is to integrate multiple decision process strategies, Value Learning (VL) and Relative Advantage Maximisation (RAM) in particular, alongside the traditional LPAA 'process rule' with behavioural refinements (i.e., risk attitudes, perceptual conditioning and overt experience), to take into account process endogeneity in choice responses. A novel approach is used to include process heterogeneity, referred to as *conditioning of random process heterogeneity*, where the mean and standard deviation of the parameters normally defined under an LPAA heuristic are conditioned by process strategies. This approach takes into account the relationship between process heterogeneity and preference heterogeneity, of particular interest in studies that integrate random parameters and process strategies. The model performance results and willingness to pay estimates are compared to those obtained when using a *probabilistic decision process* method, increasingly used in the choice literature to accommodate process heterogeneity.

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes. I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in the preparation of this thesis and sources have been acknowledged.

Acknowledgments

This investigation is the result of years of conversations, brain storming and discussions. This journey definitely would not have been possible if it was not for my main supervisor, David Hensher. Since I started my PhD in August 2014, David has always supported and motivated me, and always challenged me into thinking outside the box. It has been an honour working with such an amazing mentor and supervisor. I am also extremely grateful to my co-supervisor, Andrew Collins, who has encouraged me since the beginning and has always been willing to help me. I am forever thankful to have shared this journey with two extraordinary supervisors.

I need to thank everyone in ITLS for having been a great support during these past years, especially to Corinne Mulley with whom I have had the pleasure of working with on several occasions, and Michiel Bliemer for having helped me during the last parts of my thesis. I also want to thank my incredible MSc supervisor in Chile, Juan de Dios Ortúzar, who has continued to encourage me during these past years and who introduced me to David. To my fellow PhDs, especially Ines Osterle, Collins Teye, Mark Raadsen and Mahbub Hakim, thank you for having been an amazing part of my PhD life.

I need to recognise Chile's National Commission for Scientific and Technological Research (CONICYT) for providing me with the scholarship *Beca Chile Doctorado en el Extranjero*. This research was partially funded by ARC Discovery Project (DP): DP140100909: Integrating Attribute Decision Heuristics into Travel Choice Models that accommodate Risk Attitude and Perceptual Conditioning. I also acknowledge the facilities, and the scientific and technical assistance of the Sydney Informatics Hub at the University of Sydney and, in particular, access to the high performance computing facility Artemis which was fundamental to run all the models presented.

Finally, to those who have been with me forever and are my heart and soul: my mother, María Elina; my father, Carlos; my sisters, Carla and María Elina; all of my incredible friends, especially Isadora and Isidora; and my partner, Raimundo. This would have never been possible without you all.

Table of Contents

List of Tables

List of Figures

Notational Glossary

General Nomenclature

AIC: Akaike's Information Criterion

ANA: Attribute non-attendance

BR: Behavioural refinements (refers to risk attitudes and perceptual conditioning specifically)

CRPH: Conditioning random process heterogeneity

EC: Error components

EEUT: Extended expected utility theory

EUT: Expected utility theory

HWF: Heuristic weighting function

LPAA: Linear in the parameters and additive in the attributes

LR: Log likelihood ratio test

MML: Mixed multinomial logit model

MNL: Multinomial logit model

PDP: Probabilistic decision process

PDPC: Probabilistic decision process combined

PT: Prospect theory

RAM: Relative advantage maximisation

RP: Revealed preference

RUM: Random utility maximisation

RUT: Random utility theory

SP: Stated preference

VL: Value learning

WTP: Willingness to pay

Equations

- Individuals *q*
- Attributes *n* and *m*
- Choice situation/choice task *t*
- Alternatives *i* and *j*
- Classes *c*

Attributes and Parameters

Acronym	Description	Dataset
ASCBUS	Alternative Specific Constant Bus	Metro Rail and Northwest
ASCTRAIN	Alternative Specific Constant Train	Metro Rail and Northwest
ASCMETRO	Alternative Specific Constant Metro	Metro Rail
ASCLR	Alternative Specific Constant New Light Rail	Northwest
ASCNHR	Alternative Specific Constant New Heavy Rail	Northwest
ASCNBW	Alternative Specific Constant New Busway	Northwest
ASCBW	Alternative Specific Constant Busway	Northwest
ASCCAR	Alternative Specific Constant Car	Metro Rail and Northwest
ACTIMEPT	Access Time	Metro Rail and Northwest
COSTPT	Fare Public Transport	Metro Rail and Northwest
COSTCRTRC	Fuel + Toll Cost Car	Metro Rail and Northwest
COSTCRPC	Parking Cost Car	Metro Rail and Northwest
TTPT	Travel Time Public Transport	Metro Rail and Northwest
TTCR	Travel Time Car	Metro Rail and Northwest
EGTIME	Egress Time	Metro Rail and Northwest
TRANPT	Transfer Public Transport	Metro Rail
FREQPT	Headway Public Transport	Metro Rail and Northwest
SEATPT	% Seat Public Transport	Metro Rail
STANDPT	Density Public Transport	Metro Rail
EXPBS	Experience Bus	Metro Rail and Northwest
EXPBW	Experience Busway	Northwest
EXPTR	Experience Train	Metro Rail and Northwest
EXPCR	Experience Car	Metro Rail and Northwest
ALPHABSTT	Risk Attitudes Travel Time Bus	Metro Rail
ALPHATRTT	Risk Attitudes Travel Time Train	Metro Rail
ALPHAMTTT	Risk Attitudes Travel Time Metro	Metro Rail
ALPHAEXTT	Risk Attitudes Travel Time Currently Available Modal Facilities	Northwest
ALPHANEXTT	Risk Attitudes Travel Time New Modal Investments	Northwest
ALPHACRTT	Risk Attitudes Travel Time Car	Metro Rail and Northwest
ALPHABSCS	Risk Attitudes Cost Bus	Metro Rail
ALPHATRCS	Risk Attitudes Cost Train	Metro Rail
ALPHAMTCS	Risk Attitudes Cost Metro	Metro Rail
ALPHAEXCS	Risk Attitudes Cost Currently Available Modal Facilities	Northwest
ALPHANEXCS	Risk Attitudes Cost New Modal Investments	Northwest
ALPHACRTRCS	Risk Attitudes Fuel+Toll Car	Metro Rail and Northwest
ALPHACRPCS	Risk Attitudes Parking Car	Metro Rail and Northwest
GAMMABS	Perceptual Conditioning Bus	Metro Rail
GAMMACR	Perceptual Conditioning Car	Metro Rail

Process Homogeneity Models

General Models

	LPAA	RAM	VL	Parameters	Behavioural refinements	Experience
LPAA_MNL	Yes	No	No	Fixed	No	No
LPAA_MNL_BRExp	Yes	No	No	Fixed	Yes	Yes
LPAA_MML	Yes	No	No	Random	No	No
LPAA_MML_BRExp	Yes	No	No	Random	Yes	Yes
VL_MNL	No	No	Yes	Fixed	No	No
VL_MNL_Exp	No	No	Yes	Fixed	Yes	Yes
VL_MML	No	No	Yes	Random	No	No
VL_MML_Exp	No	No	Yes	Random	Yes	Yes
RAM_MNL	No	Yes	No	Fixed	No	No
RAM_MNL_Exp	No	Yes	No	Fixed	Yes	Yes
RAM_MML	No	Yes	No	Random	No	No
RAM_MML_Exp	No	Yes	No	Random	Yes	Yes

Error Components Models

	LPAA	VL	Parameters	Choice set sequence included?
EC_LPAA	Yes	No	Fixed	No
EC_SeqLPAA	Yes	No	Fixed	Yes
EC_VL	No	Yes	Fixed	No
EC_SeqVL	No	Yes	Fixed	Yes

Utility Functions

1) LPAA Model:

$$U_{iqt} = \theta_{i1} \cdot x_{i1qt} + \theta_{i2} \cdot x_{i2qt} + \ldots + \theta_{in} \cdot x_{inqt} + \varepsilon_{iqt}$$

θ_{in} are the estimates representing the difference between the level of attribute *n* and alternative *i* and the reference level for that same attribute *n*; and x_{inq} represents the level of attribute *n* of alternative *i* for the individual *q.*

The utility function when considering behavioural refinements can be written as:

$$U_{iqt} = \theta_{i1} \cdot x_{i1qt} + \theta_{i2} \cdot \sum_{l \in L} \left[\frac{p_{i2qt,l}{}^{\gamma}}{(p_{i2qt,l}{}^{\gamma} + (1 - p_{i2qt,l})^{\gamma})^{1/\gamma}} \cdot \frac{x_{i2qt,l}{}^{1-\alpha_2}}{1-\alpha_2} \right] + \theta_{i3} \cdot \left[\frac{x_{i3qt}{}^{1-\alpha_3}}{1-\alpha_3} \right] + \varepsilon_{iqt}$$

where γ represents the degree of curvature of the weighting function, and α represents the risk attitudes towards the second and third attribute.

2) VL Model

$$VL(\mathrm{x}_{inqt}) = \begin{cases} \left(x_{inqt} - ref_n\right)^{\varphi} & \text{if } \left(x_{inqt} - ref_n\right) \geq 0 \\ -\left[-\left(x_{inqt} - ref_n\right)\right]^{\varphi} & \text{if } \left(x_{inqt} - ref_n\right) < 0 \end{cases}$$

$$U_{iqt} = \theta_{i1} \cdot VL(\mathrm{x}_{i1qt}) + \theta_{i2} \cdot VL(\mathrm{x}_{i2qt}) + \ldots + \theta_{in} \cdot VL(\mathrm{x}_{inqt}) + \varepsilon_{iqt}$$

ref_n represents the reference level for attribute *n*; and φ represents the concavity factor used to transform the differences between the attribute levels and the reference levels as follows

3) RAM Model

The definition of advantage of an alternative *i* over an alternative *j* is equivalent to the disadvantage of the alternative *j* over alternative *i* over all attributes *m*, as follows:

$$A(\mathrm{j},\mathrm{i}) = D(i,\mathrm{j}) = \sum_{m} \ln(1+\exp(\theta_{jm}\cdot x_{jm} - \theta_{im}\cdot x_{im}))$$

The relative advantage of *i* over *j*, denoted as $R(i,\mathrm{j})$ can be written as:

$$R(i,\mathrm{j}) = \frac{A(i,\mathrm{j})}{A(i,\mathrm{j}) + D(i,\mathrm{j})}$$

The utility function for an alternative *i* will be equivalent to the utility function under an LPAA heuristic plus the relative advantage of alternative *i* over all other alternatives *j* as follows:

$$U_{iqt} = \theta_{i1}\cdot x_{i1qt} + \theta_{i2}\cdot x_{i2qt} + \ldots + \theta_{in}\cdot x_{inqt} + \sum_{j\in S} R(i,j) + \varepsilon_{iqt}$$

4) Adding experience

$$U_{iqt}^{experience} = \left(1 + \theta_{exp,i}\cdot x_{experience,iq}\right)\cdot U_{iqt}$$

where $U_{iqt}^{experience}$ is the transformed utility function; $x_{experience}$ is defined as a dummy variable equal to 1 if alternative *i* was chosen by respondent *q* in their most recent decision and 0 otherwise; and $\theta^{0}_{exp,i}$ is the associated parameter estimates as:

$$\theta_{exp,i} = \begin{cases} \theta^{0}_{exp,i} & \text{if } U_{iqt} \geq 0 \\ -\theta^{0}_{exp,i} & \text{if } U_{iqt} < 0 \end{cases}$$

Process Heterogeneity Models: including LPAA, RAM and VL

	Parameters	Behavioural refinements	Experience	Description
PDP	Fixed	No	No	Each heuristic is represented by a different class
PDP_BRExp	Fixed	Yes	Yes	Each heuristic is represented by a different class
CRPHm	Random	No	No	Interactions between the **mean** normally defined under a LPAA heuristic, and the process strategies VL and RAM
CRPHs	Random	No	No	Interactions between the **standard deviation** normally defined under a LPAA heuristic, and the process strategies VL and RAM
CRPHms	Random	No	No	Interactions between the **mean** and **standard deviation** normally defined under a LPAA heuristic, and the process strategies VL and RAM
CRPHms_BRExp	Random	Yes	Yes	Interactions between the **mean** and **standard deviation** normally defined under a LPAA heuristic, and the process strategies VL and RAM

Utility Functions

1) PDP

In this research three heuristics will be included: RAM, VL and LPAA. Therefore, the model equations for each class will be equivalent to the utility functions for each heuristic, as follows:

$$V_{iqt|C1} = \theta_{i1|C1} \cdot x_{i1qt} + \theta_{i2|C1} \cdot x_{i2qt} + \ldots + \theta_{in|C1} \cdot x_{inqt} + \sum_{j \in S} R(i, j)$$

$$V_{iqt|C2} = \theta_{i1|C2} \cdot \left[\left(x_{i1qt} - ref_n\right)\right]^{\varphi} + \theta_{i2|C2} \cdot \left[\left(x_{i2qt} - ref_n\right)\right]^{\varphi} + \ldots + \theta_{in|C2} \cdot \left[\left(x_{inqt} - ref_n\right)\right]^{\varphi}$$

$$V_{iqt|C3} = \theta_{i1|C3} \cdot x_{i1qt} + \theta_{i2|C3} \cdot x_{i2qt} + \ldots + \theta_{in|C3} \cdot x_{inqt}$$

2) CRPH

$$U_i = \sum_n \left(\left[\begin{array}{l} \theta_{in} + \lambda'^{m}_{VL,in} \cdot \left(x_{inqt} - ref_n \right)^{\varphi} + \lambda^{m}_{RAM,in} \cdot \left(\theta_{in} \cdot x_{inqt} + \sum_{j \in S} R(i,j) \right) \\ + \left[\sigma_{in} + \lambda'^{s}_{VL,in} \cdot \left(x_{inqt} - ref_n \right)^{\varphi} + \lambda^{s}_{RAM,in} \cdot \left(\theta_{in} \cdot x_{inqt} + \sum_{j \in S} R(i,j) \right) \right] \cdot v \end{array} \right] \cdot x_{inqt} \right) + \varepsilon_{iqt}$$

The parameters θ_{in} can be considered common between LPAA and RAM, but not for VL, which will have its own parameters θ_{in}^{VL}, defined as:

$$\lambda'_{VL,in} = \lambda_{VL,in} \cdot \theta_{in}^{VL}$$

where $\lambda^{m}_{VL,in}$ represents the relationship between the mean estimate and VL; $\lambda^{m}_{RAM,in}$ represents the relationship between the mean and RAM; $\lambda^{s}_{VL,inq}$ the relationship between the standard deviation and VL; $\lambda^{s}_{RAM,inq}$ the relationship between the standard deviation and RAM.

CHAPTER 1

Introduction

1.1. Research Topics and Questions

Individuals are frequently faced with decisions which are influenced by a variety of factors. Some influences are related to the information available that is relevant to a specific context, and others are related to relevant experience or knowledge that the individual has acquired that can inform a decision. The psychology and marketing literature in particular has focused on understanding the diverse ways in which an individual may reach a decision and offers different ways in which context, experience and information might relate to each other and influence choices (Bates, 1954; Payne, 1976; Gigerenzer et al., 1999). In addition, the broader multidisciplinary decision making literature suggests that there may be differences across individuals, including the ways in which people integrate context, experience and information; the risk attitude applied (risk aversion, risk taking or risk neutrality); and the way in which objective evidence is perceived (Edwards, 1961; Slavic et al., 1977). In general, the body of existing literature suggests that studying individual or group behaviour is complex with numerous possible behaviours and behavioural responses.

By way of example, imagine a man named Bob who travels from his house to work every day during peak hours. He used to take the bus, but a year ago the government built a train station near his house and he started using the train. In the route Bob uses, the trains are less frequent than the buses, but the travel times for the trains are more predictable than for the buses (i.e., there is less variability). Bob is highly risk averse towards travel time: he prefers to wait longer for the train than taking the bus because the latter mode has a higher travel time variability, even when the buses might be faster on average and more frequent. The government is now promoting a new bus service that is very regular and has the same frequencies as the old bus service. Since this new alternative reduces the travel time variability, the government would expect to attract individuals such as Bob back to the bus alternative. However, since Bob has had experience with high travel time variability in buses, he does not *believe* that the new bus alternative could be as regular as the trains, and thus perceives the bus travel times as having a high degree of variability. Hence, Bob continues using the train. This is not the case for his neighbour Alice who is risk neutral, where her decision depends only on the average perceived travel time when it is lower than 15 minutes, but if it is higher she will also take into account how crowded the mode of transport is. The train and the old bus service both took on average more than 15 minutes, so Alice used the train because it was usually less crowded than the bus. However, the new bus alternative only takes Alice 12 minutes to get to work, so she switches to this new bus service. Bob and Alice are examples of classes of individuals who make travel decisions on a daily basis, and yet they are very different in their behavioural processes. Thus, any attempt to understand the decision making of a large number of individuals is complicated because of differences between individuals.

When trying to understand individual or group (such as household or organization) behaviour, discrete choice modelling has been widely used across many disciplines such as transportation, health economics, marketing, and resource and environmental economics. In a discrete choice study, decision makers[1] are observed to choose between two or more alternatives; each of these may be described by attribute levels, and on any given choice occasion the decision maker chooses one as the preferred alternative. Such studies have wide appeal because of the range of useful outputs they can provide, including estimates of willingness to pay (WTP) for different attributes of the choice alternatives, and prediction of the influence of changes in explanatory variables on total demand and choice shares. The choices made by the decision maker may be stated in a hypothetical context or revealed in a market or natural environment. In the former case, hypothetical choice situations are used (referred as stated choice scenarios), which are generated using experimental design techniques, and in which respondents are asked to evaluate a number of sequenced choice scenarios. For each scenario, the respondent indicates their preferred alternative. Stated preference (SP) data that is rich in design and scenario options enables the analyst to study variations in attribute levels for existing alternatives as well as totally new alternatives that are not observed in real markets. In contrast, revealed preference (RP) data can be used to reveal the preferences of decision makers for alternatives available in the market, although markets

[1] This thesis will mainly refer to individuals as the decision makers, but it is important to mention that decision makers may not be individuals. For example, they could be groups of individuals such as households and firms.

may reveal a limited spectrum of preference due to limited variability in the explanatory variables. Whereas the attribute levels are designed in SP experiments, they are observed or reported by a respondent in RP studies. The former has no measurement error in the attributes (but can have perceptual conditioning, i.e., the probability of occurrence may not exactly replicate the value that individuals assign to the outcome, so there exists a perceptually conditioned probability) but does so in the choice response; by contrast the attributes in RP studies may have measurement error but the choice response is the real market response. Furthermore, the calculation of the willingness to pay (WTP) estimates is better facilitated in SP experiments because they encourage the trade-off between attributes. On the contrary, in RP data usually the cost attribute is highly correlated to some of the attributes, such as the travel time. In both study paradigms information, context and experience can and do influence how attributes are processed and hence choices are made.

Traditional discrete choice models assume that decision makers assess all the attributes that describe the alternatives as if all are relevant and tradeable, and their preferences are completely rational. These assumptions imply risk neutrality and an objective perception of the attributes presented. McFadden (2001, p. 374), amongst others, raised the necessity of including information, experience and decision processes in the traditionally used random utility maximization (RUM):

"*What lies ahead? I believe that the basic RUM theory of decision-making, with a much larger role for experience and information in the formation of perceptions and expression of preferences, and allowance for the use of rules as agents for preferences, can describe most economic choice behavior in markets, surveys, and the laboratory. If so, then this framework can continue for the foreseeable future to form a basis for microeconometric analysis of consumer behavior and the consequences of economic policy.*"

Recent literature has offered a more extensive set of behavioural assumptions linked in particular to process rules or heuristics that recognise the real possibility that heterogeneity in a population of decision makers exists and needs to be captured in ways that represent how individuals actually process information, including the extent of influence of specific attributes which varies across a population of interest (Hensher 2006; Chorus et al. 2008; Leong and Hensher 2012; Hess et al. 2012). This recent literature has not only considered preference heterogeneity but also identified process heterogeneity. This recognises a role for different processing strategies that individuals use when facing a decision, which are often referred to in the literature as *heuristics* or simply as *process rules*. The great majority of choice studies also assume that respondents have a *risk attitude* that is risk neutral and that they also perceive the levels of attributes in choice experiments in a way that suggests the absence of perceptual conditioning. Considering each in turn, there are people who are risk adverse, risk taking or risk neutral, and this heterogeneity in risk attitude does influence individuals' decisions when faced with different choice scenarios (Senna 1994; Polak et al. 2008; Hensher and Rose 2009). Heterogeneity is also present for *perceptual conditioning* in cases where there is variability in the outcomes of an attribute(s), which allows for differences between the stated probability of occurrence (or a subjective value such as beliefs) and the perceived probability used when evaluating the prospect (Tversky and Kahneman 1992; Camerer and Ho 1994; Wu and Gonzalez 1999; Hensher et al. 2013). All of these components (multiple heuristics, risk attitudes and perceptual conditioning) recognise topics that influence decision making, and to the best of our knowledge have not been analysed jointly in both the transport literature and other literatures.

The conceptual framework proposed for this thesis is shown in Figure 1-1. The elements inside the dashed square influence directly the valuation of individual attributes of the choice alternatives. When an individual is faced with making a decision, he might have risk attitudes that influence the way in which he processes the information (even if there is not any variability on the levels of the attributes, and in real or hypothetical market contexts). On the other hand, if there is variability in the levels of an alternative's attribute, degrees of risk or of uncertainty might be present. If the probabilities of occurrence of the levels are defined or known by the individual, the individual faces a decision with risk. If these probabilities are not defined or are unknown by the individual, the decision presents uncertainty or ambiguity. Typically in real decisions, the probabilities of the outcomes are not known, so decisions are made under uncertainty and ambiguity. When an individual faces a decision under risk, there might be a presence of perceptual conditioning through the interpretation of the probabilities of occurrence associated with varying levels of attributes or circumstances. The interaction between these elements can be seen in Figure 1-1, where experience and the elements inside the grey square may vary between individuals. The dashed square represents behavioural refinements that allow us to have a better understanding of individual preferences.

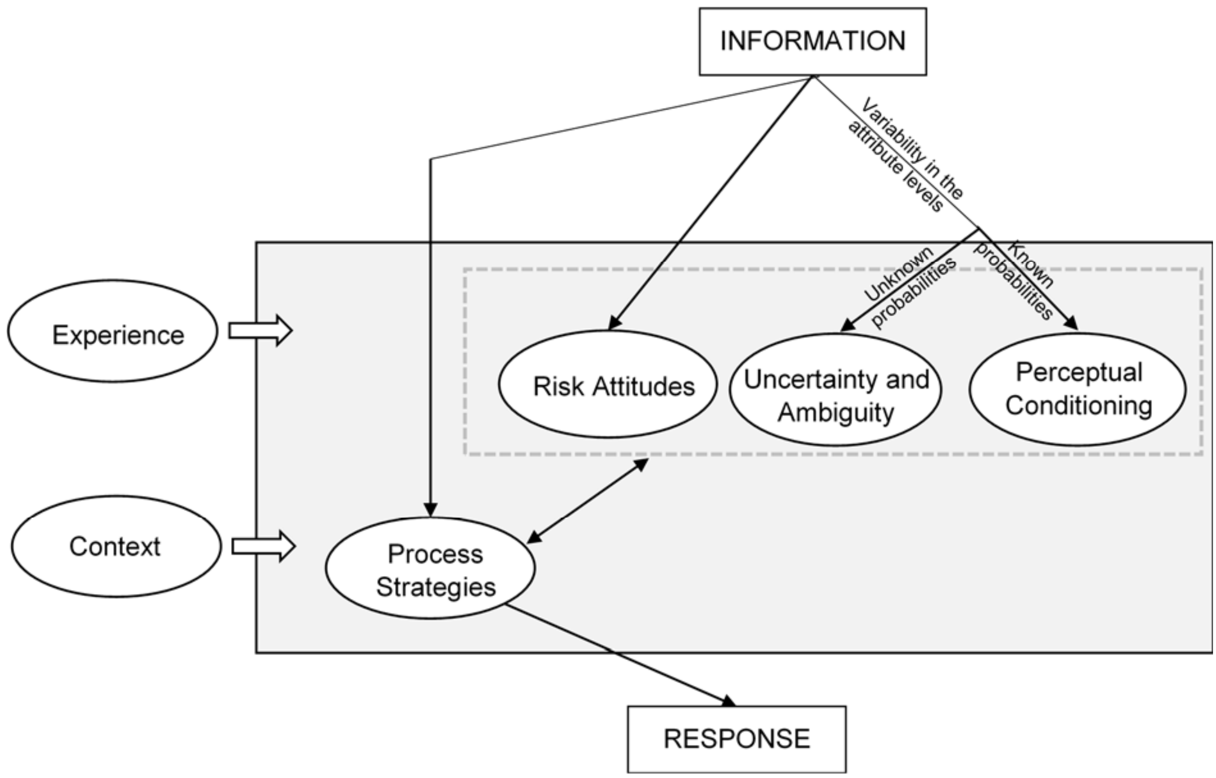


Figure 1-1: Proposed conceptual framework for decision making

Although there is research focused on studying which process strategies might be more accurate than others, they are not necessarily exclusive. More than one heuristic might be acting together in decision-making in two ways: (1) if one individual uses one process strategy and another individual uses a different one – process heterogeneity across individuals; (2) if an individual uses more than one process heuristic to make a decision – process heterogeneity within individuals. Studies that have included multiple process strategies usually consider each one as acting by itself. One of the approaches used in the literature, referred as *Probabilistic Decision Process* (PDP), considers that an individual might choose one process strategy only, whether it be a traditional or an alternative one. A different approach, referred to as Heuristic Weighting Function (HWF), considers that one individual might be partly using one process strategy and partly using another one, weighting them by a certain factor. Another approach, referred in literature as a *Hybrid Model,* considers more than one process heuristic in the utility function, either because the heuristics are essentially different in terms of what they are including, or they can be considering acting upon a different subset of the attributes. All of these approaches used in choice studies assume that there is no interaction between the process strategies when valuing the attributes. However, it might certainly be the case that the valuation of attributes considered under a traditional heuristic might be influenced by process strategies.

The research questions that this thesis seeks to answer are the following:

1. Are preferences better represented when considering multiple decision process strategies, risk attitudes, perceptual conditioning and experience? How might they work together?

2. How do decision process strategies interact with each other?

3. Is there any relationship between process heterogeneity and taste heterogeneity?

4. How do the various behavioural elements of the integrated choice process influence key behavioural outputs such marginal (dis)utilities, willingness to pay estimates and confidence intervals?

1.2. Research Approach and Contributions

The purpose of this research is to provide a better understanding of decision making by simultaneously considering risk attitudes, perceptual conditioning, and multiple decision process strategies. Although these components have been recognised as an important part of decision-making (McFadden 2001), the full set of candidate process influences have not yet been incorporated in discrete choice modelling to investigate the contribution of each behavioural element in the identification of key behavioural outputs such as willingness to pay estimates and elasticities.

There is a growing number of studies investigating the role of various heuristics in choice modelling, including multiple heuristics. However, the existing literature is still rather limited even though all of them suggest that the consideration of multiple heuristics significantly improves the goodness of fit of the models and provides a richer understanding of individuals' behaviour. The heuristics used in literature can be separated in three major categories: (1) context free heuristics, where individuals evaluate each alternative using only the attributes that define it, (2) local choice context dependent heuristics, where the attributes that define one alternative influence how individuals evaluate the competing alternatives, and (3) choice set interdependent, where the attributes that individuals previously faced influence how they make their current decisions. This thesis will focus in three heuristics to cover the different types of categories that represent different relationships between alternatives and choice sets.

The first of them is the linear in the parameter additive in the attributes heuristic (LPAA) that is the one traditionally used and is part of the context free heuristics. The second heuristic belongs to the category local choice context dependent heuristics. Specifically, this thesis considers an influence referred to as Relative Advantage Maximisation (RAM) where an individual will ponder the advantages and disadvantages of an alternative relative to the competing alternatives as a function of the difference between the alternatives' characteristics. This heuristics was chosen because it has proven to have a significant influence on how decisions are made. The third heuristic is part of the choice set interdependent heuristics, and is referred to as Value Learning (VL). This heuristic considers that the best characteristics of the alternatives that have previously been chosen will have an effect of how an individual assesses the alternatives in the current scenario. VL was selected because, even though it has proven to be significant in decision making, it has received limited attention in transportation literature.

The three heuristics considered in this thesis – LPAA, VL and RAM - can take place in real life decisions (RP) or in hypothetical decisions (SP). The LPAA and RAM heuristic can easily be pictured in both RP and SP scenarios as they only require more than one alternative with common characteristics. However, the VL heuristic requires sequential decision-making. Common SP experiments present more than one scenario to each individual in a short period of time. Hence, it is likely that they are able to remember the past alternatives' characteristics and use them to assess the current alternatives. However, in an RP context this is more arguable since individuals face the same decision with less frequency so their memories are more likely to be influenced by other factors. Nevertheless, VL can be used both in SP and RP experiments. This thesis will use two SP datasets to validate the methodologies and results. The focus of this thesis is to assess different formulations to incorporate process heterogeneity and other behavioural refinements, so it is out of the scope to compare the results to the ones from an RP experiment.

As mentioned in the previous section, different model forms have been proposed to include process heterogeneity in one model. The PDP approach is the most flexible ones in terms of allowing for different heuristics to be included, so this is the first model form that will be used in this thesis. Current literature has suggested a possible confoundment between process heterogeneity and preference heterogeneity (Hess et al., 2012; Hensher et al., 2013a; Collins et al., 2013; Campbell et al. 2014). One of the most important contributions of this thesis is a novel methodology to incorporate process heterogeneity that shows that process heterogeneity conditions random preference heterogeneity. Therefore, it allows there to be a relationship between process and preference heterogeneity, and determines the extent to which this relationship changes the standard output measures of interest to analysts, such as willingness to pay (WTP) estimates. If there is a modification, this thesis will analyse to what extent do the mean and standard deviation change, and if there is an increase or decrease. These are interesting and behaviourally relevant questions given the importance of WTP indicators in demand forecasting and user benefits in project appraisal. This original methodology will be described with more detail throughout the sections of this thesis.

Different studies have investigated the role of risk attitudes, perceptual conditioning and experience in decision-making. These studies reveal a significant improvement in the models' outcomes and in the understanding of individuals' preferences when including these behavioural refinements in the choice models. This thesis will incorporate different ways in which these components might be influencing decision making (e.g., risk attitudes when there is variability in an attribute and where there is not).

An objective of the thesis is to make a contribution to support researchers and practitioners in their consideration of the potential role of richer behavioural forms of choice models. Two different datasets are used to establish the extent to which the proposed method can be supported more generally. The two datasets vary in the way alternatives are presented, with one data source (referred to as the Northwest data) includes commonly used attributes, such as travel times and cost, and the other source (called Metro Rail) additionally including travel time reliability and public transport crowding. The datasets also vary in terms of their attribute levels and socioeconomic characteristics.

As discussed above, there is an important void in the choice modelling literature regarding the use of multiple heuristics, especially concerning a possible relationship between them, and there is also a gap regarding the influence of behavioural refinements and experience on the use of process strategies. This thesis aims to fill this gap in the discrete choice modelling literature. This approach proposed has much behavioural merit, given that both of these literatures – multiple heuristics together with risk attitudes, perceptual conditioning and experience – have been independently shown to be significant in the modelling results, suggesting valuable insights into choice making. Likewise, it shows that the decision process strategies used by individuals are also relevant in their behavioural responses and influence the willingness to pay estimates. To the best of our knowledge, there are no current studies that integrate multiple heuristics together with risk attitudes, perceptual conditioning or experience. One of the contributions of this thesis is to analyse and interpret how the consideration of both literatures together improves our understanding of respondent preferences. Another important contribution - as mentioned above - is a novel approach to include multiple heuristics by considering a relationship between process and preference heterogeneity, a subject which has been widely mentioned in the literature but not yet studied.

1.3. Thesis Outline

The chapters of this thesis are organised as follows. Chapter 2 presents the current literature on the various topics of this thesis. It starts by explaining different process strategy rules that have been proposed in the literature, divided into three categories: (1) context free heuristics – where an individual assesses an alternative without considering the competing alternatives; (2) local choice context dependent heuristics – where an individual takes into account all the possible alternatives when assessing each of them; and (3) choice set interdependent heuristics – where current decisions are influenced by similar decisions made in the past. Then the chapter reviews literature that has combined more than one heuristic. The second part of the chapter is focused on behavioural refinements and experience.

Chapter 3 presents the econometric methodology used in this thesis. This is followed by the selection of heuristics used and their formulations and implications. The methodologies used to integrate the behavioural refinements and experience are reviewed, followed by the methodologies used to incorporate process heterogeneity. The methodology proposed in this thesis is presented with a detailed analysis of its contribution and interpretation. The final part of this chapter presents how to calculate the marginal (dis)utilities and WTP estimates with their confidence intervals using the Delta method for all the models.

Chapter 4 presents a description of the two datasets used in the modelling: the Metro Rail dataset and the Northwest dataset. The first and second sections describe each survey with their attributes and attribute levels and sample sizes; whereas the third section provides a comparison of the characteristics of the surveys with regards to the attributes considered and the mean and standard deviation of the attribute levels, as well as an overview of the characteristics of respondents.

Chapter 5 and Chapter 6 provide a detailed assessment of the estimated model results using the range of studied heuristics - LPAA, VL and RAM, together with behavioural refinements and experience in the MetroRail and Northwest dataset, respectively. The first part of each chapter (Sections 2, 3 and 4) assesses the statistical performance of each model type separately, with Section 5 undertaking a comparison of the range of model specifications. This comparative assessment comprises the estimated parameters associated with behavioural refinements and overt experience; the log likelihood and AIC indicators; and the willingness to pay estimates (WTP).

The final chapter (Chapter 7) begins with the main contributions. The many model forms that investigate the role of process homogeneity, process heterogeneity, behavioural refinements and experience are assessed in the context of the research questions posed earlier in the thesis. The chapter concludes with a discussion of areas for future research together with some concluding remarks.

CHAPTER 2

Background: Process strategies and behavioural refinements

2.1. Introduction

Mainstream discrete choice modelling approaches have evolved in a setting in which some very specific behavioural assumptions are made in specifying choice models and estimation methods. While they have served the literature well and are often the ‘bread and butter’ procedures in practical applications, they are not without criticism. For example, the majority of choice studies assume that individuals assess the offered attributes using a fully compensatory decision rule which involves trading off all attributes as if all are relevant and tradeable. Another debatable assumption is that people’s preferences are ‘rational’; as proposed by Luce (1959) this means transitivity of preferences (i.e., if ‘a’ is preferred over ‘b’ and ‘b’ is preferred over ‘c’, then ‘a’ is preferred over ‘c’), logical dominance (i.e., if you introduce a preferred alternative, that would be the dominant) and indifference between formally equivalent alternatives. This assumes risk neutrality and an objective perception of the attributes presented, and as will be discussed, these assumptions might be easily violated by some individuals.

This chapter reviews the spectrum of process rules or heuristics that have been presented in the literature with a commentary on the merits of each rule. The first section of this chapter reviews the heuristics proposed in the literature, divided into three different categories: (1) those ones that are independent of the alternatives presented in the choice set – context free heuristics; (2) those that depend on all the alternatives shown in a choice set – local choice context dependent; and (3) those that depend on the multiple choice tasks shown to an individual – choice set interdependent. The next section reviews the literature that has combined more than one heuristic in explaining preferences. Section 2.7 presents the issue on confoundment between process and preference heterogeneity as has been suggested in the literature. Subsection 2.8 presents different behavioural refinements that have been studied, such as risk attitudes and perceptual conditioning. The next subsection reviews how past experiences have proven to influence decision making. The final section summarises the main findings and the research gaps that have not been addressed.

2.2. Heuristics

The great majority of choice studies assume that decision makers are rational, take into account all the attributes included in a stated choice experiment or as listed in a revealed preference model, and value the levels exactly as are presented to them. This process strategy will be referred to as linear in the parameters and additive in the attributes (LPAA). As mentioned above, the underlying assumptions of this strategy have been widely criticised in the literature as they do not necessarily represent how decisions are made (although sometimes they do). These criticisms have led to the development of several other process strategies as possible alternatives to represent choice making. In this thesis, LPAA is considered as a process strategy as it could certainly be a shortcut in processing when making a decision. Literature often treats it as an established and dominant decision process but, as will be seen later in this thesis, it could be extended in a number of directions. For example, it could represent a simplification of risk attitudes, perceptual conditioning, or other behavioural refinements. This section summarises a number of heuristics that represent the state of the art. They are divided into three major topics: (1) Context free heuristics – when valuing an alternative individuals will only consider the characteristics of it; (2) Local choice context dependent – when valuing an alternative individuals will also consider the characteristics of competing alternatives, and; (3) Choice set dependent heuristics – when valuing an alternative individuals will take into account past information (i.e., previous choice sets) they faced.

The following subsections explore each one of the three major topics mentioned above. Table 2-1 summarises all of these heuristics as a reference source. Table 2-2 presents which of the principal disciplines that investigate choice heuristics – transportation, marketing, environmental, and health - have included or discussed each specific heuristic. It shows that most of the heuristics have been discussed in marketing and transportation, with a few exceptions. Likewise, health studies also considered several of these heuristics, and environmental studies only consider five of them. The wide discussion of these heuristics in transportation studies strengthens the importance of estimating them simultaneously and studying their interaction in this area of applied research.

Table 2-1: Summary Table of the Analysed Heuristics

	HEURISTIC	REFERENCES	DESCRIPTION
CONTEXT FREE HEURISTICS	Satisficing	Simon (1955)	The individual chooses the first alternative whose attributes satisfy his payoff requirements.
	Elimination-By-Aspects (EBA)	Tversky (1972)	The individual ranks the attribute depending on the importance level. All the alternatives that fail to meet the threshold requirement for the most important attribute are eliminated. The process is repeated using the second most important attribute, and so on until only one alternative is left.
	Attribute non-attendance (ANA)	Hensher (2006); Hensher (2010); Scarpa et al. (2009); Scarpa et al. (2010)	To make a decision, the individual only evaluates a subset of the attributes, i.e., does not attend to certain attributes. The subset attended to can be inferred or stated.
	Prospect Theory	Kahneman and Tversky (1979)	The attribute levels are compared to a reference point, and are evaluated as gains or losses.
CONTEXT DEPENDENT HEURISTICS	Lexicography	Tversky (1972)	The individual makes a decision based on the level of only one attribute.
	Majority of Confirming Dimensions (MCD)	Russo and Dosher (1983); Hensher and Collins (2011)	The alternative chosen is the one that has the highest count of 'best' attribute levels compared to the other alternatives.
	Extremeness Aversion Heuristic	Tversky and Simonson (1993); Simonson and Tversky (1992)	The alternatives whose levels are extreme (too good or too bad) are avoided. Compromise Effect: An alternative whose attribute levels are 'in between' has a higher probability of being chosen. Polarisation Effect: There is a preference towards some 'extreme alternatives'. The extremeness aversion presents only towards a subset of the attributes.
	Random Regret Minimisation (RRM)	Chorus et al. (2008); Chorus (2010); Chorus (2012)	The utility functions (disutility) assigned to each alternative relies on regret, which is defined as the loss felt when a non-chosen alternatives performs better than the chosen one on certain attribute(s).
	Relative Advantage Maximisation (RAM)	Tversky and Simonson (1993)	The utility functions assigned to each alternative have an additional component (to the RUM model) which represents the relative advantage of the alternative compared to the others.
	Contextual Concavity Model (CCM)	Kivetz et al. (2004)	The utility function is the sum across attributes of concave functions of the relative advantage (or gain) of an attribute relative to the worst one.
	Extremeness Seeking Heuristic	Gourville and Soman (2007)	The alternatives whose levels are extreme (too good or too bad) have a higher probability of being chosen.
CHOICE SET INTERDEPENDENT HEURISTICS	Reference Point Heuristics		
	Reference Revision and Value Learning	DeShazo (2002); Hensher and Collins (2011); Bateman et al. (2008); McNair et al. (2011)	Over the course of an experiment, preferences are not stable and depend on starting point and attribute levels presented.
	Case Based Theory	Gilboa and Schmeidler (1995); Gilboa and Pazgal (2001)	Decision making under uncertainty is based on experiences on past cases and their outcomes
	Cost Expectations Model	Carson et al. (1994); Alberini et al. (2017)	Respondents assume that the price of the initial offer conveys information about the actual cost of the good, so any other price will violate their expectations.
	Strategic Misrepresentation (Strategic Responding)	McNair et al. (2012)	An individual may hide their true preferences, rejecting the preferred alternative within the choice task, because in a previous choice task he was presented with a better alternative. The theory relies on the individuals' hope that their preferred alternative is implemented.

HEURISTIC	REFERENCES	DESCRIPTION
‘Yea-Saying’ Model	Couch and Keniston (1960); Mitchell and Carson (1989); Cameron and Quiggin (1994)	When a respondent accepts the first question, he wants to be consistent in the follow-up questions and keeps accepting them
Attraction Effect	Huber et al. (1982)	When an asymmetrically dominated alternative is included in the choice set, the relative market share of the dominant alternative increases.
Similarity Effect	Tversky (1972)	When an alternative is included in a choice set, it tends to receive more market share from similar alternatives (decreasing their relative market share).

Table 2-2: Summary Table of Applied Studies of Heuristics Investigated in Different Disciplines

	HEURISTIC	MARKETING	TRANPORTATION	ENVIRONMENTAL	HEALTH
CONTEXT FREE HEURISTICS	Satisficing	X	X	-	X
	Elimination-By-Aspects (EBA)	X	X	-	X
	Attribute non-attendance (ANA)	X	X	X	X
	Prospect Theory	X	X	-	-
CONTEXT DEPENDENT HEURISTICS	Lexicography	X	X	X	X
	Majority of Confirming Dimensions (MCD)	X	X	-	
	Extremeness Aversion Heuristic	X	X	-	X
	Random Regret Minimisation (RRM)	X	X	X	X
	Relative Advantage Maximisation (RAM)	X	X	-	-
	Contextual Concavity Model (CCM)	X	X	-	-
	Extremeness Seeking Heuristic	-	X	-	-
CHOICE SET INTERDEPENDENT HEURISTICS	Reference Point Heuristics				
	Reference Revision and Value Learning	X	X	X	X
	Case Based Theory	X	-	-	-
	Cost Expectations Model	X	X	-	-
	Strategic Misrepresentation (Strategic Responding)	X	X	-	X
	‘Yea-Saying’ Model	X	X	X	-
	Attraction Effect	X	X	X	X
	Similarity Effect	X	-	-	-

Threshold and Cut-offs Concept

Before understanding in more depth the heuristics that will be included in this research, it is useful to understand the threshold concept since it will be referred to by several heuristics, such as satisficing, EBA and ANA. It is important to note that it is not a heuristic by itself but a concept relevant to many of them, and that is why it will be explained first. Thresholds and cut-offs define ranges of the attribute levels that establish if it is acceptable or non-acceptable. Many heuristics use this concept, stating that attitudes might change depending on the level of the attribute. The critical value(s) for which the attitude changes are determined by thresholds and cut-offs. Moreover, this critical value(s) can be assumed as equal for the entire population or to vary across individuals.

Swait (2001) proposes to incorporate attribute cut-offs when studying consumer behaviour through discrete choice modelling, using elicited cut-offs. He considers the restrictions added as ‘soft’, that is, a respondent can choose an alternative that violates some cut-offs if they believe the benefits are high enough. His results are very encouraging; however he notes the additional burden of having to collect additional information (i.e., asking the respondents to define the cut-offs used). Hensher and Rose (2012) implemented a similar method to condition the entire utility expression. Cantillo et al. (2004) incorporate thresholds in the perception of attribute level variations using discrete choice modelling. They discuss the importance of considering thresholds in the estimation process, since if not, the attributes’ estimates could be over and under estimated. For example, if under a certain level individuals do not value travel time as much and the modeller does not consider this threshold, the estimate for travel time will be overestimated for small values and underestimated for higher ones. They develop a choice model that includes random thresholds specified as particular analytical distributions (without elicited evidence) as minimum perceptible changes in attributes. Their results show that the benefits of including thresholds depend on the size of the change in the variable; when the changes increase, the benefits decrease.

There are several other papers that have dealt with the concept of thresholds or cut-offs when understanding decision-making (e.g., Hensher 2010a); the ones that considered them as part of a decision process strategy will be further analysed in the corresponding heuristic section below.

2.3. Context Free Heuristics

Context free heuristics refer to those which do not consider the influence that the other alternatives presented in the choice set may have over the assessment of one alternative. That is, the function used by individuals to assess each alternative is independent of the opposing alternatives' attribute levels.

2.3.1. Satisficing

This heuristic was initially proposed by Simon (1955), where an individual chooses the first alternative that satisfies his payoff requirements. He discusses the fact that this model does not guarantee a unique solution or even the existence of one. In traditional discrete choice models, all alternatives are evaluated and then a decision is made. However, in human decision-making, alternatives are often examined sequentially. Hence, to assure a unique solution, this heuristic states that the individual will choose the first alternative that satisfies their imposed conditions. Regarding the existence of a solution, the author proposes that when evaluating the alternatives sequentially, if the individual does not find a satisfactory alternative, he will change the pay-offs or conditions. Such changes will then guarantee the existence of a solution.

Later on, Grether and Wilde (1984), consider that individuals' environment influences the way in which they processes the attributes associated with alternatives. Therefore, based on the knowledge of respondents and experience, they assign a payoff or cut-off (i.e., threshold) function to each attribute.

Subsequently, this heuristic was studied by Gigerenzer and Todd (2008) together with other heuristics considered in the literature. Todd and Gigerenzer (2007) investigate the implications of this heuristic in an environmental study, and Chen and Sun (2003) include this heuristic when analysing the relationship between age and financial decisions. Their results show that older people have a higher tendency to use this decision process strategy.

2.3.2. Elimination by Aspects

Tversky (1972) develop a probabilistic theory of choice as a way of explaining possible uncertainty and inconsistencies in responses (when a person seems to violate one or more of the axioms of rational choice behaviour). Uncertainty and inconsistency have stimulated the development of two theoretical frameworks that treat choice as a probabilistic process; random utility models and constant utility models. The first one ascribes uncertainty and inconsistency in the determination of the attributes' values, and the second ones to the decision rule. Tversky (1972) propose a third theoretical framework called elimination models. This heuristic states that the importance of each alternatives' attributes depends on an attribute ranking, which itself is subject to the respondent cut-offs. The level of importance of each attribute can be determined probabilistically or deterministically. Respondents assign a cut-off or threshold to each attribute, which determines if they are acceptable for them. Then, the individual eliminates all the alternatives that fail to meet the threshold of that attribute. They then repeat this process using the second most important attribute, and so on until there is only one alternative left.

Many researchers have considered this heuristic as a useful tool when the number of attributes and alternatives presented in the choice task is too large, making it a complex experiment (e.g., Williams and Ortúzar 1982; Young 1984; Payne 1976; Cantillo et al. 2004). However one might argue that this reflects the way choices are made in real markets, and what one must recognise is that we do not know a great deal about which attributes matter to each respondent, and hence reducing the set prematurely (often occurs in stated choice studies) may deny the opportunity to recognise which attributes are *relevant*, given their levels, for each person. As Hensher (2006b) says, relevance is what matters and not the complexity of choice experiments, which is in the eyes of the decision maker.

Gigerenzer and Todd (2008) include this heuristic in their analysis of the most general and common heuristics used in decision making. Transportation studies such as Hess et al. (2012) studied different ways in which this heuristic can be included, revealing their influence and implications.

2.3.3. Attribute Non-Attendance

One of the most common heuristics used in transportation studies is referred to as attribute non-attendance (ANA), which suggests that individuals may not take into account every attribute when reaching a decision. This heuristic challenges the assumption that individuals are fully-compensatory, i.e., they assess all the attribute levels presented to them. This heuristic can be: context free (Section 2.3), where the individual only considers the characteristics of an alternative to decide which attributes he will attend to when evaluating it; or choice context dependent, where the individual considers the characteristics of the other alternatives when deciding which attributes he will attend to associated with an alternative. The beginning of this chapter showed some of the first papers that took the thresholds and cut-offs concept heuristic into consideration in the field of transportation. Hensher et al. (2005) were the first to recognise and account for ANA in choice analysis through supplementary questions on whether an attribute was ignored or not. The ANA heuristic can be identified through stated responses, or inferred analytically. In the former case, respondents explicitly state which of the attributes they did not attend to; in the inferred case non-attendance is inferred through an appropriately specified model. Puckett and Hensher (2009) studied whether asking respondents to determine which attributes they did not attend to *after* each choice task or after the whole experiment influence the model parameter estimates. Their results showed that respondents do not always ignore the same attributes in every choice task; thus, it is better to ask them after each choice set (which can account for the attribute levels in each choice scenario).

Essentially, inferred ANA can be considered as part of the ‘context free’ heuristics or part of the ‘local choice context dependent’ heuristics; it depends on the definition used by the modeller. For example, it is considered a ‘context free’ heuristic if the modeller defines ANA towards all the attributes other than costs and travel times (e.g., frequency). In this case, the utility function of each alternative would be independent of the characteristics of its competing alternatives. However, where ANA is defined as being influenced by the choice task structure (e.g., number of alternatives), there would be a relationship between the alternatives so it would be a ‘context dependent’ heuristic (Hensher, 2006b; Collins and Hensher, 2015). This form is included in the ‘context free’ section because most of the studies use this definition. When using stated ANA, it would normally be considered as a ‘context free’ heuristic.

Hess and Rose (2007) use an SP dataset for route choice in Australia and propose a methodology for inferring ANA using latent classes (see Section 3.5.2 for more details), such that the decision process strategies need not be stated by respondents. They present this approach by criticising the method of asking respondents to state the ANA they used directly. Their results show a significant difference between the stated ANA and the inferred one.

Hensher (2010b) studied inferred non-attendance using a dataset collected for a study of car driving commuters in Sydney in 2004. His model allowed for attribute non-attendance, without having to ask respondents to define which attributes they did not attend. His objective was to analyse the impact on the willingness to pay estimates (WTP) for travel time savings, and the results showed that when including attribute processing rules the estimates for travel time savings have a higher mean.

Hensher (2006b) was also the first to study how the design of an experiment (known as 'design of designs') might influence willingness to pay estimates through the influence of ANA, i.e., considering 'context dependence'. Later on, Weller et al. (2014) compare inferred and stated responses and analysed the influence that the same design may have on the ANA heuristic. Their results show that a higher number of alternatives increased the probability of non-attendance, both for stated and inferred specifications. Specifically, they found that for stated non-attendance a higher number of levels for the cost attribute increased the probability of non-attending it; and for inferred non-attendance, a higher number of alternatives increased the probability of not-attending to the cost attribute. Likewise, Collins and Hensher (2015) studied the influence of the experiment design on inferred attribute non-attendance using a latent class random parameter approach (more details of this approach can be found on Section 3.5.2). They also confirmed the finding that when the number of attribute levels is larger, there seems to be more ANA, and some attributes are more likely to be not attended. Also, their findings suggest the WTP estimates significantly change when considering stated vs inferred ANA.

Recent studies, such as Campbell et al. (2012; 2014) consider possible sub-sets of the choice task in a study on fish and chicken markets. They consider four possible choice sets: (1) the complete choice task, (2) all alternatives except those that have a cost above a certain value $XX, (3) all alternatives except those that have a cost above another given value $X, and so on. Campbell et al. (2014) show that even though a small number of respondents did not attend all the alternatives presented, this had a major impact on the WTP estimates. Campbell et al. (2012) show that these non-attendance heuristics are significant in the model estimates, and improved the statistical fit and associated understanding of individual behaviour. These findings suggest that controlling for the design of a choice experiment in modelling choices has great merit. In practice this is not always possible since fixed designs are used to consider a constant context. However, the findings reported in this study can be used as evidence to guide the design of a choice experiment and be explicit about the likely implication on WTP.

All of the studies mentioned above were estimated using maximum log-likelihood (refer to Section 3.2 for more details). There are other studies that have considered Bayesian Estimation (the interested reader is referred to Rossi et al. 2006 and Train 2009 for more details on this approach) to consider possible ANA. George and McCulloch (1993) proposed this methodology for attribute selection, and they referred to it as the Stochastic Search Variable Selection. Through the hierarchical Bayes model, they identified the attributes that should be included and excluded from the model. Gilbride et al. (2006) introduce a model that incorporates heterogeneous selection in which respondents differ in the attributes they take into account. Their results on an empirical application in a brand choice study shows that including ANA improves inference and predictive accuracy.

Scarpa et al. (2009) refer to the Bayesian approach for attribute selection as Stochastic Attribute Selection (STAS). In this study they use a STAS and a latent class approach to include non-attendance. Their objective was not to see which approach was superior, but to show the relevance of non-attendance. Both of the approaches revealed a significant improvement when taking into account non-attendance. The non-attendance frequencies did not change considerably within both approaches.

Many studies in different disciplines have demonstrated that not all respondents attend to all attributes. These include transportation (Hensher et al. 2005; Hensher 2006b; Hensher and Greene 2010; Scarpa et al. 2009; Zellman et al. 2010), marketing (Swait 2001; Fasolo et al. 2007; Islam et al. 2007; Moser and Raffaelli 2014) and health (McIntosh and Ryan 2002; Lancsar and Louviere 2006). There are also a few exceptions which demonstrate the contrary (for example, Rigby and Burton 2006; Campbell et al. 2008; Carlsson et al. 2008).

Furthermore, there are studies which have focused on the risk of confounding ANA and preference heterogeneity. Hess et al. (2013) argue that when using a latent class approach (for more information see Section 3.5.2) to represent ANA the results might be misguided by representing, in part, regular taste heterogeneity. Their first proposed model estimates the class-specific parameters to avoid the confounding between ANA and regular taste heterogeneity. Their second proposed model combines latent class with mixed logit model (LC-MMNL) which allows for ANA and for taste heterogeneity in the form of random parameters. The results of both models suggest that the implied rate of ANA might be much lower than what is suggested by a traditional model. Their findings reveal a potential confounding between process strategies, especifically ANA, and taste heterogeneity, which might have a significant influence on the results. Collins et al. (2013) propose a similar model but with a more general definition of ANA, called random parameter attribute non-attendance model (RPANA). Their results show significant differences in the ANA rates as the distribution of the random parameters varied, which suggests a possible confounding between ANA and taste heterogeneity.

2.3.4. Prospect Theory

Prospect Theory (PT) was developed by Kahneman and Tversky (1979) as an alternative to Expected Utility Theory (EUT). EUT states that decisions made under risk are valued non-linearly as the sum of all the possible outcomes multiplied by their probabilities of occurrence (the functional forms will be described in Section 2.7). They critiqued the theory of EUT on decisions made under risk by describing three empirical effects which seem to violate expected utility theory (EUT): (1) the Reflection Effect; (2) Probabilistic Insurance; and (3) the Isolation Effect.

- The *Reflection Effect* considers the situation where an outcome is replaced by gains to losses (i.e., has a negative prospect). They demonstrate through a study that certainty increases the averseness of losses as well as the desirability of gains, which contradicts EUT principles.
- *Probabilistic Insurance* refers to insurance policies, where people spend money to obtain insurance against both large and small losses. In case of an accident, a regular insurance covers all damages or a part of them. However, probabilistic insurance offers several alternatives each one up to a certain probability, e.g., if the accident occurs on an odd day of the month there is no cover, otherwise it covers everything. EUT implies that probabilistic insurance is superior to regular insurance. However, their results show a higher degree of risk aversion towards probabilistic insurance.
- The *Isolation Effect* refers to the situation where a person has to choose between different alternatives and will focus his attention on those attributes that differentiate the alternatives, while disregarding common characteristics. This may generate preference inconsistencies since the decomposition of the alternative components into common or dissimilar may be done differently between individuals, and this might lead to different preferences. This is what they refer to as the Isolation effect. They provide an example where there are two scenarios: one risky and one riskless. The certainty on the outcome enhances the attractiveness of this scenario, relative to a risky

scenario with the same probability and outcome. PT states that respondents evaluate attribute levels relative to a reference point, evaluating each attribute level as a gain or a loss using a weighting probability.

PT proposes that the value function of an attribute is defined relative to a reference point, and that generally gains will have a concave function while losses a convex one. Moreover, the value function for losses will be steeper than for gains, which translates into losses being weighted more heavily than gains in the utility function. The value function proposed is shown in Figure 2-1.

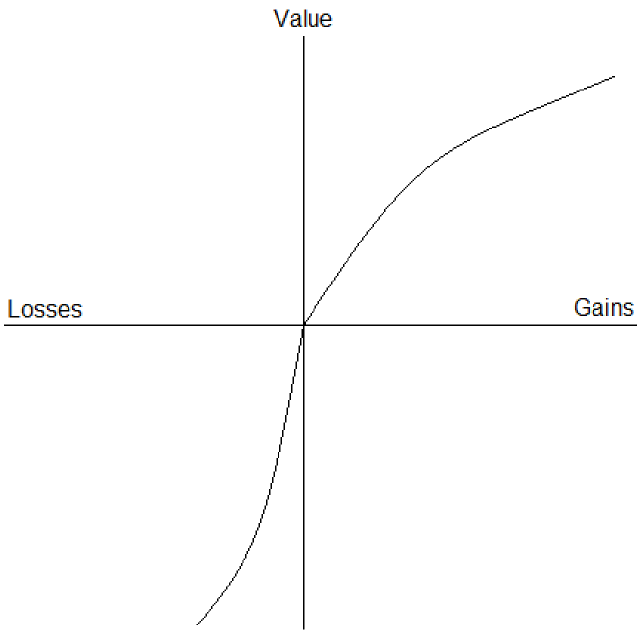


Figure 2-1: Hypothetical Value Function of Gains and Losses. Source: Kahneman and Tversky (1979).

That is, PT states that, in risky situations where each outcome is presented up to a certain probability, the utility function will be as follows:

$$U = w(p_1) \cdot u(x_1) + w(p_2) \cdot u(x_2) + ... + w(p_n) \cdot u(x_n) \qquad (2.1)$$

where $u(x_n)$ is the value function of attribute x_n; and $w(p_n)$ represents a weighting function of the probability of occurrence p_n.

Quiggin (1982) detected a limitation of the PT approach in its inability to assign a separate weighting probability. If two different outcomes have the same probability of occurrence, their transformed probability would also be the same, regardless of the nature of the outcome. This led to the development of Cumulative Prospect Theory (CPT), which considers a weighting probability that is defined over the cumulative probability distribution, and not over the probability by itself. To define the cumulative function, it is necessary to order the possible outcomes x_n (with its corresponding probabilities p_n) from worst to best: $x_1, x_2, x_3, ..., x_n$, where $x_{k-1} < x_k$ for k=2,…,n.

The anticipated utility function, as they define it, is as follows:

$$V = h(p) \cdot U(x) = \sum_n h_n(p) \cdot U(x_n) \tag{2.2}$$

where $h(p)$ is the vector of decision weights; $h_n(p)$ depends on all the probabilities and not only on p_n, where $\sum_n h_n(p) = 1$. If $p_k = p_m$ it does not mean that $h_k(p) = h_m(p)$ (for k=1,2,…,n and m=1,2,…,n). However, if $p_k = 0$ then $h_k(p) = 0$; and for simplicity of scaling if $p_k = 1$ then $h_k(p) = 1$. They also assume that if there are only two possible outcomes, if $p_k = 1/2$ then $h_k(p) = 1/2$, that is, if the probabilities are 50%-50% they will not be distorted. The rest of the decision weights are defined as follows:

- If the outcome x_k is a positive one then the decision weight $h_k(p)$ is the difference between the outcome being at least as good as x_k and of it being strictly better than x_k:

$$h_k(p)^+ = w(p_k + p_{k+1} + ... + p_n) - w(p_{k+1} + ... + p_n) \quad \text{for } k = 1, 2, ..., \text{n} - 1 \tag{2.3}$$

- If the outcome x_k is a negative one then the decision weight $h_k(p)$ is the difference between the outcome being at least as bad as x_k and of it being strictly worse than x_k:

$$h_k(p)^- = w(p_1 + ... + p_{k-1} + p_k) - w(p_1 + ... + p_{k-1}) \quad \text{for } k = 2, ..., \text{n} \tag{2.4}$$

Senbil and Kitamura (2004) apply PT and CPT in decision making using empirical data. The dataset was collected in Otsu City in Shiga Prefecture, Japan in 2002, where respondents were asked to record the departure and arrival history for three consecutive commute days. Their model incorporates a value function over the departure time decision. Their results show that commuters evaluate the outcomes very differently, and the reference point indeed determines how they are evaluated.

Xu et al. (2011) study CPT in route choice behaviour where the travel times are made under uncertainty, where each outcome was presented with a probability of occurrence. They define the reference point as the effective reserved time, which is the period used by commuters to arrive at their destination and achieve an effective trip. This measure is objective and can be calculated for any trip. However, as the origin-destination (O-D) matrix is so large, a uniform method is required to define the reserved time of generic trips. One of the highlights of their study is that CPT is generalizable to route choice modelling, and their results show that the CPT based framework is more appropriate than one based in EUT.

Nicolau (2011) study PT in airline demand, in the context of price sensitivity among low cost, regular and charter airlines. Their results show significant differences between the reference point and the actual prices, revealing the importance of reference dependence. Their results align with PT principles regarding gains and losses, and showed that people react more strongly to price increases than to decreases.

Several other studies have incorporated PT and CPT heuristics in discrete choice modelling contexts primarily to consider decisions made under risk. The results are encouraging and reveal the importance of considering decision weights and weighting functions in different contexts (e.g., Bell and Lattin 2000; Li and Hensher 2011; Hensher and Li 2012; Rasouli and Timmermans 2014).

2.4. Local Choice Context Dependent Heuristics

This section will summarise the current literature on local choice context dependent heuristic. These heuristics consider that the utility function of a certain alternative not only takes into account the characteristics of that alternative, but also the characteristics of its competing alternatives.

2.4.1. *Lexicography*

This heuristic was introduced by Tversky (1972), where an individual selects a most important attribute based on his experience and/or knowledge, and makes a decision based solely on that attribute. That is, an alternative is chosen if it has the 'best' level on that attribute. If two alternatives have the 'best' level in the choice set, then these two are compared based on the second most important attribute. This heuristic has been referenced as a processing rule when there are many attributes and alternatives, as a way of simplifying the choice process (Payne, 1976).

Gigerenzer and Todd (2008) suggest a variant called 'one-reason decision making', where individuals reach a decision using only one reason such as only comparing one attribute (e.g., travel time), which is the lexicographic heuristic. Hess et al. (2012) estimate a lexicographic model in a transportation study involving rail travel behaviour. They use a latent class model (which will be referred to in Section 3.5.2) to allow for taste heterogeneity with three classes: the first class represents a LPAA (traditional fully compensatory) heuristic; the second one represents lexicographic behaviour towards the travel time; and the third one lexicographic behaviour towards the travel costs. They estimate a latent class multinomial mixed logit model (LC-MMNL) using the same classes as before but considering random parameters in the first class. Their results show that the probability of belonging to the first class is the highest, followed by the second one. They find that lexicographic inclusion improves the overall model goodness to fit. Their results also show there is a reduction on the random taste heterogeneity (in class 1) when considering more process heuristics, as opposed to when considering only the traditional LPAA heuristic in a mixed logit model. Their findings suggest a possible confounding between process heuristics and taste heterogeneity through random parameters, which will be further analysed in Section 2.7.

2.4.2. *Majority of Confirming Dimensions*

The *majority of confirming dimensions* (MCD) heuristic considers that individuals assess each alternative through how many 'best' attribute levels exist. For each attribute, a respondent looks for the alternative that has the 'best' level and marks that one as the best performing attribute level. The alternative that has the larger number of best performing attribute levels is the chosen one. Russo and Dosher (1983) conducted a study where respondents were asked to determine the decision process strategy used. Their eye movements were recorded to compare these to the stated strategy used. They propose this heuristic as a form of reducing the processing effort required by the choice task. Thus, instead of comparing the attribute level values, they assign a value -1 to +1 to the 'worst' and 'best' performing attribute level. Russo and Dosher (1983)'s results showed that this criteria was not always correct, and almost half of the times it was wrong, which could be suggesting process heterogeneity. They mention that this could have been caused by small differences among the attribute levels.

Hensher and Collins (2011) include this heuristic by identifying the number of best performing attribute levels as an additional component in the utility function. The estimated parameter for this component was highly significant and positive, suggesting that when the number of best performing attribute levels is higher, the probability of that alternative being chosen increases. Hensher and Collins (2011) asked respondents to state which attributes they did not attend to. These responses enabled them to test the *majority of confirming dimensions* heuristic using only the attributes that were attended to, which improved the model performance even more. In their conclusions, the authors recommend further inquiry into underlying sources of process heterogeneity directly in the utility expressions that represent individual preferences.

2.4.3. Extremeness Aversion Heuristics

Simonson and Tversky (1992) propose two hypotheses of how context might influence respondent decisions. The first hypothesis states that the attractiveness of an alternative depends on whether the trade-offs within the choice set are favourable towards that alternative. The second hypothesis, which they refer to as *extremeness aversion*, states that an alternative is more attractive if it is an intermediate option within the choice set. Hence, the extreme options are less attractive to respondents. They define the extreme alternatives within a choice set as the ones that perform best on certain attributes, but worse on others. Hence, the extreme alternatives are the ones that have high advantages and high disadvantages relative to each other, and have small advantages and small disadvantages relative to the intermediate alternative. The intermediate alternatives are the ones that have small advantages and small disadvantages relative to the extremes (there are two extreme alternatives).

Simonson and Tversky (1992) developed the *extremeness aversion* heuristic within the framework of loss aversion, according to which individuals assign a higher weight to losses than they do to gains (Tversky and Kahneman 1991). The losses or gains of an alternative are measured using a reference point, which in this case is considered to be the other alternatives presented in the choice set. Their proposition on how the context influences individual behaviour extends the notion that disadvantages are more heavily weighted than advantages. Therefore, this heuristic considers that respondents tend to avoid the alternatives with extreme levels, which can have two forms: (1) a compromise effect (CE), or (2) a polarisation effect (PE). CE suggests that extremeness aversion holds for every attribute; thus, the alternative that has a higher probability of being chosen is the one whose every attribute has an intermediate level relative to the other alternatives. PE relaxes the condition on certain attributes; that is, individuals have an extremeness aversion towards a subset of the attributes that describe the alternatives.

Tversky and Simonson (1993) describe an initial situation with only two alternatives. When a third alternative is included, one of the existing alternatives becomes the intermediate alternative, and hence the probability of choosing the intermediate alternative increases. This heuristic violates the rational principle of regularity proposed by Luce (1959) which states that the probability of choosing an alternative should never increase with the addition of a new alternative. However, in such a case individual behaviour might not necessarily equate with a preference regarding usage, but instead be conditioned by social elements (e.g., voting strategically to deny an outcome they do not like).

Leong and Hensher (2012) discuss the concept of extremeness aversion within a discrete choice modelling context, stating that it is highly feasible to test this heuristic with datasets which do not include additional questions (e.g., questions regarding their decision process strategies), but that there is still lack of literature on this topic. Sharpe et al. (2008) investigated how the compromise effect influences consumers' decisions on soft drinks; their results show the existence of extremeness aversion, increasing the market share of intermediate size drinks when smaller or larger drink size alternatives are introduced in the choice set. Chernev (2004) studied the extremeness aversion heuristic and developed an attribute-balance hypothesis which assumed that extremeness aversion is also influenced by the dispersion of the attribute levels within the choice set. That is, if two alternatives have similar attribute levels, none of this will be considered as the intermediate alternative regarding this attribute.

Several theoretical models were developed to accommodate the compromise and polarisation effect, such as: Contextual Concavity Model, Relative Advantage Maximisation, and Random Regret Minimisation.

2.4.4. *Random Regret Minimisation*

Chorus et al. (2008) and Chorus (2010) develop an alternative discrete choice modelling approach to the random utility maximisation model (RUM), called random regret minimisation (RRM). This model considers that individuals make their decision to avoid negative emotions. They define regret as the loss felt when a non-chosen alternative performs better than the chosen one on certain attributes. It corresponds to the sum of the regret felt for all the alternatives attributes. Chorus (2012) develops this theoretical framework further to study risky situations, where risk is defined as the probability of occurrence of the different possible states of the world. The author includes risky situations in the RRM model by considering that the level of regret of an individual varies according to the state, where each state has a probability of occurring. Hence, there may be varied levels of regret up to different probabilities.

The classical regret function (Chorus et al., 2008) between an alternative i and its competitor alternatives j, for all attributes m, was defined as follows:

$$R(\mathrm{i},\mathrm{j}) = \sum_{m} \max\left\{0, \theta_m \cdot (\mathrm{x}_{\mathrm{jm}} - \mathrm{x}_{\mathrm{im}})\right\} \tag{2.5}$$

The total regret for alternative i was equal to the maximum regret between alternative i and all of its competitors alternatives j, as follows:

$$RR_i = \max\left\{R(i,1),...,R(i,j)\right\} \tag{2.6}$$

The extended regret function (Chorus, 2010) between an alternative i and its competitor alternatives j, for all attributes m considers a Logsum function to smoothen the binary regret (i.e., avoid the discontinuity in 0 of the classical RRM). It can be written as follows (for a certain individual n):

$$RR_{in}^{RRM} = \sum_{j \neq i} \sum_{m} \ln(1 + \exp(\theta_m \left[x_{jmn} - x_{imn} \right])) \tag{2.7}$$

The error term in the classical and extended RRM are assumed distributed i.i.d. of extreme value type I with a variance of $\pi^2 / 6$ (McFadden, 1973), so the choice probabilities are equivalent to the logit model. Since regret has to be minimised, it would be the same as maximising the negative or the regret as follows:

$$P_{in} = \frac{\exp^{-RR_{in}}}{\sum_{j} \exp^{-RR_{jn}}} \tag{2.8}$$

Several transportation studies have adopted the random regret minimisation model, but when comparing the results with the ones obtained from a traditional RUM model, the differences on prediction performance are not highly significant (Rasouli and Timmermans 2014; Leong and Hensher 2015). As stated in Chorus *et al.* (2008) the RRM model was proposed as a complimentary tool to the traditional RUM model for the modeller to consider, and its aim is not to replace it. Hensher et al. (2015) support this idea by saying that using the RRM model with the RUM model simultaneously, the researcher or policy maker may obtain a broader perspective of choice behaviour.

van Cranenburgh et al. (2015) propose a new RRM model, where there is an added parameter that estimates the shape of the attribute level regret function, μ . This parameter informs on the profundity of regret, i.e., on the degree of regret minimization behaviour imposed by the model. If μ is relatively large, it means a mild profundity of regret, while a low μ represents a strong profundity of regret. The regret function for this new μRRM model can be written as follows (the choice probability function remains unchanged):

$$RR_{in}^{\mu RRM} = \sum_{j \neq i} \sum_{m} \mu_{\mathrm{m}} \cdot \ln(1 + \exp(\frac{\theta_m}{\mu_{\mathrm{m}}} \left[x_{jmn} - x_{imn} \right])) + \varepsilon_{in} \tag{2.9}$$

They re-analyse ten datasets that had been previously used to compare the RUM model with the classical RRM model. Their results suggest that the degree of regret minimising behaviour imposed by the classical RRM model is very limited and could explain the relatively low differences between the RUM and classic RRM models’ performance. In four out of the ten datasets analysed, the μRRM model significantly improves the model fit as compared to the RUM.

2.4.5. *Relative Advantage Maximisation*

The relative advantage maximisation (RAM) model was introduced by Tversky and Simonson (1993) to consider how individuals compare the attribute levels across alternatives taking into account the compromise and polarisation effect, which they referred to as the 'componential context model'. They suggest that each attribute level is an advantage or disadvantage relative to the other alternatives, and therefore, the utility function for each alternative is the sum of its advantages and disadvantages. Tversky and Simonson (1993) define a model with two components, one that is a combination of the "intrinsic" weight of that attribute and the weight induced by the background beta (what is normally included in a LPAA), and the effect of a choice set.

The utility function (deterministic part) for alternative i over a subset S of alternatives, considering $v_{im}(x_{im})$ as the utility of attribute m of alternative i, is as follows:

$$V(i,\mathrm{S}) = \sum_m \theta_m \cdot v_{im}(x_{im}) \; + \theta_{m+1} \cdot g(i,S) \tag{2.10}$$

As can be seen, the first part of the equation is as traditionally studied in discrete choice models, and the second one depends on the choice set. They define the advantage of an alternative i over j for all attributes m, can be written as follows:

$$A(i,\mathrm{j}) = \begin{cases} \sum_m v_{im}(x_{im}) - v_{jm}(x_{jm}) & \text{if } \; v_{im}(x_{im}) > v_{jm}(x_{jm}) \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

And the disadvantage would be

$$D(i,\mathrm{j}) = \begin{cases} \sum_m v_{jm}(x_{jm}) - v_{im}(x_{im}) & \text{if } \; v_{im}(x_{im}) < v_{jm}(x_{jm}) \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

The relative advantage of i over j, denoted as $R(i,\mathrm{j})$, would be:

$$R(i,\mathrm{j}) = \frac{A(i,\mathrm{j})}{A(i,\mathrm{j}) + D(i,\mathrm{j})} \tag{2.13}$$

Tversky and Simonson (1993) do not introduce a local context (choice set) component $g(i,S)$ if there are only two alternatives:

$$\mathrm{g}(i,\mathrm{S}) = \begin{cases} \sum_{j \in S} R(i,j) & \text{if } S > 2 \\ 0 & \text{if } S \leq 2 \end{cases} \tag{2.14}$$

Resulting on the following utility function:

$$V(i,\mathrm{S}) = \sum_{m} \theta_m \cdot v_{im}(x_{im}) + \theta_{m+1} \cdot \sum_{j \in S} R(i,j) \tag{2.15}$$

Kivetz et al. (2004) are the authors that first refer to this model formulation as the 'relative advantage model'. Their starting point is the Tversky and Simonson (1993) model described above, but they define the disadvantage of an alternative i with respect to alternative j on each attribute m as an increasing convex function of the corresponding advantage of j relative to i. That is, using the following function:

$$D(i,\mathrm{j}) = \begin{cases} \sum_{m} \left[A_m(j,\mathrm{i}) + L_m \times A_m(j,i)^{\psi_m} \right] & \text{if } v_{im}(x_{im}) < v_{jm}(x_{jm}) \\ 0 & \text{otherwise} \end{cases} \tag{2.16}$$

where L_m are the loss aversion parameters and ψ_m the power parameters. The disadvantages are a convex function of the corresponding advantages. Without convexity in the disadvantage function, the relative advantages of any set of options with equal context-independent utility would be identical despite considering loss aversion ($L_m \neq 0$). Besides the disadvantage (Equation (6) is transformed into Equation (10)), the rest of the formulations are equivalent to the model proposed by Tversky and Simonson (1993).

Leong and Hensher (2014) and Leong and Hensher (2015) propose a new version of the RAM model, which has some desirable properties from the classical RRM functional forms, and has the symmetry between advantage and disadvantage as proposed by Tversky and Simonson (1993). They define the disadvantage using the RRM formulation as follows:

$$D(i,\mathrm{j}) = \sum_{m} \ln(1 + \exp(\theta_{jm} \cdot x_{jm} - \theta_{im} \cdot x_{im})) \tag{2.17}$$

With the asymmetry condition between advantages and disadvantages:

$$A(\mathrm{j},\mathrm{i}) = D(i,\mathrm{j}) = \sum_{m} \ln(1+\exp(\theta_{jm}\cdot x_{jm} - \theta_{im}\cdot x_{im})) \tag{2.18}$$

where the utility function equals:

$$\mathrm{V_i} = \sum_{m} \theta_m \cdot v_{im}(x_{im}) + \sum_{j\in S} R(i,\, j) \tag{2.19}$$

Leong and Hensher (2014) were the first to introduce RAM in a binary choice data setting. As was initially proposed by Tversky and Simonson (1993), RAM required at least three alternatives. However, Leong and Hensher (2014) show that, oppositely to the RRM model, the RAM model can still take into account the relative advantage/regret effects in binary choice data.

Leong and Hensher (2015) results showed, that the differences in these model fits were quite small. However, they state that this decision process requires further investigation and has the real potential to improve the performance of choice models.

2.4.6. Contextual Concavity Model

Kivetz et al. (2004) propose a model to capture the compromise effect called the ‘contextual concavity model’ (CCM). They suggest that the compromise effect can be included in a model by combining the notions of concavity and context dependence. Specifically, the deterministic component of the utility of alternative i equals the sum across attributes m of concave functions of the gains between this alternative and the one with the worst attribute level. Considering $v_{im}(x_{im})$ as the utility of attribute m of alternative i, S the total set of attributes, and $\mathrm{c_m}$ the concavity parameter of attribute m, the overall utility function of alternative i would be equal to:

$$U_i^{S} = \sum_{m} (v_{im}(x_{im}) - v^{S}_{\min,\mathrm{m}}(x^{S}_{\min,\mathrm{m}}))^{\mathrm{c_m}} + \varepsilon_i \tag{2.20}$$

If we assume that the error ε_i is distributed i.i.d. of the extreme value type I, the choice probabilities will follow the multinomial logit model (McFadden, 1973). If all the $C_m = 1$ then the model would collapse to a RUM multinomial logit model.

Kivetz et al. (2004) also defined a 'normalized contextual concavity model' (NCCM), where they normalize the concave gain of each attribute by the attribute's weight. That is, considering the relative position of each attribute. As in the CCM, the choice probabilities will follow the multinomial logit model. The equation can be written as follows:

$$U_i^S = \sum_m \left[(v_{im}(x_{im}) - v^S_{\mathrm{min,m}}(x^S_{\mathrm{min,m}})) \cdot \left(\frac{(v_{im}(x_{im}) - v^S_{\mathrm{min,m}}(x^S_{\mathrm{min,m}}))}{(v_{\mathrm{max},m}(x_{im}) - v^S_{\mathrm{min,m}}(x^S_{\mathrm{min,m}}))} \right)^{C_m} \right] + \varepsilon_i \quad (2.21)$$

Hensher et al. (2017) estimate a model that jointly account for extremeness aversion in the type of CCM and a traditional LPAA model adding behavioural refinements. They estimate a probability of relevance of each process rule, and their results show that the CCM extremeness aversion heuristic has, on average, a 0.63 probability of relevance compared to a 0.27 probability of relevance of the other process rule. Hence, they show that CCM extremeness aversion is a very appealing process rule that handles degrees of attribute risk.

2.4.7. Extremeness Seeking Heuristic

Gourville and Soman (2007) propose that the attitude towards extremeness depends on the type of choice sets structured in an experiment. Their theory, however, contradicts the rational principle of regularity proposed by Luce (1959). They propose that individuals will behave as extremeness seekers in a scenario where the alternatives have multiple and non-compensatory dimensions, i.e. they are described by non-alignable attributes. That is, one alternative is described by one set of attributes (attractive or unattractive), while another alternative is described by a different set of attributes (Gourville and Soman 2005). Hence, the alternative's attributes cannot be compared directly. The alternatives are defined as extreme or intermediate depending on their characteristics. For example, the extreme alternatives could be a basic one (less expensive) and a fully loaded alternative (most expensive), and the choice set may contain other intermediate alternatives. Each one is described by a different set of attributes. In this type of choice set, Gourville and Soman (2007) propose that extreme alternatives (basic and fully loaded alternatives) will have a higher probability of being chosen.

Gourville and Soman (2007) provide a choice set example that contains three products within the same brand; one is a basic alternative, another is a fully-loaded alternative, and there are one or more intermediate alternatives. In this situation, they say that if the number of intermediate alternatives increases, customers will tend to move towards the extremes.

To confirm their hypothesis, Gourville and Soman (2007) present three studies. The first two consist of making respondents choose between laptops. The first study presents alternatives whose attributes are non-alignable. As they expected, the results show that when adding more alternatives, the probability of choosing the extreme one increases. The second study collected two types of data: one presented non-alignable attributes, and the other alignable ones. Their results show that in the first type of survey, the behaviour is extremeness seeking, and in the second one extremeness averse. The third study presented four different markets to study the robustness of their theory: cable television service, digital cameras, wireless telephone service, and Italian vacations. All of the choice sets contained alternatives described by non-alignable attributes. Their results on the three different markets confirmed the theory of extremeness aversion behaviour.

2.5. Choice Set Interdependent Heuristics

These type of heuristics propose that one choice set might influence decisions made in future choice sets. Different studies have addressed this topic suggesting that the principles of response independence, response veracity and preference stability are sometimes violated when discrete choice questions are asked sequentially. Hanemann (1991) showed statistically that the estimated parameters could not be considered common across two sequential discrete questions. Cameron and Quiggin (1994) show that the estimates and willingness to pay (WTP) for an initial question are significantly different than for the follow-up question, but they do not provide a unique explanation for this phenomena. In this context several studies have proposed decision rules to explain these anomalies and they will be discussed in the following sections.

2.5.1. Reference Points

Several studies have proposed, in one way or another, decision rules in which alternatives are evaluated relative to a reference point. However, the use and/or origin of the reference point changes across the proposed processing rules. For instance, the reference point might be used as a rational comparison to evaluate the attribute levels of the alternative presented (*Revision Heuristic and Value Learning*), might be past experiences that influence the expected outcome (*Case Based Decision Theory*), or might be considered as "real" market levels to evaluate the fairness of the alternatives (*Cost Expectations Model*). The following subsections will discuss each one with more detail.

2.5.2. Revision Heuristic and Value Learning

'Anchoring' or 'Starting Point Bias' states a concern that the starting price or initial offer in an experiment may serve as an anchor for respondents, making them believe that it provides the real value of the goods. There are a number of papers that provide evidence for this pattern of behaviour (Tversky and Kahneman, 1974 ;Thayer, 1981; Herriges and Shogren 1996; Ariely et al. 2003; Holmes and Boyle 2005; Carlsson and Martinsson 2006; Ladenburg and Olsen 2008). This concern has proven to be relevant and process rules have been developed to explain this bias, namely 'reference revision heuristic' and 'value learning'. These heuristics have a similar nature, as they both suggest the interdependence between choice sets through their characteristics. The main difference is that with value learning, preferences might change through the experiment, while in the reference revision heuristic preferences might be stable but the utility functions that individuals try to maximise change. Their differences are not very clear and, for example, Hensher and Collins (2011) state that the reference revision heuristic is a special case of value learning. McNair et al. (2011) include both the reference revision heuristic and value learning, but mention in their conclusions that their approach did not allow further discrimination between the heuristics (their approaches and results will be analysed in each of the sections below). Henceforth, the separation between these heuristics is not very clear as they share fundamental characteristics. The sections below analyse the heuristics as were referred to as in the studies.

'Reference Revision Heuristic'

The 'reference revision heuristic' (RRH) was initially proposed by DeShazo (2002), based on Prospect Theory which states that an individual faced with a choice set will form a reference point up to a probability (Kahneman and Tversky, 1979). This reference point is an increment of expected utility and the respondent will frame the second choice set as one with gains or losses relative to it. Essentially, this heuristic assumes that preferences are fixed, and only the reference point changes (as preference revision).

The experiment conducted in DeShazo (2002) refers to willingness to pay (WTP) questions, namely, if the person would accept or reject a certain good for an offered price. It considers that the respondent will have an initial reservation value (or reference) of *C*, and a subjective probability that the good will be provided of *p*. If the first choice set presents the good at a price \$B, the expected gain will be as follows:

$$E(utility)_{\text{initial}} = p \cdot (C - \$B) \tag{2.22}$$

Their theory proposes that if respondents accept the good at the offered price, they expect the exchange to be concluded so they will form a reference point equal to their updated expected gain (equation (2-22)). Conversely, if respondents reject the good at the offered price they will not form a reference point. If they accepted the good in the first question and the follow-up question offers a higher price \$A, they will frame it as a loss defined by:

$$Loss_{\text{follow-up}} = p \cdot (\$B - \$A) \tag{2.23}$$

And the expected gain for the question will be:

$$E(utility)_{\text{follow-up}} = p \cdot \left[(C - \$B) + (\$B - \$A)\right] \tag{2.24}$$

Alternatively, if the follow-up question offers a lower price, respondents will answer by optimising relative to their *ex ante* level of welfare rather than the induced reference point. If respondents had rejected the first question, they will not frame the follow-up question either as a loss or as a gain. Therefore, this model predicts a downward bias in the WTP when the questions are framed as ascending in terms of the offered price but not when they are framed as descending.

Day and Pinto (2010) study if a particular order of choice task influences decision making. That is, if the choice tasks presented later are worse than the ones presented before, they are worsening choice tasks, and they are improving choice tasks if they get better. Their experiment relates to valuation of health, where respondents were asked to imagine they had been diagnosed with a medical problem and they had to choose between two medical treatments:

(1) A treatment provided by the hospital which was free of charge but it led to feeling sick (mild or severely sick) for the first two months of treatment.

(2) To purchase (low, medium or high price) an alternative medicine from a pharmacy that would avoid feeling sick.

The results show that when the price of alternative 2 was always getting worse, the alternative was chosen with a significantly lower rate. However, the difference was insignificant when the price was always improving. In the case of alternative 1, when the level of sickness kept getting better (worse) the alternative was chosen at a higher (lower) rate. Therefore, the sequence used to show the alternatives did affect the results in all the cases except when the price was always improving.

McNair et al. (2011) study the reference point revision heuristic by testing the relationship between the relative positions of the choice task with the cost sensitivity. That is, they incorporate interactions between the cost variable and the variable (effects-coded) that represents the order in which choice tasks are presented: q_1 for the first position; q_2 for the second one; q_3 the third; q_4 the fourth that is represented when all the other equal -1. The cost sensitivity for choice task t can be written as follows:

$$CostSens_t = \frac{-\partial U}{\partial cost} = -(\theta_{cost} + \theta_{q_1 \times cost} \cdot q_{1t} + \theta_{q_2 \times cost} \cdot q_{2t} + \theta_{q_3 \times cost} \cdot q_{3t}) \tag{2.25}$$

Their results show that $\theta_{q_1 \times cost}$ is significantly higher than the others, which indicates that the cost sensitivity is significantly lower for the first question (WTP higher) relative to the later questions in the experiment. Their conclusion is that the order in which alternatives are presented does have an influence in the WTP. This paper combines results with a value learning heuristic which will be analysed below.

'Value Learning'

This heuristic determines a situation where the principle of preference stability is violated, and assumes that throughout the experiment individuals discover their preferences. This heuristic underlies a theory that individuals have weak preferences, which can be influenced by the alternatives shown to them. This was originally proposed by Plott (1996) as a discovery preference hypothesis (DPH) that argues that stable and theoretically consistent preferences are formed by practice and repetition. This was later analysed in a discrete choice experiment context by Bateman et al. (2006), Bateman et al. (2008), McNair et al. (2011), and Hensher and Collins (2011).

Bateman et al. (2006) state that repetition was the key for learning, and so the starting point bias, i.e., the comparison to the initial level, is reduced as the respondent is presented with more decisions or choice sets. The anchoring parameter (γ) proposed by Herriges and Shogren (1996) has an influence over the willingness to pay (WTP) estimate, and can be expressed as follows:

$$WTP_r = (1-\gamma)\cdot WTP_0 + \gamma \cdot b_1 \qquad (2.26)$$

where WTP_0 is the prior WTP; WTP_r is the revised WTP (considering the anchoring effect); and b_1 is the initial level. The level for the second response b_2 will be transformed into:

$$b_{2r} = (b_2 - \gamma \cdot b_1)/(1-\gamma) \qquad (2.27)$$

This allowed the authors to test whether there is a reduction in the initial anchoring effects, suggesting a DPH. Their results are very interesting as they show that while there is a significantly high anchoring in responses on the first choice set, there is almost a null anchoring effect on the final choice set.

McNair et al. (2011) test whether choices are affected when a respondent faces four choice tasks instead of one, and their focus is in the cost attribute. They incorporate interactions between the cost attribute and a variable (effects-coded) that represents the order in which the choice tasks were presented. This variable represents the relative position of the cost attribute shown in this choice task relative to all the ones presented up to this point in the experiment: (1) if it is both the maximum and minimum level shown so far (m_{11}); (2) if it is the minimum, but not the maximum level shown so far (m_{10}); (3) if it is the maximum, but not the minimum level shown so far (m_{01}); (4) if it is neither the minimum not maximum level presented so far (m_{00} , which is represented when all the other variables equal -1). These variables enable the researcher to study the relationship between the cost sensitivity and the relative positioning of the cost level. Hence, the cost sensitivity in choice task t for this model can be written as follows:

$$CostSens_t = \frac{-\partial U}{\partial cost} = -(\theta_{cost} + \theta_{m_{11}\times cost} \cdot m_{11t} + \theta_{m_{10}\times cost} \cdot m_{10t} + \theta_{m_{01}\times cost} \cdot m_{01t}) \qquad (2.28)$$

Their results show that $\theta_{m_{10}\times cost}$ is the highest, which implies that the cost sensitivity is lowest (and WTP highest) when the cost level shown is the minimum level shown in the experiment. The second highest is $\theta_{m_{11}\times cost}$, followed by the implicit value. Additionally, Day and Pinto (2010) estimate these interactions in a random parameter model, where they include the interaction variables as part of the heterogeneity (i.e., in the standard deviation). The cost sensitivity for individual q can be written as:

$$\begin{aligned} CostSens_{tq} = \frac{-\partial U}{\partial cost} &= -(\theta_{cost} + \theta_{m_{11}\times cost} \cdot m_{11t} + \theta_{m_{10}\times cost} \cdot m_{10t} + \theta_{m_{01}\times cost} \cdot m_{01t}) \\ &+ (\sigma_{cost} + \sigma_{m_{11}\times cost} \cdot m_{11t} + \sigma_{m_{10}\times cost} \cdot m_{10t} + \sigma_{m_{01}\times cost} \cdot m_{01t}) \cdot \eta_q \qquad (2.29) \end{aligned}$$

where η_q is normally distributed with mean 0 and variance 1. There is a significant increase in the variance for m_{10} relative to m_{11}, which suggests that most respondents answered differently when they were presented with a level that was minimum but not maximum, than if this level had been presented in the first choice task (neither maximum nor minimum).

Hensher and Collins (2011) study a revision of the reference alternative as a value learning heuristic. They create a dummy variable that is equal to 1 if the alternative was chosen in the previous choice set, and 0 otherwise. The experiment was structured with one alternative representing the status-quo situation for each individual (which was asked at the beginning of the experiment) and two other alternatives. The utility function (deterministic component) for alternative i, attributes m and choice task t was as follows:

$$V_{i,t} = \sum_m \theta_m \cdot v_{im}(x_{im}) + \theta_{ref} \cdot dummy_{i,t} \tag{2.30}$$

In the first choice task $dummy_{SQ,1} = 1$, and $dummy_{Alt1,1} = dummy_{Alt2,1} = 0$. If in the first choice task the individual chose an alternative other than the status-quo, for example alternative 1, then the 'reference' alternative was updated so that in the next choice set $dummy_{Alt2,2} = 1$ and $dummy_{SQ,2} = dummy_{Alt2,2} = 0$, and so on. Their results show that θ_{ref} was positive and highly significant, which implies that when a 'reference' alternative is revised, the utility of that alternative increases in the next choice set.

McNair et al. (2012) jointly estimate this heuristic along with another choice set dependent heuristic (strategic misinterpretation which will be explained below). They focus on the role of the cost levels in value learning since they are considered as one of the most influential in the value learning process. They specify the alternative-specific preference in a way that it varies with the average cost levels presented during the entire experiment (i.e., considers all the previous cost levels presented and the current one). The utility function for alternative i; individual q; and choice task t is as follows:

$$U_{iqt} = \theta_0 + \theta_1 \cdot x_{1,iqt} + \ldots + \theta_m \cdot x_{\mathrm{m},iqt} + \delta \cdot (z^0_{iqt} - \bar{z}) \tag{2.31}$$

where θ_0 is the alternative specific constant; θ_m is the parameter for attribute m; δ is the parameter that represents the influence of the average cost levels presented during the experiment; $z^0_{it,ALT}$ the average of cost levels observed up to and including the current choice set; and $\bar{z}$ the average cost levels in the sample (considering all respondents and choice tasks). This latter was included to normalise the average observed cost to ensure that its sample mean is zero given that in this particular study the value learning heuristic was included in a latent class structure and the other classes represented other heuristics. Their results show that indeed the current choices are influenced by the cost levels seen during the entire experiment. In the study, they tested a weighted form to assign a higher importance to the most recent observations, but it did not appear to be statistically significant.

2.5.3. Case Based Decision Theory

Gilboa and Schmeidler (1995) state that decision making under uncertainty is based on experiences in past cases which is referred as Case-Based Decision Theory (CBDT). Their theory says that respondents remember past choices and their outcomes, and these memories influence their current choices. Later on, they extend this framework to treat cases where the person experienced different but similar scenarios in the past. Moreover, Gilboa and Pazgal (2001) present a case-based decision theory form where preferences for each alternative are affected by memories of past experiences with similar choices, and they are stochastically updated when an alternative is chosen.

2.5.4. Cost Expectations Model

The '*Cost Expectations Model,'* states that respondents assume that the price of the initial offer conveys information about the actual cost of the good, so any other price will violate their expectations. In the case of higher prices in the follow-up questions, respondents might see it as an attempt by the government to acquire more money than is needed (Carson et al., 1994; Alberini et al., 2017).

2.5.5. Strategic Misinterpretation

This heuristic proposes a situation where the principle of preference veracity is violated, which assumes truthfulness in individual responses. Samuelson (1954) established that respondents may hide their true preferences if that would enable them to obtain a public good at a lower cost. Other studies have strengthened this theory in relation to stated choice surveys (e.g., Mirrlees 1971; Hurwicz 1972; Mitchell and Carson 1989; Carson and Groves 2007) in a public policy context. The heuristic considers a situation where the respondent chooses a status quo situation over an alternative that he actually prefers, because in a previous choice task he was presented with a similar good at a lower cost. Hensher (2010a) explains how the truthfulness in responses is more problematic in a public policy setting when considering contingent valuation, where respondents are asked to accept or reject an alternative, or to state their maximum willingness to pay. When using contingent valuation, WTP are usually exaggerated relative to choice experiments and for public goods allow more easily (that stated choice experiments) for strategic choosing. Recent literature argues that answering in a strategic way is not necessarily clouding real preferences, but might be necessary in stated choice experiments (Carson and Groves, 2011; Johnston et al., 2017). We suggest this issue is more problematic with public policy issues; whereas the thesis focusses on private (use-related) choices.

Bateman et al. (2008) establish a difference between strong and weak strategic misinterpretation (SM). In a strong SM, individuals will always prefer a *status quo* situation to a preferred alternative if they had been presented with a similar one, but with a lower cost, on a previous choice task. Weak SM suggests that when an individual encounters this situation, he will make a trade-off between rejecting the preferred alternative and obtaining a similar outcome at a lower cost.

The extant literature has proven the existence of response patterns associated with the strategic misinterpretation heuristic by itself (e.g., Carson et al. 2004; Carson et al. 2009); or by jointly estimating this heuristic together with the value learning heuristic (e.g., Hensher and Collins, 2011; McNair et al., 2012), which will be analysed in Section 2.6. In these latter papers, results show that considering the heuristics together improved the performance of the models.

2.5.6. ‘Yea-Saying Model’

The ‘*Yea-Saying Model*’, proposes that when a respondent accepts the first question, he wants to be consistent in the follow-up questions and keeps accepting them (Couch and Keniston, 1960; Mitchell and Carson, 1989; Cameron and Quiggin, 1994).

2.5.7. Attraction Effect

Huber et al. (1982) present a new effect that could violate the Luce (1959) rational principle of regularity. An asymmetrically dominated alternative is one that is weaker across all of its attributes than only one other alternative in the choice set, and this latter is called the dominant alternative. The authors say that when an asymmetrically dominated alternative is included in the choice set, then the relative probability of the dominant alternative being chosen increases. They collected student responses on choices among six product categories: cars, restaurants, beers, lotteries, film and television sets. Each choice set contained either two or three alternatives, where each was described by two attributes. The three alternative choice sets contained one asymmetrically dominated alternative, as was designed in the choice experiment. Each individual faced choice sets with two alternatives, and with three alternatives, which enabled the authors to compare market share changes within and across respondents. In both cases the results showed an increase in the relative market share of the dominant alternative when including an asymmetrically dominated one. Moreover, when comparing the choices within respondents, this increase was lower that when comparing the results across respondents.

2.5.8. Similarity Effect

Tversky (1972) develops a framework that relates the choice probability across similar alternatives. Similar alternatives are those which have relatively similar attribute levels in the choice set. He proposes that when a new alternative is included in the choice set, it receives more market share from similar alternatives (i.e., with similar characteristics). Therefore, the probability of being chosen for them is reduced.

2.6. Multiple Heuristics

The previous sections reviewed the main heuristics that have been studied in the literature. The evidence reviewed thus far suggests that the traditionally used LPAA heuristic is not always adequate, and alternative decision process strategies are sometimes better at describing decision-making. However, the majority of the literature on alternative heuristics has focused on explaining decision-making using only one decision process strategy. Some studies have studied the possibility that individuals might be using different heuristics in decision making by including multiple heuristics in the modelling, even though the literature on this topic is rather limited. Section 3.5 details the different approaches used to jointly include multiple heuristics; the current section will focus on the way the existing literature handles this topic. This section will only mention the approaches existent in the literature, which are (1) considering different heuristics directly in the utility function (e.g., as acting upon a subset of the attributes); (2) the probabilistic decision process approach (PDP), and the (3) heuristic weighting function (HWF). It is important to mention that this section will refer to the studies that consider more than one heuristic besides an LPAA heuristic. There are other studies that have considered one heuristic together with a LPAA heuristic (Swait and Adamowicz 2001; Hess et al. 2012; Campbell et al. 2012; Weller et al. 2014), but will not be explained in more detail below.

Hensher and Collins (2011) jointly study two heuristics (one context free and one local choice context dependent) - majority of confirming decisions (MCD) and stated attribute non-attendance (ANA). Stated ANA was subject to the response of each individual and the majority of confirming decisions considered an additional parameter that represented how many attributes of the alternative had the 'best' levels within the choice set. They estimated several models including the process heuristics directly in the utility function of a traditional LPAA: considering only ANA, only MCD; LPAA plus MCD; both of them; and none. Their results show that the including ANA and MCD improves the overall performance of the model - this model considers that both heuristics are used together when reaching a decision. The authors estimate another model using a PDP approach that suggests some individuals use only MCD to reach a decision and others use only ANA (up to a certain probability). The results suggest that more than 80% of individuals use ANA while the rest use MCD. Regarding the approaches used, the PDP model had a significantly better overall performance than when considering both heuristics directly in the utility function. In the final sections of this study they include a third heuristic directly in the utility function referred to as value learning (VL), which is defined as an additional dummy variable equal to 1 if the individual chose that alternative in the previous choice set, and 0 otherwise. Their results show that this additional parameter significantly improves the overall performance of the model relative to the one that included ANA and MCD directly in the utility function. The authors did not test VL using the PDP approach.

McNair et al. (2012) incorporate two choice set interdependent heuristics: value learning and strategic misinterpretation (SM), together with a LPAA using a PDP approach. Both heuristics state that decisions are influenced by previous choice sets, so the authors tested their formulation by changing the order of the choice sets and re-estimating the models. The results show that both heuristics were adequately formulated in the stand alone heuristic models that considered choice set interdependence, as they did not seem to be significant when changing the order of the choice sets. Their models included socioeconomic characteristics such as the household income and age group, both of which were significant in the models. The PDP model results showed that there was a higher probability of belonging to the VL and SM classes than to the LPAA class. These models had a significantly better overall performance than a standard LPAA model. However, the WTP estimates were not statistically different from each other. The authors state that the WTP results might be influenced by many factors, such as the data source, so these models should be tested in other experiments.

Leong and Hensher (2012b) estimate a joint model including a reference revision (choice set interdependent) heuristic together with the majority of confirming dimensions (local choice context dependent) heuristic. The reference revision is included through a dummy variable that is equal to 1 if the alternative was chosen in the previous choice set, and 0 otherwise. The MCD is included as the number of attributes in that alternative that have the 'best' levels. They use the heuristic weighting functions approach, which considers that all heuristics are used by an individual up to a percentage, i.e., they are multiplied by a weight (all the weights sum to 1). In this study, they considered the weight as a function of socioeconomic characteristics age and income (for more information on this approach refer to Section 3.5.3). They estimate two models combining two heuristics: the first model considers the standard LPAA heuristic and a combination of the LPAA model plus the reference revision parameter (LPAA+Ref heuristic); and the second model considers the LPAA heuristic and a combination of the LPAA, MCD and reference revision heuristic (LPAA+MCD+Ref heuristic). They compare these models with a traditional LPAA model and other models that consider some non-linearities. Their results show that the models that consider more than one process heuristic have a better overall performance, and the preferred model is the one that considers LPAA, MCD and reference revision heuristic. Their results also show that the value of travel time savings decrease when considering more than one heuristic.

Leong (2014) studied the incorporation of multiple heuristics in discrete choice modelling in the context of transportation. He focuses his research on majority of confirming dimensions (local choice context dependent), extremeness aversion (local choice context dependent) and reference revision (choice set interdependent heuristic). As part of the extremeness aversion heuristic he analyses three formulations: non-linear worst level referencing, RRM, and RAM. He proposes the heuristic weighting function to include multiple heuristics, and compares the results to the PDP where both approaches have different advantages. The results show that there is a combination of extremeness seeking and extremeness aversion attitudes. The value of travel time savings (VTTS) are not statistically different for each of the models comparing them to a linear additive RUM model. He suggests for future research to include the treatment of risk, uncertainty and probabilities of occurrence.

As can be noted, even though the consideration of multiple heuristics has shown to significantly improve the statistical performance of the discrete choice models, this topic is relatively new and there is still a large space for further analysis. The consideration of multiple heuristics has helped researchers to further understand individual behaviour by differentiating the influence of several heuristics. Some of the studies analysed above have shown significant influences on the WTP estimates (in their mean and standard deviation) when considering multiple heuristics. However, other did not find significant differences compared to a MML model. Nevertheless, the choice set specific preferences suggested by the different heuristics produce different behavioural insights which lead to a richer interpretation of the trade-offs which is equally as relevant in decision-making.

2.7. Confoundment between Process and Preference Heterogeneity

The possibility arises that the preference heterogeneity captured by random parameters in a LPAA choice model may be, in part, associated with (or explained by) an underlying process heuristic, and hence there may be some relationship between preference heterogeneity specified by a random parameter distribution and a systematic explanation offered through an interaction with an underlying process rule that is not assumed to be LPAA. This possibility has been hinted at in a number of studies for some time. In a recent study, Hensher et al., (2013a) considered two process rules: attribute non-attendance and aggregation of common metric attributes, and embedded them into a mixed latent class model structure. Their results showed a small improvement in statistical performance when adding the random parameters into the latent class structure (i.e., overlaying the discrete distribution assumption for latent classes with a within-class random parameter structure). In their conclusions, they contemplated whether there may exist confoundment between attribute processing as a mechanism to reveal process heterogeneity, and random parameters, the common representation of (unsystematic) preference heterogeneity[2] (Hensher et al. 2013a, p. 1017):

"… a random parameter treatment … may be confounding with attribute processing; and that including attribute processing in the absence of continuously distributed random parameters is preferred to including continuously distributed random parameters in the absence of

---

[2] We recognise that there also may exist potential confoundment between preference heterogeneity and scale heterogeneity (see Fiebig et al., 2010; Hensher et al., 2015; Hess and Train, 2017); however this has not be accounted for in this study. Whether the inclusion of process heuristics as conditioning effects on scale heterogeneity (via the error variance structure) is significant is in itself an interesting additional research theme.

attribute processing. This is an important finding that might suggest the role that attribute processing rules play in accommodating attribute heterogeneity, and that random parameters within class are essentially a potential confounding effect."

Collins et al., (2013) found a large improvement in model fit once a random parameter logit model was extended to include latent classes that explicitly modelled attribute non-attendance (ANA). They also found that stated ANA responses only partly aligned with inferred ANA behaviour, suggesting that self-reporting of heuristics might contain some valuable information, but is not free from error. Collins (2012) used simulated data to demonstrate the biasing impact that ANA behaviour can have on both the mean and the variance of random parameters. He also showed that modelling ANA but not preference heterogeneity, when both exist in a dataset, can bias both the ANA incidence rates, and the WTP of those that do attend to an attribute. The motivation therefore is to appropriately model and capture process heuristics, and conventional LPAA choice.

A few studies have incorporated more than one decision process strategy together with random parameters (see Leong and Hensher, 2012a, 2012b). Hess et al., (2012) propose latent mixed logit models, where each one includes a standard random utility maximisation (RUM) rule[3] together with one of the following process heuristics: lexicography based models, models with multiple reference points, elimination by aspects models and random regret minimisation. Their results show significant statistical gains when considering more than one decision rule. When they allow for random parameters in the RUM class, the share of the other class (the one that considers another decision rule) is reduced. They also observe a reduction in the degree of random heterogeneity in the RUM class compared to a simple multinomial mixed logit model by itself as a one class model. They suggest that what might be retrieved as taste heterogeneity in traditional models may be heterogeneity in decision rules, leading to a questioning of whether there is confoundment, which they suggested should be investigated further in future research.

Campbell et al. (2014) studied different definitions of Elimination by Aspects (EBA) together with a LPAA heuristic in the same model using a latent class approach (LC). They define seven classes: the first class represents the LPAA heuristic, the next five represent EBA by

---

[3] This class is equivalent to what is referred to in this research as the LPAA heuristic.

restricting the choice set to include alternatives that cost less than a certain value (different values for each class), and the last one is confined to the alternative ‘buy none’. They estimate a simple MNL model, a MML model, a LC model with fixed parameters, and a LC model with random parameters referred to as LC-RP. Some of the classes did not seem to be significant in the model estimation including class 7, which implies that there are few respondents that preferred not to buy. Their results show that the majority of individuals (89% for LC and 83% for LC-RP) use the LPAA heuristic to reach a decision. However, there are significant influences on the WTP estimates when using the LC approach, although there is no clear pattern of decrease or increase in the estimates for all the attributes. The authors acknowledge a possible confounding between preference and process heterogeneity, but do not test it any further.

2.8. Behavioural Refinements

Alternative behavioural paradigms to RUM were developed to represent decision-making under risky situations, such as Expected Utility Theory (EUT) and Prospect Theory (PT), briefly presented in Section 2.3.4. EUT was originally developed by Bernoulli (1738) which recognised that choices were made under uncertainty or risk, i.e., deals with situations in which there are several prospects with associated probabilities. EUT models postulate a non-linear representation of the utility function as a representation of risk attitudes (Von Neumann, J. and Morgenstern, 1947; Harrison and Rutström, 2009), where the outcomes are weighted by their probabilities:

$$U = p_1 \cdot u(x_1) + p_2 \cdot u(x_2) + \ldots + p_n \cdot u(x_n) \tag{2.32}$$

where x_n represents the value of the outcome (or attribute), $u(x_n)$ represents a value function of the outcome and p_n the probability of the outcome. In EUT, the risk aversion is equivalent to the concavity of the utility function and, therefore, strictly depends on the value function $u(x_n)$. In EUT, risk attitudes are represented by the marginal (dis)utility $u'(x_n)$.

Kahneman and Tversky (1979) defined a framework as an alternative to EUT called Prospect Theory (PT) to study decisions made under risk. The differences between PT and EUT were explained in Section 2.3.4, and this Section will focus on their differences in representing the probabilities. In PT, the value of each outcome is multiplied by a decision weight rather than by the probability of the outcome:

$$U = w(p_1)\cdot u(x_1) + w(p_2)\cdot u(x_2) + ... + w(p_n)\cdot u(x_n) \quad (2.33)$$

where $w(p_n)$ represents a weighting function of the probability of occurrence. Hence, risk is measured as the value function and weighting function. It is important to note that the concepts of value function and weighting function are not exclusive to PT and have been used in other frameworks, as proposed by Hensher et al. (2011a) and Hensher et al. (2013) as an Extended EUT (EEUT).

Choice under risk implies that the choice maker will have to (1) evaluate information about the attributes describing the risky outcomes through the value function, which will be referred in this thesis as *risk attitudes*, and (2) evaluate the information about the probability of occurrence of each outcome through the weighting function, which is the starting point of what will be referred in this thesis as *perceptual conditioning.* The next sections will explain each of these concepts separately and refer to the literature on each one. However, there are some studies that consider them together (Harrison and Rutström, 2009; Hensher et al., 2011; Hensher et al., 2013; Balbontin et al., 2017).

2.8.1. Risk Attitudes

The value function $u(x_j)$ presented above takes into account possible risk attitudes relative to outcomes. The great majority of choice studies assume that respondents have a risk attitude that is risk neutral. Considering each in turn, there are people who are risk adverse, risk taking or risk neutral, and this heterogeneity in risk attitude does influence individuals' decisions when faced with different choice scenarios. In general, we consider that people are risk averse if their utility function is concave and risk takers if their utility function is convex (Kahneman and Tversky, 1979). Senna (1994) explains these concepts numerically stating that if the utility of the expected value is equal to the expected value of its utility, there is risk neutrality. If the utility of the expected value is smaller than the expected value of its utility, there is risk aversion; and if it is larger, there is risk propensity. An illustrative example of how these attitudes are represented is shown in Figure 2-2.

Eeckhoudt et al. (2005) define risk aversion as "the rate at which marginal utility decreases when wealth is increased by 1%". There are three possible risk attitude outcomes which have to be modified when considering costs rather than wealth, which generate a disutility instead of utility:

Risk taking: the rate at which the marginal **utility** increases when **wealth** is increased by 1%. Or, equivalently, the rate at which the marginal **disutility** increases when **cost** is increased by 1%.

Risk aversion: the rate at which the marginal **utility** decreases when **wealth** is increased by 1%. Or, equivalently, the rate at which the marginal **disutility** decreases when **cost** is increased by 1%.

Risk neutrality: the marginal **utility/disutility** remains constant when wealth/cost is increased by 1%.

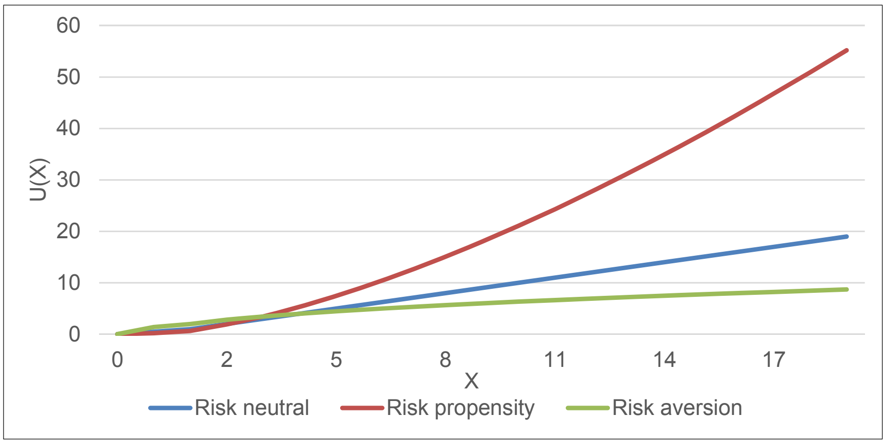


Figure 2-2: Illustrative example of risk attitudes

Table 2-3 presents several functional forms for the risk aversion that have been studied in literature (Stott, 2006; Rasouli and Timmermans, 2014).

Table 2-3: Summary Table of Risk Attitudes Functional Form

Name	Equation
Linear	$v(x) = x$
Logarithmic	$v(x) = \ln(a + x)$
Power	$v(x) = x^a$
Quadratic	$v(x) = ax - x^2$
Exponential	$v(x) = 1 - \exp^{-ax}$
Bell	$v(x) = bx - \exp^{-ax}$
Hyperbolic Absolute Risk Aversion (HARA)	$v(x) = -(b + x)^a$

Arrow (1965) studied the probability of a risk averse individual to accept a gamble, and proposed a measure of local risk aversion which is known as the Arrow-Pratt measure of absolute risk aversion:

$$v(x) = \frac{u''(x)}{u'(x)} \tag{2.34}$$

This equation is used to derive the value functions. In this thesis we focus on the Hyperbolic Absolute Risk Aversion (HARA) functional form, which has proven to be the most adequate in representing risk attitudes in different contexts (Ingersoll, 1987; Andersen et al., 2012; Hensher et al., 2013). There are two special cases of HARA that have been widely used (are derived using equation (2.34)): Constant Relative Risk Aversion (CRRA) that occurs when b=0; and Constant Absolute Risk Aversion (CARA) which occurs when a=0 and are shown in Table 2-4. Blanchard and Fischer (1989, p. 44) explain the CARA and CRRA, and explain that sometimes analytically it is more convenient to use the CRRA.

Table 2-4: Summary Table of Risk Attitudes CRRA and CARA

Name	Equation	Interpretation	References
Constant Relative Risk Aversion (CRRA)	$u(x)=\frac{x^{1-\alpha}}{1-\alpha}$	α represents the risk attitude[4]: If α = 0 risk neutral If α < 0 risk taker If α > 0 risk averse	Arrow (1963); Pratt (1964); Holt and Laury (2002); Andersen et al. (2012); Hensher et al. (2013b), (2011a); Rasouli and Timmermans (2014)
Constant Absolute Risk Aversion (CARA)	$u(x)=-\exp^{-\alpha x}$	α represents the risk attitude: If α = 0 risk neutral If α < 0 risk taker If α > 0 risk averse	Arrow (1963); Pratt (1964); Polak et al. (2008); Rasouli and Timmermans (2014)

Holt and Laury (2002) study risk attitudes in lottery choices using different value function formulations. They use a combined functional form of exponential and power, which increases relative risk aversion and decreases absolute risk aversion. Their results show that most people are risk averse and a few are risk loving, and when high payoffs are paid in cash, respondents become more risk averse. One of the main conclusions of their results is that there is clear evidence for risk aversion which suggests the danger of not taking it into account when analysing behaviour.

Polak et al. (2008) was one of the first studies in transportation to test CARA in a stated preference data on mode choice under travel time variability. They estimate two type of models; the first one considers a EUT form, and the second one estimates risk attitudes using CARA. The utility function on mode *i* for the latter model can be written as follows:

$$u_i=\sum_{k=1}^{K}\frac{\left(1-\exp^{-\alpha\cdot v_{i,\mathrm{k}}}\right)}{\alpha}p_{\mathrm{k}} \tag{2.35}$$

where α is the risk attitudes parameter; $v_{i,k}$ is the valuation for outcome k; and p_k is the probability assigned to outcome k. They estimated α as fixed and as random (mixed logit).

[4] The special case where α=1, collapses to 1/0 which is undefined (it goes to infinity); however the occurrence of such a case is very unlikely.

Their results show that including risk attitudes significantly improves the overall performance of the model relative to a simple EUT. The estimates reveal that respondents are, on average, risk averse towards travel time variability. Their mixed logit model results show that there is systematic and random heterogeneity in risk aversion across the sample.

Hensher et al. (2011) study risk attitudes through the CRRA specification shown in Table 2-4 in a route choice experiment with travel time variability. Their results on a multinomial logit model show a significant influence of risk attitudes in all the models estimates, revealing that respondents are, on average, risk averse as the α parameter is always positive. They estimate a number of models which are mixed logit models and estimate α as a random parameter. Their results are consistent with respondents being risk averse, and show a high heterogeneity between respondents in their degree of risk aversion (as evidenced by a significant standard deviation)[5].

Andersen et al. (2012) study lottery choices including risk attitudes as a constant relative risk aversion. The first utility function defines it as follows:

$$U(m_k) = m_k{}^{\alpha} \tag{2.36}$$

where α represents the constant relative risk aversion. If the value lies between 0 and 1 it represents risk aversion, if it is equal to 1 it represents risk neutrality and if it is larger than 1 it represents risk taking. If it is negative, it implies respondents violate the assumption of non-satiation. They estimate an α parameter for everyone and another one that interacts with gender; as well as models that estimated α as random (with a mean and standard deviation). Their results show that it is appropriate to estimate α as a random parameter, since there is significant heterogeneity across the sample. The interactions with gender were found to be significant, revealing that females, on average, were more risk averse and behaved more differently between them (larger standard deviation).

The second utility function, defined in Andersen et al. (2012), follows the CRRA equation shown in Table 2-4 (together with its interpretation) and also includes discount rates in a

---

[5] The results reported by the authors are not the same as in this paragraph because there was an error in the interpretation of the alpha parameter in the paper, which was corrected in later studies by the same authors. The interpretation of results included in this paragraph is the correct one.

sample of lottery choices and time-delay choices. They estimate α as a random parameter. The results suggest a concave utility function with a positive α, i.e., risk aversion, and a significant heterogeneity between respondents (standard deviation).

Risk attitudes can influence decision making when there is variability in the levels of an attribute since the probabilities of occurrence might induce a certain behaviour in people that are risk averse or risk taking. However, risk attitudes could also be influential when a person faces a decision without any variability in the levels of an attribute, assuming he is comparing the levels presented with a reference point (similarly as what is proposed in the PT framework). In this case, the interpretation would be exactly the same as the definition proposed by Eeckhoudt et al., (2005). Balbontin et al. (2017) estimate risk attitudes in the CRRA form towards the cost of road pricing reforms (namely, cordon-based charging and distance-based charging), which is shown in the experiment as a fixed value. They estimated the α as a random parameter to allow heterogeneity in risk attitudes. Their results show significant risk attitudes towards the costs, and on average there was a risk aversion towards them. The experiment used presented the road pricing reform's cost as a unique value, nonetheless individuals did have a risk attitude towards it. That is, the marginal disutility, on average, decreased when cost was increased by 1%.

2.8.2. Perceptual Conditioning

The definition of perceptual conditioning used in this research refers to outcome conditioning by an attribute influenced by perception. Hereby, the weighting function proposed by Kahneman and Tversky (1979) is a specific case of perceptual conditioning, where the outcome is conditioned by its probability of occurrence, and this probability is subject to perception (through the weighting function).

Most studies in the literature have included perceptual conditioning as a weighting function of the probabilities of occurrence. This is a phenomenon of special interest where attributes exhibit variability over repeated occasions (e.g., the commuting trip across a week). Whether the analyst uses monitored (revealed preference) attribute levels (e.g., travel time on each of 7 days) or provides travel times with associated occurrences over a given period in a stated choice experiment, a respondent's perception of 'objective' probabilities often translates into an over- or under-weighting of such probabilities, especially at the extremities.

The function used for perceptual conditioning allows for a deviation between the subjective value or probability and the one that individuals actually use when evaluating the outcome. Different functional forms have been developed to consider this deviation (Stott, 2006; Rasouli and Timmermans, 2014) and a summary of them is provided in Table 2-5.

Table 2-5: Summary Table of Functional Forms of Perceptual Conditioning

Name	Equation	References
Linear form with discontinuous end points	Example: $w(p)=\begin{cases}(1-r)p & \text{if } p<1\\ 1 & \text{otherwise}\end{cases}$	Loomes et al. (2002); (Abdellaoui et al., 2010)
Power	$w(p)=p^r$	Simplification of the models: Goldstein and Einhorn (1987); Quiggin (1982); Tversky and Kahneman (1992)
Goldstein-Einhorn	$w(p)=\dfrac{sp^\gamma}{sp^\gamma+(1-p)^\gamma}$	Goldstein and Einhorn (1987)
Tversky and Kahneman	$w(p)=\dfrac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{1/\gamma}}$	Quiggin (1982); Tversky and Kahneman (1992); Camerer and Ho (1994); Tversky and Fox (1995)
Wu-Gonzalez	$w(p)=\dfrac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^s}$	Wu and Gonzalez (1999)
Prelec I	$w(p)=\exp^{-(-\ln p)^\gamma}$	Prelec (1998)
Prelec II	$w(p)=\exp^{-s(-\ln p)^\gamma}$	Prelec (1998)
Exponential-power	$w(p)=\exp^{-\frac{r}{s}(1-p^s)}$	Prelec (1998)
Hyperbolic-logarithm	$w(p)=(1-r\cdot\ln p)^{-s/r}$	Prelec (1998)

The Tversky and Kahneman functional form has proven to be the most adequate in several discrete choice studies (Camerer and Ho, 1994; Hensher et al., 2011). This function is an inverse S-shape where γ represents the degree of curvature of the weighting function. An estimated parameter γ with a value between 0 and 1 suggests that individuals will over-weight low belief probabilities and under-weight medium to high belief probabilities.

Camerer and Ho (1994) study the probability of gambles, comparing different approaches: the betweenness axiom, where a probability mixture of two gambles should be between them in preference; the disappointment aversion theory, which relaxes the independence between probabilities but obeys betweenness; and the probability weighting function using the Tversky and Kahnemann functional form. Their results show that the disappointment aversion theory and the weighting function are superior in model fit. The authors were encouraged by the weighting function results since they were similar to other results using the same methods (Tversky and Kahneman, 1992), saying that low probabilities are overweighted and higher probabilities are underweighted.

Hensher et al. (2011) study travel time variability by incorporating a weighting function in their probabilities of occurrence. They compare results using four functional forms: Tversky and Kahnemann; Goldstein-Einhorn; Prelec I; and Prelec II. Their results show a superior model fit using Goldstein-Einhorn and Prelec II weighting functions, which they attribute to the extra parameters these functions estimated that appeared to be highly significant. However, the Tversky and Kahnemann functional form was the one that seemed more aligned with theory, i.e., low probabilities being overweighted and higher probabilities underweighted. Thereby, and due to willingness to pay estimates analysis they select the Tversky and Kahnemann functional form as their preferred model.

Balbontin et al. (2017) include perceptual conditioning by using individual beliefs. The dataset used was collected as part of a road pricing reform study, where respondents were asked to choose between their status quo situation, a cordon-based charging regime, and a distance-based charging regime, each one described by several attributes of time and costs. Additionally, the (two) road pricing alternatives contained five attributes that represented the allocation of the revenue raised (e.g., improving public transport, improving existing and construct new roads). After each choice set respondents were asked “to what extent do you think that each of these schemes will make you better/worse off” and they had to answer using a score from 0 (totally worse off) to 100 (totally better off). This was referred to as belief, and was used to condition the revenue raised allocation attributes. They used the Tversky and Kahneman (1992) formulation as follows:

$$\begin{aligned} U_{RoadPricing,i} &= \theta_0 + \theta_{TT} \cdot TT + \theta_{\$\$} \cdot \$\$ \\ &+ \left[\frac{belief^{\gamma}}{\left[belief^{\gamma} + (1 - belief)^{\gamma} \right]^{\frac{1}{\gamma}}} \right] \cdot \left(\theta_{RevenueAlloc1} \cdot RevenueAlloc_1 + \ldots \right) + \varepsilon_i \end{aligned} \qquad (2.37)$$

where TT represent the travel times; $$ represent the cost attributes; $RevenueAlloc_1$ represent the first option for revenue allocation (and the dots represent the other four options); β are the estimated parameters; and γ the weighting curvature parameter. Their results show a significant influence of the belief towards revenue allocation, and suggested that individuals under-weight low belief probabilities and over-weight high belief probabilities when interacting with the allocation of the revenue raised. This study is particularly interesting because it incorporates the concept of perceptual conditioning even though the experiment did not include variability in the attributes with assigned probabilities of occurrence (also see Hensher et al., 2013).

Li and Hensher (2017) provide an interesting overview of the literature and develop a multivariate method for discrete choice analysis with risky prospects. They propose a cumulative weighting function that does not have to be defined a priori and can be derived from the modelling. Thus, they develop a general functional form to include perceptual conditioning that enables the data to determine which form is appropriate for it.

2.9. Experience

Several studies mention the importance that previous knowledge has on decisions (including McFadden's quote at the beginning of this chapter), and many of the heuristics previously analysed try to integrate it in the utilities. For example, Kahneman and Tversky (1979) state that people value differently gains than losses and these depend on the reference point of an individual. Others, such as the Case-Based Decision Theory (Gilboa and Schmeidler, 1995) consider that past experiences with similar choices and their outcomes affect decisions. Many of them try to include respondents' understanding of the choices, past experiences, or something similar. Evidence suggests that what an individual did in the past affects his/her decision today. In discrete choice models, this has been included in different ways. The previous section showed that this concept has been included through different decision process strategies, such as into the reference revision heuristic, value learning, among others. This section will analyse studies that have included it directly in the utility function.

Hensher (1975) define a 'habit period' as when an individual does not see other alternatives as directly relevant, i.e., he/she is not willing to change; and the 'decision period' when the individual analyses other alternatives by considering a change. He proposes to include this in choice models. Hensher (1976) propose a model consistent with inertia theory to measure it:

$$C=\theta_0+\theta_1\cdot X_1+\theta_2\cdot X_2+...+\theta_n\cdot X_n \tag{2.38}$$

where C is the monetary benefit of mode choice, which is equivalent to the monetary choice of the usual mode and the alternative mode, plus a transfer payment TP_c, that is: $C_u+TP_c-C_a$; X_n represent the difference in levels of attribute *n* between the usual and alternative mode, where the first attribute X_1 represents travel time: $X_1=t_u-t_a$; and θ_n are the estimated parameters for each attribute *n*. Therefore, the author proposes the following model to be empirically tested:

$$C_u+TP_c-C_a=\theta_0+\theta_1\cdot X_1+...+\theta_n\cdot X_\mathrm{n} \tag{2.39}$$

where θ_0 represents inertia, as the equivalent amount of money an individual is willing to outlay in order to maintain indifference between alternatives, for reasons other than the explanatory variables. If we consider mode characteristic differences only in the travel times (i.e., $X_2=...=X_n=0$), the equation could be written as follows:

$$C_u+TP_c-C_a=\theta_0+\theta_1\cdot\left(t_u-t_a\right) \tag{2.40}$$

Goodwin (1977) define 'habit' as a source of resistance to change and also incorporate it in choice models. He presents a model in which it can be included as deterministic or probabilistic. Ben-Akiva and Morikawa (1990) incorporate the 'habit', also referred to as 'inertia', in a mode choice study by asking respondents if they would be willing to switch from their currently used mode to the new mode presented. They estimate the model in such a way that a respondent would switch mode if the utility function of the new alternative is superior to the utility of the current mode by more than a threshold value. They use revealed and stated preference data, and their results in the stated preference data estimated a negative threshold value, indicating that respondents overstate their switching to the new mode.

Cantillo et al. (2007) study travel behaviour including inertia. The inertia is a function of the previous valuation of alternatives, so it is considered as random within the population. They use panel data; the first instance is simulated data and the second one a mixed stated and revealed preference experiment. One of the most important conclusions of their study is that not taking into account inertia where it exists may lead to biases and significant response errors.

Hensher and Collins (2011) incorporate the value learning heuristic. In doing so, they include in their model a dummy variable in the utility function that, in the first choice set, equals to 1 if the individual chose that mode in his most recent trip, or 0 otherwise. This variable is updated when an individual makes a choice other than their status-quo situation. Even though the objective of this study was not to study the influence of past experiences, it is mentioned in this section because it includes overt experience on the modes of transportation directly in the utility function and it did show to be significant. We could consider another word than experience or overt experience to avoid any implication of lagged effects, since most studies recognise the mode that respondents recently or currently use; however we like the word 'experience' but it could also be referred to as the reference alternative, which is common in studies that incorporate gains and losses relative to a reference alternative.

Hensher and Ho (2016) analyse the influence of overt experience in a commuter modal choice study. Overt experience is defined by individual familiarity with the mode of transportation, i.e., if they have or have not used it. They include overt experience as an additional parameter that conditions the utility expression and they include a lognormal transformation of the frequency of use of the modes as a representation of experience. The utility function can be written as follows:

$$U_i(\theta, \mathrm{X})' = \exp\left[\gamma_i \ln\left(FR_{q,i} + 1\right)\right] \cdot U_i(\theta, \mathrm{X}) \tag{2.41}$$

where $U_i(\theta, \mathrm{X})$ is the traditional utility function for mode *i*, described by the vector of attributes X with their parameters θ; $U_i(\theta, \mathrm{X})'$ is the transformed utility; $FR_{q,i}$ is the usage frequency for mode i and individual q; and γ_i the estimated parameter for overt experience. The model form is consistent with the overall utility maximisation assumption with the conditioning component (referred to as heteroscedastic conditioning) derived from information associated with the variance of the unobserved effects, effectively delimiting an observable and systematic way of accommodating heterogeneous scaling (Hensher and Ho 2016).

Their results show a significant influence of the usage frequency, and both for the car and public transport γ_i was negative. To analyse the results they studied the relationship between the frequency of use of each mode and their utilities. Their findings show that, if a commuter has car experience, the car is significantly less attractive in comparison to a commuter that does not have any experience with using the car. Conversely, if a car user has commuter experience, commuting is significantly less attractive relative to a car user who has no commuter experience. They show graphically the differences in the utility functions when recognising experience and when not recognising it, revealing the importance of experience.

The studies analysed in this section showed that considering the role of experience in decision making has a significant influence in the overall performance of the models and in the results. If experience is neglected, it might lead to biases in the results. The recognition of the role of past experiences and how do they influence current behaviour is fundamental for an adequate interpretation of decision making.

2.10. Conclusions

The first part of this chapter presented the extensive literature on decision process strategies, separating them into three categories: context free; local choice context dependent; and choice set interdependent. Most of the studies analysed one heuristic and compared the results to the ones obtained through a simple LPAA heuristic. Section 2.6 focused on the studies that have considered more than one heuristic and the results show a significant improvement in the overall performance of the models. Even though not all the studies show a significant influence over the WTP estimates when considering multiple heuristics (although some do), the influence on preferences produce different behavioural insights which leads to a richer interpretation of the results and decision-making. The results of these studies are very encouraging and suggest a significant improvement in the model results and interpretation when including multiple heuristics, however the literature is still rather limited.

The issue of a possible confoundment between preference and process heterogeneity has been raised in several studies, yet none of them have been able to test the existent relationship between random parameters and multiple heuristics. The phenomenon of degrees of potential substitution or complementarity between the non-systematic representation of preference heterogeneity through random parameters and multiple heuristics is a fundamental issue that might completely change the way process rules are integrated, and it has been acknowledged by the majority of studies that have included them together.

The second part of this chapter focused on studying behavioural refinements that have been analysed in the literature. The first one corresponds to experience and most studies consider that the most recent alternative an individual chose will influence their current decision. Experience has been included in models through different formulations and all of them have shown to have a significant influence on preferences. The evidence suggests that not taking into account individuals' experience might result in misleading estimates. There are very few studies that have incorporated experience together with alternative decision process strategies, even though they have resulted in improved and richer models.

Other behavioural refinements, such as risk attitudes and perceptual conditioning have been studied in literature. Both have shown to improve the overall model performance and also provide important insights on how preferences are made. Several studies have used them in choice studies; however there is limited knowledge on how these act together with multiple heuristics.

This chapter has provided a review of what has been done in the field of process heuristics and other behavioural refinements. The numerous studies have shown a significant influence on the model estimates and the interpretation derived from them. However, there are still many research gaps concerning multiple heuristics, especially considering their relationship with random parameters, experience, risk attitudes and perceptual conditioning. These gaps will be addressed in the next chapters.

CHAPTER 3

Research Methodology

3.1. Introduction

The previous chapters have explained the motivation and objectives of this research, reviewed the current literature on heuristics, behavioural refinements and experience, and identified gaps in this literature. This chapter will detail the methodologies used in this study, including existent and original methodologies.

As was mentioned in Chapter 2, choice studies may replicate real life decisions, which are referred to as revealed preferences (RP), or may present hypothetical scenarios, which represent stated preference (SP) experiments. Each of them have different advantages and disadvantages. The main advantage of the RP data approach is that it represents real decisions, but has the limitation that the modeller is not able to observe all the alternatives and attributes taken into account by an individual. The main advantage of an SP experiment is that it is able to assess alternatives that are not present in real life (e.g., a new metro line) and that the modeller is able to control the attributes being evaluated by an individual through varying the levels associated with observed alternatives which might not be those observed in real markets. Another advantage of the SP experiment is that respondents can be presented with multiple choice tasks (unless there exists an RP panel) so a larger number of observations can be obtained providing greater variability in data and increased opportunity to better understand individual preferences. It has the disadvantage that they are not real decisions and might not exactly replicate how individuals would behave in real life. Some of the models estimated as part of this research include three process strategies together with experience and behavioural refinements. These models require an increased variance in the dataset to enable inquiry into the different elements. This is often not possible using RP data since the attribute levels of variation are very restricted relative to a SP data.

When trying to understand individual preferences under the stated preference experiment paradigm, respondents are presented with a choice scenario with different alternatives described by certain attributes and they have to make a decision. This scenario is referred to as a choice task or choice set. The different values an attribute can have (e.g., travel time of 3, 5 or 15 minutes) are called attribute levels; and each alternative is typically described by one level for each attribute. A choice experiment can be labelled or unlabelled: an unlabelled choice experiment is one where the alternatives have generic titles (e.g., Route 'A' or 'B'): and a labelled choice experiment has alternative-specific titles (e.g., 'Car' or 'Bus').

The first two sections will explain the Multinomial Logit (MNL) and Mixed Logit (MML) models that have been widely used in choice studies and are the basis of the models discussed in this research. Section 3.4 will present the selection of heuristics for this research, where the subsections present each heuristic with their formulation and implications. Section 3.5 shows how the models will take into account multiple heuristics, behavioural refinements and experience. Section 3.5.1 presents one of the most popular methods used to include process heterogeneity, referred to as the Probabilistic Decision Process (PDP), and also presents an extension of this methodology called Probabilistic Decision Processes Combined (PDPC). The second Subsection 3.5.2 presents the Hybrid RUM-RRM and Heuristic Weighting Function. Section 5.3 proposes a new method to include process heterogeneity, called Conditioning of Random Process Heterogeneity (CRPH). Section 3.5.4 shows how the behavioural refinements and experience are included in the models. The next section presents the marginal (dis)utilities that can be obtained for each model. Section 7 presents the method to estimate the Willingness to Pay (WTP). The final section reviews the main conclusions of this chapter.

3.2. Random Utility Theory (RUT) and the Multinomial Logit (MNL) Model

The traditional discrete choice model is an econometric framework within which we specify a series of expressions designed to capture increased behavioural richness in the choice making process. Estimation of choice models provides a way of establishing the role of the various attributes and behavioural strategies in explaining, in statistical terms, specific choices, and in obtaining behavioural derivatives such as willingness to pay estimates for specific attributes.

The models are linked to the theory of utility maximisation, but set out within a theory of random utility (McFadden, 1974) which assumes that each individual behaves rationally and it is assumed that they act as if they choose the alternative that maximises utility. The utility of each alternative is defined as a function of its attributes, where each attribute is weighted by an estimated coefficient as a measure of the marginal (dis)utility associated with a specific attribute. We can describe this utility for each individual q with a utility function U_{iq} assigned to each alternative A_i, as described in equation (3.1). This utility function has a systematic component which is measurable, V_{iq}, that is a function of the alternatives' attributes measurable by the modeller. It also has a random component, ε_{iq}, which allows for possible preference heterogeneity and for potential errors induced by the modeller (i.e., in measuring or observing) including excluded relevant attributes. It is important to understand that the modeller is likely only able to see or account for an undefined subset of the attributes considered by sampled individuals when reaching a decision, and hence is not able to include all the attributes that could influence the respondent choice.

$$U_{iq} = V_{iq} + \varepsilon_{iq} \tag{3.1}$$

The systematic part described above in its simplest form can be written as follows (Hensher et al., 2015b; Ortúzar and Willumsen, 2011):

$$V_{iq} = \theta_{i1} \cdot x_{i1q} + \theta_{i2} \cdot x_{i2q} + \theta_{i3} \cdot x_{i3q} + \ldots + \theta_{in} \cdot x_{inq} \tag{3.2}$$

θ_{in} represents the estimate for attribute *n*, which is assumed to be fixed for all individuals but can vary between alternatives i, and x_{inq} represents the level of attribute *n* of alternative *i* for the individual *q*.

As explained above, a person will choose the alternative *i* (A_i) only if his utility is larger than (or equal to) that associated with each and every other alternative j (A_j).

$$U_{iq} \geq U_{jq} \qquad \forall A_j \in A(q) \tag{3.3}$$

where $A(q)$ represents the set of alternatives available for individual *q*. Hence,

$$V_{iq} - V_{jq} \geq \varepsilon_{jq} - \varepsilon_{iq} \qquad \forall A_j \in A(q) \tag{3.4}$$

Since the modeller is unable to capture all sources of utility is the observed component, he is not able to determine with certainty if the person would choose the alternative. However, he is able to estimate the probability of choosing alternative *i*, P_{iq}, as shown in equation (3.5).

$$\mathrm{P}_{iq} = \mathrm{Prob}\left[\varepsilon_{jq} \leq \varepsilon_{iq} + V_{iq} - V_{jq}, \quad \forall A_j \in A(q)\right] \tag{3.5}$$

The log likelihood is represented through the following equation:

$$l(\theta) = \ln L(\theta) = \sum_{q=1}^{Q} \sum_{A_j \in A(q)} y_{jq} \ln \mathrm{P}_{jq} \tag{3.6}$$

where y_{jq} is a dummy variable that takes the value one if the alternative is chosen and zero in other cases. The overall log likelihood at convergence compared to various definitions of knowledge of market shares only or absence of this knowledge (i.e., no observed alternative-specific constants known) represents the goodness of fit of a model[6]. The closer to zero the log likelihood after controlling for the number of degrees of freedom (linked to the number of parameter estimates), the better the overall goodness of fit in a statistical sense.

The assumptions imposed on the distribution of the unobserved effects or random errors associated with each alternative define the functional form of a specific choice model. The simplest form is the *multinomial logit* (MNL) model which assumes that the random component in the utility function is distributed independent and identically (IID) Gumbel with mean 0 and standard deviation σ (Domenecich and McFadden, 1975). In this case the choice probability will be as follows:

$$P_{iq} = \frac{e^{\lambda \cdot V_{iq}}}{\sum_{A_j \in A(q)} e^{\lambda \cdot V_{jq}}} \tag{3.7}$$

where the parameter λ is related with the standard deviation as follows:

---

[6] Note that it depends on the number of observations in the sample. In Section 3.8 some basic tests that compare the log likelihoods will be shown. These can only be used when the models use the same observations from the sample.

$$\lambda = \frac{\pi}{\sigma\sqrt{6}} \tag{3.8}$$

V_{iq} is usually assumed to have a linear form as shown in equation (3.2), with linearity in the parameters and in the attributes. However, this is not essential, and we can have a non-linear MNL model which we discuss in the next sections.

3.3. Mixed Multinomial Logit (MML) Model

A more complex model widely used in transportation studies, is the mixed logit (ML) model. The current form was proposed by two research groups: Ben-Akiva and Bolduc (1996) and McFadden and Train (2000) (see also Hensher and Greene 2003). The ML form consists of any model for which choice probabilities can be expressed as an open form (Ortúzar & Willumsen 2011):

$$E(P_{iq}) = \int P_{iq}(\theta) \cdot f(\theta) \; d\theta \tag{3.9}$$

$P_{iq}(\theta)$ represents the standard MNL probabilities evaluated at a set of parameters θ, and $f(\theta)$ is their density function (also known as a 'mixing distribution'). If the density function is degenerate at fixed parameters θ, the choice probability is equivalent to the MNL form shown in Equation (3.6). If, however, the density function is discrete, the ML becomes a *latent class model*. This latter is useful when there are different segments in the population that have their own choice behaviour.

The MML model can be estimated on a single cross-section or a panel dataset. In the single cross-section dataset, each respondent responded to the choice task one time, as it commonly would in an RP study or and SP experiment with only one choice set. The panel dataset considers that each individual is faced with multiple choice sets and that an individual's preferences remain invariant across those choice sets.

Random Parameters

A specification of the MML model considers *random parameter estimates*. That is, the parameters are allowed to vary over individuals *q* (but not over choice sets *t*) with density function $f(\theta)$ in order to capture preference heterogeneity. Not all the parameters θ have to be considered random; some can be specified as preference homogeneous. The utility function of alternative *i* for individual *q* in choice situation *t* (*t*=1,…,*T*) is defined as follows:

$$U_{iqt} = \theta_q \cdot X_{iqt} + \varepsilon_{iqt} \tag{3.10}$$

where the parameter θ_q can be decomposed in its mean, θ^m, which is common for all individuals *q*, and a standard deviation, σ; and a distribution v specified by the modeller:

$$U_{iqt} = \theta^m \cdot X_{iqt} + \sigma \cdot v \cdot X_{iqt} + \varepsilon_{iqt} = \left(\theta^m + \sigma \cdot v\right) \cdot X_{iqt} + \varepsilon_{iqt} \tag{3.11}$$

For example, if the modeller assumes a normal distribution then v will have a mean 0 and standard deviation of 1, $v \sim N[0,1]$. The choice probabilities are equivalent to an MNL model but considering the different choice situations *t*:

$$P_{iqt} = \frac{e^{\lambda \cdot V_{iqt}}}{\sum_{A_j \in A(q)} e^{\lambda \cdot V_{jqt}}} \tag{3.12}$$

The log likelihood is represented through the following equation:

$$l(\theta) = \ln L(\theta) = \sum_{q=1}^{Q} \sum_{t \in T} \sum_{A_j \in A(q)} y_{jqt} \ln \mathrm{P}_{jqt} \tag{3.13}$$

Latent Class Model

There is a special case where the modeller considers random parameters in a mixed logit model, in which the density function is not discrete, called a *latent class mixed logit* model. The underlying theory is that individual behaviour depends on observable attributes and on a latent heterogeneity unknown to the modeller, and this heterogeneity can be analysed through discrete parameter variation. It considers that individuals can be distributed into *C* classes - each class with specific parameters - but the modeller does not know who belongs to which class. The utility functions for each class *c* are equivalent to a MNL model:

$$V_{iqt|c} = \theta_{i1|c} \cdot x_{i1qt} + \ldots + \theta_{in|c} \cdot x_{inqt} \tag{3.14}$$

The choice probability for individual *q* in choice situation *t* and alternative *i* considering he belongs to class *c* is as follows (Greene and Hensher, 2003):

$$P_{iqt|c} = \frac{e^{\lambda \cdot V_{iqt|c}}}{\sum_{A_j \in A(q)} e^{\lambda \cdot V_{jqt|c}}} \tag{3.15}$$

The unknown class assignment for class *c* can be considered in different forms (Greene, 2001), but the one that has been commonly used is one that considers a MNL model form as follows:

$$H_{qc} = \frac{e^{h_c}}{\sum_{i \in C} e^{h_i}} \tag{3.16}$$

where H_{qc} is the probability that individual *q* belongs to class *c*; *C* is the total number of classes; and h_c a class assignment function defined by the modeller, for example, using a set of *m* observable characteristics z_m (such as individual socioeconomic characteristics) as follows:

$$h_c = \mu_{0c} + \mu_{1c} \cdot z_1 + \ldots + \mu_{mc} \cdot z_m \tag{3.17}$$

μ_{mc} represents the parameters associated with the observable characteristics, and μ_{0c} the class specific constant. Similarly to the MNL model, *C*-1 class specific constants can be estimated due to identification issues and one of them has to be considered the base (i.e., equal to zero). The same is the case when z_m remains constant for all of an individual's alternatives (e.g., age). A special case of the class assignment model form considers all μ_{mc} equal to zero except for the class specific constants. In this case, class assignment would be independent of any observable characteristics.

The contribution of the choice probabilities to the likelihood function are given by[7]:

[7] In a panel data set t>1 and in a single cross-section data set t=1, but the equation is equivalent.

$$P_{q|c} = \prod_{t \in T} P_{iqt|c}{}^{g^{iqt}} \tag{3.18}$$

where g^{iqt} is a dummy variable equal to 1 if individual *q* chooses alternative *i* in choice situation *t*, and 0 otherwise. Therefore, the likelihood for individual *q* is equal to the sum of the class specific contributions:

$$P_q = \sum_{c=1}^{C} H_{qc} \cdot P_{q|c} \tag{3.19}$$

The log likelihood for the sample can be written as follows:

$$l(\theta) = \ln L(\theta) = \sum_{q=1}^{Q} P_q = \sum_{q=1}^{Q} \ln \left[\sum_{c \in C} H_{qc} \cdot \left(\prod_{t \in T} P_{iqt|c}{}^{g^{iqt}} \right) \right] \tag{3.20}$$

Error Components

The second specification of MML refers to *error components,* which has been proposed as a way to include highly flexible substitution patterns between alternatives. This considers that the error term has two elements; one that corresponds to the error term of the MNL probability ε_{iqt} , and another one represented by $\varpi_{iqt} \cdot Y_{iqt}$ in the following equation:

$$U_{iqt} = \theta_{it} \cdot X_{iqt} + \varpi_{iqt} \cdot Y_{iqt} + \varepsilon_{iqt} \tag{3.21}$$

where the parameters θ_{it} are fixed, i.e., do not vary across individuals[8]; X_{iqt} are the observed attributes; ϖ_{iqt} is a vector of random elements with a distribution specified by the modeller, with zero mean and unknown covariance; and Y_{iqt} is a vector of attributes unknown to the modeller.

[8] The parameters θ can be considered to vary between individuals, i.e., random parameters, and this would be considered as a combination between random parameters and error components.

Panel Data

Both the random parameters and error component specifications can be specified so that it acknowledges the panel nature of the data. This is important in stated preference (SP) experiments where each individual is faced with multiple choice situations, so there are multiple observations for each individual. In the random parameters specification, the parameters are not allowed to vary over the choice situations of the same individual, but only across individuals. In the error components specification, the additional error term chosen by the modeller is considered to be the same within the responses of each individual.

3.4. Heuristics Selection

Based on a review of the available literature on heuristics in Chapter 2, three heuristics have been selected for investigation in the present research. The first heuristic considered is one that has been traditionally used in choice studies, referred in this research as linear in the parameters and additive in the attributes, as it provides a reference point of what is currently being done in most choice studies.

The most important criteria used to select the other two heuristics was to take into account both local choice context dependence and choice set interdependence (see Sections 2.4 and 2.5 respectively). Hence, one of each was chosen. Considering local choice set dependence, the line of heuristics that have proven to be most successful and significant in the literature are those ones that reflect extremeness aversion. From these ones, different formulations were tested and the Relative Advantage Maximisation (RAM) heuristic was chosen since it is relatively simple in its formulation yet adds a significantly richer interpretation of decision making. Regarding the choice set interdependent heuristics, the ones that consider a reference point are the ones that have proven to be most informative in choice studies. Different formulations were tested and Value Learning (VL) seemed to be the most interesting as it allows for an individual learning process but has not been as widely investigated as others. The following subsections explain each chosen heuristic together with their properties and model form.

3.4.1. Linear in the Parameters and Additive in the Attributes (LPAA)

Traditional studies assume that individuals take into account to all the attributes that describe an alternative (i.e., all are deemed relevant), and they evaluate them exactly as they are presented to them. This decision process strategy will be referred to as linear in the parameters and additive in the attributes (LPAA). The utility function is as follows:

$$U_{iqt} = \theta_{i1} \cdot x_{i1qt} + \theta_{i2} \cdot x_{i2qt} + \ldots + \theta_{in} \cdot x_{inqt} + \varepsilon_{iqt} \tag{3.22}$$

3.4.2. Value Learning (VL) Proposal

As seen in the previous chapter (Section 2.5.1.1), Value Learning assumes that throughout the experiment individuals' preferences change. The reason for their preferences to change can be caused by the attribute levels presented to them, or often referred to as "rules of the market" (i.e., institutional learning), or simply because they are gaining knowledge and discovering their preferences. In a stated preference study, both can be influencing preferences and the way in which they change is defined in each study. In this research we propose that when valuing the alternatives, individuals compare each of the alternatives' attribute levels to a reference level. For example, there is one reference level for the travel time (whether it is car or public transport) and individuals will compare the travel times of the alternatives to that reference level. When an individual faces a new decision, an attribute's reference level is updated only if the attribute level of the chosen alternative is better than the current reference level. In a stated preference experiment, the starting reference level corresponds to the attribute levels of the mode they use in real life.

One of the simplest model formulations of Value Learning is to consider directly the difference between the attribute levels and reference level, without any type of transformation. In this case, the observed part of the utility function for alternative *i* can be written as follows:

$$U_{iqt} = \theta_{i1} \cdot \left(x_{i1qt} - ref_1\right) + \theta_{i2} \cdot \left(x_{i2qt} - ref_2\right) + \ldots + \theta_{in} \cdot \left(x_{inqt} - ref_n\right) + \varepsilon_{iqt} \tag{3.23}$$

θ_{in} are the estimates representing the difference between the level of attribute *n* and alternative *i* and the reference level for that same attribute *n*; x_{inq} represents the level of attribute *n* of alternative *i* for the individual *q*; and ref_n represents the reference level for attribute *n*. This is a very simple model formulation and has some restrictions. For example, it collapses to a simple MNL model when the choice context is unlabelled, or when the same attributes are present in all the alternatives and their parameters are considered generic in a labelled experiment. In these cases, the reference levels will be the same for all the alternatives, and since the MNL models are estimated based on the differences, the reference levels will be nulled. Moreover, as has been widely mentioned in literature, the valuation of gains and losses represented by $\left(x_{inqt} - ref_n\right)$ may not be linear. Therefore, a concavity factor φ can be estimated to transform the differences between the attribute levels and the reference levels as follows:

$$U_{iqt} = \theta_{i1} \cdot \left(x_{i1qt} - ref_n\right)^{\varphi} + \theta_{i2} \cdot \left(x_{i2qt} - ref_n\right)^{\varphi} + \ldots + \theta_{in} \cdot \left(x_{inqt} - ref_n\right)^{\varphi} + \varepsilon_{iqt} \qquad (3.24)$$

The difference between the attribute levels and the reference levels can be positive or negative. As φ has to be a continuous parameter (with decimal points), to avoid any estimation problems when the difference is negative, the transformation explained above will be considered as follows:

$$VL(\mathrm{x}_{inqt}) = \begin{cases} \left(x_{inqt} - ref_n\right)^{\varphi} & \text{if } \left(x_{inqt} - ref_n\right) \geq 0 \\ -\left[-\left(x_{inqt} - ref_n\right)\right]^{\varphi} & \text{if } \left(x_{inqt} - ref_n\right) < 0 \end{cases} \qquad (3.25)$$

$$U_{iqt} = \theta_{i1} \cdot VL(\mathrm{x}_{i1qt}) + \theta_{i2} \cdot VL(\mathrm{x}_{i2qt}) + \ldots + \theta_{in} \cdot VL(\mathrm{x}_{inqt}) + \varepsilon_{iqt} \qquad (3.26)$$

Figure 3-1 presents the Value Learning transformation considering different values for φ, where $\varphi > 0$.

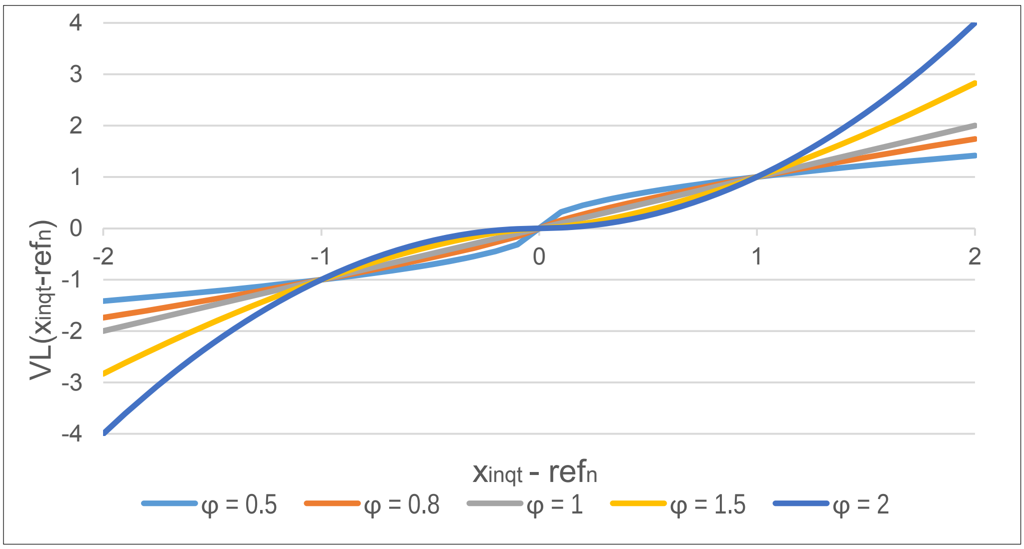


Figure 3-1: Value learning transformation of attribute levels

3.4.2.1. Choice Set Correlation

The relationship between choice sets that is generated by including VL in a stated choice experiment is of interest and should be investigated in more detail. This relationship can be studied through error components to see if the unobserved part of the utility function depends on the choice set number. Similarly to equation (3.21), the utility function can be written as:

$$U_{iq} = V_{iq} + f(\mathrm{i},\mathrm{q},t) \cdot Y_{iqt} + \varepsilon_{iq} \tag{3.27}$$

where the error term is divided in the traditional ε_{iq} and an additional error term that varies across individuals *q*, alternatives *i*, and choice sets *t*, $f(\mathrm{i},\mathrm{q},t) \cdot Y_{iqt}$; where $f(\mathrm{i},\mathrm{q},\mathrm{t})$ is an error component function and Y_{iqt} is a vector of attributes unknown to the modeller, and V_{iq} is the observed part of the utility function.

It is important to note that there might be a significant relationship between the error components and the choice set order number, regardless of the process strategy being considered. However, since the proposition of VL creates a dependence between choice sets it is of special interest to see how the choice set sequence number plays a role in the error components function. Under VL an individual will have evaluated his/her preferences more times as the experiment progresses, although not necessarily updated them. Therefore, the interest of including error components is to see if, by considering a VL process strategy, the error term is being significantly influenced by the progression of the experiment.

The number of a choice set can be included as a continuous variable:

$$f(\mathrm{i},\mathrm{q},t) = \varpi_{iqt} + \delta_i \cdot t \tag{3.28}$$

where ϖ_{iqt} is the part of the error that varies across individuals but it is the same within an individual; t equals the sequence number of a choice set – 1 (i.e., equals 1 for choice set number 2, 2 for choice set number 3, and so on); and δ_i is the parameter associated with the sequence number of a choice set for alternative *i*.

The parameter δ_i represents the relation between the choice set sequence numbers and the error components. This means that, if it has the same sign as ϖ_{iqt}, as the experiment progresses (i.e., the choice set sequence number increases) the standard deviation of the error term increases. This relationship can also be significant when considering a traditional LPAA heuristic, so it will be tested both in a LPAA and VL for comparison. Therefore, the relationship between the choice set sequence numbers and error components will be compared using the values of: ϖ_{iqt}; δ_i; and the quotient between them, δ_i / ϖ_{iqt}, which will represent the influence of the choice set sequence numbers on the error component.

3.4.3. Relative Advantage Maximisation (RAM)

The relative advantage maximisation (RAM) heuristic that will be used in this research is the one proposed by Leong and Hensher (2014) and Leong and Hensher (2015) seen in Section 2.4.3.2. It considers some properties from the classical RRM functional form, and has the symmetry between advantage and disadvantage as proposed by Tversky and Simonson (1993). The definition of advantage of an alternative *i* over an alternative *j* is equivalent to the disadvantage of the alternative *j* over alternative *i* over all attributes *m*, as follows:

$$A(\mathrm{j},\mathrm{i}) = D(i,\mathrm{j}) = \sum_m \ln(1+\exp(\theta_{jm}\cdot x_{jm} - \theta_{im}\cdot x_{im})) \tag{3.29}$$

The relative advantage of *i* over *j*, denoted as $R(i,\mathrm{j})$ can be written as:

$$R(i,\mathrm{j}) = \frac{A(i,\mathrm{j})}{A(i,\mathrm{j}) + D(i,\mathrm{j})} \tag{3.30}$$

The utility function for an alternative *i* will be equivalent to the utility function under an LPAA heuristic plus the relative advantage of alternative *i* over all other alternatives *j* as follows:

$$U_{iqt} = \theta_{i1}\cdot x_{i1qt} + \theta_{i2}\cdot x_{i2qt} + \ldots + \theta_{in}\cdot x_{inqt} + \sum_{j\in S} R(i,j) + \varepsilon_{iqt} \tag{3.31}$$

3.5. Inclusion of Multiple Heuristics, Risk Attitudes, Perceptual Conditioning and Experience

This research opts to not include the subjective self-reported responses to explicit questions regarding heuristics, and instead infers the heuristics and behavioural refinements. The following subsections will describe how these will be inferred. The first subsection presents how the behavioural refinements and experience will be included in the models. The next two subsections present two model forms used in the literature to incorporate process heterogeneity, referred to as Probabilistic Decision Process (PDP) strategy and a Weighting Function. The final subsection will present the model form proposed in this research, called Conditioning of Random Process Heterogeneity (CRPH).

3.5.1. Inclusion of Behavioural Refinements and Experience

Behavioural Refinements

In Section 2.8 the literature on risk attitudes and perceptual conditioning was reviewed. In this research both will be tested by themselves and together. As was defined under Prospect Theory (PT) by Kahneman and Tversky (1979), the utility function for decisions made under risk can be defined by:

$$U = w(p_1)\cdot u(x_1) + w(p_2)\cdot u(x_2) + ... + w(p_n)\cdot u(x_n) \tag{3.32}$$

where $w(p_n)$ represents a weighting function of the probability of occurrence; and $u(x_n)$ represents a value function of the outcome for attribute $\underline{n}$. Hence, risk is measured as the value function and weighting function. In this research, as has been already presented in the literature, it will be used within an EUT framework (defined as EEUT by Hensher et al., 2011a and Hensher et al., 2013).

Perceptual conditioning will be considered for those attributes that are presented with different levels of variation, each with a probability of occurrence. On the other hand, risk attitudes will be considered towards the attributes that are presented with levels of variation and also towards cost attributes. Several functional forms have been proposed to define the weighting and value function (presented in Subsection 2.8). This research will consider those that have proven to significantly improve the model results and understanding of preferences. Namely, CRRA for risk attitudes (value function) and Tversky and Kahneman for perceptual conditioning (weighting function).

Risk attitudes and perceptual conditioning will be included in a LPAA heuristic in two cases:

(1) Attributes that are presented with levels of variation: will take into account risk attitudes and perceptual conditioning.
(2) Cost attributes: will take into account risk attitudes.

If none are included, the LPAA utility function can be written as:

$$U_{iqt} = \theta_{i1} \cdot x_{i1qt} + \theta_{i2} \cdot x_{i2qt} + \ldots + \theta_{in} \cdot x_{inqt} + \varepsilon_{iqt} \tag{3.33}$$

For explanation purposes, we will consider a utility function that has three attributes: a cost attribute, x_{i1qt} ; an attribute presented with *L* levels of variation, x_{i2qt} ; and a third one with no levels of variation and which does not represent cost , x_{i3qt} . The utility function can be written as:

$$U_{iqt} = \theta_{i1} \cdot x_{i1qt} + \sum_{l \in L} w(p_{i2qt,l}) \cdot u(x_{i2qt,l}) + \theta_{i3} \cdot x_{i3qt} + \varepsilon_{iqt} \tag{3.34}$$

where $x_{i2qt,l}$ is the level *l* of variation for attribute x_{i2qt} and $p_{i2qt,l}$ is the associated probability of occurrence; $w(p_{i2qt,l})$ is the weighting function which will be referred as perceptual conditioning; and $u(x_{i2qt,l})$ the value function which is referred to as risk attitudes.

The utility function when considering behavioural refinements can be written as:

$$U_{iqt} = \theta_{i1} \cdot x_{i1qt} + \theta_{i2} \cdot \sum_{l \in L} \left[\frac{p_{i2qt,l}{}^{\gamma}}{(p_{i2qt,l}{}^{\gamma} + (1 - p_{i2qt,l})^{\gamma})^{1/\gamma}} \cdot \frac{x_{i2qt,l}{}^{1-\alpha_2}}{1-\alpha_2} \right] + \theta_{i3} \cdot \left[\frac{x_{i3qt}{}^{1-\alpha_3}}{1-\alpha_3} \right] + \varepsilon_{iqt} \tag{3.35}$$

where γ represents the degree of curvature of the weighting function, and α represents the risk attitudes towards the second and third attribute. The degree of curvature of the weighting function, γ is a positive parameter. An estimated parameter γ with a value between 0 and 1 suggests that individuals will overweight low probabilities and underweight medium to high probabilities; and a parameter larger than 1 suggests that individuals will underweight low to medium probabilities and overweight high probabilities. Figure 3-2 graphically shows the weighting function for different values of γ .

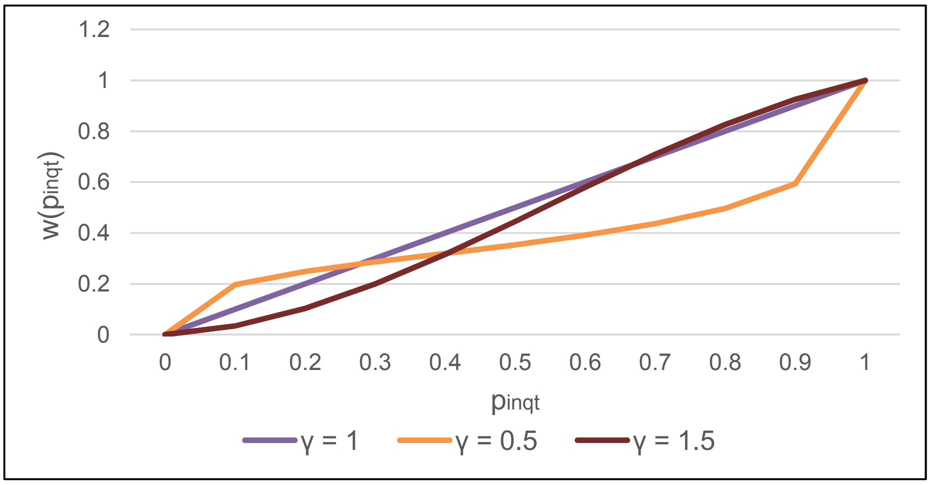


Figure 3-2: Tversky and Kahneman functional form for perceptual conditioning

The interpretation of the risk attitudes for different values of α are presented in Figure 3-3. An estimated parameter of $\alpha = 0$ would represent risk neutrality (i.e., the transformation is not necessary); $\alpha < 0$ represents a risk taking attitude; and $\alpha > 0$ risk aversion.

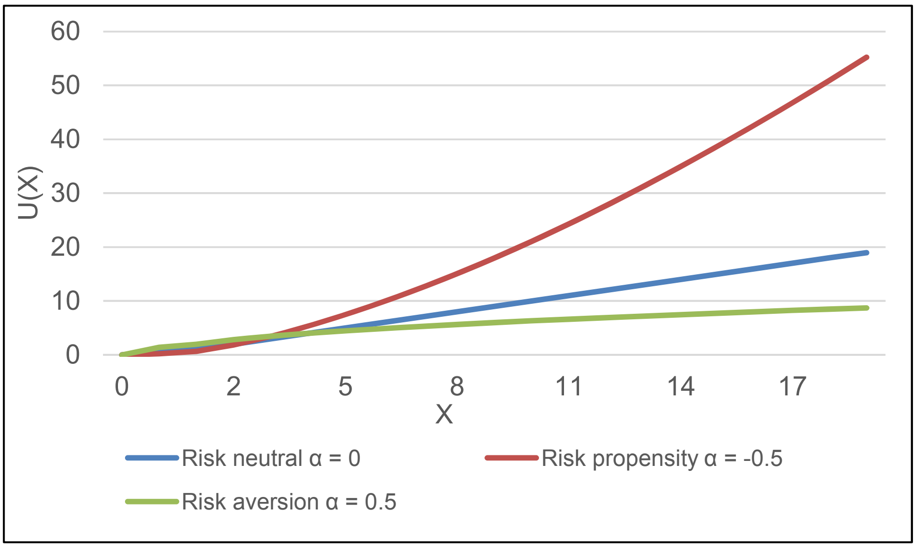


Figure 3-3: CRRA functional form for risk attitudes

Experience

This research will consider experience as the alternative chosen in their most recent decision, which has been proven to be significant in different studies analysed in Section 2.9. In a transportation context, this refers to the most recent mode used by the respondent. It will be included by conditioning the utility function, as follows:

$$U_{iqt}^{experience} = \left(1 + \theta_{exp,i}^{0} \cdot x_{experience,iq}\right) \cdot U_{iqt} \tag{3.36}$$

where $U_{iqt}^{experience}$ is the transformed utility function; $x_{experience}$ is defined as a dummy variable equal to 1 if alternative *i* was chosen by respondent *q* in their most recent decision and 0 otherwise; and $\theta_{exp,i}^{0}$ is the associated parameter. As $x_{experience}$ is always positive, then a negative value of $\theta_{exp,i}^{0}$ would represent a reduction in the utility of mode *i* when individual *q* chose it in his most recent decision; and if it is positive it would represent an increase in utility. However, the interpretation will depend on the sign of the utility, U_{iqt}. If it represents a dis-utility (i.e., negative utility), then a negative parameter of $\theta_{exp,i}^{0}$ would improve the dis-utility of alternative *i* by reducing it, and a positive parameter would worsen alternative *i* by increasing its negative value.

Table 3-1: Experience parameter interpretation

Parameter Sign	Utility sign	If individual *q* chooses alternative *i* in his most recent trip, then…
$\theta_{exp,i}^{0} < 0$	$U_{iqt} < 0$	He is more likely to choose it
	$U_{iqt} > 0$	He is less likely to choose it
$\theta_{exp,i}^{0} > 0$	$U_{iqt} < 0$	He is less likely to choose it
	$U_{iqt} > 0$	He is more likely to choose it
$\theta_{exp,i}^{0} = 0$	Any case	Does not have any implication on his current decision

As we know, the utility function of mode *i* can be negative in one scenario and positive in another scenario depending on its characteristics. Thereby, the modeller needs to define a new parameter with a consistent interpretation across scenarios. The transformed parameter, $\theta_{exp,i}$, will be estimated as follows:

$$\theta_{exp,i} = \begin{cases} \theta_{exp,i}^{0} & \text{if } U_{iqt} \geq 0 \\ -\theta_{exp,i}^{0} & \text{if } U_{iqt} < 0 \end{cases} \tag{3.37}$$

where the utility function will become:

$$U_{iqt}^{experience} = \left(1 + \theta_{exp,i} \cdot x_{experience,iq}\right) \cdot U_{iqt} \tag{3.38}$$

In transportation, one would expect that if an individual uses a certain mode he is more likely to choose it again. Hence, the hypothesis is that if the respondent used the mode in a most recent trip, he will scale the (dis)utility for that mode in such a way that it becomes (smaller) larger, so that $0<\theta_{exp,i}<1$. If the opposite is true, where an individual is less likely to choose the same alternative he used in his most recent trip, then $-1<\theta_{exp,i}<0$. However, the absolute value of $\theta_{exp,i}$ has to be smaller than 1, otherwise it would imply a change in the sign of the utility when it is negative.

This variable will introduce a form of heteroscedastic conditioning in the utility function to account for the individual-specific experience. It also introduces correlation between choice sets, which is taken into consideration through a panel data treatment that relaxes the assumption of independence across choice sets within a respondent. As this conditioning is a form of scaling the utility, it will not change the distribution of the error term.

3.5.2. Probabilistic Decision Process (PDP) and Combination

One of the methods to include multiple heuristics is through a latent class model. Every class represents a different processing strategy, and every individual can belong to each class with a certain probability. This assigned probability could be considered as a function of other characteristics, such as the socioeconomic characteristics of respondents. However, the modeller implicitly assumes that each person only uses one decision process strategy. Several choice studies have used this approach to include multiple processing strategies (Swait and Adamowicz, 2001; Hensher and Collins, 2011; Campbell et al., 2012; Hess et al., 2012; Weller et al., 2014).

In this research three heuristics will be included: RAM, VL and LPAA. Therefore, the model equations for each class will be equivalent to the utility functions for each heuristic, as follows:

$$V_{iqt|C1} = \theta_{i1|C1} \cdot x_{i1qt} + \theta_{i2|C1} \cdot x_{i2qt} + \ldots + \theta_{in|C1} \cdot x_{inqt} + \sum_{j \in S} R(i, j) \tag{3.39}$$

$$V_{iqt|C2} = \theta_{i1|C2} \cdot \left[\left(x_{i1qt} - ref_n \right) \right]^{\varphi} + \theta_{i2|C2} \cdot \left[\left(x_{i2qt} - ref_n \right) \right]^{\varphi} + \ldots + \theta_{in|C2} \cdot \left[\left(x_{inqt} - ref_n \right) \right]^{\varphi} \tag{3.40}$$

$$V_{iqt|C3} = \theta_{i1|C3} \cdot x_{i1qt} + \theta_{i2|C3} \cdot x_{i2qt} + \ldots + \theta_{in|C3} \cdot x_{inqt} \tag{3.41}$$

The class assignment model will be considered using the MNL form presented in Section 3.3. The class assignment can be formulated as only class specific constants or can be subject to other variables, such as the socioeconomic characteristics of respondents.

The combined approach of the Probabilistic Decision Process (PDPC) is an extension of this framework and, in contrast to PDP, which is also specified with a latent class structure, has each class representing a combination of heuristics, instead of only one heuristic. The underlying theory of the PDPC is that multiple heuristics are not mutually exclusive and a person might be using more than one simultaneously. That is, an individual might not necessarily be choosing between one heuristic or the other, but might be choosing a combination of them (Balbontin et al., 2017b). This approach will not be considered in this research, and only the PDP will be included.

To incorporate behavioural refinements and experience in this model form, the modeller has to decide which heuristic will be subject to them. In this thesis, we will consider behavioural refinements in the LPAA heuristic, given that the other process strategies take into account a type of risk attitude. However, experience will be considered as influencing decisions made under any of the process strategies.

3.5.3. Hybrid RUM-RRM and Heuristic Weighting Function (HWF)

The hybrid RUM-RRM model considers each heuristic as influencing a subset of the attributes. The heuristics are included directly in the utility function, and each one only considers a group of the attributes. It is worth noting that each attribute is only related to one heuristic. This form was proposed by Chorus et al. (2013), where they combine a RRM with a RUM model. In the utility function of each alternative, they suggested that the RRM was applicable to only a subset of the attributes, with the other attributes treated as a standard RUM model. When considering this hybrid model structure, their results show an improvement in the fit in comparison to a model that considers RUM or RRM by themselves. However, the difference between the models is not large.

Another way to include multiple heuristics directly in the utility function is by weighting multiple decision process strategies directly in the utility function (Leong and Hensher, 2012b; Hensher et al., 2013; Hensher et al., 2017). The weighting value can be considered as a function of the individual's socioeconomic characteristics or of other context factors. The heuristics may have different weighting functions. In general terms, this methodology allows each heuristic to contribute to the overall utility function, where the contribution is proportional to its weighting value. If we want to include *H* heuristics in the utility function, the model form would be as follows:

$$V_{iqt} = W_1 \cdot U_{H1,iqt} + \ldots + W_h \cdot U_{Hh,iqt} \tag{3.42}$$

where W_h is the weight for heuristic *h* and $U_{Hh,iqt}$ is the utility function of heuristic *h* for alternative *i*, individual *q* and choice situation *t*. The relationship between the weights has to be defined by the modeller and could, for example, be:

$$W_1 + \ldots + W_h = 1 \tag{3.43}$$

Equivalent to the PDP approach, when considering behavioural refinements the modeller has to decide in which heuristic to include them. Experience can be incorporated in each heuristic utility or in the overall utility function.

3.5.4. Conditioning of Random Process Heterogeneity (CRPH)

As the literature on process heuristics grows in interest within a discrete choice setting, and especially where preference heterogeneity is increasingly accommodated by a random parameter specification, the question arises of whether there is a systematic relationship between random parameters as a representation of preference heterogeneity and one or more process heuristics. That is, is there a relationship between preference heterogeneity and process heterogeneity such that process heterogeneity, as represented by specific heuristics, conditions the distribution of preferences in a sampled population in such a way that it adds a systematic (in contrast to random) explanation of preference heterogeneity?

This thesis develops a framework to investigate these questions, called Conditioning of Random Process Heterogeneity (CRPH). The approach recognises that the parameters defined under LPAA may be conditioned by a process strategy. It analyses the degrees of potential substitution or complementarity between the non-systematic representation of preference heterogeneity through random parameters and a systematic representation through a conditioning of the heterogeneous preference distribution, where the latter may offer up a behaviourally richer (and statistically improved) explanation of the choice process.

The random parameter specification, as was analysed previously, decomposes parameter θ_q in its mean, θ^m, and standard deviation, $\sigma \cdot v$:

$$U_{iqt} = (\theta + \sigma \cdot v) \cdot X_{iqt} + \varepsilon_{iqt} \tag{3.44}$$

To incorporate process heuristics using the CRPH approach, the mean and standard deviation of attribute x_{inqt} under an LPAA mixed logit model have to be a function of the process heuristics. The utility can be written as follows:

$$U_i = \sum_n \left(\begin{bmatrix} \theta_{in} + \lambda^m_{VL,in} \cdot VL(x_{inqt}) + \lambda^m_{RAM,in} \cdot RAM(x_{inqt}) \\ + \left[\sigma_{in} + \lambda^s_{VL,in} \cdot VL(x_{inqt}) + \lambda^s_{RAM,in} \cdot RAM(x_{inqt}) \right] \cdot v \end{bmatrix} \cdot x_{inqt} \right) + \varepsilon_{iqt} \tag{3.45}$$

where $VL(x_{inqt})$ represents the transformation of x_{inqt} for the VL heuristic; $RAM(x_{inqt})$ for the RAM heuristic; $\lambda^m_{VL,in}$ represents the relationship between the mean estimate and VL; $\lambda^m_{RAM,in}$ represents the relationship between the mean and RAM; $\lambda^s_{VL,inq}$ the relationship between the standard deviation and VL; $\lambda^s_{RAM,inq}$ the relationship between the standard deviation and RAM.

The parameters estimated in RAM are equivalent to the ones in LPAA, since both are related to the direct values of the attribute: x_{inqt}. However, they are not equivalent to the ones estimated in a VL heuristic since they represent the difference between the attribute level and the reference level: $(x_{inqt} - ref_n)^\varphi$. Therefore, the parameters θ_{in} can be considered common between LPAA and RAM, but not for VL, which will have its own parameters θ^{VL}_{in}. The transformations of x_{inqt} associated with VL and RAM are as follows:

$$VL(x_{inqt}) = \theta^{VL}_{in} \cdot (x_{inqt} - ref_n)^\varphi \tag{3.46}$$

$$RAM\left(x_{inqt}\right) = \theta_{in} \cdot x_{inqt} + \sum_{j \in S} R(i,j) \tag{3.47}$$

Merging Equation (3.38) with (3.39) and (3.40) the expression for CRPH to include VL and RAM results in the following form:

$$U_i = \sum_n \left(\left[\begin{array}{l} \theta_{in} + \lambda^m_{VL,in} \cdot \left(\theta^{VL}_{in} \cdot \left(x_{inqt} - ref_n \right)^{\varphi} \right) + \lambda^m_{RAM,in} \cdot \left(\theta_{in} \cdot x_{inqt} + \sum_{j \in S} R(i,j) \right) \\ + \left[\sigma_{in} + \lambda^s_{VL,in} \cdot \left(\theta^{VL}_{in} \cdot \left(x_{inqt} - ref_n \right)^{\varphi} \right) + \lambda^s_{RAM,in} \cdot \left(\theta_{in} \cdot x_{inqt} + \sum_{j \in S} R(i,j) \right) \right] \cdot v \end{array} \right] \cdot x_{inqt} \right) + \varepsilon_{iqt} \tag{3.48}$$

As can be noted, the part of the equation that considers the VL transformation includes $\lambda_{VL,in} \cdot \theta^{VL}_{in}$. As both are alternative and attribute specific, it is not possible to estimate them both. Therefore, a parameter including the multiplication will be estimated equal to:

$$\lambda^{'}_{VL,in} = \lambda_{VL,in} \cdot \theta^{VL}_{in} \tag{3.49}$$

$$U_i = \sum_n \left(\left[\begin{array}{l} \theta_{in} + \lambda^{'m}_{VL,in} \cdot \left(x_{inqt} - ref_n \right)^{\varphi} + \lambda^m_{RAM,in} \cdot \left(\theta_{in} \cdot x_{inqt} + \sum_{j \in S} R(i,j) \right) \\ + \left[\sigma_{in} + \lambda^{'s}_{VL,in} \cdot \left(x_{inqt} - ref_n \right)^{\varphi} + \lambda^s_{RAM,in} \cdot \left(\theta_{in} \cdot x_{inqt} + \sum_{j \in S} R(i,j) \right) \right] \cdot v \end{array} \right] \cdot x_{inqt} \right) + \varepsilon_{iqt} \tag{3.50}$$

The λ_{in} can be considered common between the attributes or specific. If they are considered common then the relationship between the alternative process strategies and the mean or standard deviation estimate will be the same for all the attributes, which is what is assumed under the PDP and HWF. Hence, one of the major advantages of this approach is that the λ_{in} parameters can be considered as attribute-specific (i.e., depend on *n*) to allow for individuals to use alternative process heuristics for some attributes but not for all of them. If this is the case, the attributes that are not being influenced by a process heuristic would simply have a $\lambda^m_{in} = \lambda^s_{in} = 0$ (in its mean and standard deviation). It also allows process strategies to have an influence over the mean but not standard deviation of an attribute with $\lambda^m_{in} = 0$ and $\lambda^s_{in} \neq 0$ or, oppositely, over its standard deviation but not over its mean with $\lambda^m_{in} \neq 0$ and $\lambda^s_{in} = 0$.

Another advantage of this approach is that behavioural refinements and experience are considered to be independent of the process strategies. That is, risk attitudes and perceptual conditioning will only influence x_{inqt} and experience will affect the entire utility function.

All the models in this thesis were estimated using PythonBiogeme (Bierlaire, 2016). The optimisation algorithm used was SQP method (Lawrence et al., 1994) and the models were estimated using 100 draws. Several of the models were also estimated using 500 draws to test for stability in the results, and the results were equivalent. For the CRPH models, the mixed logit model parameter estimates were used as starting values.

3.6. Willingness to Pay Estimates

This section focuses on the estimation of the marginal (dis)utilities, willingness to pay (WTP) estimates and their confidence levels. It will start by explaining the generalised concepts, and the subsections will describe the specific cases for each process strategy (with fixed and random parameters) and for the process heterogeneity approaches (PDP and CRPH).

Marginal (dis)utilities can be explained as the change in the utility due to a one unit increase in a certain attribute. In transportation, many of the attributes describing a mode represent a dis-utility, e.g., travel time, cost, number of transfers, etc. That is, when they increase there is a negative influence on the utility and that is why we refer to them as marginal disutilities. For other attributes, such as % probability of finding a seat, they will represent marginal utilities. Numerically, the marginal (dis)utilities equals the derivative of the utility relative to attribute *n*: $\frac{\partial U_{iqt}}{\partial x_{inqt}}$.

The willingness to pay estimates (WTP) represent how much is a person willing to pay for a one unit increase (decrease) in a certain attribute. The WTP estimates for an attribute x_{inqt} can be expressed as follows:

$$WTP\left(x_{inqt}\right) = \frac{\partial U_{iqt} / \partial x_{inqt}}{\partial U_{iqt} / \partial x_{\$.iqt}} \tag{3.51}$$

where $x_{\$,iqt}$ represents the cost attribute; and $\partial U_{iqt} / \partial x_{inqt}$ the marginal (dis)utilities.

In some cases, as will be seen later in this thesis, costs are described using more than one attribute. For example, the costs of using a car be described by fuel costs and parking costs. If this is the case, a weighted average will be calculated for the marginal (dis)utilities of the cost attributes which will be used to calculate the WTP (Hensher et al., 2012; Hensher et al., 2013a; Leong and Hensher, 2015):

$$\frac{\partial U_{iqt}}{\partial x_{\$,iqt}} = \frac{\partial U_{iqt} / \partial x_{\$1,iqt} \cdot x_{\$1,iqt} + \partial U_{iqt} / \partial x_{\$2,iqt} \cdot x_{\$2,iqt}}{x_{\$1,iqt} + x_{\$2,iqt}} \tag{3.52}$$

The majority of the models presented in this thesis are non-linear and, therefore, their derived WTP estimates are subject to the value of the attributes and other relevant parameters. A WTP is estimated for each individual given the attribute levels presented to them.

3.6.1. Linear in the Parameters and Additive in the Attributes

3.6.1.1. Simple model

In a LPAA heuristic without any behavioural refinements, the marginal (dis)utilities are obtained straight from the model and are equal to the parameter estimates that describe the attributes. For alternative *i* and attribute *n* the marginal (dis)utility will be equal to θ_{in}, therefore, the WTP will be equal to $\theta_{in} / \theta_{i\$}$, where $\theta_{i\$}$ represents the cost attribute.

3.6.1.2. With Behavioural Refinements

Behavioural refinements will be considered when including a LPAA heuristic. If risk attitudes and perceptual conditioning are included, the utility function is non-linear, hence the derivative changes. As seen in Subsection 3.5.1, perceptual conditioning will be considered for those attributes that are presented with different levels of variation, which in this research are the travel times. On the other hand, risk attitudes will be considered towards travel times and costs. The marginal (dis)utilities will change depending on which of these is considered (risk attitudes and/or perceptual conditioning), and the different cases will be analysed below. The WTP for attribute *n* will be the marginal (dis)utility of attribute *n* divided by the marginal (dis)utility of the cost attribute.

Only Risk Attitudes

If an attribute *n* for individual *i* and choice set *t,* x_{inqt} , is presented with *L* levels of variation each with a certain probability of occurrence $p_{inqt,l}$, it can be written as:

$$x_{inqt} = \sum_{l \in L} w(p_{inqt,l}) \cdot u(x_{inqt,l}) \tag{3.53}$$

where $x_{inqt,l}$ represents the level *l* for attribute *n*; $w(p_{inqt,l})$ represents the weighting function for the probability of occurrence for attribute *n* level *l*, i.e., perceptual conditioning; and $u(x_{inqt,l})$ represents the transformation for the value of attribute *n* level *l*, i.e., risk attitudes. If an attribute considers risk attitudes in the form of CRRA (presented in Section 3.8.1), the marginal (dis)utilities are equal to:

$$\frac{\partial U_{iqt}}{\partial x_{inqt}} = \theta_{in} \cdot x_{inqt}^{-\alpha_n} \tag{3.54}$$

Only Perceptual Conditioning

If an attribute is conditioned on perceptual conditioning (regardless of its functional form), but not on risk attitude, then the marginal (dis)utility is equal to:

$$\frac{\partial U_{iqt}}{\partial x_{inqt}} = \theta_{in} \cdot \sum_{l \in L} w(p_{inqt,l}) \tag{3.55}$$

Risk Attitudes and Perceptual Conditioning

If an attribute is specified with risk attitude and perceptual conditioning, the marginal (dis)utility will depends on the particular functional form used. In this research, risk attitude is included using CRRA and perceptual conditioning is included using the Tversky and Kahneman form. The marginal (dis)utility will be as follows:

$$\frac{\partial U_{iqt}}{\partial x_{inqt}} = \theta_{in} \cdot \sum_{l \in L} w(p_{inqt,l}) \cdot x_{inqt,l}{}^{-\alpha_n} = \theta_{in} \cdot \sum_{l \in L} \frac{p_{inqt,l}{}^{\gamma_i}}{\left[p_{inqt,l}{}^{\gamma_i} + \left(1 - p_{inqt,l}\right)^{\gamma_i} \right]^{\frac{1}{\gamma_i}}} \cdot x_{inqt,l}{}^{-\alpha_n} \quad (3.56)$$

3.6.2. Relative Advantage Maximisation

When considering Relative Advantage Maximisation (RAM) as the process strategy, the marginal (dis)utility is more complex than when considering LPAA since the utility function is non-linear. Recalling the utility function for this heuristic:

$$U_{iqt} = \theta_{i1} \cdot x_{i1qt} + \theta_{i2} \cdot x_{i2qt} + \ldots + \theta_{in} \cdot x_{inqt} + \sum_{j \in S} \frac{A(i,\mathrm{j})}{A(i,\mathrm{j}) + D(i,\mathrm{j})} \quad (3.57)$$

where $A(i,\mathrm{j})$ is the advantage of *i* over *j*, and $D(i,\mathrm{j})$ the disadvantage of *i* over *j*, defined as:

$$A(\mathrm{j},\mathrm{i}) = D(i,\mathrm{j}) = \sum_{n} \ln(1 + \exp(\theta_{jn} \cdot x_{jn} - \theta_{in} \cdot x_{in})) \quad (3.58)$$

the marginal (dis)utility is equal to:

$$\frac{\partial U_{iqt}}{\partial x_{inqt}} = \theta_{in} + \sum_{j \in S} \frac{D(i,\mathrm{j}) \cdot \frac{\partial A(i,\mathrm{j})}{\partial x_{inqt}} - A(i,\mathrm{j}) \cdot \frac{\partial D(i,\mathrm{j})}{\partial x_{inqt}}}{\left[A(i,\mathrm{j}) + D(i,\mathrm{j})\right]^2} \quad (3.59)$$

where the derivatives for the advantage and disadvantage of alternative *i* over alternative *j*, relative to attribute *n* are:

$$\frac{\partial A(i,\mathrm{j})}{\partial x_{inqt}} = \frac{\theta_{in}}{1 + \exp(-\theta_{in} \cdot x_{in} + \theta_{jn} \cdot x_{jn})} \quad (3.60)$$

$$\frac{\partial D(i,\mathrm{j})}{\partial x_{inqt}} = \frac{-\theta_{in}}{1 + \exp(\theta_{in} \cdot x_{in} - \theta_{jn} \cdot x_{jn})} \quad (3.61)$$

The WTP will be equivalent to the marginal (dis)utility of attribute *n* divided by the marginal (dis)utility of the cost attribute.

3.6.3. *Value Learning*

The Value Learning (VL) heuristic considers a concavity factor φ transforming the difference between the attribute level and the reference level. If we remember the utility function for VL presented in Subsection 3.4.2 is as follows:

$$U_{iqt} = \theta_{i1} \cdot \left(x_{i1qt} - ref_n\right)^{\varphi} + \theta_{i2} \cdot \left(x_{i2qt} - ref_n\right)^{\varphi} + \ldots + \theta_{in} \cdot \left(x_{inqt} - ref_n\right)^{\varphi} + \varepsilon_{iqt} \tag{3.62}$$

The marginal (dis)utility for attribute x_{inqt} is represented by:

$$\frac{\partial U_{iqt}}{\partial x_{inqt}} = \theta_{in} \cdot \left[\varphi \cdot \left(x_{inqt} - ref_n\right)^{\varphi - 1}\right] \tag{3.63}$$

If $\varphi = 0$ the marginal (dis)utility would be equal to 0. If $\varphi = 1$, then it would be equivalent to the LPAA model, that is equal to θ_{in}. The WTP will be equivalent to the marginal (dis)utility of attribute *n* divided by the marginal (dis)utility of the cost attribute.

3.6.4. *Models with Experience*

If any of the heuristics above include experience as explained in Subsection 3.5.1, the marginal (dis)utilities will be transformed as follows:

$$\left[\frac{\partial U_{iqt}}{\partial x_{inqt}}\right]' = \left(1 + \beta_{i,\exp} \cdot x_{iq,\exp}\right) \cdot \left[\frac{\partial U_{iqt}}{\partial x_{inqt}}\right] \tag{3.64}$$

where $x_{iq,\exp}$ is a dummy variable equal to 1 if individual *q* chooses alternative *i* in his most recent experience, i.e., in his most recent trip; and $\beta_{i,\exp}$ is the associated parameter. The WTP estimate will not be affected by the experience parameter, since all the attributes in the same alternative will be conditioned by the same experience function, so they will be nulled and the expected WTP equation (but not the numerical estimates) will be equivalent to the model that does not consider experience.

3.6.5. Process Heterogeneity

When considering more than one heuristic, the marginal (dis)utilities change and have to consider all the different forms. The two approaches that will be included is the Probabilistic Decision Process (PDP) strategy and Conditioning of Random Process Heterogeneity (CRPH).

3.6.5.1. Probabilistic Decision Process

The Probabilistic Decision Process (PDP) approach estimates a different process strategy in each class. To obtain the total marginal (dis)utility we have to weigh the marginal (dis)utility by the probability of choosing that class. If we have *C* classes, then the total marginal (dis)utility can be written as follows:

$$\frac{\partial U_{iqt}}{\partial x_{inqt}} = \Pr(\mathrm{Class}_1)\cdot\left[\frac{\partial U_{iqt}}{\partial x_{inqt}}\right]_{Class1} + \ldots + \Pr(\mathrm{Class}_C)\cdot\left[\frac{\partial U_{iqt}}{\partial x_{inqt}}\right]_{ClassC} \tag{3.65}$$

3.6.5.2. Conditioning of Random Process Heterogeneity

The Conditioning of Random Process Heterogeneity (CRPH) approach conditions the mean and standard deviation of the parameter estimates by the process heuristics. The mean of the CRPH utility function that can be extracted from Equation (3.38) (remembering $v \sim N\left[0,1\right]$) equals to:

$$U_i = \sum_n \left(\left[\theta_{in} + \lambda_{VL,in}^{'m}\cdot VL\left(x_{inqt}\right) + \lambda_{RAM,in}^{m}\cdot RAM\left(x_{inqt}\right)\right]\cdot x_{inqt}\right) \tag{3.66}$$

where:

$$VL\left(x_{inqt}\right) = \theta_{in}^{VL}\cdot\left(x_{inqt} - ref_n\right)^{\varphi} \tag{3.67}$$

$$RAM\left(x_{inqt}\right) = \theta_{in}\cdot x_{inqt} + \sum_{j\in S} R(i,j) \tag{3.68}$$

Hence, the mean marginal (dis) utility can be written as:

$$\frac{\partial U_i}{\partial x_{inqt}} = \sum_n \left(\begin{array}{l} \theta_{in} + \lambda'^{m}_{VL,in} \cdot VL\left(x_{inqt}\right) + \lambda^{m}_{RAM,in} \cdot RAM\left(x_{inqt}\right) \\ + \left[\theta_{in} + \lambda'^{m}_{VL,in} \cdot \frac{\partial VL\left(x_{inqt}\right)}{\partial x_{inqt}} + \lambda^{m}_{RAM,in} \cdot \frac{\partial RAM\left(x_{inqt}\right)}{\partial x_{inqt}} \right] \cdot x_{inqt} \end{array} \right) \tag{3.69}$$

where $\frac{\partial VL\left(x_{inqt}\right)}{\partial x_{inqt}}$ represents the marginal (dis)utility obtained when considering the VL heuristic shown in Equation (3.63); and $\frac{\partial RAM\left(x_{inqt}\right)}{\partial x_{inqt}}$ when considering the RAM heuristic shown in Equation (3.59).

3.7. Willingness to Pay Confidence Intervals

The willingness to pay (WTP) estimates are one of the most important outcomes of choice studies. The equations presented above are used to obtain the expected WTP estimate for the sample. Frequently, we see studies that do not present confidence intervals for the WTP estimates, even though they are as important as the value itself because they represent how robust the estimates are.

This thesis will use the Delta method to calculate the standard error of the WTP estimates (Oehlert, 1992; Scarpa and Rose, 2008; Bliemer and Rose, 2013; Hensher et al. 2015). This method can be used to calculate the standard error of any function of the parameters, as are the marginal (dis)utilities. The WTP in a linear MNL model are equivalent to the ratio between two parameters, so the latter one will be explained in more detail since it is a bit more complex, but the same method will be used to calculate the confidence intervals of the marginal (dis)utilities. The delta method states that if $\hat{\theta}$ is asymptotically distributed, then a function of $f(\hat{\theta})$ is asymptotically normally distributed with a mean of $f(\theta)$ and a variance of:

$$\nabla_\theta f(\theta)^{\mathrm{T}} \Omega_\theta \nabla_\theta f(\theta) \tag{3.70}$$

where $\nabla_\theta f(\theta)$ denotes the Jacobian of $f(\theta)$. In a linear MNL model (LPAA), the WTP estimate for attribute *n* and alternative *i* would be equal to $WTP\left(x_{inqt}\right) = w_{in} = {}^{\theta_{in}}\!/_{\theta_{i\$}}$, where $\theta_{i\$}$ represents the parameter estimate for the cost attribute. The WTP, $\hat{w}_{in}$, would be distributed as:

$$\hat{w}_{in} \sim N\left(\frac{\theta_{in}}{\theta_{i\$}}, \begin{pmatrix} \nabla_{\theta_{in}} w_{in} \\ \nabla_{\theta_{i\$}} w_{in} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{var}(\theta_{in}) & \mathrm{cov}(\theta_{in}, \theta_{i\$}) \\ \mathrm{cov}(\theta_{in}, \theta_{i\$}) & \mathrm{var}(\theta_{i\$}) \end{pmatrix} \begin{pmatrix} \nabla_{\theta_{in}} w_{in} \\ \nabla_{\theta_{i\$}} w_{in} \end{pmatrix}\right)$$

$$\sim N\left(\frac{\theta_{in}}{\theta_{i\$}}, \begin{pmatrix} {}^{1}\!/_{\theta_{i\$}} \\ -{}^{\theta_{in}}\!/_{\theta_{i\$}^{2}} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{var}(\theta_{in}) & \mathrm{cov}(\theta_{in}, \theta_{i\$}) \\ \mathrm{cov}(\theta_{in}, \theta_{i\$}) & \mathrm{var}(\theta_{i\$}) \end{pmatrix} \begin{pmatrix} {}^{1}\!/_{\theta_{i\$}} \\ -{}^{\theta_{in}}\!/_{\theta_{i\$}^{2}} \end{pmatrix}\right) \tag{3.71}$$

The standard error according to the Delta method would be equal to:

$$\frac{1}{\theta_{i\$}} \sqrt{\mathrm{var}(\theta_{in}) - 2w_{in}\,\mathrm{cov}(\theta_{in}, \theta_{i\$}) + w_{in}^{2}\,\mathrm{var}(\theta_{i\$})} \tag{3.72}$$

The example provided above considers a linear MNL model (LPAA); however, in this study we will estimate non-linear mixed logit models, where all random parameters are normally distributed. Bliemer and Rose (2013) explain the methodology to use the Delta method to obtain confidence intervals in linear mixed logit models. The first thing is to re-write the parameter estimates $\hat{\theta}$ equivalent to equation (3.11) in terms of the distributional parameters $\theta^{m}{}_{in}$ and σ_{in} , which together will be referred to as ϑ , and a parameter-free distribution v, as follows:

$$\theta_{in} = \theta(v_{in} \mid \vartheta_{in}) = \theta^{m}{}_{in} + \sigma_{in} \cdot v_{in} \tag{3.73}$$

In this thesis, all random parameters will be normally distributed[9]. The Jacobians (first derivatives) for normal distributed parameters are the following:

---

[9] Different distribution assumptions were tested (e.g., triangular and lognormal) and the normal distribution gave the best fit and behaviourally appealing results.

$$\nabla_{\vartheta}\theta = \begin{pmatrix} \nabla_{\theta^{m}{}_{in}}\theta \\ \nabla_{\sigma_{in}}\theta \end{pmatrix} = \begin{pmatrix} 1 \\ \nu \end{pmatrix} \quad \text{and} \quad \nabla_{\nu}\theta = \sigma \tag{3.74}$$

In a linear mixed logit model (LPAA), where both the cost attribute and attribute *n* are random, the WTP estimate would be equal to:

$$WTP\left(x_{inqt}\right) = w_{in}(\nu_{in}, \nu_{i\$} \mid \vartheta_{in}, \vartheta_{i\$}) = \frac{\theta_{in}(\nu_{in} \mid \vartheta_{in})}{\theta_{i\$}(\nu_{i\$} \mid \vartheta_{i\$})}$$

and its distribution would be:

$$\hat{w}_{in}(\nu_{in}, \nu_{i\$}) \sim N\left(w_{in}, \begin{pmatrix} \nabla_{\vartheta_{in}} w_{in} \\ \nabla_{\vartheta_{i\$}} w_{in} \\ \nabla_{\nu_{in}} w_{in} \\ \nabla_{\nu_{i\$}} w_{in} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \Omega_{\vartheta} & 0 \\ 0 & diag(1,...,1) \end{pmatrix} \begin{pmatrix} \nabla_{\vartheta_{in}} w_{in} \\ \nabla_{\vartheta_{i\$}} w_{in} \\ \nabla_{\nu_{in}} w_{in} \\ \nabla_{\nu_{i\$}} w_{in} \end{pmatrix} \right) \tag{3.75}$$

where $\nabla_{\vartheta_{in}} w_{in}$ and $\nabla_{\vartheta_{i\$}} w_{in}$ represent the first derivatives (Jacobian) of the WTP estimate of attribute x_{inqt} relative to ϑ_{in} and $\vartheta_{i\$}$, respectively; and $\nabla_{\nu_{in}} w_{in}$ and $\nabla_{\nu_{i\$}} w_{in}$ are relative to ν_{in} and $\nu_{i\$}$, respectively. Ω_{ϑ} is the submatrix of the variances and covariances of the distributional parameters ϑ_{in} and $\vartheta_{i\$}$; 0 represents a matrix with zeros and magnitude of $\left(\#\vartheta_{in} + \#\vartheta_{i\$}\right) \times \left(\#\nu_{in} + \#\nu_{i\$}\right)$, where # represent the number of elements of ϑ_{in}, $\vartheta_{i\$}$, ν_{in} and $\nu_{i\$}$; and diag(1,…,1) is a diagonal matrix with ones and magnitude of $\left(\#\nu_{in} + \#\nu_{i\$}\right) \times \left(\#\nu_{in} + \#\nu_{i\$}\right)$.

The Jacobians can be calculated as follows:

$$\begin{aligned} \nabla_{\vartheta_{in}} w_{in} &= \nabla_{\theta_{in}} w_{in} \cdot \nabla_{\vartheta_{in}} \theta_{in} = \frac{1}{\theta_{in}} \cdot \nabla_{\vartheta_{in}} \theta_{in} = \frac{1}{\theta_{in}} \cdot \begin{pmatrix} 1 \\ \nu_{in} \end{pmatrix} \\ \nabla_{\vartheta_{i\$}} w_{in} &= \nabla_{\theta_{i\$}} w_{in} \cdot \nabla_{\vartheta_{i\$}} \theta_{i\$} = -\frac{\theta_{in}}{\theta_{i\$}{}^{2}} \cdot \nabla_{\vartheta_{i\$}} \theta_{i\$} = -\frac{\theta_{in}}{\theta_{i\$}{}^{2}} \cdot \begin{pmatrix} 1 \\ \nu_{i\$} \end{pmatrix} \end{aligned} \tag{3.76}$$

$$\nabla_{v_{in}} w_{in} = \nabla_{\theta_{in}} w_{in} \cdot \nabla_{v_{in}} \theta_{in} = \frac{1}{\theta_{in}} \cdot \nabla_{v_{in}} \theta_{in} = \frac{1}{\theta_{in}} \cdot \sigma_{in}$$

$$\nabla_{v_{i\$}} w_{in} = \nabla_{\theta_{i\$}} w_{in} \cdot \nabla_{v_{i\$}} \theta_{i\$} = -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \cdot \nabla_{v_{i\$}} \theta_{i\$} = -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \cdot \sigma_{i\$}$$

Therefore, equation (3.75) can be written as:

$$\hat{w}_{in}(v_{in}, v_{i\$}) \sim N\left(w_{in}, \begin{pmatrix} \frac{1}{\theta_{in}} \\ \frac{1}{\theta_{in}} \cdot v_{in} \\ -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \\ -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \cdot v_{i\$} \\ \frac{1}{\theta_{in}} \cdot \sigma_{in} \\ -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \cdot \sigma_{i\$} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \Omega_{\vartheta} & 0 \\ 0 & diag(1,...,1) \end{pmatrix} \begin{pmatrix} \frac{1}{\theta_{in}} \\ \frac{1}{\theta_{in}} \cdot v_{in} \\ -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \\ -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \cdot v_{i\$} \\ \frac{1}{\theta_{in}} \cdot \sigma_{in} \\ -\frac{\theta_{in}}{\theta_{i\$}^{\ 2}} \cdot \sigma_{i\$} \end{pmatrix} \right) \tag{3.77}$$

If both parameters are fixed, then $v_{in} = v_{in} = \sigma_{in} = \sigma_{i\$} = 0$ and Equation (3.77) would collapse to (3.72). The expected WTP estimate $\hat{w}_{in}$ is:

$$\hat{w}_{in} = \int_{v_{in}} \int_{v_{i\$}} \hat{w}_{in}(v_{in}, v_{i\$}) \,\mathrm{d}\, F_{in}(v_{in}) \,\mathrm{d}\, F_{i\$}(v_{i\$}) \tag{3.78}$$

where $F_{in}(v_{in})$ and $F_{i\$}(v_{i\$})$ are the cumulative distribution functions of the standard distributed v_{in} and $v_{i\$}$, respectively. Since normal distributions are defined on the complete domain of $(-\infty, +\infty)$, the integral would be undefined when $\theta_{i\$} = 0$. Daly et al. (2012) demonstrate that for the moments to be finite, the probability of observing $\theta_{i\$} = 0$ should be zero as in, for example, a lognormal distribution. Bliemer and Rose (2013) propose to use the median instead of the mean, as it would represent a more robust estimator that will not vary as much as a mean when the cost attribute is normally distributed. When estimating mixed logit models, all the attributes will be considered as normally distributed. The ratio between two normal distributions is a Cauchy distribution, also referred to as Lorentzian distribution or Lorentz distribution.

The expected WTP can be approximated by Monte Carlo simulation as:

$$\hat{w}_{in} \approx \frac{1}{R}\sum_{r=1}^{R} \hat{w}_{in}(v_{in}{}^{(r)}, v_{i\$}{}^{(r)}) \tag{3.79}$$

where $r = 1,\ldots,R$ are pseudo random draws such as Halton sequences to ensure more uniform coverage over the distribution (Train, 1999). The approximation will be more accurate if more draws are used. We will use 25,000 pseudo random draws.

In non-linear models, detailed in the subsections below, the function $\hat{w}_{in}(v_{in}{}^{(r)}, v_{i\$}{}^{(r)})$ involve sample data (e.g., attribute levels). In these cases, they can be evaluated using the mean levels of the data, or they can be averaged over the observations. This thesis will use both methods to verify that the results do not change significantly, although the results presented will consider the first method. The levels for the expected WTP with a confidence level of α, where $se(\hat{w}_{in})$ is the standard error, will be as follows:

$$\left(\hat{w}_{in} - t_{1-\alpha/2} \cdot se(\hat{w}_{in}), \hat{w}_{in} + t_{1-\alpha/2} \cdot se(\hat{w}_{in})\right) \tag{3.80}$$

The previous section presented the equations for the willingness to pay estimates, so the following subsections will show the jacobians/derivatives that have to be included for use in the Delta method to estimate the confidence intervals. This thesis used PythonBiogeme to calculate the derivatives (of the WTP and the Jacobians) for each model (Bierlaire, 2016).

3.7.1. Linear in the Parameters and Additive in the Attributes

The confidence intervals for a simple LPAA model (without behavioural refinements) will be exactly as shown above. However, this will change when considering risk attitudes, perceptual conditioning and/or experience.

The risk attitudes will be considered both towards attribute *n* and the cost attribute using the model form explained in Section 3.5.1. Perceptual conditioning is only considered towards travel time attribute as explained in Section 3.5.1. Experience will be considered towards all the modes available. The risk attitudes, perceptual conditioning and experience parameters will be fixed and the attribute estimates are normally distributed.

Applying the Delta method for a model that includes risk attitudes, perceptual conditioning and experience, we arrive at:

$$\hat{w}_{in}(v_{in},v_{i\$}) \sim N\left(w_{in}, \begin{pmatrix} \nabla_{\theta_{in}^m} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\alpha_n} w_{in} \\ \nabla_{\gamma_n} w_{in} \\ \nabla_{\theta_{i\$}^m} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\alpha_\$} w_{in} \\ \nabla_{v_n} w_{in} \\ \nabla_{v_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \Omega_{\theta^m,\sigma,\alpha,\gamma,\beta_{i,\exp}} & 0 \\ 0 & diag(1,...,1) \end{pmatrix} \begin{pmatrix} \nabla_{\theta_{in}^m} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\alpha_n} w_{in} \\ \nabla_{\gamma_n} w_{in} \\ \nabla_{\theta_{i\$}^m} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\alpha_\$} w_{in} \\ \nabla_{v_n} w_{in} \\ \nabla_{v_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix} \right) \tag{3.81}$$

where α_n and $\alpha_\$$ represent the risk attitudes for attribute *n* and for the cost attribute; γ_n represents perceptual conditioning for attribute *n* when it is equal to the travel time; and, $\beta_{i,\exp}$ represents the parameter estimated for experience on mode *i*. If a model does not consider risk attitudes, perceptual conditioning, or experience then the corresponding derivative will be equal to zero, which would be equivalent to removing it from equation (3.82).

3.7.2. Value Learning

Assuming the concavity factor for the VL heuristic is fixed and common between attributes, and the attribute estimates are all normally distributed, the Delta method would result in:

$$\hat{w}_{in}(\nu_{in},\nu_{i\$}) \sim N\left(w_{in}, \begin{pmatrix} \nabla_{\theta_{in}^m} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\theta_{i\$}^m} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\varphi} w_{in} \\ \nabla_{\nu_n} w_{in} \\ \nabla_{\nu_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \Omega_{\theta^m,\sigma,\varphi,\beta_{i,\exp}} & 0 \\ 0 & diag(1,...,1) \end{pmatrix} \begin{pmatrix} \nabla_{\theta_{in}^m} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\theta_{i\$}^m} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\varphi} w_{in} \\ \nabla_{\nu_n} w_{in} \\ \nabla_{\nu_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix} \right) \quad (3.82)$$

Equation (3.82) is equivalent to equation (3.75) but adding the derivative of the WTP relative to the concavity factor, φ and the experience parameter on mode *i,* $\beta_{i,\exp}$. If either of them are not significant (i.e., $\beta_{i,\exp} = 0$ or $\varphi = 1$), then the corresponding derivative would be equal to zero which is the same as removing them from equation (3.85).

3.7.3. Relative Advantage Maximisation

The RAM heuristic considers the same number of parameters as an LPAA heuristic with experience; however the marginal (dis)utility function of RAM (presented in equation (3.59)) is highly non-linear. Applying the Delta method while assuming the attribute estimates are all normally distributed, we have that:

$$\hat{w}_{in}(\nu_{in},\nu_{i\$}) \sim N\left(w_{in}, \begin{pmatrix} \nabla_{\theta_{in}^m} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\theta_{i\$}^m} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\nu_n} w_{in} \\ \nabla_{\nu_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \Omega_{\theta^m,\sigma,\beta_{i,\exp}} & 0 \\ 0 & diag(1,...,1) \end{pmatrix} \begin{pmatrix} \nabla_{\theta_{in}^m} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\theta_{i\$}^m} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\nu_n} w_{in} \\ \nabla_{\nu_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix} \right) \quad (3.83)$$

which is equivalent to equation (3.75) adding experience on mode *i,* $\beta_{i,\exp}$. If the experience is not significant (i.e., $\beta_{i,\exp} = 0$), then the corresponding derivative would be equal to zero which is the same as removing them from equation (3.91).

3.7.4. Probabilistic Decision Process

The PDP approach considers a latent class structure, where each class is not latent but represents a pre-defined heuristic. Applying the Delta method we have that:

$$\hat{w}_{in}(v_{in}, v_{i\$}) \sim N\left(w_{in}, \begin{pmatrix} \nabla_{\theta_{in|C1}} w_{in} \\ \nabla_{\theta_{in|C2}} w_{in} \\ \nabla_{\theta_{in|C3}} w_{in} \\ \nabla_{\alpha_n} w_{in} \\ \nabla_{\gamma_n} w_{in} \\ \nabla_{\theta_{i\$|C1}} w_{in} \\ \nabla_{\theta_{i\$|C2}} w_{in} \\ \nabla_{\theta_{i\$|C3}} w_{in} \\ \nabla_{\alpha_\$} w_{in} \\ \nabla_{\varphi} w_{in} \\ \nabla_{\beta_{i,\exp|C1}} w_{in} \\ \nabla_{\beta_{i,\exp|C2}} w_{in} \\ \nabla_{\beta_{i,\exp|C3}} w_{in} \\ \nabla_{Cte_{C1}} w_{in} \\ \nabla_{Cte_{C2}} w_{in} \\ \nabla_{Cte_{C3}} w_{in} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \Omega_{\theta,\alpha,\gamma,\varphi,\beta_{i,\exp},Cte_{Class}} & 0 \\ 0 & diag(1,...,1) \end{pmatrix} \begin{pmatrix} \nabla_{\theta_{in|C1}} w_{in} \\ \nabla_{\theta_{in|C2}} w_{in} \\ \nabla_{\theta_{in|C3}} w_{in} \\ \nabla_{\alpha_n} w_{in} \\ \nabla_{\gamma_n} w_{in} \\ \nabla_{\theta_{i\$|C1}} w_{in} \\ \nabla_{\theta_{i\$|C2}} w_{in} \\ \nabla_{\theta_{i\$|C3}} w_{in} \\ \nabla_{\alpha_\$} w_{in} \\ \nabla_{\varphi} w_{in} \\ \nabla_{\beta_{i,\exp|C1}} w_{in} \\ \nabla_{\beta_{i,\exp|C2}} w_{in} \\ \nabla_{\beta_{i,\exp|C3}} w_{in} \\ \nabla_{Cte_{C1}} w_{in} \\ \nabla_{Cte_{C2}} w_{in} \\ \nabla_{Cte_{C3}} w_{in} \end{pmatrix} \right) \tag{3.84}$$

The classes consider specific and fixed parameters for each attribute, as shown in Section 3.5.2. Since the parameters are fixed, $\theta_{in} = \theta^{m}{}_{in}$. Considering $\theta_{in|C1}$, $\theta_{in|C2}$, and $\theta_{in|C3}$ represent the estimate for attribute *n* alternative *i* for the RAM heuristic (class 1), VL heuristic (class 2) and LPAA (class 3), respectively; α_n and $\alpha_\$$ represent the risk attitudes for attribute *n* and for the cost attribute, respectively, for the LPAA heuristic (class 1); γ_n represents perceptual conditioning for attribute *n* (it is different from zero when it refers to the travel time attribute) for the LPAA heuristic (class 1); φ represents the concavity factor for the VL heuristic (class 2); $\beta_{i,\exp|C1}, \beta_{i,\exp|C2}, \beta_{i,\exp|C3}$ represents the parameter estimated for experience on mode *i* for each class; and $Cte_{C1}, Cte_{C2}, Cte_{C3}$ represent the class specific constants for the class assignment.

3.7.5. Conditioning of Random Process Heterogeneity

The CRPH approach considers interactions between the process strategies and the mean and/or standard deviation of the estimates, as presented in Section 3.5.4. Applying the Delta method while assuming the attribute estimates are all normally distributed, we have that:

$$\hat{w}_{in}(\nu_{in}, \nu_{i\$}) \sim N\left(w_{in}, \begin{pmatrix} \nabla_{\theta_{in}^{m}} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\alpha_n} w_{in} \\ \nabla_{\gamma_n} w_{in} \\ \nabla_{\lambda_{VL,in}^{m}} w_{in} \\ \nabla_{\lambda_{RAM,in}^{m}} w_{in} \\ \nabla_{\lambda_{VL,in}^{s}} w_{in} \\ \nabla_{\lambda_{RAM,in}^{s}} w_{in} \\ \nabla_{\theta_{i\$}^{m}} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\alpha_\$} w_{in} \\ \nabla_{\lambda_{VL,i\$}^{m}} w_{in} \\ \nabla_{\lambda_{RAM,i\$}^{m}} w_{in} \\ \nabla_{\lambda_{VL,i\$}^{s}} w_{in} \\ \nabla_{\lambda_{RAM,i\$}^{s}} w_{in} \\ \nabla_{\varphi} w_{in} \\ \nabla_{\nu_n} w_{in} \\ \nabla_{\nu_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \Omega_{\theta,\alpha,\gamma,\lambda,\varphi,\beta_{i,\exp}} & 0 \\ 0 & diag(1,...,1) \end{pmatrix} \begin{pmatrix} \nabla_{\theta_{in}^{m}} w_{in} \\ \nabla_{\sigma_{in}} w_{in} \\ \nabla_{\alpha_n} w_{in} \\ \nabla_{\gamma_n} w_{in} \\ \nabla_{\lambda_{VL,in}^{m}} w_{in} \\ \nabla_{\lambda_{RAM,in}^{m}} w_{in} \\ \nabla_{\lambda_{VL,in}^{s}} w_{in} \\ \nabla_{\lambda_{RAM,in}^{s}} w_{in} \\ \nabla_{\theta_{i\$}^{m}} w_{in} \\ \nabla_{\sigma_{i\$}} w_{in} \\ \nabla_{\alpha_\$} w_{in} \\ \nabla_{\varphi} w_{in} \\ \nabla_{\lambda_{VL,i\$}^{m}} w_{in} \\ \nabla_{\lambda_{RAM,i\$}^{m}} w_{in} \\ \nabla_{\lambda_{VL,i\$}^{s}} w_{in} \\ \nabla_{\lambda_{RAM,i\$}^{s}} w_{in} \\ \nabla_{\nu_n} w_{in} \\ \nabla_{\nu_\$} w_{in} \\ \nabla_{\beta_{i,\exp}} w_{in} \end{pmatrix} \right) \quad (3.85)$$

where α_n and $\alpha_\$$ represent the risk attitudes for attribute *n* and for the cost attribute, respectively; γ_n represents perceptual conditioning for attribute *n* (it is different from zero when it refers to the travel time attribute); $\lambda^{m}_{VL,in}$ and $\lambda^{m}_{VL,i\$}$ represents the relationship between the mean estimate and VL for attribute *n* and the cost attribute; $\lambda^{m}_{RAM,in}$ and $\lambda^{m}_{RAM,i\$}$ represents the relationship between the mean and RAM for attribute *n* and the cost attribute; $\lambda^{s}_{VL,inq}$ and $\lambda^{s}_{VL,i\$q}$ the relationship between the standard deviation and VL for attribute *n* and the cost attribute; $\lambda^{s}_{RAM,inq}$ and $\lambda^{s}_{RAM,i\$q}$ the relationship between the standard deviation and RAM for attribute *n* and the cost attribute; φ represents the concavity factor for the VL heuristic; and $\beta_{i,\exp}$ represents the parameter estimated for experience on mode *i.*

3.8. Results Analysis and Models Comparison

3.8.1. Comparing Two Parameters using t-test

An important part of the model analysis will consist of determining if there are significant differences between the WTP estimates for the different models. The null hypothesis of t-test states that an estimated parameter θ is equivalent to a hypothesized value θ_0, and is rejected if:

$$t = \frac{\theta - \theta_0}{se(\theta)} > t_{\mathrm{df},\%} \tag{3.86}$$

where $se(\theta)$ is the standard error of parameter θ; *df* degrees of freedom equal to the number of observations - 1 (usually very large so considered as infinite) [10]; and % the confidence level.

When testing if two WTP estimates are significantly different from each other, the test is adapted to fit in a parameter difference:

$$t = \frac{(\theta_1 - \theta_2) - 0}{se(\theta_1 - \theta_2)} > t_{\mathrm{df},\%} \tag{3.87}$$

where the standard error of θ_1 and θ_2 are σ_1 and σ_2, respectively, and n_1 and n_2 the number of observations in the sample. In this thesis, the WTP obtained through different

[10] When the sample represents all the population, the degrees of freedom are equal to the number of observations.

models will be compared in the same sample so $n_1 = n_2 = n$. The standard error for the difference will be equal to:

$$se(\theta_1 - \theta_2) = \sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}} = \sqrt{\frac{\sigma_1^{\,2} + \sigma_2^{\,2}}{n}} \qquad (3.88)$$

3.8.2. Comparing Two Models

3.8.2.1. AIC Indicator

The Akaike's Information Criterion (AIC) was proposed by Akaike (1974) and it will be used to compare the models in regards to their overall fit. It takes into account the log likelihood $l(\theta)$ of a model while penalising the number of parameters estimated, $\#Params$. The indicator also takes into account the number of observations $\#Obs$ and it is calculated as follows:

$$AIC = \frac{-2 \cdot l(\theta) + 2 \cdot \#Params}{\#Obs} \qquad (3.89)$$

3.8.2.2. Nested Models

The *Loglikelihood Ratio test* (LR) test can be used to compare two models where one is the restricted version of the other. When a model is the restricted version of a more general one, it can be obtained by adding restrictions with the number of restrictions included representing the degrees of freedom for the test. The null hypothesis of this test states that both models are equivalent. The null hypothesis is rejected if:

$$LR = -2 \cdot \left[l\left(\theta_{restricted}\right) - l\left(\theta_{general}\right) \right] > \chi^2_{df,\%} \qquad (3.90)$$

where $l\left(\theta_{restricted}\right)$ represents the log likelihood of the restricted model, $l\left(\theta_{general}\right)$ the log likelihood of the general model, *df* the degrees of freedom, and % the confidence level. If the test is rejected, there is not enough evidence to say that both models are equivalent; hence, the general model seems to be more adequate. If the test is not rejected, then there is not enough evidence to say that both models are statistically different; thus, the additional parameters estimated in the general model do not seem to improve the model significantly.

3.8.2.3. Not Nested Models

When we want to compare two models that are not nested (i.e., one is not the restricted version of the other), we use the Vuong Statistic (Vuong, 1989). An example of two non-nested models is the model obtained using the CRPH model form, with one obtained using the PDP model form. This statistic is defined as follows:

$$V = \frac{\bar{m}\sqrt{N}}{S_n} \tag{3.91}$$

$\bar{m}$ represents the mean of the difference between the log likelihood of the models: $l(model1) - l(model2)$; N the sample size; and S_n the standard deviation. A value of this statistic greater than the critical t-test value (1.96 with 95% confidence level) indicates that the test favours the first model; and if it is less than the negative of the critical t-test (-1.96 with 95% confidence level) the test favours the second model. Values in between are inconclusive.

3.9. Conclusions

This section has presented the econometric theory of discrete choice modelling that is used to incorporate multiple process strategies as well as behavioural refinements. The first novel part of this section is the proposal to incorporate the Value Learning heuristic and the method to test if it is inducing a choice set correlation. Secondly, it shows how to integrate behavioural refinements with process heterogeneity. Finally, it proposes a new method referred to as *Conditioning Random Process Heterogeneity* (CRPH) to include process heterogeneity that allows the parameter estimates commonly defined under an LPAA approach be conditioned by the process strategies, while considering the behavioural refinements as independent of these strategies. The last two sections detail how the WTP estimates and their confidence intervals will be calculated in these highly non-linear mixed logit models.

CHAPTER 4

Datasets

4.1. Introduction

This chapter presents and analyses the characteristics of the two datasets that will be used in the research. Sections 4.2 and 4.3 present the characteristics, attributes and attribute levels of the Metro Rail Sydney and Northwest datasets, respectively. Section 4.1 compares the characteristics and attributes of the datasets, followed by Section 4.2 that compares the socioeconomic characteristics and travel behaviour of the respondents.

4.2. Dataset #1: Metro Rail Sydney

During 2009, the New South Wales government announced that they proposed to evaluate a new Metro rail system for Sydney. As part of this feasibility study, Hensher et al. (2011) collected stated choice and revealed preference data to estimate a modal choice model to be used in predicting potential patronage for the new Metro. In this thesis, the stated preference data will be used to estimate the models, and the revealed preference data will be used to include experience and the starting point for the value learning heuristic. The proposed Metro lines are presented in Figure 4-1. The sampled areas were selected to reflect travel across the catchment area, which includes the central business district (CBD), and geographical locations within the metropolitan area that extend as far west as Westmead and the north west of Sydney. In the survey, respondents had to compare their currently available alternatives together with a metro option.

The survey was constructed through a Bayesian-efficient design using D-error as the optimality criteria. The collection of the final dataset was conducted using a computer aided personal interview (CAPI). 1,756 respondents were interviewed with each individual choosing to answer questions for a specific trip purpose, one of commuter, non-commuter, other work related trips and trips made inside the central business district area (intra CBD). The candidate modes for these trips are bus, light rail, train, or car. The intra CBD purpose trips also include walking and taxi as alternatives. In this thesis the intra CBD trips were excluded as they are very different from the other trips. The dataset for our analysis has 1,578 respondents (90% of the total sample). Each individual was given six sequential stated choice sets to assess and make a choice, giving 9,468 observations.

Each alternative was described by access, main mode and egress trip attributes. The attributes and attribute levels for each mode are presented in Appendix A1. The description, acronyms and general statistics of the attributes are presented in Table 4-1. The main mode attributes include travel times, costs (parking, running, fares and tolls), reliability and crowding. Reliability, represented as a travel time variability over repeated trips, is associated with bus and car. Three levels of travel time were presented (i.e., quickest, average and slowest travel time) to describe travel time variability, each associated with a probability of occurrence which also varied in the experimental design.

An illustrative screenshot of a choice experiment is shown in Figure 4-2. At the beginning of the experiment, respondents were asked to describe their most recent trip and the attribute levels associated with the trip were used as base levels about which the levels presented in the choice scenarios pivoted. Data were also collected on the socioeconomic characteristics of each respondent, some of which are described in Section 4.4.2 along with the profile of travel behaviour.

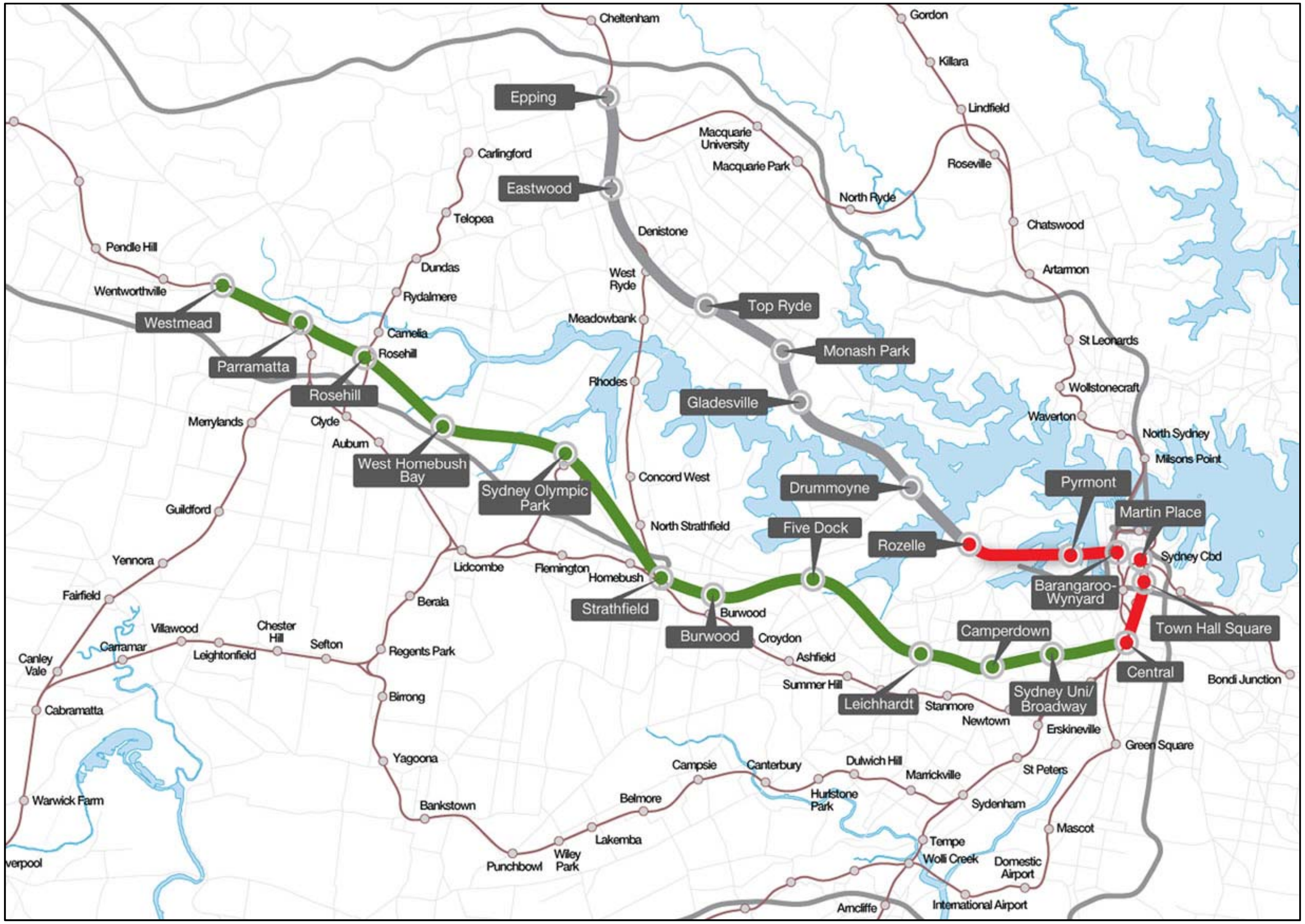


Figure 4-1: Proposed new metro rail line

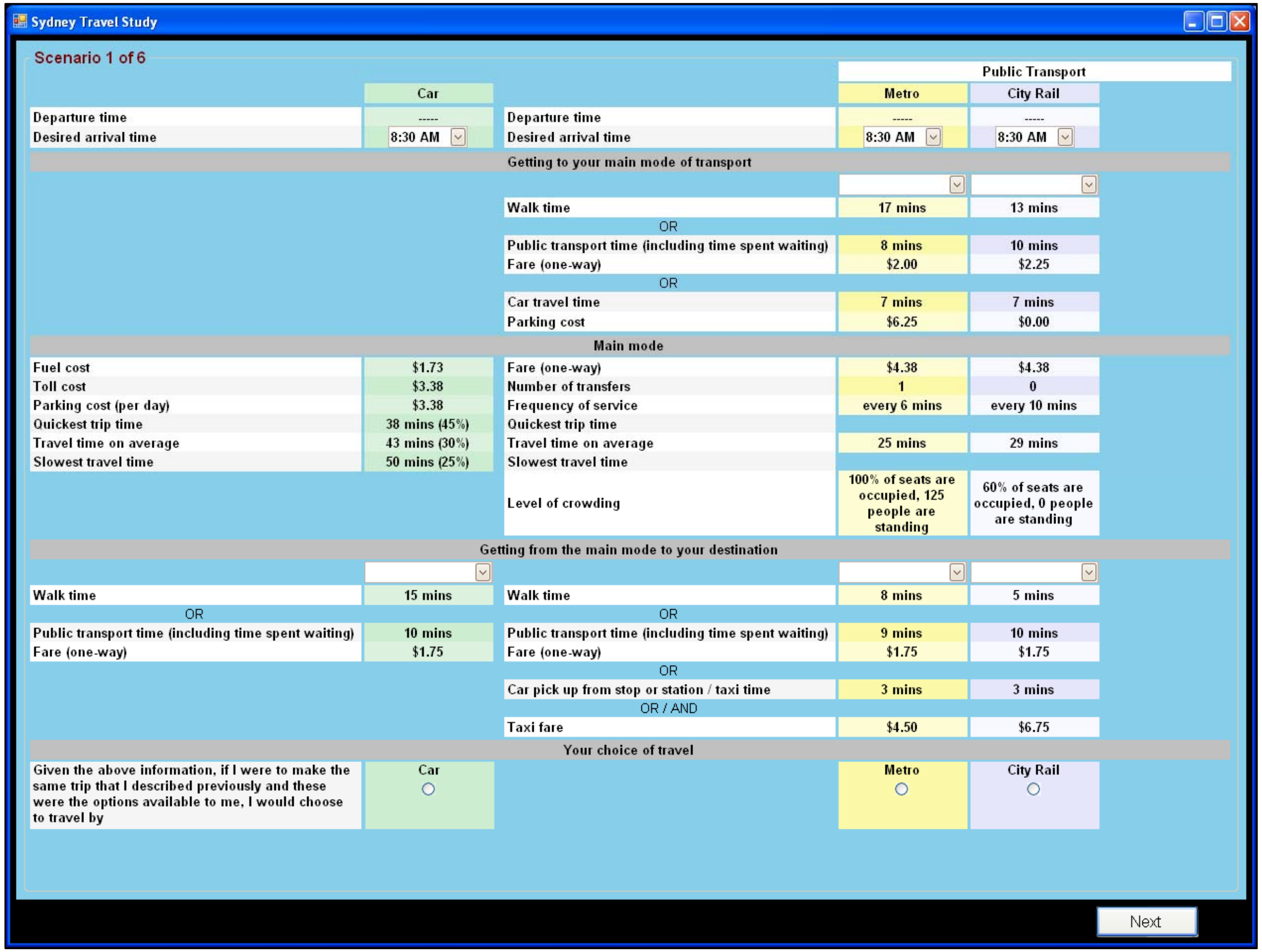


Figure 4-2: Illustrative screenshot of Metro Rail Sydney choice experiment

Table 4-1: Attributes description, acronyms and general statistics Metro Rail data

	Attribute	Unit	Acronym	Mode	Mean	Std Dev
Main Mode Attributes	Travel Time	Minutes	TTPT	Bus	33.97	23.39
				Train	31.83	17.47
				Metro	16.72	8.22
			TTCR	Car	43.70	23.07
	Fare	AUD$	COSTPT	Bus	3.27	1.77
				Train	3.86	1.91
				Metro	3.64	1.62
	Fuel cost	AUD$	COSTCRTRC	Car	2.26	1.19
	Toll cost			Car	4.34	3.00
	Parking cost		COSTCRPC	Car	13.84	10.95
	Transfers	Number of transfers	TRANPT	Bus	1.01	0.83
				Train	1.01	0.82
				Metro	0.50	0.50
	Headway	Minutes	FREQPT	Bus	28.14	18.92
				Train	30.92	18.35
				Metro	6.90	3.38
	% of Seats Available	%	SEATPT	Bus	0.18	0.23
				Train	0.17	0.23
				Metro	0.17	0.23
	Standees density	Standees per square metre	STANDPT	Bus	1.10	1.40
				Train	1.19	1.46
				Metro	1.17	1.45
Access Attributes	Access Time	Minutes	ACTIMEPT	Bus	5.97	4.87
				Train	9.77	7.10
				Metro	11.99	8.61
	Access Fare	AUD$	ACFARE	Bus	3.97	2.70
				Train	3.16	3.12
				Metro	3.73	3.72
Egress Attributes	Egress Time	Minutes	EGTIME	Bus	7.08	4.96
				Train	8.62	6.11
				Metro	9.29	7.09
				Car	5.92	4.22
	Egress Fare	AUD$	EGFARE	Bus	12.26	4.50
				Train	4.39	4.39
				Metro	3.89	3.46
				Car	4.71	6.06

Figure 4-3 summarises the revealed preference modal availability as a percentage of the total number of observations, and the percentage of times each mode was chosen when it was available. The metro alternative was always included in the choice scenarios and was chosen 62% of the time. The bus was available in 32% of the choice sets and was chosen in 25% of the offered scenarios. The train was available in 63% of the choice sets and it was chosen 28% of the time while the car was available in 38% of the choice sets and chosen 32% of the time.

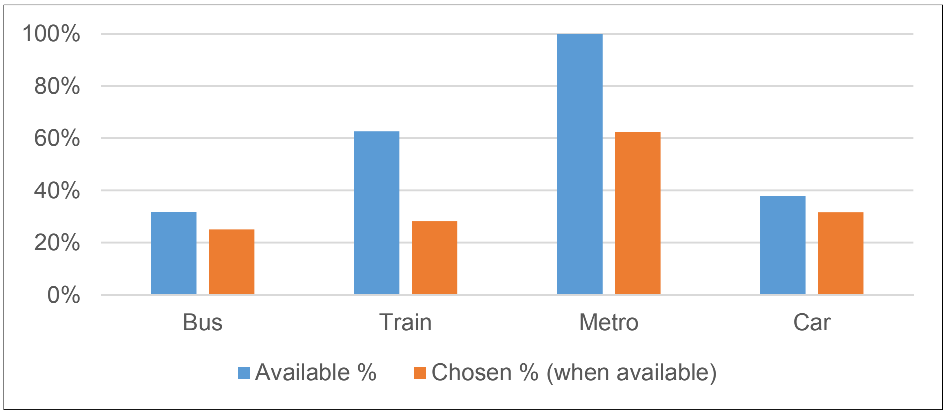


Figure 4-3: Mode availability and % of times they were chosen (when available)

4.3. Dataset #2: Northwest Data

This dataset on travel preferences of commuters and non-commuters was collected as part of a larger study to evaluate public transport investment options (train and bus) in the north west of Sydney, one of the fastest growing areas in Sydney (Hensher and Rose, 2007). The projects under consideration included variations of new heavy rail systems into the area presented in Figure 4-4, as well as new light rail and dedicated busway systems along the same corridor. The stated choice experimental design was constructed using a D-optimal design. The dataset was collected in July 2003 in the north-west catchment of metropolitan Sydney using a computer aided personal interview (CAPI). The sample covered residents that made trips within the region (intra-regional) and outside of the region (inter-regional). Individuals were first asked for a recent trip during the last week (not necessarily the most recent one) and the characteristics of it. The characteristics were used to pivot the attribute levels presented to them in the survey.

If an individual made intra-regional trips, the survey presented three public transport modes: new light rail, new heavy rail and bus, plus a car alternative if it was available for him. If an individual made inter-regional trips, the survey included five public transport modes: new light rail, new heavy rail, bus, existing M2 busway and existing train line, plus a car alternative if available. Each alternative was described by access, egress and main mode attributes. An illustrative screenshot of the North West Sydney data is shown in Figure 4-5. In each choice scenario, respondents had to choose their preferred main and access mode. The attributes and attribute levels for each mode and trip type are presented in Appendix A2.

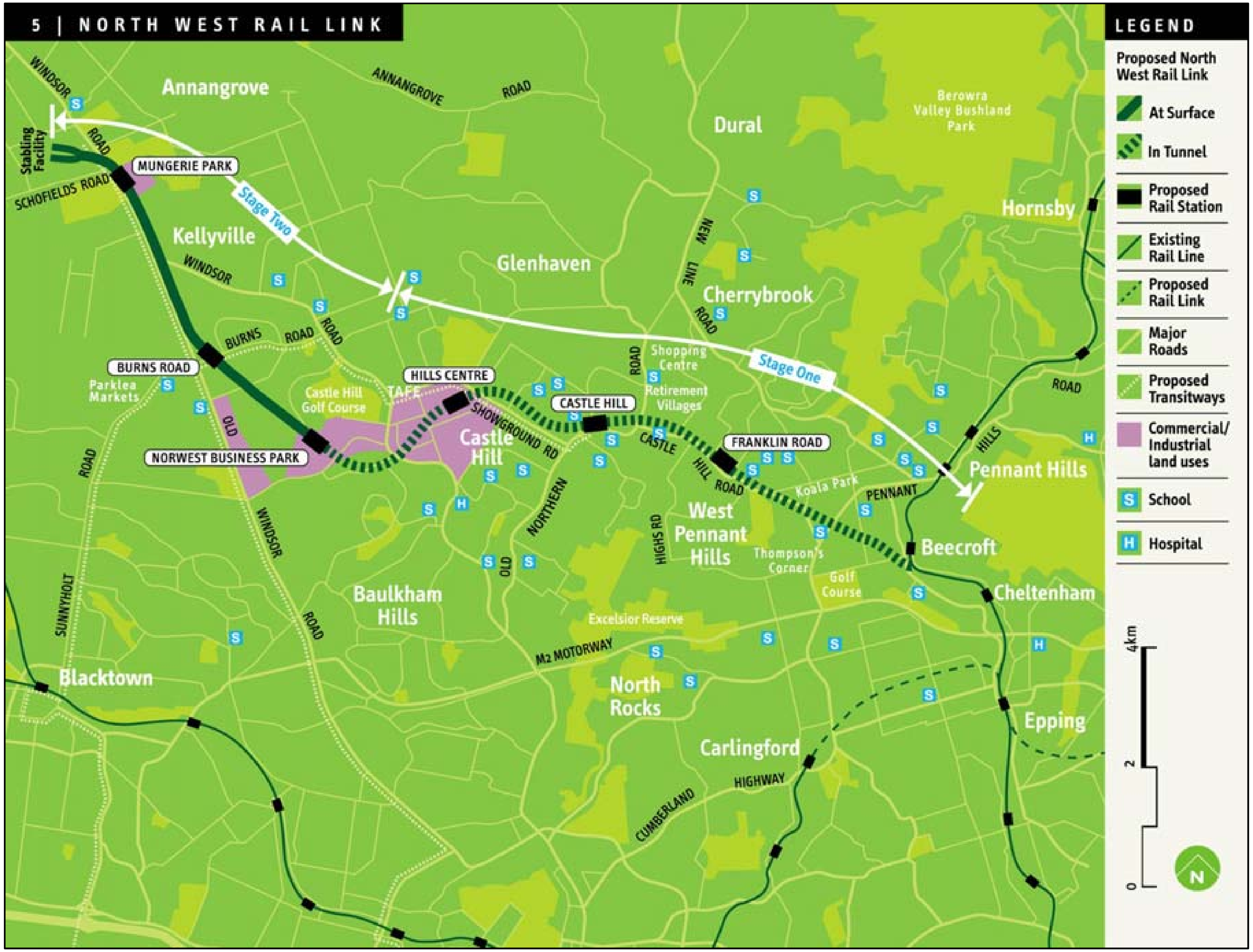
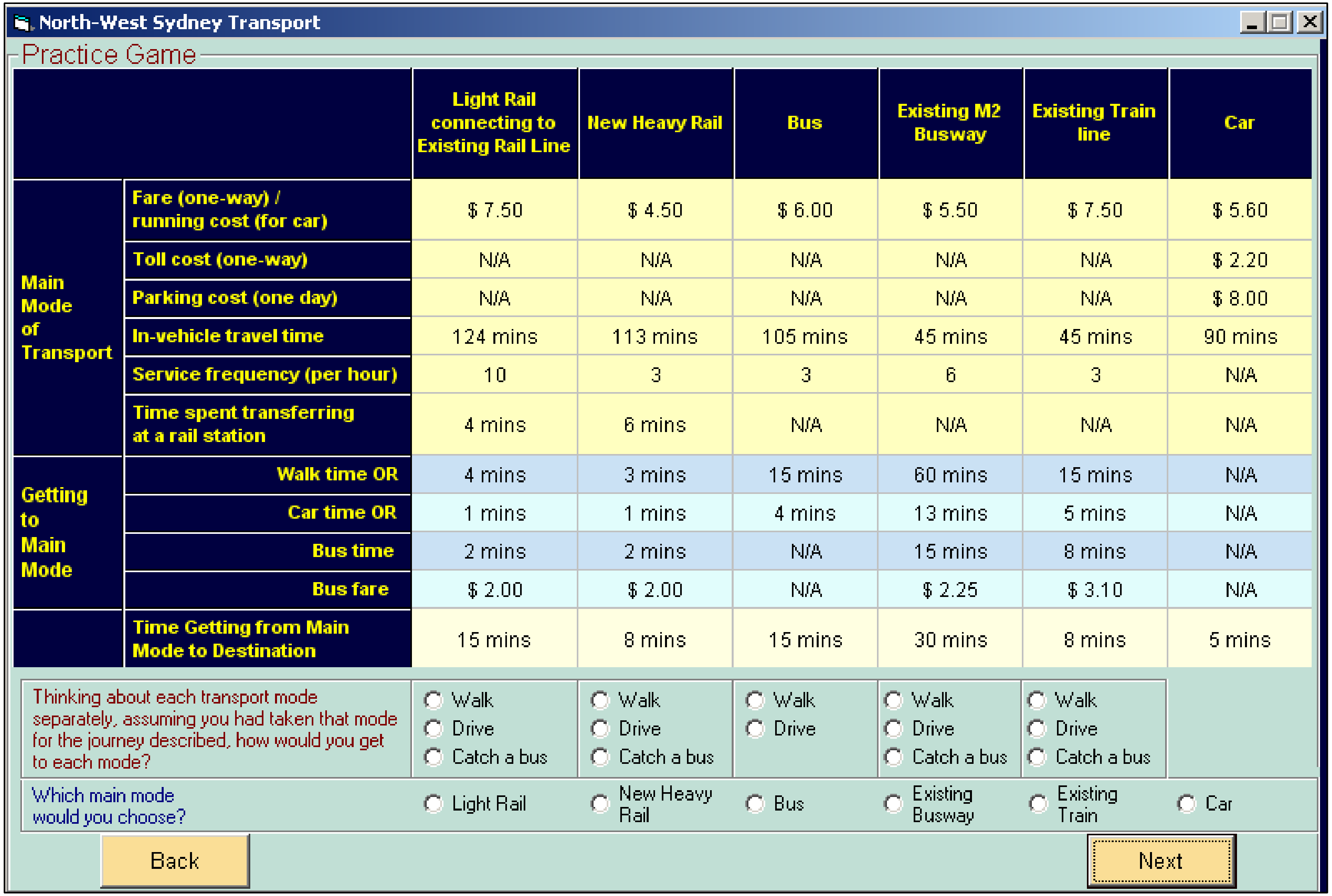


Figure 4-4: Proposed new rail system in the Northwest sector


North-West Sydney Transport

Practice Game

		Light Rail connecting to Existing Rail Line	New Heavy Rail	Bus	Existing M2 Busway	Existing Train line	Car
Main Mode of Transport	Fare (one-way) / running cost (for car)	$ 7.50	$ 4.50	$ 6.00	$ 5.50	$ 7.50	$ 5.60
	Toll cost (one-way)	N/A	N/A	N/A	N/A	N/A	$ 2.20
	Parking cost (one day)	N/A	N/A	N/A	N/A	N/A	$ 8.00
	In-vehicle travel time	124 mins	113 mins	105 mins	45 mins	45 mins	90 mins
	Service frequency (per hour)	10	3	3	6	3	N/A
	Time spent transferring at a rail station	4 mins	6 mins	N/A	N/A	N/A	N/A
Getting to Main Mode	Walk time OR	4 mins	3 mins	15 mins	60 mins	15 mins	N/A
	Car time OR	1 mins	1 mins	4 mins	13 mins	5 mins	N/A
	Bus time	2 mins	2 mins	N/A	15 mins	8 mins	N/A
	Bus fare	$ 2.00	$ 2.00	N/A	$ 2.25	$ 3.10	N/A
	Time Getting from Main Mode to Destination	15 mins	8 mins	15 mins	30 mins	8 mins	5 mins
Thinking about each transport mode separately, assuming you had taken that mode for the journey described, how would you get to each mode?		Walk / Drive / Catch a bus	Walk / Drive / Catch a bus	Walk / Drive	Walk / Drive / Catch a bus	Walk / Drive / Catch a bus	
Which main mode would you choose?		Light Rail	New Heavy Rail	Bus	Existing Busway	Existing Train	Car

Back

Next

Figure 4-5: Illustrative screenshot of North West Sydney choice experiment

The survey sample size is 453 individuals, where 92% had inter-regional trips and 8% intra-regional trips. Each respondent faced 10 choice sets giving a total sample size of 4,530 observations. The description, acronyms and general statistics of the attributes are presented in Table 4-2. Individuals were also asked for their socioeconomic characteristics analysed in Section 4.4 along with the profile of travel behaviour.

Table 4-2: Attributes description, acronyms and general statistics Northwest data

	Attribute	Unit	Acronym	Mode	Mean	Std Dev
Main Mode Attributes	Travel Time	Minutes	TTPT	Light Rail	57.39	27.33
				New Heavy Rail	52.00	25.70
				New Busway	67.11	26.27
				Bus	66.36	30.33
				Busway	52.95	24.52
				Train	46.99	20.06
			TTCR	Car	56.66	27.67
	Fare	AUD$	COSTPT	Light Rail	5.21	3.33
				New Heavy Rail	5.25	3.56
				New Busway	7.84	3.33
				Bus	7.62	4.26
				Busway	7.29	3.43
				Train	4.99	3.35
	Fuel cost	AUD$	COSTCRTRC	Car	2.93	1.28
	Toll cost			Car	5.04	2.76
	Parking cost		COSTCRPC	Car	14.43	14.45
	Transfers	Number of transfers	TRANPT	Light Rail	1.83	0.38
				New Heavy Rail	1.80	0.40
				New Busway	1.82	0.38
				Bus	0.73	0.65
				Busway	0.92	0.71
				Train	1.21	0.45
	Headway	Minutes	FREQPT	Light Rail	11.50	2.71
				New Heavy Rail	4.50	1.12
				New Busway	4.50	1.12
				Bus	3.36	2.58
				Busway	4.04	2.81
				Train	3.82	2.41
Access Attributes	Access Time	Minutes	ACTIMEPT	Light Rail	9.44	6.03
				New Heavy Rail	9.52	6.04
				New Busway	9.48	5.56
				Bus	7.34	7.35
				Busway	9.74	8.25
				Train	18.35	9.80
Egress Attributes	Egress Time	Minutes	EGTIME	Light Rail	12.54	10.54
				New Heavy Rail	12.08	10.37
				New Busway	11.80	10.80
				Bus	10.16	8.88
				Busway	11.95	10.96
				Train	11.99	9.78
				Car	7.83	7.82

Figure 4-6 summarises the mode availability in each choice set as a percentage of the total number of observations, and the percentage of times each mode was chosen when it was available. The bus alternative was always included in the survey and was chosen 14% of the times. The busway and train were available in 92% of the choice sets and chosen 16% and 21% of those times, respectively. The car was available in 86% of the choice sets and chosen in 13% of them. New light rail, new heavy rail and new busway were included 80%, 72% and 48% of the times and were chosen 22%, 30% and 5% times, respectively. Aggregating the new alternatives, the results show that they were chosen 41% of the times with currently available alternatives 59% of the times.

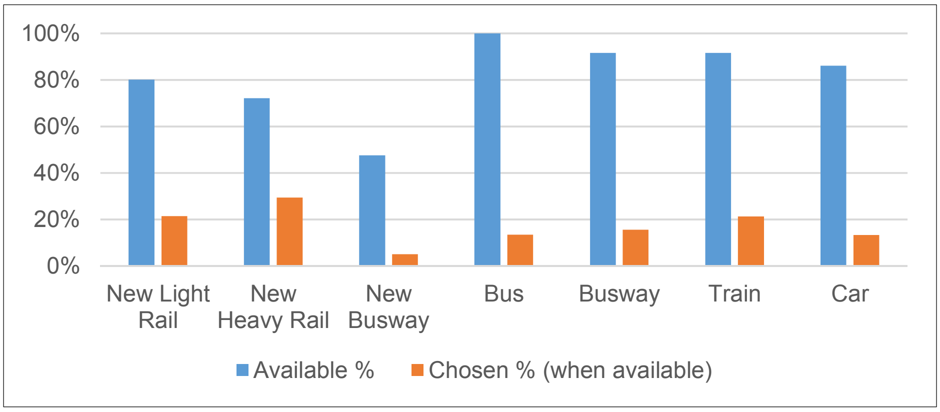


Figure 4-6: Mode availability and % of times they were chosen (when available)

4.4. Dataset Comparison

4.4.1. Choice Task Attributes

The stated preference (SP) part of the surveys have several attributes in common. Figure 4-7 presents the mean of the common attribute for both SP datasets. The main difference between the mean attribute levels is the public transport (PT) travel time, where the Northwest survey has a significantly higher mean than the Metro Rail survey, and the mean of the headway which is significantly higher in the Metro Rail survey. Since these levels were pivoted using the respondents' recent trip, the mean travel time reflects the trip length differences of the two datasets. The rest of the attributes have similar mean attribute levels.

Figure 4-8 presents the quotient between the standard deviation and the mean, which will be referred to as a relative standard deviation (or coefficient of variation). There is no clear pattern between the two datasets. The Metro Rail survey has a significantly higher relative standard deviation than the Northwest survey for the public transport travel time and the number of transfers. Contrarily, the Northwest survey has a significantly higher relative standard deviation than the Metro Rail survey for the egress time in car and the parking cost. This shows that both datasets are rich in terms of the attribute levels' variance. This is important when trying to estimate complex models, especially ones that include risk attitudes or experience. A higher variation in the attribute levels provides an opportunity to investigate sample preferences on attributes with noticeably wide ranges across the sample.

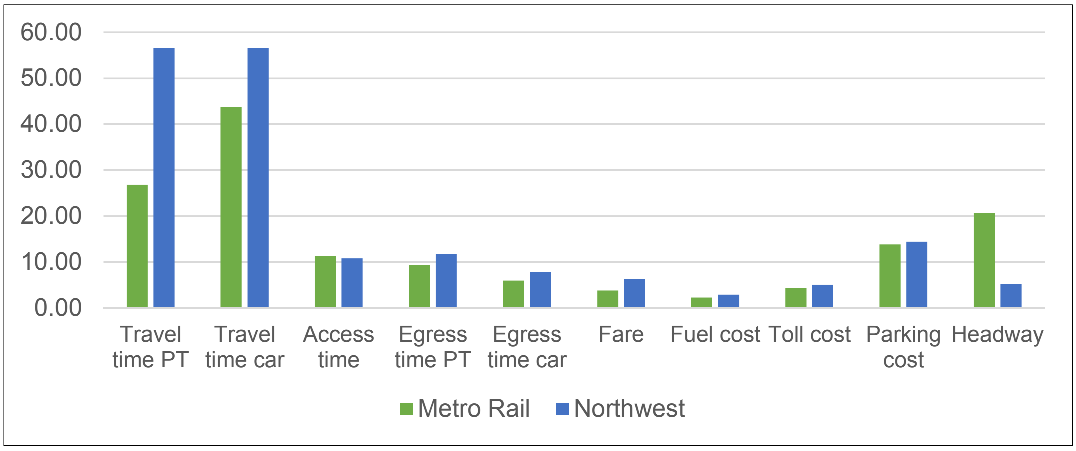


Figure 4-7: Comparison of the mean attribute levels for the Metro Rail and Northwest datasets

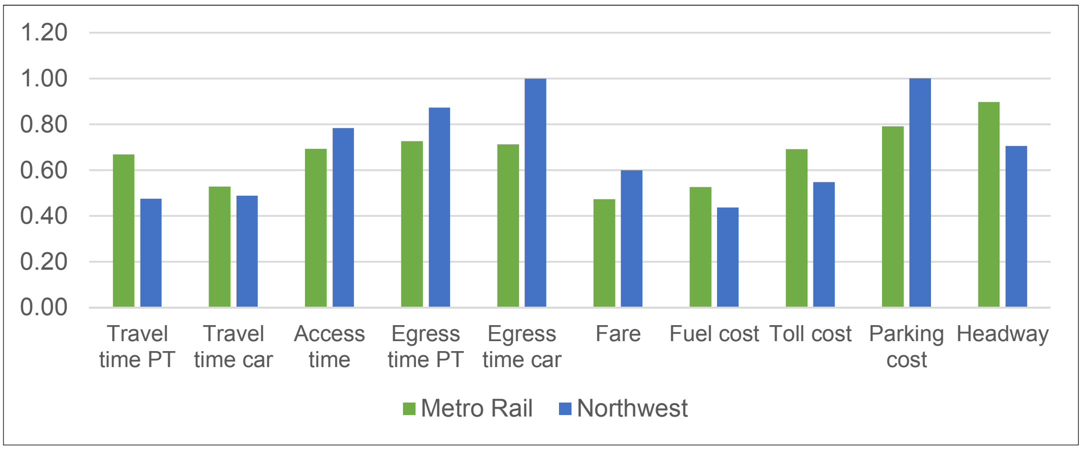


Figure 4-8: Comparison of the relative standard deviation (standard deviation/mean) of the attribute levels for the Metro Rail and Northwest datasets

4.4.2. Socioeconomic Characteristics and Travel Behaviour

The Northwest data survey was collected in a specific north-western region of metropolitan Sydney area, while the Metro Rail dataset covered a larger area of Sydney as shown in Figure 4-9. Figure 4-10 presents the age distribution of both samples. The distributions are relatively similar, although the Northwest dataset has a higher percentage of individuals under 24 years old, but a lower percentage between 25 and 34. There is a similar percentage of individuals between 45 and over. The average age for the Northwest data is 42.85 years and for the Metro Rail data is 43.28 years.

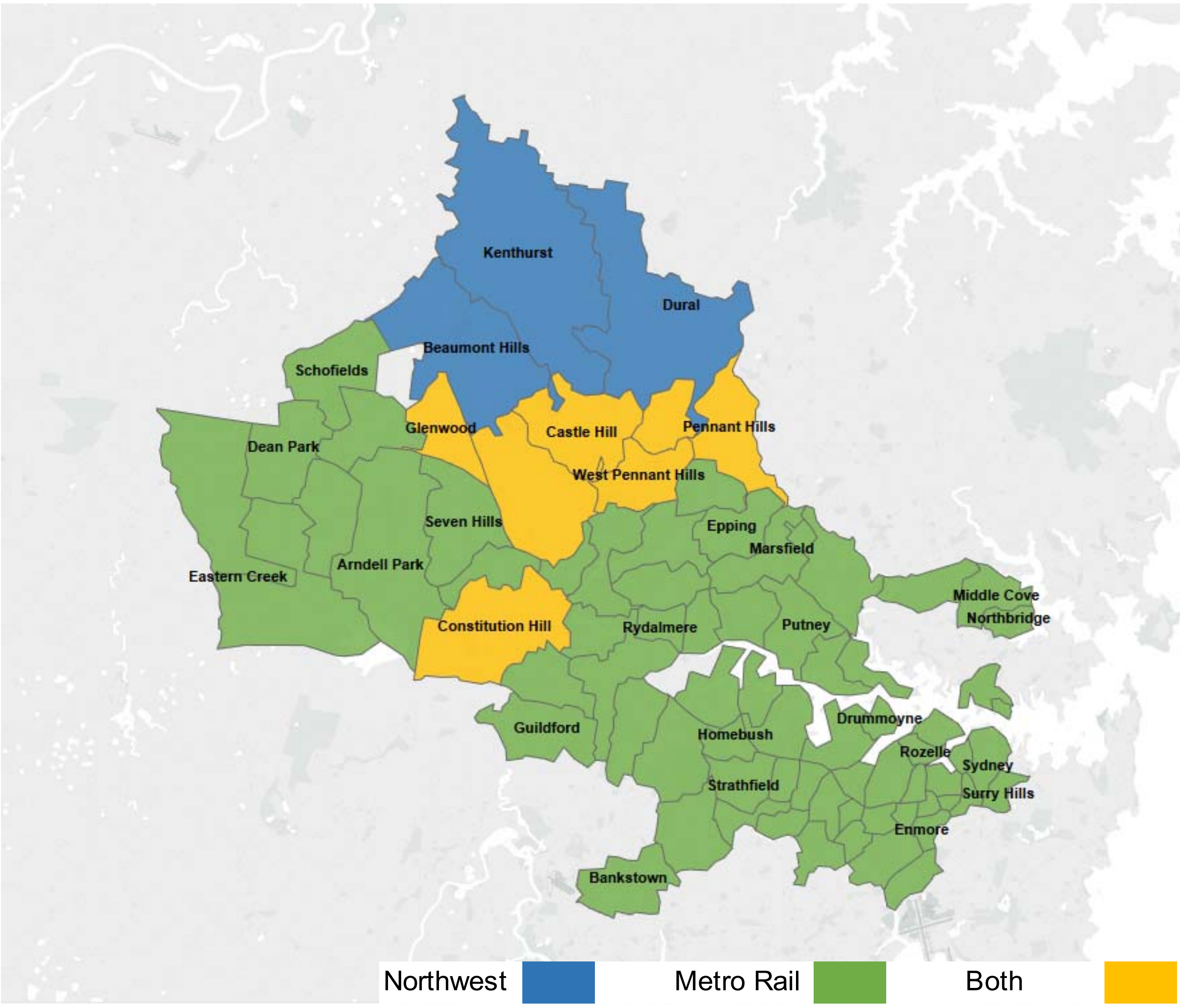


Figure 4-9: Survey areas in Sydney map

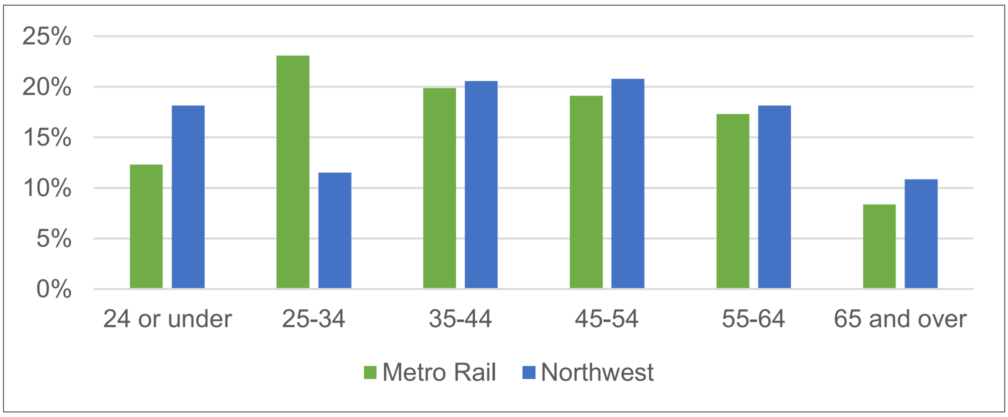


Figure 4-10: Respondents' age for the Metro Rail and Northwest datasets

Figure 4-11 presents the percentage of female and male respondents. 56 percent of respondents in the Metro Rail data are female, and for the Northwest data it is 43%. Figure 4-12 presents the annual personal income distribution relative to year 2009[11]. There is a significantly higher percentage of individuals in the Northwest data that have an income of $10,000 or lower. The average annual personal income in the Northwest data is of $43,380, which is significantly lower than in the Metro Rail data, which has an average annual income of $52,800.

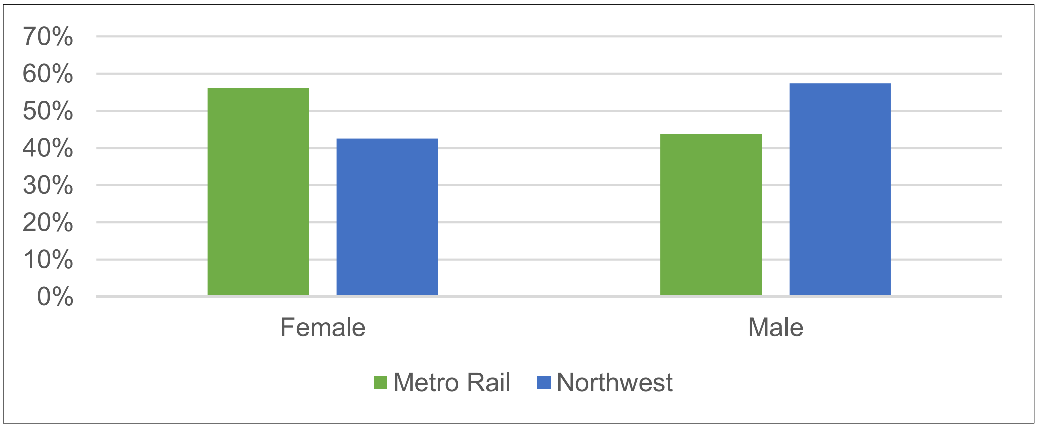


Figure 4-11: Respondents' gender for the Metro Rail and Northwest datasets

[11] Calculated using the annual inflation rate provided by the Reserve Bank of Australia. Respondents were asked for the interval that best represented their income, so a uniform distribution was assumed within each intervals to include inflation.

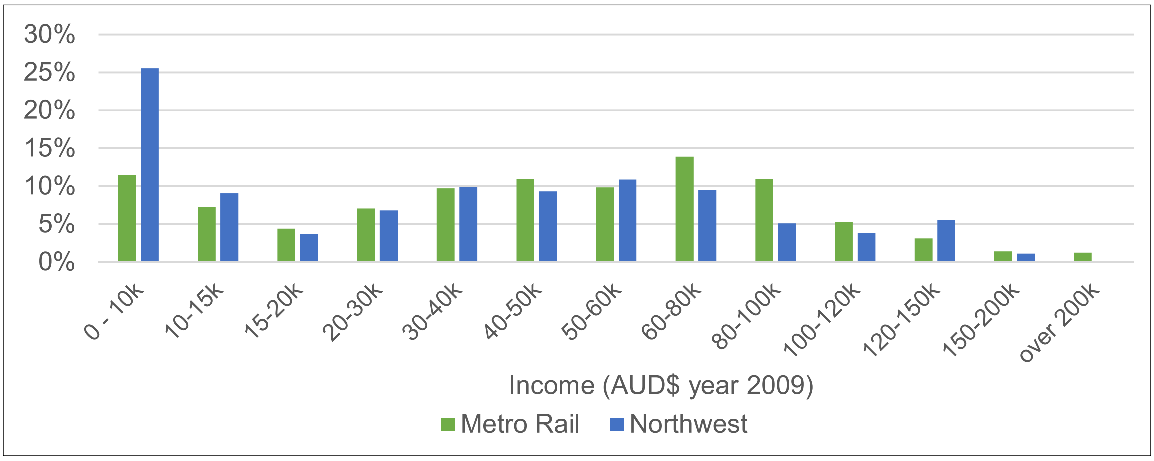


Figure 4-12: Respondents' income for the Metro Rail and Northwest datasets

Figure 4-13 presents the work status of respondents. The percentages are relatively similar in both samples, except for the respondents that have not worked in the last month, which represents a 25% of respondents in the Northwest data and a 15% in the Metro Rail data.

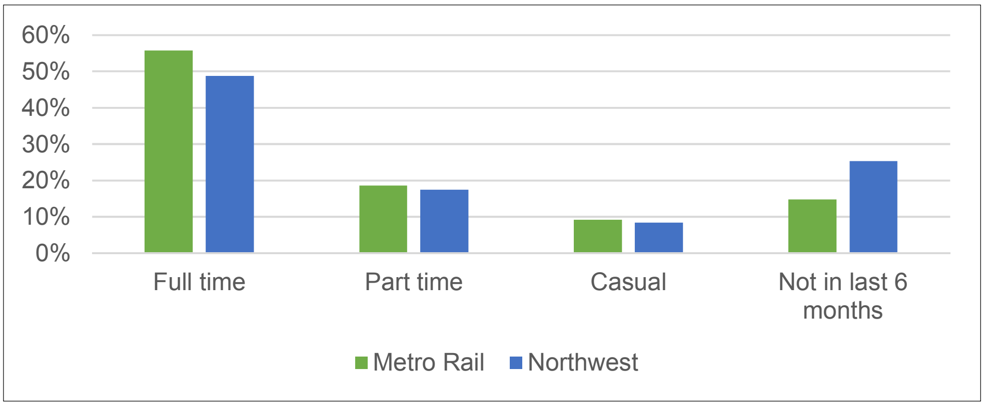


Figure 4-13: Respondents' work status for the Metro Rail and Northwest datasets

The modes used in respondents' most recent trips (which were used as the *experience* variable) are presented in Figure 4-14 and Figure 4-15 for the Metro Rail and Northwest data, respectively. The usage of bus (22% for both) and car (27% for Metro Rail and 30% for Northwest) are very similar across the datasets. However, the train usage in the Metro Rail dataset (51%) is significantly higher than for the Northwest dataset (27%). If we add the busway and train usage in the Northwest data (48%) it would be equivalent to the train usage in the Metro Rail data.

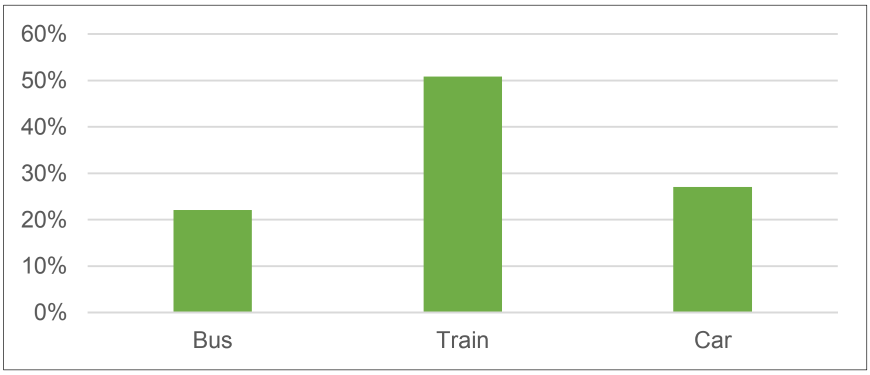


Figure 4-14: Recent trip mode of the Metro Rail survey respondents

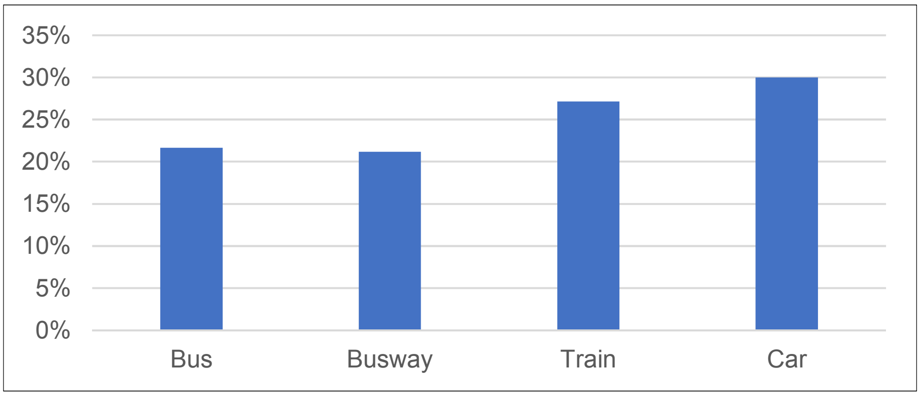


Figure 4-15: Recent trip mode of the Northwest survey respondents

Figure 4-16 presents the characteristics of respondents' most recent trips. The majority of the characteristics are equivalent, except for the travel time in public transport (PT) and toll costs, which are significantly higher in the Northwest data.

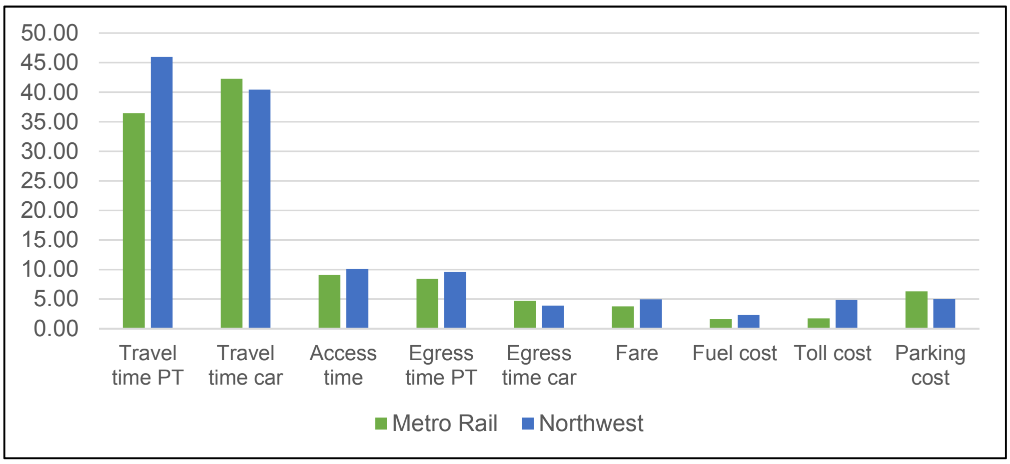


Figure 4-16: Characteristics of the recent trips

4.5. Conclusions

This chapter provided a description of the datasets that will be used in the empirical analysis in the following chapters. Both surveys studied travel behaviour through labelled choice experiments including new and existing modal alternatives. The characteristics used to describe the alternatives included access, egress and main mode attributes in both datasets. However, the Metro Rail dataset included reliability (trip time variability) and crowding attributes that were not considered in the Northwest data. Moreover, the Northwest data considered seven different alternatives, while the Metro Rail data only four.

The profile of respondents had some relevant differences, such as the average income which was significantly lower in the Northwest data, and work status which was 10% higher in the category not worked in the last six months for the Northwest data. The modal split between the datasets was relatively similar for the car and bus, but for the train it was very different. Northwest respondents had an additional mode available (the busway). The characteristics of respondents' recent trips was different in terms of their travel time in public transport and the toll costs, which were significantly higher in the Northwest data. Since these levels were used to pivot the attribute levels presented in the survey, the public transport travel times in the survey were also higher in the Northwest data.

In conclusion, even though both datasets were collected in Sydney, they do represent different geographical catchment areas with informative differences in the profile of respondents. This is very important as it will provide an opportunity to study preferences in settings that are sufficiently different to inform the extent to which there are common behavioural traits in travel choice making.

CHAPTER 5

Results Metro Rail

5.1. Introduction

This chapter presents the model results on the MetroRail dataset described in Chapter 4 using the heuristics LPAA, VL and RAM, together with behavioural refinements and experience. For the utility expressions used in this chapter, the reader is referred to the 'Notational Glossary' at the beginning of the thesis. The purpose will be, first, to study each model type separately and understand their findings and interpretation. After this, the different types of models will be compared. In Section 2 we present the model results for the process homogeneity models. Section 3 will present the process heterogeneity model results using the PDP approach, and the fourth section will present the process heterogeneity model results using the CRPH approach. Section 5 will compare the models separated in three main topics: the results on behavioural refinements and experience; the log likelihood and AIC indicators; and the willingness to pay estimates (WTP). The subsection that refers to the WTP is separated into evidence on the value of travel time savings and then all the other attributes. This chapter ends by presenting the main findings.

5.2. Simple MNL and MML Model Results

5.2.1. Linear Parameters and Additive Attributes

The heuristic that considers linear parameters and additive attributes (LPAA) has been widely used in choice studies. In its simplest form it estimates all parameters as fixed, denoted here as LPAA_MNL, which will represent the basic reference model to compare all others. A slightly more complex model that has also been traditionally used in choice studies considers the LPAA heuristic but estimates all the parameters as randomly distributed (mixed logit); this model will be referred to as LPAA_MML. The results for these models are presented in Table 5-1. The other two models presented in this table represent the LPAA heuristic with behavioural refinements and experience influencing decision-making. All the possible combinations of risk attitudes, perceptual conditioning and experience were tested but not all of them were found to be statistically significant. The last two columns in Table 5-1 present the final models for: (1) a fixed parameter LPAA with behavioural refinements and experience, called LPAA_MNL_BRExp; and (2) a random parameter LPAA with behavioural refinements and experience, called LPAA_MML_BRExp. As expected, there is a significant improvement in the overall fit when allowing for random parameters and/or when including experience and behavioural refinements.

Results show that experience associated with each mode of transportation available (remembering that metro was not available) has a noticeable influence on preferences. The utility conditioning form, which includes experience, as explained in Section 3.5.1, is appropriate as a way of representing individual decision-making. All the experience parameters were positive, which indicate that individuals are more likely to choose the mode they used in their most recent trip. That is, they increase (decrease) the positive (negative) utility function of the mode they used in their most recent trip.

Regarding behavioural refinements, not all of them appear to be significant. Perceptual conditioning has a significant influence in the LPAA_MML_BRExp model towards the probabilities of the car travel times. Figure 5-1 graphically presents this estimate, which shows that individuals tend to underweight low to medium probabilities (below 80%). Individuals tend to underweight in a more significant way probabilities between 20% and 30%. Probabilities higher than 80% are slightly overweight but this is almost insignificant as can be seen in Figure 5-1.

Model LPAA_MNL_BRExp showed that there are risk aversion attitudes towards the bus, train and metro travel times. In contrast, there is risk neutrality towards the travel times when estimating random parameters in the LPAA_MML_BRExp model. Figure 5-2 shows the concave transformation of the risk attitudes in the MNL model with behavioural refinements and experience, representing risk aversion. Both the MNL and MML models obtained estimated parameters suggesting significant risk aversion attitudes towards parking cost, and in addition the MNL model also estimated risk taking attitude towards bus fares. The cost transformations due to risk attitudes are shown in Figure 5-3. As seen, the risk attitudes for the parking costs have a concave shape, whereas the risk attitudes for the bus fares have a convex shape.

Table 5-1: LPAA MNL and MML models (t-values in brackets)

			LPAA_MNL	LPAA_MML		LPAA_MNL_BRExp	LPAA_MML_BRExp	
Number of Parameters Estimated			**14**	**25**		**24**	**31**	
Log Likelihood at convergence			**-6,326.73**	**-5,024.55**		**-6,133.04**	**-4,958.17**	
Log likelihood at zero			**-13,125.44**					
AIC			**1.339**	**1.067**		**1.301**	**1.054**	
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Std Dev**	**Mean**	**Mean**	**Std Dev**
Alternative Specific Constant Bus	ASCBUS	Bus	1.22 (7.61)	2.39 (4.55)	-	1.46 (2.78)	1.47 (2.59)	-
Alternative Specific Constant Train	ASCTRAIN	Train	1.24 (8.02)	2.38 (4.57)	-	1.76 (3.75)	2.45 (5.10)	-
Alternative Specific Constant Metro	ASCMETRO	Metro	1.51 (11.01)	2.28 (4.55)	-	2.01 (4.12)	2.63 (5.78)	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Bus, Train, Metro	-0.05 (18.17)	-0.10 (10.78)	0.15 (10.22)	-0.05 (18.63)	-0.10 (10.83)	0.14 (12.95)
Fare Public Transport	COSTPT	Bus, Train, Metro	-0.21 (12.17)	-0.60 (13.18)	0.52 (9.30)	-0.23 (12.41)	-0.64 (13.88)	0.56 (11.59)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.01 (1.06)	-0.16 (4.18)	0.28 (7.28)	-0.05 (2.23)	-0.24 (3.80)	0.35 (6.03)
Parking Cost Car	COSTCRPC	Car	-0.04 (8.58)	-0.22 (13.02)	0.20 (14.44)	-0.23 (10.02)	-0.58 (5.08)	0.53 (5.73)
Travel Time Public Transport	TTPT	Bus, Train, Metro	-0.03 (15.67)	-0.09 (14.61)	0.08 (13.44)	-0.16 (2.33)	-0.09 (14.42)	-0.08 (12.77)
Travel Time Car	TTCR	Car	-0.02 (11.99)	-0.06 (5.20)	0.03 (4.14)	-0.04 (11.80)	-0.16 (4.74)	0.06 (3.64)
Egress Time	EGTIME	All Alternatives	-0.06 (17.94)	-0.13 (10.88)	0.14 (6.81)	-0.06 (16.84)	-0.12 (10.06)	0.15 (9.63)
Transfer Public Transport	TRANPT	Bus, Train, Metro	-0.11 (4.21)	-0.26 (5.71)	0.41 (4.21)	-0.11 (3.88)	-0.29 (5.79)	0.38 (3.23)
Headway Public Transport	FREQPT	Bus, Train, Metro	-0.01 (6.82)	-0.03 (9.13)	0.05 (12.87)	-0.01 (6.96)	-0.03 (9.50)	0.05 (12.65)
% Seat Public Transport	SEATPT	Bus, Train, Metro	0.39 (4.24)	1.00 (5.78)	1.57 (5.03)	0.44 (4.43)	1.07 (5.90)	1.60 (4.90)
Density Public Transport	STANDPT	Bus, Train, Metro	-0.18 (12.25)	-0.37 (12.12)	0.39 (9.44)	-0.19 (11.93)	-0.37 (11.98)	0.42 (9.12)
Experience Bus	EXPBS	Bus	-	-	-	0.30 (9.37)	0.34 (10.52)	-
Experience Train	EXPTR	Train	-	-	-	0.09 (3.70)	0.07 (2.74)	-
Experience Car	EXPCR	Car	-	-	-	0.44 (16.36)	0.54 (15.22)	-
Risk Attitudes Travel Time Bus	ALPHABSTT	Bus	-	-	-	0.44 (3.73)	-	-
Risk Attitudes Travel Time Train	ALPHATRTT	Train	-	-	-	0.46 (3.72)	-	-

			LPAA_MNL	LPAA_MML		LPAA_MNL_BRExp	LPAA_MML_BRExp	
Risk Attitudes Travel Time Metro	ALPHAMTTT	Metro	-	-	-	0.60 (3.91)	-	-
Risk Attitudes Travel Time Car	ALPHACRTT	Car	-	-	-	-	-	-
Risk Attitudes Cost Bus	ALPHABSCS	Bus	-	-	-	-	-0.23 (2.53)	-
Risk Attitudes Cost Train	ALPHATRCS	Train	-	-	-	-	-	-
Risk Attitudes Cost Metro	ALPHAMTCS	Metro	-	-	-	-	-	-
Risk Attitudes Fuel+Toll Car	ALPHACRTRCS	Car	-	-	-	-	-	-
Risk Attitudes Parking Car	ALPHACRPCS	Car	-	-	-	0.65 (8.79)	0.20 (2.06)	-
Perceptual Conditioning Bus	GAMMABS	Bus	-	-	-	-	-	-
Perceptual Conditioning Car	GAMMACR	Car	-	-	-	-	1.83 (2.96)	-

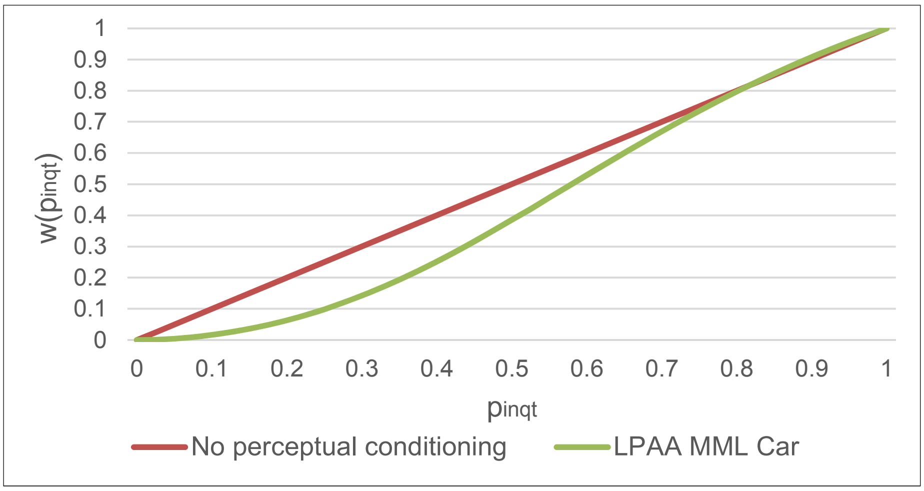


Figure 5-1: Perceptual conditioning in LPAA_MML_BRExp

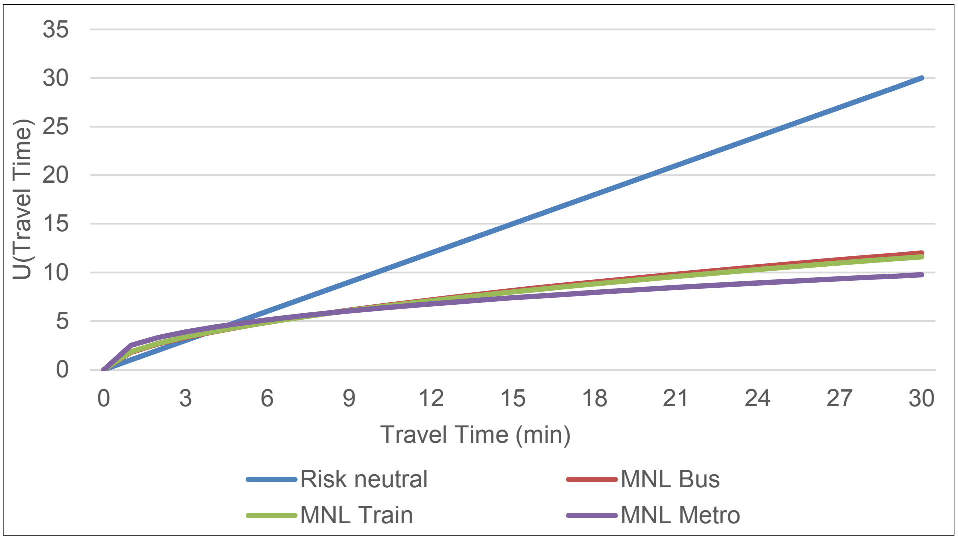


Figure 5-2: Risk attitudes towards the travel times in LPAA_MNL_BRExp

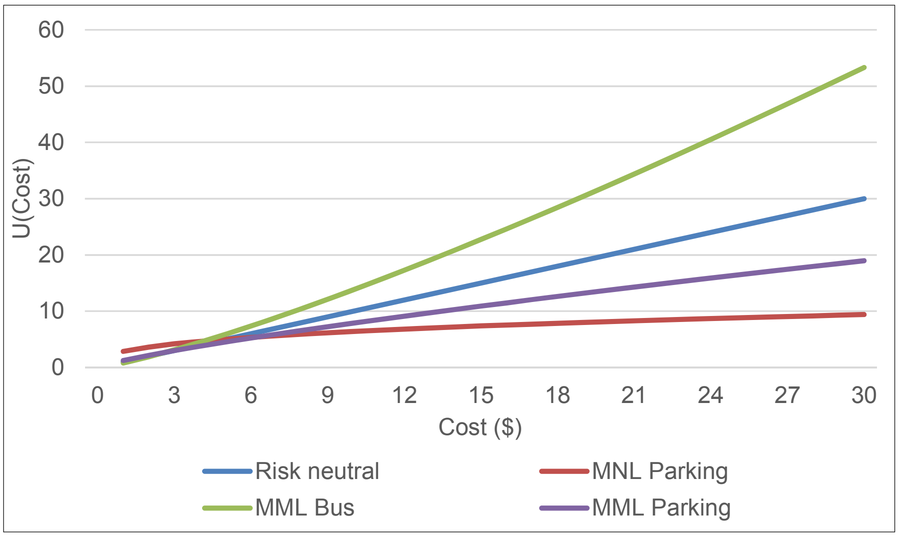


Figure 5-3: Risk attitudes towards costs in LPAA_MNL_BRExp and LPAA_MML_BRExp

These models are nested, with LPAA_MML_BRExp the most general and the LPAA_MNL the most restricted. They can be compared using the log likelihood ratio test (explained in Section 3.8.2). This comparison has two purposes: (1) identifying if the parameters can be treated as random is appropriate, and (2) assessing if the inclusion of behavioural refinements and experience is appropriate. For the first purpose, the models that include random parameters (LPAA_MML and LPAA_MML_BRExp) will be compared to the models with a similar structure but that consider all parameters as fixed (LPAA_MNL and LPAA_MNL_BRExp). For the second purpose, models that include behavioural refinements and experience (LPAA_MNL_BRExp and LPAA_MML_BRExp) will be compared to the equivalent models that do not (LPAA_MNL and LPAA_MML). The results are shown in Table 5-2. As can be seen, all the null hypotheses that state that the models are equivalent are rejected, suggesting that the inclusion of random parameters, behavioural refinements and experience significantly improve the statistical performance (as well as behavioural integrity[12]) of the models. The preferred model is the LPAA_MML_BRExp as it provides a better understanding of how preferences are formed and influenced by preference heterogeneity, risk attitudes, perceptual conditioning and experience.

[12] Behavioural integrity is the capability of a model to include/estimate behavioural refinements.

Table 5-2: Log likelihood ratio test results for the LPAA models

	Random parameters		Behavioural refinements and experience	
	LPAA_MML vs. LPAA_MNL	LPAA_MML_BRExp vs. LPAA_MNL_BRExp	LPAA_MNL_BRExp vs. LPAA_MNL	LPAA_MML_BRExp vs. LPAA_MML
LR	2604.366	2349.748	387.388	132.77
Degrees of freedom	11	7	10	6
$\chi^2_{d.f.;0.001}$	31.264	24.322	29.588	22.458
Result	Reject null	Reject null	Reject null	Reject null

5.2.2. Value Learning

MNL and MML

The value learning (VL) heuristic proposes that individuals value each attribute level relative to a reference level, which might be updated throughout the experiment as explained in Section 3.5.2. The parameter estimates for the models that assume that all respondents use a value learning (VL) heuristic are shown in Table 5-3. The first model, referred to as VL_MNL considers all parameters as fixed across the sample. The second one, Model VL_MML, estimates all parameters as random. Model VL_MNL_Exp considers all the parameters as fixed and includes experience conditioning of the utility function. Model VL_MML_Exp also considers experience but estimates all parameters as random. Results show that when considering an MML structure (i.e., estimating random parameters) there is a significant improvement in the model fit in terms of the log-likelihood and AIC.

Model VL_MNL_Exp does not include bus experience because, in the exploratory model results, the parameters was not significant and thereby the final model for fixed parameters does not consider it. Experience on train was statistically significant in the fixed parameters model, but with a negative sign. That is, individuals who made their most recent trip by train were less likely to choose it when using a VL heuristic. This is very interesting as it is opposite from what was found in the models that considered an LPAA heuristic. This implies that when taking into account the previous choice sets' attribute levels and their effect in current decisions, experience on the train behaves oppositely. By contrast, results show that experience is using a car is significant and positive, i.e., a person that used the car in his most recent trip is more likely to choose the car, even after taking into consideration value learning effects.

The concavity factors are presented graphically in Figure 5-4. When considering random parameters, the concavity factor was not statistically significant indicating a linear evaluation of the difference between the attribute level and the reference level. Models VL_MNL and VL_MNL_Exp both estimate a relatively similar concavity parameter, which tends to overweight higher differences as shown in the Figure.

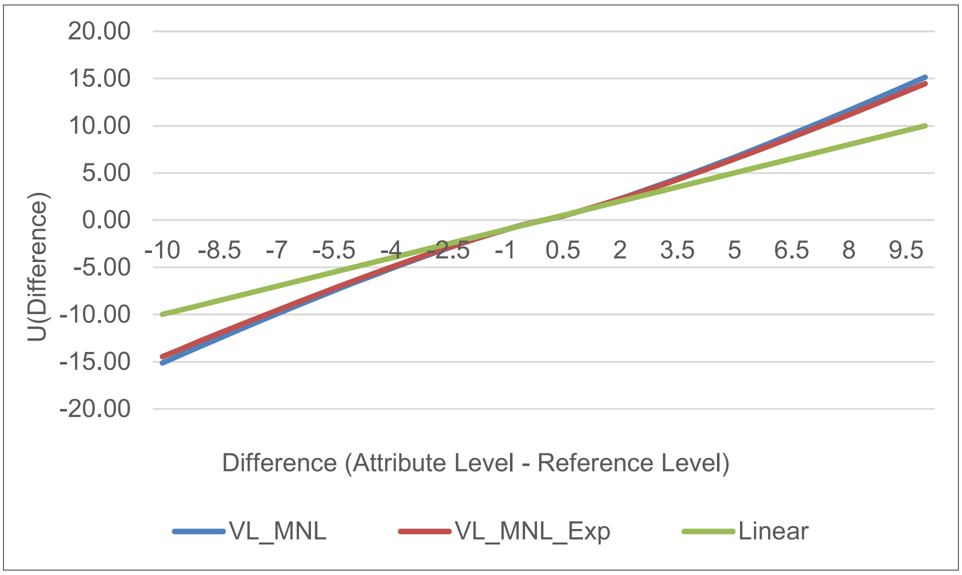


Figure 5-4: Concavity factor for VL models

When considering all parameters as random, the experience attribute behaved differently. This time, the experience on train was not significant for the models. The experience that individuals had on the bus and car were significant and both positive. This is aligned with what was found in the LPAA models, where individuals were more likely to choose the mode they had used in their most recent trip.

Table 5-3: VL MNL and MML models (t-values in brackets)

			VL_MNL	VL_MML		VL_MNL_Exp	VL_MML_Exp	
Number of Parameters Estimated			**15**	**25**		**17**	**27**	
Log Likelihood at convergence			**-5,973.30**	**-5,066.10**		**-5,944.52**	**-5,043.60**	
Log likelihood at zero			**-13,125.44**					
AIC			**1.265**	**1.075**		**1.259**	**1.071**	
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Std Dev**	**Mean**	**Mean**	**Std Dev**
Alternative Specific Constant Bus	ASCBUS	Bus	0.38 (3.31)	1.34 (5.57)	-	0.39 (3.16)	1.19 (4.68)	-
Alternative Specific Constant Train	ASCTRAIN	Train	0.38 (3.55)	1.30 (5.88)	-	0.46 (4.37)	1.57 (6.91)	-
Alternative Specific Constant Metro	ASCMETRO	Metro	0.71 (7.51)	1.21 (5.96)	-	0.74 (7.39)	1.47 (6.98)	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Bus, Train, Metro	-0.03 (9.52)	-0.09 (10.58)	0.14 (10.27)	-0.04 (9.53)	-0.09 (9.83)	0.15 (8.76)
Fare Public Transport	COSTPT	Bus, Train, Metro	-0.25 (14.45)	-0.55 (13.00)	0.56 (8.95)	-0.24 (13.65)	-0.56 (13.12)	0.51 (7.60)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.05 (4.85)	-0.14 (3.79)	0.28 (6.19)	-0.08 (4.90)	-0.21 (4.26)	0.39 (6.96)
Parking Cost Car	COSTCRPC	Car	-0.04 (8.03)	-0.17 (10.25)	0.18 (7.94)	-0.05 (7.50)	-0.21 (8.56)	0.20 (7.67)
Travel Time Public Transport	TTPT	Bus, Train, Metro	-0.01 (7.68)	-0.08 (15.55)	0.07 (13.16)	-0.01 (7.66)	-0.09 (15.59)	0.08 (12.88)
Travel Time Car	TTCR	Car	-0.01 (6.55)	-0.06 (7.05)	0.07 (7.75)	-0.02 (6.52)	-0.07 (7.11)	0.08 (7.92)
Egress Time	EGTIME	All Alternatives	-0.04 (9.34)	-0.13 (11.82)	0.15 (9.6)	-0.04 (9.23)	-0.14 (11.93)	0.15 (9.95)
Transfer Public Transport	TRANPT	Bus, Train, Metro	-0.09 (3.66)	-0.23 (5.27)	0.23 (1.6)	-0.09 (3.86)	-0.24 (5.24)	0.30 (2.65)
Headway Public Transport	FREQPT	Bus, Train, Metro	-0.01 (6.07)	-0.03 (9.95)	0.05 (14.15)	-0.01 (5.90)	-0.03 (9.55)	0.05 (12.89)
% Seat Public Transport	SEATPT	Bus, Train, Metro	0.28 (3.07)	0.93 (5.85)	1.61 (5.93)	0.30 (3.29)	0.99 (5.99)	1.84 (7.34)
Density Public Transport	STANDPT	Bus, Train, Metro	-0.16 (12.14)	-0.37 (12.54)	0.37 (8.39)	-0.16 (11.89)	-0.37 (12.36)	0.39 (9.30)
Experience Bus	EXPBS	Bus	-	-	-	-	0.29 (4.88)	-
Experience Train	EXPTR	Train	-	-	-	-0.11 (2.13)	-	-
Experience Car	EXPCR	Car	-	-	-	0.43 (8.99)	0.30 (3.24)	-
Concavity VL	CONC	All Alternatives	1.18 (6.04)	-	-	1.16 (5.35)	-	-

The VL models are nested, where the VL_MML_Exp has the most general form and the VL_MNL the most restricted form. The log likelihood ratio test is used to establish if: (1) the random parameter specification is appropriate, and (2) if the inclusion of experience is supported. To see if the random parameters specification significantly improves the model form, models VL_MML and VL_MML_Exp will be compared to the equivalent model forms without including random parameters, VL_MNL and VL_MNL_Exp, respectively. To analyse if the inclusion of experience significantly improves the models, models VL_MNL_Exp and VL_MML_Exp will be compared to models VL_MNL and VL_MML, respectively. The results are shown in Table 5-4. They show that both the inclusion of random parameters and experience significantly improves the models. Thus, the preferred model on statistical and behavioural grounds is VL_MML_Exp.

Table 5-4: Log likelihood ratio test results for the VL models

	Random parameters		**Experience**	
	VL_MML vs. VL_MNL	VL_MML_Exp vs. VL_MNL_Exp	VL_MNL_Exp vs. VL_MNL	VL_MML_Exp vs. VL_MML
LR	1814.416	1801.828	57.574	44.986
Degrees of freedom	10	10	2	2
$\chi^2_{d.f.;0.001}$	29.588	29.588	13.816	13.816
Result	Reject null	Reject null	Reject null	Reject null

Choice set correlation

As was explained in Section 3.4.2.1, it is important to analyse if the model form proposed for VL induces any type of relationship between the choice sets. Table 5-5 presents the results for the models used to study the unobserved part of the utility function, and see if the VL heuristic induces a relationship with the choice set sequence (or number). As defined in Chapter 3, the model form used to study this type of relationship is as follows:

$$U_{iq} = V_{iq} + \left[\varpi_{iqt} + \delta_i \cdot t\right] \cdot Y_{iqt} + \varepsilon_{iq} \tag{5.1}$$

where V_{iq} is the deterministic part of the utility function; ε_{iq} the traditional error term; and $\left[\varpi_{iqt} + \delta_i \cdot t\right] \cdot Y_{iqt}$ is an additional error term that varies across individuals *q*, alternatives *i*, and choice sets *t.* This additional error term is defined by ϖ_{iqt}, which is the part of the error that varies across individuals but not within individuals; t represents the sequence number of a choice set – 1; and δ_i is the parameter associated with the sequence number of a choice set for alternative *i*.

Models EC_LPAA and EC_VL represent a MNL model for the LPAA and VL heuristic, respectively, adding an error component common within individuals and different across individuals. These models do not consider a relationship between the choice set sequence and the error term; i.e., $\delta_i = 0$. As can be seen in Table 5-5, both models have similar parameter estimates and show there is a significant part of the unobserved utility function that varies across - but not within - individuals.

Models EC_SeqLPAA and EC_SeqVL include the traditional error component and the additional error component that depends on the choice set sequence (included as a continuous variable) shown in equation (5.1). Model EC_SeqLPAA shows that the choice set number significantly influences the unobserved part of the utility function for train and the metro, but not for the bus and car – when considering the traditional process strategy LPAA. It is interesting to note that when estimating the similar model but with the VL heuristic, only the train mode seems to be significantly influenced by the choice set number (with a 95% confidence level). In both models the traditional error components are statistically significant.

Therefore, results show that VL does not induce any relationship between the error term and the choice set sequence relative to a traditional LPAA heuristic. In fact, it reduces the influence of the continuous variable choice set number in one of the modes. These findings suggests that the proposed model form for VL is appropriate.

Table 5-5: Choice set correlation model results (t-values in brackets)

			EC_LPAA	EC_SeqLPAA	EC_VL	EC_SeqVL
Number of Parameters Estimated			**18**	**22**	**19**	**23**
Log Likelihood at convergence			**-5,042.73**	**-5,035.32**	**-5,033.40**	**-5,027.83**
Log likelihood at zero			**-13,125.44**			
AIC			**1.069**	**1.068**	**1.067**	**1.067**
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Mean**	**Mean**
Alternative Specific Constant Bus	ASCBUS	Bus	1.66 (3.70)	1.55 (3.54)	0.64 (2.42)	0.70 (2.64)
Alternative Specific Constant Train	ASCTRAIN	Train	1.97 (4.67)	1.74 (4.17)	0.84 (3.66)	0.83 (3.57)
Alternative Specific Constant Metro	ASCMETRO	Metro	2.06 (5.09)	1.92 (4.83)	0.95 (4.31)	0.98 (4.43)
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-
Access Time	ACTIMEPT	Bus, Train, Metro	-0.08 (11.39)	-0.08 (11.70)	-0.07 (7.51)	-0.07 (7.45)
Fare Public Transport	COSTPT	Bus, Train, Metro	-0.42 (13.60)	-0.41 (13.37)	-0.36 (11.15)	-0.36 (11.13)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.09 (3.51)	-0.10 (3.84)	-0.10 (3.44)	-0.10 (3.39)
Parking Cost Car	COSTCRPC	Car	-0.12 (11.91)	-0.13 (11.97)	-0.15 (8.25)	-0.15 (8.23)
Travel Time Public Transport	TTPT	Bus, Train, Metro	-0.05 (14.25)	-0.05 (13.68)	-0.07 (7.62)	-0.07 (7.57)
Travel Time Car	TTCR	Car	-0.04 (6.59)	-0.05 (7.47)	-0.04 (5.42)	-0.04 (5.42)
Egress Time	EGTIME	All Alternatives	-0.09 (10.28)	-0.10 (11.07)	-0.11 (8.76)	-0.12 (8.78)
Transfer Public Transport	TRANPT	Bus, Train, Metro	-0.23 (6.34)	-0.24 (6.29)	-0.22 (5.83)	-0.22 (5.78)
Headway Public Transport	FREQPT	Bus, Train, Metro	-0.02 (8.68)	-0.02 (8.69)	-0.02 (5.85)	-0.02 (5.86)
% Seat Public Transport	SEATPT	Bus, Train, Metro	0.73 (5.53)	0.75 (5.58)	0.71 (5.76)	0.71 (5.73)
Density Public Transport	STANDPT	Bus, Train, Metro	-0.30 (14.00)	-0.31 (13.91)	-0.32 (11.75)	-0.33 (11.74)
Concavity VL	CONC	All Alternatives	-	-	0.93 (2.47)	0.93 (2.47)
Error Components ASC Bus	EC_BUS	Bus	1.66 (8.44)	1.96 (5.56)	-1.74 (8.48)	-1.86 (5.08)
Error Components Sequence Bus	EC_BUS_SEQ	Bus	-	-0.08 (0.96)	-	0.04 (0.47)
Error Components ASC Train	EC_TRAIN	Train	0.62 (1.68)	-1.84 (7.05)	0.78 (2.97)	1.68 (6.58)
Error Components Sequence Train	EC_TRAIN_SEQ	Train	-	0.22 (4.67)	-	-0.21 (4.21)
Error Components ASC Metro	EC_METRO	Metro	1.90 (15.54)	1.41 (6.40)	-1.83 (14.61)	-1.58 (7.50)
Error Components Sequence Metro	EC_METRO_SEQ	Metro	-	0.10 (2.22)	-	-0.07 (1.55)
Error Components ASC Car	EC_CAR	Car	3.38 (13.74)	3.67 (11.49)	2.18 (10.92)	2.12 (6.93)
Error Components Sequence Car	EC_CAR_SEQ	Car	-	-0.04 (0.71)	-	0.02 (0.29)

5.2.3. Relative Advantage Maximisation

The relative advantage maximisation (RAM) heuristic proposes that decision making will be influenced by the relative advantage of one alternative over the competing ones, as explained in Section 3.5.3. Four models were estimated that consider RAM as the only process strategy being used in the sample, and the results are shown in Table 5-6.

Table 5-6: RAM MNL and MML models (t-values in brackets)

			RAM_MNL	RAM_MML		RAM_MNL_Exp	RAM_MML_Exp	
Number of Parameters Estimated			**14**	**25**		**17**	**28**	
Log Likelihood at convergence			**-6,340.49**	**-5,018.34**		**-6,140.35**	**-4,963.32**	
Log likelihood at zero			**-13,125.44**					
AIC			**1.342**	**1.065**		**1.301**	**1.054**	
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Std Dev**	**Mean**	**Mean**	**Std Dev**
Alternative Specific Constant Bus	ASCBUS	Bus	0.93 (5.57)	2.39 (4.56)	-	0.43 (2.04)	1.90 (3.28)	-
Alternative Specific Constant Train	ASCTRAIN	Train	0.96 (5.99)	2.39 (4.65)	-	0.79 (4.68)	2.95 (6.07)	-
Alternative Specific Constant Metro	ASCMETRO	Metro	1.24 (8.65)	2.30 (4.64)	-	1.20 (8.57)	3.07 (6.25)	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Bus, Train, Metro	-0.05 (18.26)	-0.09 (10.16)	0.14 (9.84)	-0.05 (19.00)	-0.09 (10.22)	0.17 (14.12)
Fare Public Transport	COSTPT	Bus, Train, Metro	-0.15 (10.48)	-0.55 (13.29)	0.46 (10.31)	-0.21 (12.51)	-0.56 (13.48)	0.44 (9.11)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.02 (2.00)	-0.17 (4.54)	0.36 (8.97)	-0.07 (3.56)	-0.25 (4.43)	0.38 (7.36)
Parking Cost Car	COSTCRPC	Car	-0.03 (8.84)	-0.17 (9.52)	0.13 (8.32)	-0.06 (8.16)	-0.38 (10.64)	0.38 (11.62)
Travel Time Public Transport	TTPT	Bus, Train, Metro	-0.02 (14.87)	-0.07 (14.66)	0.06 (13.93)	-0.03 (14.80)	-0.08 (13.20)	0.05 (9.99)
Travel Time Car	TTCR	Car	-0.02 (12.52)	-0.05 (6.12)	0.03 (3.39)	-0.04 (12.10)	-0.07 (7.10)	0.07 (9.55)
Egress Time	EGTIME	All Alternatives	-0.05 (17.20)	-0.12 (9.72)	0.13 (6.28)	-0.05 (16.26)	-0.11 (10.65)	0.10 (4.98)
Transfer Public Transport	TRANPT	Bus, Train, Metro	-0.10 (4.34)	-0.23 (5.81)	0.06 (0.39)	-0.11 (4.07)	-0.24 (5.72)	0.11 (0.64)
Headway Public Transport	FREQPT	Bus, Train, Metro	-0.01 (7.15)	-0.03 (9.94)	0.04 (13.76)	-0.01 (7.06)	-0.03 (9.89)	0.05 (12.89)
% Seat Public Transport	SEATPT	Bus, Train, Metro	0.34 (4.06)	0.86 (5.45)	1.44 (4.88)	0.39 (4.28)	0.93 (5.64)	0.91 (1.73)
Density Public Transport	STANDPT	Bus, Train, Metro	-0.17 (12.22)	-0.35 (12.36)	0.36 (9.16)	-0.18 (12.04)	-0.35 (12.13)	0.37 (8.62)
Experience Bus	EXPBS	Bus	-	-	-	0.37 (9.07)	0.42 (10.06)	-
Experience Train	EXPTR	Train	-	-	-	0.11 (3.97)	0.09 (2.11)	-
Experience Car	EXPCR	Car	-	-	-	0.65 (23.54)	0.65 (13.97)	-

Equivalently to what was presented for the other heuristics, the model referred to as RAM_MNL considers RAM with all parameters fixed, and model RAM_MML considers RAM and every parameter as random. Model RAM_MNL_Exp is equivalent to RAM_MNL model but with experience included; and model RAM_MML_Exp is equivalent to RAM_MML model also with experience. The results show that there is significant preference heterogeneity across the sample; allowing for random parameters improves the overall model fit significantly.

When considering experience together with RAM in the two final models, experience was statistically significant in decision making for every mode (bus, train and car) with all estimates positive. This implies that respondents are more likely to choose the same mode they used in their most recent trip.

The log likelihood ratio test is used to compare the models and the results are shown in Table 5-7. When comparing models RAM_MML and RAM_MML_Exp with models RAM_MNL and RAM_MNL_Exp, respectively, the results suggest that the consideration of random parameters instead of fixed ones significantly improves the model fit. The comparison between the models RAM_MNL_Exp and RAM_MML_Exp with the models RAM_MNL and RAM_MML, respectively, suggest that the inclusion of experience represents a significant improvement in the overall performance of the models. Therefore, the preferred model for the RAM heuristic is the one that considers random parameters and experience, RAM_MML_Exp.

Table 5-7: Log likelihood ratio test results for the RAM models

	Random parameters		**Experience**	
	RAM_MML vs. RAM_MNL	RAM_MML_Exp vs. RAM_MNL_Exp	RAM_MNL_Exp vs. RAM_MNL	RAM_MML_Exp vs. RAM_MML
LR	2644.314	2354.056	400.288	110.03
Degrees of freedom	11	11	3	3
$\chi^2_{d.f.;0.001}$	31.264	31.264	16.266	16.266
Result	Reject null	Reject null	Reject null	Reject null

5.3. Probabilistic Decision Process Model Results: LPAA, VL and RAM together with behavioural refinements and experience

The probabilistic decision process (PDP) method has been used in the literature to allow for process heterogeneity. Therefore, the models from this section will be fundamental to compare the method proposed in this thesis that allows for process and preference heterogeneity. It is important to remember that the PDP approach considers that each individual will use a process strategy up to a probability with multiple processing rules available to consider. The models have a latent class structure (although each class is not latent but a specific pre-defined rule), where each class represents a heuristic and the probabilities are represented by the class membership. The results of the models are presented in Table 5-8. In both models the first class represents the RAM heuristic, the second class VL, and the third one the LPAA assumption. The first model has a simple structure without any behavioural refinements or experience, referred to as PDP. The second model considers behavioural refinements in the LPAA heuristic (class 3) and experience in all the classes, referred to as PDP_BRExp.

Model PDP suggests that that there is a 0.50 probability that respondents use the RAM heuristic, 0.32 probability of using the VL heuristic and only 0.18 probability of using the LPAA heuristic. Some of the parameters were not found to be statistically significant, especially for the LPAA class where the fuel plus toll costs, number of transfers, and headway were not significant. In the RAM class the % of finding a seat available was not significant either. In the VL class every attribute was significant.

The PDP_BRExp model results show that there is a 0.37 probability that respondents use the RAM heuristic, with experience, a 0.18 probability of using the VL heuristic with experience and a 0.45 probability associated with the LPAA heuristic with behavioural refinements and experience. In the RAM heuristic the fuel plus toll costs and the number of transfers were not statistically significant. The other heuristics considered every attribute as significant. The only behavioural refinement that appeared to be significant in the LPAA class is towards the car parking costs, and there is a risk aversion attitude towards it. The graphical representation of this parameter is shown in Figure 5-5 where the transformation for the parking cost has a concave shape. Experience in travelling by bus was significant in the VL and LPAA heuristic and by car in the RAM and VL heuristic. Experience in the train did not show to be significant in any of the classes.

Table 5-8: PDP models (t-values in brackets)

			PDP			PDP_BRExp		
Number of Parameters Estimated			**40**			**47**		
Log Likelihood at convergence			**-5,072.33**			**-5,007.53**		
Log likelihood at zero			**-13,125.44**					
AIC			**1.080**			**1.068**		
Class Identification			**Class 1**	**Class 2**	**Class 3**	**Class 1**	**Class 2**	**Class 3**
Heuristic			RAM	VL	LPAA	RAM	VL	LPAA
Behavioural Refinements			N	N	N	Y	N	N
Experience			N	N	N	Y	Y	Y
Class Membership (%)			50%	32%	18%	37%	18%	45%
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**
Alternative Specific Constant Bus	ASCBUS	Bus	3.40 (5.74)	-3.44 (7.16)	1.92 (3.19)	-2.18 (3.51)	-2.08 (4.23)	2.54 (5.15)
Alternative Specific Constant Train	ASCTRAIN	Train	2.52 (4.35)	-0.85 (2.61)	1.72 (2.98)	-1.45 (2.78)	-1.03 (2.88)	2.86 (6.41)
Alternative Specific Constant Metro	ASCMETRO	Metro	3.33 (5.93)	-0.54 (1.99)	-0.13 (0.24)	0.14 (0.32)	-3.11 (7.88)	2.67 (6.5)
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Bus, Train, Metro	-0.07 (9.50)	-0.04 (4.36)	-0.09 (7.29)	-0.05 (5.09)	-0.07 (5.45)	-0.06 (10.01)
Fare Public Transport	COSTPT	Bus, Train, Metro	-0.38 (8.87)	-0.39 (7.57)	-0.19 (4.09)	-0.25 (3.93)	-0.24 (4.15)	-0.42 (10.74)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.12 (3.38)	-0.05 (1.54)	-	-	-0.12 (2.13)	-0.23 (5.29)
Parking Cost Car	COSTCRPC	Car	-0.13 (5.26)	-0.08 (5.37)	-0.12 (5.14)	-0.34 (5.71)	-0.12 (4.80)	-0.24 (4.42)
Travel Time Public Transport	TTPT	Bus, Train, Metro	-0.05 (11.41)	-0.02 (3.58)	-0.03 (6.1)	-0.05 (6.57)	-0.02 (4.23)	-0.05 (11.29)
Travel Time Car	TTCR	Car	-0.04 (4.75)	-0.03 (5.63)	-0.05 (4.22)	-0.14 (6.49)	-0.06 (5.03)	-0.03 (4.81)
Egress Time	EGTIME	All Alternatives	-0.07 (8.82)	-0.08 (6.57)	-0.05 (4.33)	-0.08 (6.05)	-0.04 (2.82)	-0.09 (10.51)
Transfer Public Transport	TRANPT	Bus, Train, Metro	-0.23 (4.36)	-0.17 (2.23)	-	-	-0.18 (1.81)	-0.25 (4.86)
Headway Public Transport	FREQPT	Bus, Train, Metro	-0.02 (6.32)	-0.01 (1.72)	-	-0.02 (3.04)	-0.01 (2.21)	-0.02 (5.41)
% Seat Public Transport	SEATPT	Bus, Train, Metro	-	0.79 (3.53)	0.98 (2.97)	0.58 (1.88)	0.80 (2.51)	0.63 (3.36)
Density Public Transport	STANDPT	Bus, Train, Metro	-0.21 (7.52)	-0.52 (7.69)	-0.12 (2.15)	-0.43 (7.40)	-0.19 (3.35)	-0.24 (7.65)
Experience Bus	EXPBS	Bus	-	-	-	-	0.78 (6.88)	0.26 (4.59)
Experience Train	EXPTR	Train	-	-	-	-	-	-
Experience Car	EXPCR	Car	-	-	-	0.73 (20.61)	0.34 (2.85)	-
Risk Attitudes Travel Time Bus	ALPHABSTT	Bus	-	-	-	-	-	-
Risk Attitudes Travel Time Train	ALPHATRTT	Train	-	-	-	-	-	-
Risk Attitudes Travel Time Metro	ALPHAMTTT	Metro	-	-	-	-	-	-
Risk Attitudes Travel Time Car	ALPHACRTT	Car	-	-	-	-	-	-
Risk Attitudes Cost Bus	ALPHABSCS	Bus	-	-	-	-	-	-
Risk Attitudes Cost Train	ALPHATRCS	Train	-	-	-	-	-	-
Risk Attitudes Cost Metro	ALPHAMTCS	Metro	-	-	-	-	-	-
Risk Attitudes Fuel+Toll Car	ALPHACRTRCS	Car	-	-	-	-	-	-
Risk Attitudes Parking Car	ALPHACRPCS	Car	-	-	-	-	-	0.75 (5.88)
Perceptual Conditioning Bus	GAMMABS	Bus	-	-	-	-	-	-
Perceptual Conditioning Car	GAMMACR	Car	-	-	-	-	-	-
Concavity VL	CONC	All Alternatives	-	-	-	-	-	-

It is notable how the inclusion of behavioural refinements and experience resulted in a major shift of class memberships, with a significant increase in the use of the LPAA heuristic (class 3). This finding emphasises the importance of including behavioural refinements which appear to have a statistically significant influence on preferences.

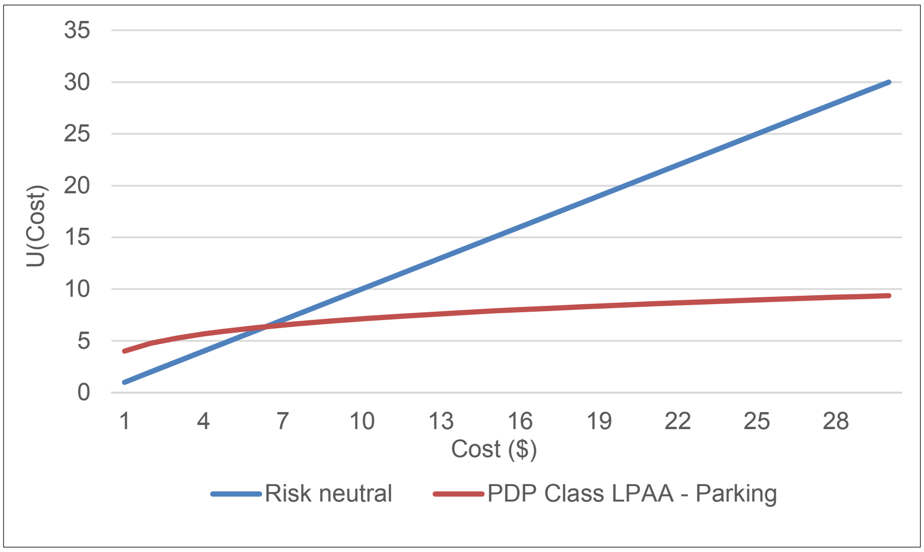


Figure 5-5: Risk attitudes towards parking cost in PDP_BRExp for the LPAA class

Both models had some attributes that were found to not be significant in a number of the heuristics. This is especially intriguing considering that in the models analysed before, where each heuristic was considered by itself, all the attributes were very significant. This might be suggesting that this type of model requires a larger sample size or that the PDP approach somehow confounds process heterogeneity with preference heterogeneity, and that is why some attributes are not significant in every class. These models are nested, so they can be compared using the log likelihood ratio test and the results are shown in Table 5-8. As can be seen, the incorporation of experience and behavioural refinements significantly improves the overall performance of the models. Therefore, the preferred model for this section is the PDP_BRExp.

Table 5-9: Log likelihood ratio test results for the PDP models

	Behavioural refinements and experience PDP_BRExp vs. PDP
LR	129.584
Degrees of freedom	7
$\chi^2_{d.f.;0.001}$	24.322
Result	Reject null

5.4. Conditioning Random Process Heterogeneity Model Results: LPAA, VL and RAM together with behavioural refinements and experience

The conditioning random process heterogeneity (CRPH) approach proposed in this thesis considers that there is a relationship between the process strategies and the mean and standard usually defined under an MML LPAA approach (for more details refer to Section 3.3). These relationships will also be referred to as interactions between the process strategies and the mean and standard deviation estimates. In this subsection, four models will be presented: (1) considering interactions between the process strategies and the mean estimates, referred to as CRPHm; (2) considering interactions between the process strategies and the standard deviation estimates, denoted as CRPHs; (3) considering the interactions both with the mean and standard deviation estimates, referred to as CRPHms; and (4) same as CRPHms but considering behavioural refinements (i.e., risk attitudes and perceptual conditioning as defined in Section 3.5.1) and experience.

As explained in Section 3.5.4, the utility expression for this model can be written as follows:

$$U_i = \sum_n \left(\begin{bmatrix} \theta_{in} + \lambda^m_{VL,in} \cdot VL\left(x_{inqt}\right) + \lambda^m_{RAM,in} \cdot RAM\left(x_{inqt}\right) \\ + \left[\sigma_{in} + \lambda^s_{VL,in} \cdot VL\left(x_{inqt}\right) + \lambda^s_{RAM,in} \cdot RAM\left(x_{inqt}\right) \right] \cdot v \end{bmatrix} \cdot x_{inqt} \right) + \varepsilon_{iqt} \qquad (5.2)$$

Table 5-10 and 5-11 show the parameter estimates for the different types of models.

Table 5-10: CRPH models with interactions in the mean or standard deviation estimates (t-values in brackets)

			CRPHm				CRPHs			
Number of Parameters Estimated			**28**				**33**			
Log Likelihood at convergence			**-4,993.87**				**-4,988.00**			
Log likelihood at zero			**-13,125.44**							
AIC			**1.061**				**1.061**			
Parameters	**Acronym**	**Alternatives**	θ	σ	λ_{VL}^{m}	λ_{RAM}^{m}	θ	σ	λ_{VL}^{s}	λ_{RAM}^{s}
Alternative Specific Constant Bus	ASCBUS	Bus	2.11 (4.19)	-	-	-	2.47 (5.34)	-	-	-
Alternative Specific Constant Train	ASCTRAIN	Train	2.10 (4.24)	-	-	-	2.40 (5.39)	-	-	-
Alternative Specific Constant Metro	ASCMETRO	Metro	1.99 (4.17)	-	-	-	2.27 (5.28)	-	-	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-	-	-
Access Time	ACTIMEPT	Bus, Train, Metro	-0.11 (10.56)	-0.16 (11.59)	-	-	-0.09 (11.95)	-0.22 (12.19)	-	-0.05 (7.01)
Fare Public Transport	COSTPT	Bus, Train, Metro	-0.83 (9.19)	-0.44 (7.99)	-	-0.03 (3.80)	-0.55 (14.75)	-0.57 (8.73)	-0.07 (3.01)	-0.12 (6.26)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.04 (3.78)	0.25 (6.45)	-	-0.32 (7.43)	-0.15 (4.42)	0.47 (9.68)	-0.04 (2.23)	-
Parking Cost Car	COSTCRPC	Car	-0.26 (9.93)	-0.19 (11.91)	-	-	-0.24 (11.96)	-0.24 (11.29)	0.00 (1.68)	-
Travel Time Public Transport	TTPT	Bus, Train, Metro	-0.09 (15.42)	0.07 (12.37)	-	-	-0.09 (15.48)	0.08 (13.49)	-	-
Travel Time Car	TTCR	Car	-0.06 (6.01)	0.02 (3.63)	-	-	-0.06 (6.46)	0.02 (4.03)	-	-
Egress Time	EGTIME	All Alternatives	-0.13 (11.29)	0.12 (7.11)	-	-	-0.12 (11.47)	0.11 (6.83)	-	-
Transfer Public Transport	TRANPT	Bus, Train, Metro	-0.27 (5.82)	0.42 (4.70)	-	-	-0.26 (5.85)	0.42 (4.42)	-	-
Headway Public Transport	FREQPT	Bus, Train, Metro	-0.03 (9.26)	-0.05 (13.22)	-	-	-0.03 (9.97)	-0.07 (9.92)	-	-0.03 (4.29)
% Seat Public Transport	SEATPT	Bus, Train, Metro	1.54 (6.56)	1.33 (4.46)	-1.19 (3.13)	-	0.97 (5.92)	1.36 (4.81)	-	-
Density Public Transport	STANDPT	Bus, Train, Metro	-0.34 (10.84)	0.38 (9.53)	-	-	-0.35 (12.18)	0.53 (8.45)	-	0.26 (3.33)
Experience Bus	EXPBS	Bus	-	-	-	-	-	-	-	-
Experience Train	EXPTR	Train	-	-	-	-	-	-	-	-
Experience Car	EXPCR	Car	-	-	-	-	-	-	-	-
Concavity VL	CONC	All Alternatives	-	-	-	-	-	-	0.73 (1.74)	-

Table 5-11: CRPH models with interactions in the mean and standard deviation estimates (t-values in brackets)

			CRPHms						CRPHms_BRExp					
Number of Parameters Estimated			**43**						**35**					
Log Likelihood at convergence			**-4,939.53**						**-4,922.41**					
Log likelihood at zero			**-13,125.44**											
AIC			**1.052**						**1.047**					
Parameters	**Acronym**	**Alternatives**	θ	σ	λ_{VL}^{m}	λ_{RAM}^{m}	λ_{VL}^{s}	λ_{RAM}^{s}	θ	σ	λ_{VL}^{m}	λ_{RAM}^{m}	λ_{VL}^{s}	λ_{RAM}^{s}
Alternative Specific Constant Bus	ASCBUS	Bus	2.36 (4.98)	-	-	-	-	-	1.60 (2.63)	-	-	-	-	-
Alternative Specific Constant Train	ASCTRAIN	Train	2.33 (5.03)	-	-	-	-	-	2.44 (4.41)	-	-	-	-	-
Alternative Specific Constant Metro	ASCMETRO	Metro	2.23 (4.99)	-	-	-	-	-	2.75 (5.27)	-	-	-	-	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-	-	-	-	-	-	-
Access Time	ACTIMEPT	Bus, Train, Metro	-0.08 (10.85)	-0.20 (8.56)	-	-0.01 (2.06)	-	-0.07 (5.67)	-0.08 (8.93)	-0.22 (10.81)	-	-	-	-0.05 (3.63)
Fare Public Transport	COSTPT	Bus, Train, Metro	-0.81 (8.87)	-0.55 (7.02)	-	-0.11 (3.62)	-0.03 (4.20)	-0.09 (4.64)	-0.92 (10.36)	-0.82 (9.12)	-	-	-0.03 (6.84)	-0.05 (6.12)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.05 (4.55)	0.38 (6.77)	-	-0.03 (2.01)	-0.32 (7.77)	-	-	0.18 (3.49)	-	-	-0.27 (5.53)	-
Parking Cost Car	COSTCRPC	Car	-0.22 (6.15)	-0.13 (6.73)	-0.01 (3.35)	-0.01 (1.93)	-0.01 (2.66)	-	-0.33 (6.77)	0.13 (3.55)	-0.002 (2.05)	-	-0.01 (4.74)	-
Travel Time Public Transport	TTPT	Bus, Train, Metro	-0.08 (8.66)	0.06 (9.67)	-0.001 (2.21)	-	-0.003 (2.37)	-	-0.10 (9.06)	0.03 (6.86)	-	-	-0.003 (3.53)	-
Travel Time Car	TTCR	Car	-0.04 (3.13)	0.02 (3.85)	-0.001 (2.33)	-	-	-	-0.09 (7.47)	0.07 (8.10)	-	-	-	-
Egress Time	EGTIME	All Alternatives	-0.13 (12.07)	0.11 (8.14)	-	-	-	-	-0.13 (10.52)	0.12 (7.53)	-	-	-	-
Transfer Public Transport	TRANPT	Bus, Train, Metro	-0.26 (5.79)	0.38 (4.16)	-	-	-	-	-0.26 (5.82)	-	-	-	-	-
Headway Public Transport	FREQPT	Bus, Train, Metro	-0.03 (9.36)	-0.09 (8.76)	-	0.002 (2.10)	-	-0.02 (2.10)	-0.03 (10.00)	0.09 (11.10)	-	-0.001 (5.09)	-	-
% Seat Public Transport	SEATPT	Bus, Train, Metro	1.50 (6.49)	1.01 (3.41)	-0.97 (3.06)	-	-	-	1.59 (6.79)	-	-1.45 (3.88)	-	-	-
Density Public Transport	STANDPT	Bus, Train, Metro	-0.32 (10.78)	0.52 (8.19)	-	-	-	0.26 (3.13)	-0.33 (10.90)	0.53 (8.15)	-	-	-	0.26 (3.19)
Experience Bus	EXPBS	Bus	-	-	-	-	-	-	0.29 (9.94)	-	-	-	-	-
Experience Train	EXPTR	Train	-	-	-	-	-	-	0.09 (3.56)	-	-	-	-	-
Experience Car	EXPCR	Car	-	-	-	-	-	-	0.37 (7.49)	-	-	-	-	-
Concavity VL	CONC	All Alternatives	-	-	0.66 (4.72)	-	0.66 (4.72)	-	-	-	-	-	-	-

The results show that none of the behavioural refinements appear to be statistically significant when allowing for process heterogeneity in this form. Different combinations were tested, but none of them were statistically significant. However, experience was statistically significant when conditioning the entire utility expression. Model CRPHms_BRExp included experience using the bus, train and car with all of the estimates being positive. This suggests that, if an individual used any of these modes on his most recent trip, he will be more likely to choose the same mode again.

The concavity factor for the VL process strategy was statistically significant in the CRPHs and CRPHms model. The CRPHm and CRPHms_BRExp models had a concavity factor equal to 1, which represents a linear consideration of the differences between the attribute levels and the reference level. The concavity factors are graphically shown in Figure 5-6, which show that in model CRPHs and model CRPHms individuals tend to underweight higher differences.

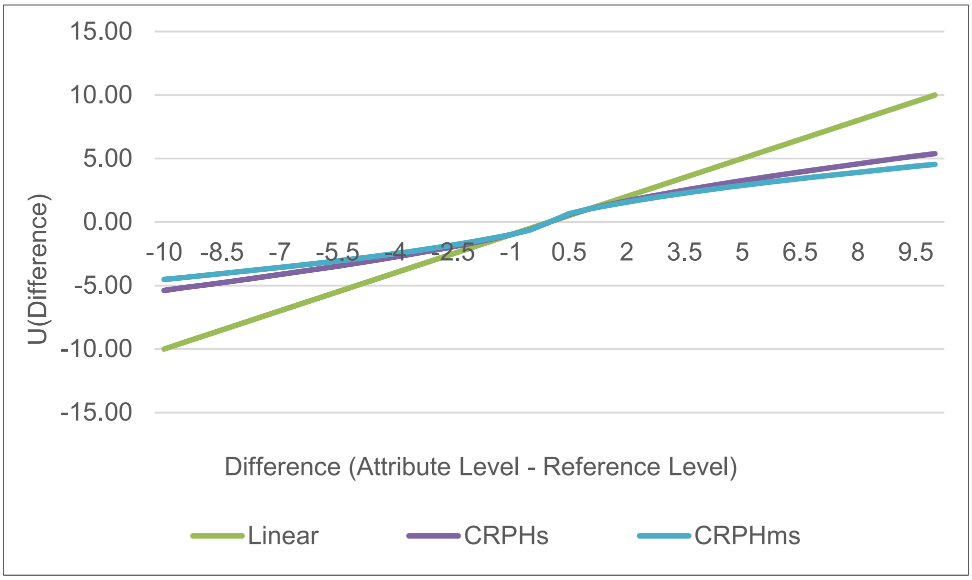


Figure 5-6: Concavity factor for CRPH models

Results show that the interactions that were significant in models CRPHm and CRPHs were also statistically significant in model CRPHms. Since models CRPHm and CRPHs are restricted versions of model CRPHms, they can be compared using the log likelihood ratio test. It is interesting to note that a greater number of interactions with the mean and with the standard deviation estimates were significant in the CRPHms model in comparison to the restricted models. Table 5-12 shows the results of the log likelihood ratio test, which suggest that model CRPHms is not equivalent to models CRPHm and CRPHs. Hence, including interactions with process strategies both in the mean and standard deviation estimates significantly improves the model fit.

Table 5-12: Log likelihood ratio test results for CRPH models

	Interactions	
	CRPHms vs. CRPHm	CRPHms vs. CRPHs
LR	108.674	96.952
Degrees of freedom	15	10
$\chi^2_{d.f.;0.001}$	37.697	29.588
Result	Reject null	Reject null

When including experience in model CRPHms, several of the interactions between the process strategies and the mean and/or standard deviation estimates were no longer statistically significant. These models are not nested and can only be compared using the Vuong statistic. The results of this test are summarised in Table 5-13. The Vuong statistic is larger than 2.60, and thus this test favours the CRPHms_BRExp model with a 99% confidence level. Therefore, our preferred model using the CRPH approach is CRPHms_BRExp.

Table 5-13: Vuong statistic test result for CRPH models

	CRPHms_BR vs. CRPHms
Mean	0.0581
Std Dev	0.7640
Sample Size	9468
Vuong Statistic	7.404
Result	Favours CRPHms_BR model

It is also important to look at the interactions that seem to be statistically significant in each model. The results are shown in Table 5-14. The model that includes only interactions with the mean estimates, CRPHm, found only three statistically significant interactions: fare public transport and fuel plus toll cost car with the RAM process strategy, and % of seats available with the VL process strategy. When considering the interactions only with the standard deviation estimates, seven interactions were statistically significant: four with the RAM heuristic and three with the VL heuristic.

A greater number of interactions were statistically significant when considering them in the mean and standard deviations for the CRPHms model. This result suggests that the most appropriate way of considering process heterogeneity using the CRPH approach includes interactions with both the mean and the standard deviation. This model includes four interactions between the VL process strategy and the mean estimate: parking cost, travel time in public transport and car, and % of seats available. Five interactions between the RAM heuristic and the mean estimate: access time, public transport fare, fuel + toll cost, parking cost, and headway. Four interactions are statistically significant between the VL heuristic and the standard deviation: public transport fare, fuel + toll cost, parking cost, and public transport travel time. Four interactions are statistically significant between the RAM heuristic and the standard deviation: access time, public transport fare, headway and standing density in public transport. The findings suggest that the travel times are mainly influenced by the VL heuristic. The fares/costs are influenced by the RAM heuristic in their mean, and by the RAM and VL heuristic in their standard deviation.

Table 5-14: CRPH models interactions with process strategies

	CRPHm		**CRPHs**		**CRPHms**				**CRPHms_BRExp**			
	Mean		Std Dev		Mean		Std Dev		Mean		Std Dev	
	VL	RAM	VL	RAM	VL	RAM	VL	RAM	VL	RAM	VL	RAM
Access Time				X		X		X				X
Fare Public Transport		X	X	X		X	X	X			X	X
Fuel + Toll Cost Car		X	X			X	X				X	
Parking Cost Car			X		X	X	X		X		X	
Travel Time Public Transport					X		X				X	
Travel Time Car					X							
Egress Time												
Transfer Public Transport												
Headway Public Transport				X		X		X		X		
% Seat Public Transport	X				X				X			
Density Public Transport				X				X				X
Total # of Interactions	**3**		**7**		**17**				**10**			

5.5. Comparison of the Models

This section compares the preferred models using general indicators, such as the log likelihood and AIC, and more importantly, the willingness to pay (WTP) estimates. We include both the LPAA_MNL and LPAA_MML models; even though they were not the preferred model in the LPAA section, they are an excellent reference point given their dominant use in choice studies.

5.5.1. Behavioural Refinements, Experience and Concavity Factor

Risk attitudes, perceptual conditioning and/or experience were included in the preferred models presented above. However, the results varied when using different process strategies or process heterogeneity approaches.

In the process homogeneity models, experience was significant towards all the modes of transportation when using the LPAA and the RAM processing rules. However, when using the VL strategy, experience towards the train was not statistically significant so the final model only includes experience associated with the bus and car. These findings appear to be suggesting that experience - to a certain level – is confounded with the process strategies. The VL heuristic takes into account reference levels with starting values equal to the characteristics of the mode individuals used on their most recent trip (i.e., experience). Therefore, it could be expected that this heuristic had a different interaction with respondents' experience.

If there is no confounding between experience and process heterogeneity, one would not expect the experience parameters to change when taking into account process heterogeneity relative to a process homogeneity model. However, when considering process heterogeneity through the PDP approach, experience was statistically significant only towards the bus in the LPAA class, towards the bus and car in the RAM class, and only towards the car in the VL class. These results suggest that the experience that is statistically significant under a process homogeneity approach is partly explaining process heterogeneity under a PDP approach. On the other hand, when using the CRPH method to include process heterogeneity, experience towards all the modes was statistically significant. The CRPH method essentially estimates the parameters as considered under an LPAA traditional process strategy adding interactions with alternative heuristics. It might thus be expected that the findings align with the traditional

LPAA process homogeneity model, and this is exactly what the results suggest. The CRPH model was significantly influenced by experience in the same way that the LPAA process homogeneity approach was.

Regarding behavioural refinements, in the LPAA process homogeneity model (LPAA_MML_BRExp), risk attitudes were found to be statistically significant towards the bus cost and parking cost: risk aversion and risk taking attitudes, respectively. In the PDP approach, the LPAA class had a significant risk taking attitude towards the parking cost, similarly for the LPAA_MML_BRExp model. This result is as expected, since it was expected that the risk attitudes were partially taking into account process heterogeneity. When using the CRPH approach, none of the risk attitudes seem to be statistically significant, which shows that this approach is taking into consideration other things not considered in the PDP approach. In the case of perceptual conditioning, it was only considered statistically significant in the LPAA_MML_BRExp model towards the car. It was not statistically significant in any of the other models, which suggests that - contrary to what was expected – perceptual conditioning is influenced by the consideration of process heterogeneity.

One of the main findings when comparing the influence of experience, risk attitudes and perceptual conditioning in the different models is that they seem to be more significant when considering process homogeneity than when considering process heterogeneity. This is a crucial finding that suggests that the importance of including additional behavioural components is reduced when considering process heterogeneity.

For the concavity factor, only four models estimated it as significant: VL_MNL, VL_MNL_Exp, CRPHs and CRPHms. All the other models consider a linear evaluation of the difference between the attribute level and the reference level. The estimated concavity factors are shown in Table 5-7. It is interesting to note that when considering fixed parameters and process homogeneity, individuals tend to overweight higher differences, while in the process heterogeneity CRPHs and CRPHms, individuals tend to underweight higher differences. This result reveals important differences between these models, although a majority of the models suggests a linear evaluation of the differences.

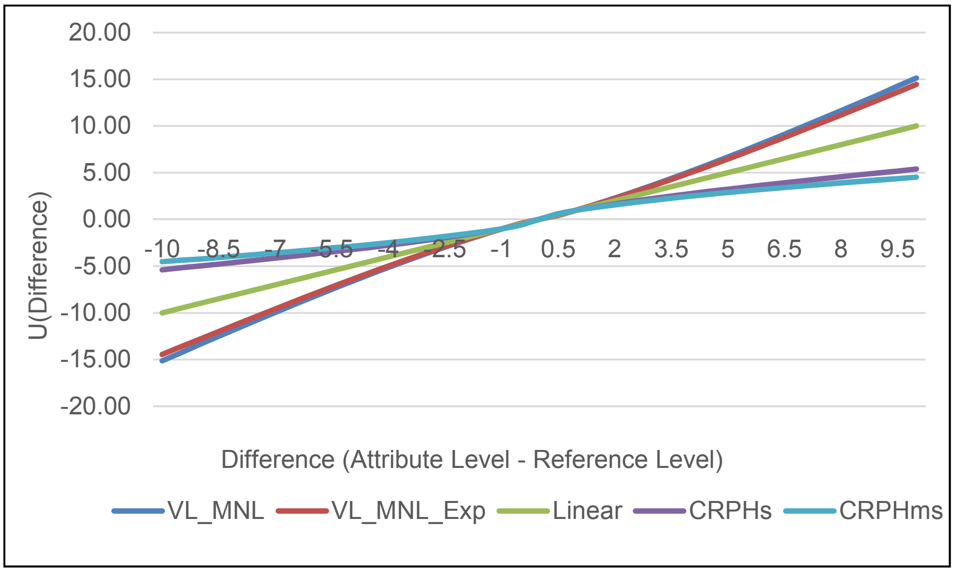


Figure 5-7: Concavity factors' comparison

5.5.2. Log likelihood and AIC

The models presented will be compared in terms of general indicators, such as their log likelihood and AIC indicators. As explained in Section 3.8.2, the AIC indicator values the log likelihood while penalising the number of parameters estimated, so it is a very helpful statistic for an initial appreciation of the models. The results are shown in Table 5-15 and the AIC is presented in the last column using a colour scale, where a darker green colour represents a better AIC indicator, and darker red colour represents a worse AIC indicator. As can be seen, the models that have a superior AIC are the CRPHms_BRExp and the CRPHms. These are followed by the LPAA_MML_BRExp and RAM_MML_Exp models. The most inferior model in terms of the AIC is the RAM_MNL and LPAA_MNL model. Interestingly, with all else being held constant, when preference but not process heterogeneity is taken into account (i.e., MML models), the LPAA_MML model outperforms the VL_MML model, but the RAM_MML model outperforms the other two. Moreover, in absence of preference and process heterogeneity, the VL_MNL outperforms the other two, and the LPAA_MNL model seems more appropriate than the RAM_MNL. These results may be indicating some sort of confounding between preference heterogeneity and the process strategies.

Table 5-15: General indicators

	Number of Parameters Estimated	Log Likelihood at convergence	Log likelihood at zero	AIC
LPAA_MNL	14	-6,326.73	-13,125.44	1.339
LPAA_MML	25	-5,024.55		1.067
LPAA_MML_BRExp	31	-4,958.17		1.054
VL_MNL	15	-5,973.30		1.265
VL_MML	25	-5,066.10		1.075
VL_MML_Exp	27	-5,043.60		1.071
RAM_MNL	14	-6,340.49		1.342
RAM_MML	25	-5,018.34		1.065
RAM_MML_Exp	28	-4,963.32		1.054
PDP	40	-5,072.33		1.080
PDP_BRExp	47	-5,007.53		1.068
CRPHm	28	-4,993.87		1.061
CRPHs	33	-4,988.00		1.061
CRPHms	43	-4,939.53		1.052
CRPHms_BRExp	35	-4,922.41		1.047

5.5.3. Willingness to Pay Estimates

The willingness to pay estimates (WTP) are often the most important outcome in a choice study. Previously this section presented different types of models, some of them including only one heuristic, and others that consider process heterogeneity.

The mean or median and confidence intervals of the WTP are used in economic appraisal of transport projects as a crucial measure of user benefits, so it is extremely important to have an adequate estimate. Even though the preferred models consider behavioural refinements and/or experience, it is essential to understand how the different components of the models affect the WTP. The range of models will be included in this section to:

1) Study how the WTP estimates vary when using different heuristics in the process homogeneity models
2) Analyse the influence on WTP of considering different process heterogeneity approaches: PDP versus CRPH
3) Understand how process heterogeneity affects WTP through comparing these models with the ones that allow for process homogeneity, and
4) Study how behavioural refinements and/or experience affect the WTP estimates.

Table 5-16 and 5-17 present the WTP estimates for the models that consider process homogeneity and process heterogeneity, respectively. As mentioned in Section 3.7, if the cost attribute in a MML model is considered as random and normally distributed it will necessarily have a value very close to zero for certain draws. This will cause the WTP estimate to be close to infinity (or negative infinity) and, therefore, the mean of the WTP estimates will be very unstable and will depend heavily on the draws. The median estimate will be more robust, so this one will be used for comparison as was proposed by Bliemer and Rose (2013). In the models where the cost attribute is fixed, the mean will be equal to the median.

Table 5-16 and 5-17 are not easy to interpret directly. Therefore, the next subsections will focus on comparing the median, standard error, and stability of the WTP estimates of a subset of models to address the issues mentioned above. The first part will focus on the value of travel time savings (VTTS) for public transport and car because it is considered one of the most important metrics in user time benefit calculations, and the next part will present the WTP for of the other attributes.

The stability of the estimates will be studied by plotting the WTP for each of the 25,000 different draws. The WTP estimates are dependent on the draws only for the models that consider random parameters. Therefore, the models that consider all parameters as fixed will have a completely stable WTP estimate and will not be included in these graphs. The standard deviation of these graphs will represent the stability of the WTP estimate in the models considering different draws. It is important to mention that this measure of stability is not the same as the WTP standard error as this is calculated using the Delta method (as explained in Section 3.7).

Table 5-16: Willingness to pay estimates for the models with process homogeneity (t-values in brackets)

		LPAA_MNL		LPAA_MML		LPAA_MML_BRExp		VL_MML		VL_MML_Exp		RAM_MML		RAM_MML_Exp	
		Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error
Travel Time $/person hour	**Bus**	7.84	0.90	6.20	19.29	4.50	12.63	5.59	19.52	6.43	19.90	5.95	17.79	6.37	21.66
	Train	7.84	0.90	6.20	19.29	6.03	16.88	5.59	19.52	6.43	19.90	5.97	17.77	6.38	21.70
	Metro	7.84	0.90	6.20	19.29	6.03	16.88	5.59	19.52	6.43	19.90	5.96	17.82	6.33	21.76
	Car	50.17	10.45	14.35	14.43	15.26	13.02	13.66	30.79	13.94	33.81	13.76	18.76	8.80	14.59
Access Time $/person hour	**Bus**	14.94	1.56	7.04	19.29	4.88	12.63	5.98	19.52	6.46	19.90	6.64	17.79	7.46	21.66
	Train	14.94	1.56	7.04	19.29	6.53	16.88	5.98	19.52	6.46	19.90	6.64	17.77	7.47	21.70
	Metro	14.94	1.56	7.04	19.29	6.53	16.88	5.98	19.52	6.46	19.90	6.64	17.82	7.48	21.76
Egress Time $/person hour	**Bus**	17.93	1.82	9.11	18.48	5.94	14.06	8.54	21.89	9.89	21.45	9.79	19.47	9.01	14.86
	Train	17.93	1.82	9.11	18.48	7.95	18.81	8.54	21.89	9.89	21.45	9.80	19.48	9.03	14.87
	Metro	17.93	1.82	9.11	18.48	7.95	18.81	8.54	21.89	9.89	21.45	9.85	19.65	9.06	14.94
	Car	125.89	24.26	28.51	53.81	18.08	37.94	31.81	75.32	26.96	55.93	33.07	65.06	14.74	26.28
Headway $/person minute	**Bus**	0.05	0.01	0.03	0.10	0.03	0.08	0.03	0.11	0.04	0.12	0.04	0.10	0.04	0.10
	Train	0.05	0.01	0.03	0.10	0.04	0.11	0.03	0.11	0.04	0.12	0.04	0.10	0.04	0.10
	Metro	0.05	0.01	0.03	0.10	0.04	0.11	0.03	0.11	0.04	0.12	0.04	0.10	0.04	0.10
Seat To increase % of seating time by 100%	**Bus**	1.90	0.47	1.15	3.36	0.86	2.42	0.99	3.72	1.17	4.15	1.06	3.17	1.23	2.61
	Train	1.90	0.47	1.15	3.36	1.15	3.22	0.99	3.72	1.17	4.15	1.06	3.16	1.23	2.60
	Metro	1.90	0.47	1.15	3.36	1.15	3.22	0.99	3.72	1.17	4.15	1.06	3.17	1.23	2.61
Stand To reduce density by 1 standees per square metre	**Bus**	0.89	0.11	0.44	0.88	0.30	0.65	0.41	0.92	0.44	0.95	0.44	0.84	0.47	0.85
	Train	0.89	0.11	0.44	0.88	0.40	0.86	0.41	0.92	0.44	0.95	0.44	0.84	0.47	0.85
	Metro	0.89	0.11	0.44	0.88	0.40	0.86	0.41	0.92	0.44	0.95	0.44	0.84	0.47	0.85
Transfers To reduce the number of transfers by 1	**Bus**	0.54	0.13	0.30	0.87	0.23	0.59	0.25	0.67	0.28	0.74	0.32	0.51	0.34	0.56
	Train	0.54	0.13	0.30	0.87	0.31	0.79	0.25	0.67	0.28	0.74	0.32	0.51	0.34	0.56
	Metro	0.54	0.13	0.30	0.87	0.31	0.79	0.25	0.67	0.28	0.74	0.31	0.51	0.34	0.56

Table 5-17: Willingness to pay estimates for the models with process heterogeneity

		PDP		PDP_BRExp		CRPHm		CRPHs		CRPHms		CRPHms_BRExp	
		Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error
Travel Time $/person hour	**Bus**	6.90	1.35	8.50	1.80	6.43	15.98	6.00	30.50	5.78	26.06	4.53	24.83
	Train	6.89	1.35	8.37	1.79	6.55	16.85	6.90	25.50	6.84	23.42	5.20	22.71
	Metro	6.84	1.36	8.36	1.79	6.43	15.94	5.84	26.24	5.76	24.40	5.87	21.86
	Car	18.15	4.85	23.60	3.93	4.75	10.10	13.97	11.35	8.58	7.35	13.74	30.70
Access Time $/person hour	**Bus**	10.26	1.35	11.25	1.80	7.82	15.98	5.49	30.50	4.23	26.06	4.02	24.83
	Train	10.24	1.35	11.33	1.79	7.97	16.85	6.23	25.50	5.00	23.42	4.10	22.71
	Metro	10.34	1.36	11.34	1.79	7.81	15.94	5.10	26.24	3.68	24.40	3.83	21.86
Egress Time $/person hour	**Bus**	12.49	2.01	14.48	2.43	9.61	12.96	8.35	17.90	8.25	13.90	7.47	13.77
	Train	12.49	2.01	14.32	2.35	9.78	13.71	9.60	17.47	9.80	13.69	8.12	14.39
	Metro	12.59	2.03	14.31	2.35	9.58	12.91	8.01	18.02	8.55	13.50	7.52	13.45
	Car	42.40	6.93	32.89	5.50	8.58	54.09	28.99	43.71	9.33	32.24	14.44	47.56
Headway $/person minute	**Bus**	0.04	0.01	0.06	0.01	0.04	0.08	0.02	0.14	0.03	0.09	0.04	0.08
	Train	0.04	0.01	0.06	0.01	0.04	0.09	0.03	0.13	0.03	0.09	0.04	0.08
	Metro	0.04	0.01	0.06	0.01	0.04	0.08	0.03	0.24	0.03	0.19	0.03	0.14
Seat To increase % of seating time by 100%	**Bus**	1.12	0.26	1.96	0.49	1.94	2.52	1.07	3.20	1.87	2.58	1.82	1.81
	Train	1.12	0.26	1.98	0.48	2.01	2.68	1.23	3.15	2.22	2.50	2.00	1.87
	Metro	1.13	0.26	1.98	0.48	1.95	2.52	1.04	3.20	1.91	2.47	1.81	1.76
Stand To reduce density by 1 standees per square metre	**Bus**	0.79	0.11	0.96	0.14	0.43	0.67	0.35	1.59	0.32	1.25	0.30	1.27
	Train	0.79	0.11	0.94	0.14	0.43	0.71	0.41	1.52	0.39	1.19	0.33	1.26
	Metro	0.79	0.11	0.94	0.14	0.42	0.67	0.34	1.81	0.33	1.38	0.30	1.39
Transfers To reduce the number of transfers by 1	**Bus**	0.47	0.09	0.42	0.09	0.33	0.73	0.29	0.96	0.28	0.75	0.29	0.29
	Train	0.47	0.09	0.43	0.10	0.33	0.77	0.33	0.95	0.34	0.76	0.32	0.30
	Metro	0.48	0.09	0.43	0.10	0.33	0.73	0.28	0.96	0.29	0.73	0.29	0.28

5.5.3.1. Value of Travel Time Savings

5.5.3.1.1. Process Homogeneity using Different Heuristics

Figure 5-8 graphically presents the VTTS median and the 95% confidence intervals for public transport and car for the three models that represent process homogeneity: LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp.

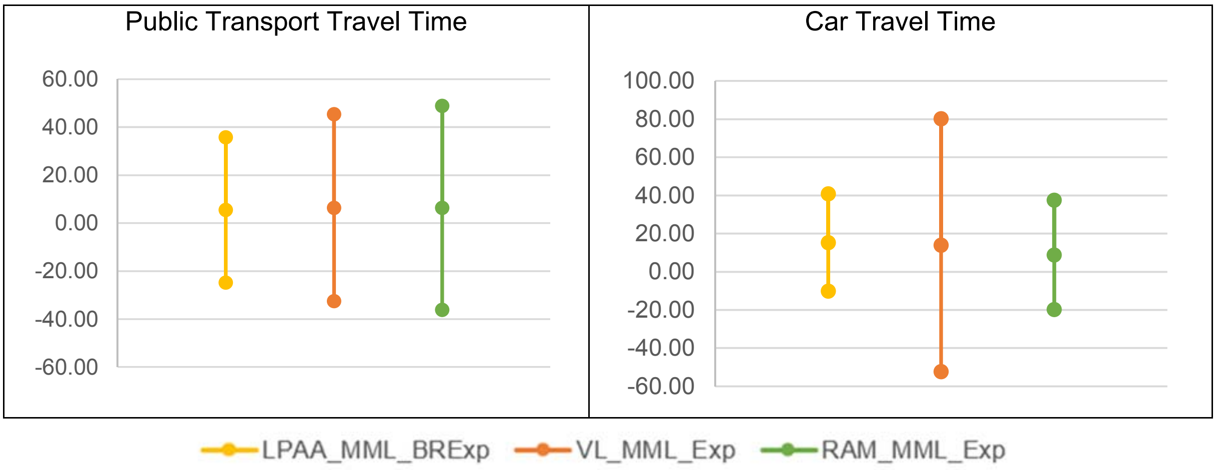


Figure 5-8: VTTS median and the 95% confidence intervals for the models with process homogeneity

The median VTTS for public transport is similar when using different process strategies (around $6 per person hour). The standard error is largest for the RAM heuristic ($21.71) with the LPAA heuristic having the smallest standard error ($15.46). The median VTTS when using the car is very similar for the LPAA model and the VL model ($15.26 and $13.94 per person hour, respectively), although the standard error for the RAM model is significantly higher ($33.81 for the VL model versus $13.02 for the LPAA model). For the RAM model, the median VTTS in car is significantly lower ($8.80 per person hour) with the standard error ($14.59) closer to the one estimated using the LPAA model.

For each mode of transportation, the median VTTS were compared across the process homogeneity models using the t-test with the median and standard error (Section 3.8.1) to see if they were statistically different from each other. Table 5-18 presents the results, where an absolute value larger than 1.96 represents statistically different estimates at a 95% confidence level. As can be seen, the median VTTS for the bus and car under the VL and RAM model are statistically different from the median VTTS of the LPAA model. However, there is not enough evidence to suggest they are different for the train and metro. The median VTTS for the car of the RAM and VL model are statistically different, but there is also not enough evidence to suggest they are statistically different for the public transport modes.

Table 5-18: Comparison of attributes' median VTTS for models with process homogeneity using t-test

Travel Time	VL_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs VL_MML_BRExp
Bus	***4.49***	***4.10***	-0.11
Train	1.19	1.00	-0.12
Metro	1.51	1.08	-0.32
Car	***-2.18***	***-19.78***	***-8.36***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 5-9 shows the distribution of the VTTS considering 25,000 different draws using the models that represent process homogeneity, which is useful in order to analyse the stability of VTTS. The results show that the stability of the VTTS for public transport is similar when using different process strategies, but it is slightly worse in the VL model and better in the RAM model. The car VTTS' stability is better in the LPAA model followed closely by the RAM model, and it is significantly worse in the VL model.

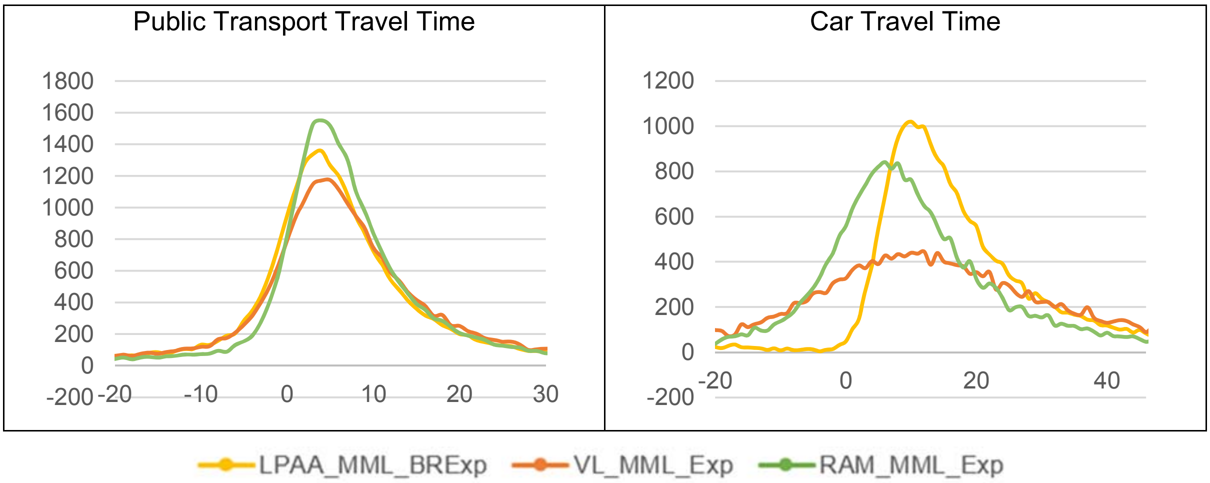


Figure 5-9: VTTS distribution for different draws using the models with process homogeneity

The results of this section suggest that the estimates for the public transport VTTS for models that use different process strategies are not considerably different in terms of their median or stability, although there are meaningful differences in the standard errors of the models. For the car VTTS, the differences are significant in terms of the median estimate, standard error and stability. Therefore, the results suggest that the process strategy considered has an important influence on the estimates and conclusions.

Summary: Comparison of the preferred process homogeneity models LPAA_MML_BRExp; RAM_MML_Exp and VL_MML_Exp

Public transport VTTS

- median (lower to higher): LPAA, RAM, VL
- standard error (lower to higher): LPAA, RAM, VL

Car VTTS

- median (lower to higher): RAM, VL, LPAA
- standard error (lower to higher): LPAA, RAM, VL

Significant differences? The bus and car estimates for the models that consider different process strategies are significantly different from each other (except for the bus VTTS between the VL and RAM model). The train and metro median VTTS are similar across the models.

Stability of the VTTS estimates? Similar stability for the public transport VTTS. The car VTTS stability is better in the LPAA model and worse in the VL model.

Conclusion: The use of LPAA, RAM or VL as the sole process strategy being used by individuals has a significant influence over the VTTS estimates.

5.5.3.1.2. Probabilistic Decision Process versus Conditioning Random Parameter Heterogeneity

Figure 5-10 compares the two approaches used to include process heterogeneity: PDP and CRPH. As explained previously the PDP approach considers three classes, where each one represents a process strategy, and every parameter is estimated within a class as fixed. On the contrary, the CRPH approach considers interactions between the process strategies and the mean and standard deviation of the parameters normally defined under an LPAA heuristic. Therefore, it is expected that the levels of variation are higher in the CRPH models since they estimate every parameter as random and normally distributed. As can be seen, the median VTTS for public transport is higher for the PPD_BRExp model ($8.41 per person hour) and lower for the CRPHms_BRExp model ($5.20 per person hour). There are significant changes in the estimate when including process heterogeneity using the different models. It is important to mention that the median VTTS for public transport is always lower when using the CRPH approach that when using the PDP approach. Regarding the car median VTTS, the lowest value is under CRPHm ($4.75 per person hour), followed by the CRPHms ($8.58 per person hour). The higher median VTTS estimated is obtained when using the PDP_BRExp ($23.60 per person hour), followed by the PDP model ($18.15 per person hour).

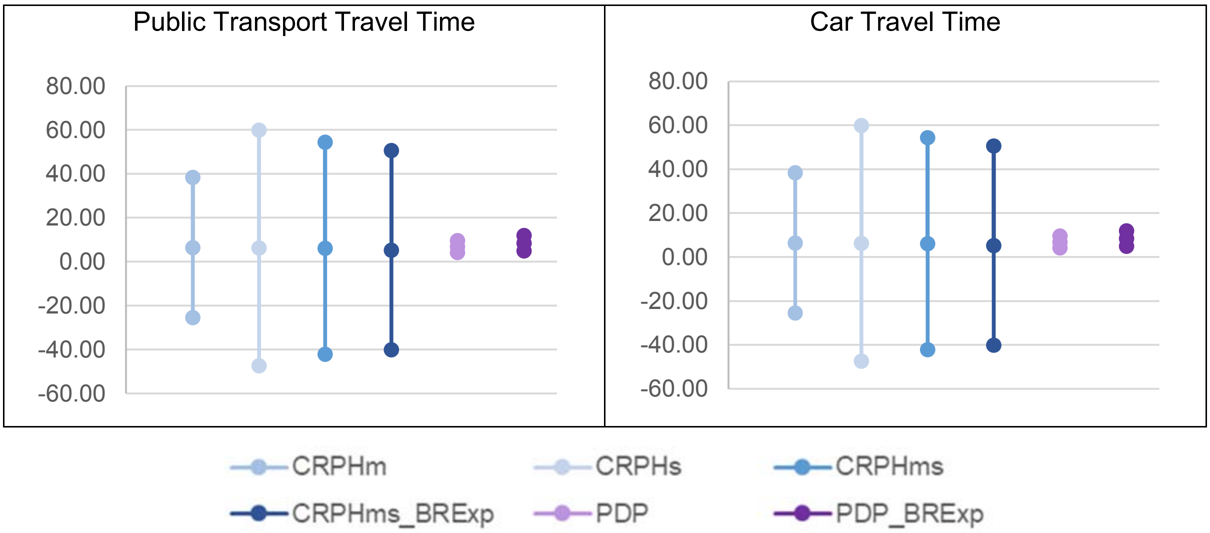


Figure 5-10: VTTS median and 95% confidence intervals for the models with process heterogeneity

Table 5-19 presents the results when comparing the median VTTS using the median and standard errors in t-test (Section 3.8.1) for each mode using the PDP and CRPH model (without and with behavioural refinements and experience). The results show that all the estimates are statistically different (absolute value larger than 1.96 with a 95% confidence level) when considering the CRPH method relative to the PDP method, and always lower. The only exception is the train median VTTS for the models without behavioural refinements and experience, as the results show that the estimate for the CRPHms model and the PDP model are not statistically different. In conclusion, the results show that the median VTTS for public transport and car are significantly different when considering the CRPH and PDP approach, and the median is always lower when using the CRPH approach.

Table 5-19: Comparison of attributes' median VTTS for models with process heterogeneity using t-test

Travel Time	CRPHms vs PDP	CRPHms_BRExp vs PDP_BRExp
Bus	***-2.36***	***-8.73***
Train	-0.17	***-10.71***
Metro	***-4.30***	***-11.05***
Car	***-65.07***	***-19.08***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 5-11 presents the stability of the VTTS for the CRPH models. The PDP is not included in these figures because they were not affected by the different draws since, given that the estimated parameters are not random. The graphs show that, for public transport, the VTTS stability is higher in the CRPHms_BRExp model, and the other ones have a relatively similar stability. In the case of VTTS for the car, the highest stability is in the CRPHms model followed by the CRPHm. The lowest stability for the VTTS using the car is for the CRPHms_BRExp model.

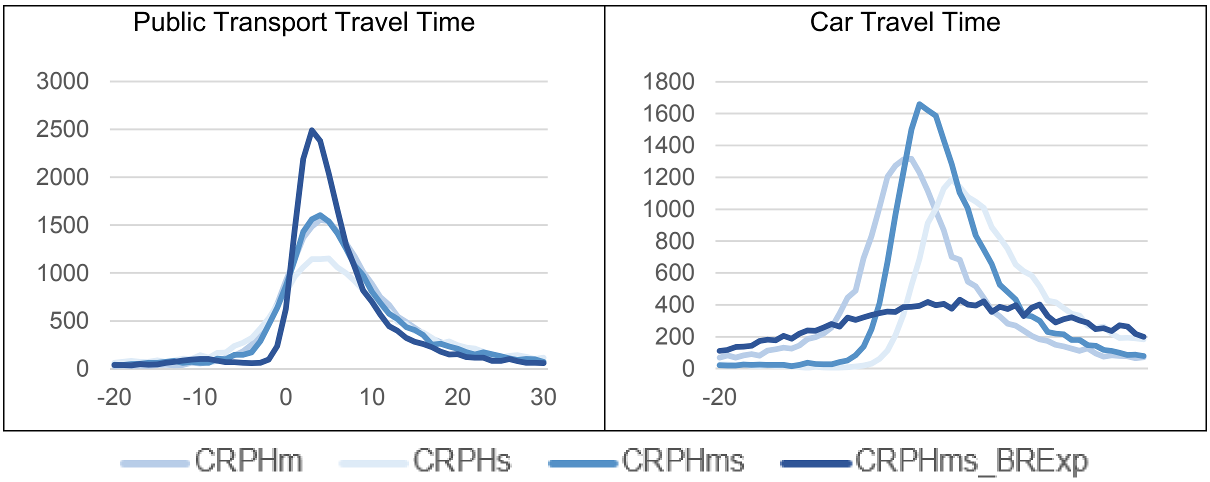


Figure 5-11: VTTS distribution for different draws using the models with process heterogeneity

Summary: Comparison of the two approaches to include process heterogeneity, PDP and CRPH

Public transport and car VTTS

- median (lower to higher): CRPH, PDP
- standard error (lower to higher): PDP, CRPH

Significant differences? All the median VTTS estimates are significantly different from each other when using the two approaches: PDP or CRPH.

Stability of the VTTS estimates? PDP is completely stable since every parameter is estimated as fixed.

Conclusion: The PDP approach produces significantly larger median VTTS, but lower standard errors as it considers all fixed parameters.

5.5.3.1.3. Process Heterogeneity versus Process Homogeneity

Figure 5-12 presents the median VTTS and 95% confidence intervals for the three preferred models with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp, and RAM_MML_Exp) and two models for each of the process heterogeneity methods (CRPHms and CRPHms_BRExp; PDP and PDP_BRExp). As was mentioned above, the standard error for the PDP models are lower because it does not estimate any parameter as random, contrary to all the other models. The lowest median VTTS for public transport is estimated using the CRPH_BRExp model ($5.20 per person hour), followed by the LPAA_MML_BRExp model ($5.52 per person hour). The highest median VTTS for public transport is estimated using the PDP_BRExp model ($8.41 per person hour), followed by the PDP model ($6.88 per person hour). For the car median VTTS, the differences are larger than for public transport, as can be seen in Figure 5-12. The lowest car median VTTS is estimated using the CRPHms model ($8.58 per person hour) followed by the RAM_MML_Exp model ($8.80 per person hour). The highest car median VTTS is estimated using the PDP_BRExp model ($23.60 per person hour) followed by the PDP model ($18.15 per person hour). The largest standard errors are found in the VL_MML and the CRPHms_BRExp models.

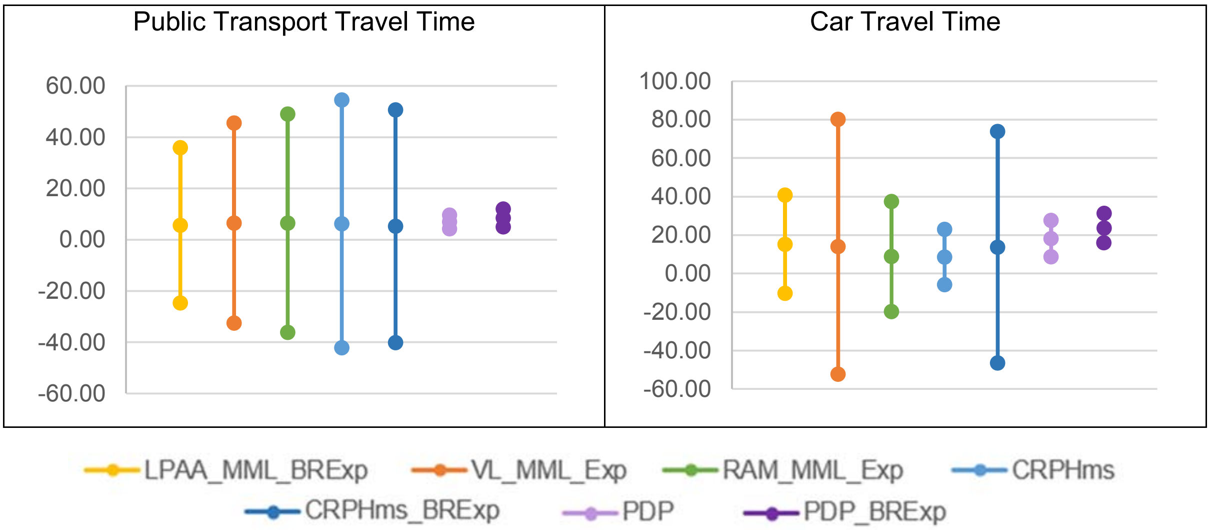


Figure 5-12: VTTS median and 95% confidence intervals for the models with process homogeneity versus process heterogeneity

Table 5-20 presents the comparison of the median VTTS using t-test (including the mean and standard error estimates) for the preferred models with process heterogeneity and homogeneity. The first three columns compare the model CRPHms_BRExp with the three preferred models of process homogeneity. The results show that every median VTTS is lower in the CRPH model than in the process homogeneity models, except for the car median VTTS for the RAM model. The majority of the median VTTS estimates using the CRPH final model are significantly different (absolute value larger than 1.96 with a 95% confidence level) from the ones estimated using the process homogeneity models. The exceptions are for the bus median VTTS for the LPAA model, metro median VTTS for all the process homogeneity models, and car median VTTS for the VL model, where there is not enough evidence to suggest they are statistically different to the CRPH preferred model. When comparing the PDP_BRExp model with the process homogeneity models, the median VTTS estimates are always larger for the PDP model. The results also show that the PDP median VTTS estimates are all statistically different from the process homogeneity median VTTS estimates.

Table 5-20: Comparison of attributes median VTTS for models with process heterogeneity and process homogeneity using t-test

Travel Time	CRPHms_BRExp vs			PDP_BRExp vs		
	LPAA_MML_BRExp	VL_MML_Exp	RAM_MML_Exp	LPAA_MML_BRExp	VL_MML_Exp	RAM_MML_Exp
Bus	↑ 0.07	↓ ***-3.27***	↓ ***-3.06***	↑ ***17.19***	↑ ***5.68***	↑ ***5.36***
Train	↓ ***-2.24***	↓ ***-3.12***	↓ ***-2.89***	↑ ***10.65***	↑ ***7.49***	↑ ***7.04***
Metro	↓ -0.56	↓ -1.85	↓ -1.47	↑ ***13.37***	↑ ***9.39***	↑ ***9.03***
Car	↓ ***-2.73***	↓ -0.27	↑ ***8.70***	↑ ***36.72***	↑ ***16.99***	↑ ***58.66***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

In conclusion, the results suggest that the CRPH median VTTS estimates are, in the majority, lower than the estimates derived from the process homogeneity models. Oppositely, the PDP median VTTS estimates are always higher than the process homogeneity models. This is a crucial finding as it suggests significant differences under alternative behavioural assumptions. This becomes a challenge in policy setting in selecting a preferred set of WTP estimates to use in practice.

Figure 5-13 summarises the stability findings of the VTTS for the different models. The VTTS for public transport has a relatively similar level of stability for the models, however the stability is superior in the CRPHms_BRExp model followed by the CRPHms model, and inferior in the VL model. In the case of the VTTS in the car, the level of stability is superior in the CRPHms model followed by the LPAA model. The worst level of stability is found in the CRPHms_BRExp and the VL_MML_Exp models. There is no clear pattern on the level of stability of the VTTS in public transport versus car regarding process heterogeneity. However, in both attributes the worst level of stability was found in the VL model.

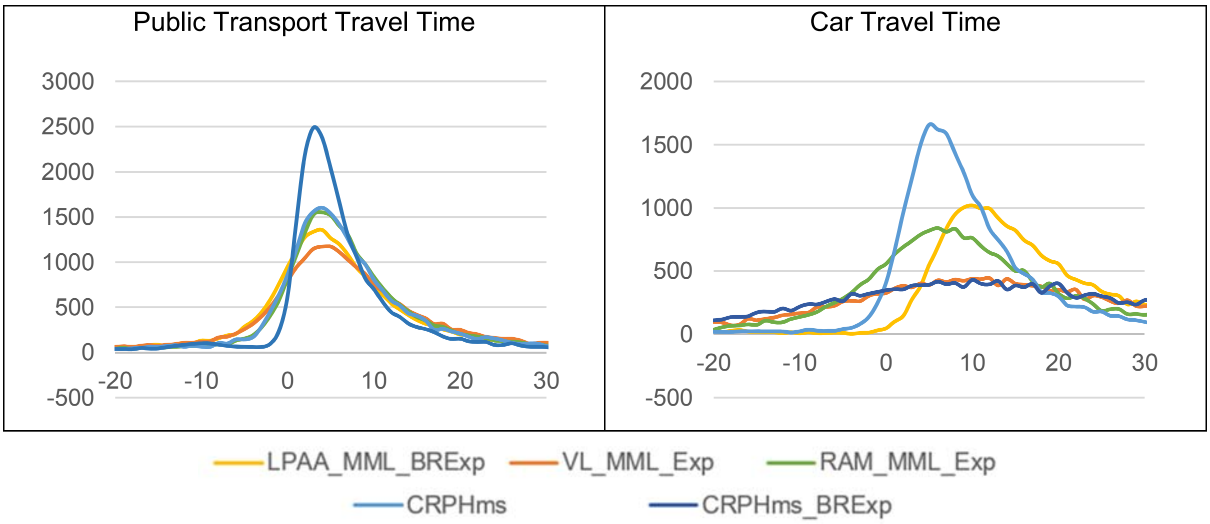


Figure 5-13: VTTS distribution for different draws using the models with process homogeneity versus process heterogeneity

Summary: Comparison of the preferred models with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp) with the preferred models with process heterogeneity (PDP_BRExp, CRPHms_BRExp)

Public transport VTTS

- median (lower to higher): CRPH, process homogeneity models, PDP
- standard error (lower to higher): PDP, process homogeneity models, CRPH

Car VTTS

- median (lower to higher): RAM, CRPH, VL, LPAA, PDP
- standard error (lower to higher): PDP, RAM, LPAA, CRPH, VL

Significant differences? The majority of the median VTTS are statistically different when considering process heterogeneity instead of process homogeneity.

Stability of the VTTS estimates? For the public transport VTTS the preferred CRPH model has a superior stability and for the car VTTS the preferred LPAA.

Conclusion: The process heterogeneity method directly influences the results. While the PDP approach estimates the largest median and lower standard errors for the VTTS relative to the process homogeneity models, the CRPH usually estimates the lowest median and largest standard error.

5.5.3.1.4. Behavioural Refinements and Experience

Figure 5-14 shows the median with their 95% confidence intervals for the VTTS for the preferred models with and without behavioural refinements and/or experience. When including behavioural refinements and/or experience for the median VTTS in public transport, there is a decrease in the value in the LPAA MML model (from $6.20 to $5.52 per person hour), an increase in the VL MML model (from $5.59 to $6.43 per person hour), an increase in the RAM MML model (from $5.96 to $6.36 per person hour), a decrease in the CRPHms MML model (from $6.13 to $5.20 per person hour), and an increase in the PDP model (from $6.88 to $8.41 per person hour). This suggests that the median estimate decreases when considering behavioural refinements and experience for the LPAA MML model and the CRPH model considering process heterogeneity, and decreases for the RAM MML, VL MML and PDP models. The standard error VTTS increases when adding behavioural refinements and/or experience in the VL MML, RAM MML and PDP models, and decreases in the LPAA MML and CRPHms models.

For the median car VTTS, there is an increase in the LPAA MML model (from $14.35 to $15.26 per person hour), an increase in the VL MML model (from $13.66 to $13.94 per person hour), a decrease in the RAM MML model (from $13.76 to $8.80 per person hour), an increase in the CRPHms model (from $8.58 to $13.74 per person hour) and an increase in the PDP model (from $18.15 to $23.60 per person hour). We see that median VTTS estimate increases when considering behavioural refinements and experience for all the models except the RAM MML model. The standard error VTTS increases when adding behavioural refinements and/or experience in the VL MML and CRPHms models, and decreases in the LPAA MML, RAM MML and PDP models. This increase is particularly high for the CRPHms model (from $7.35 to $30.70 per person hour).

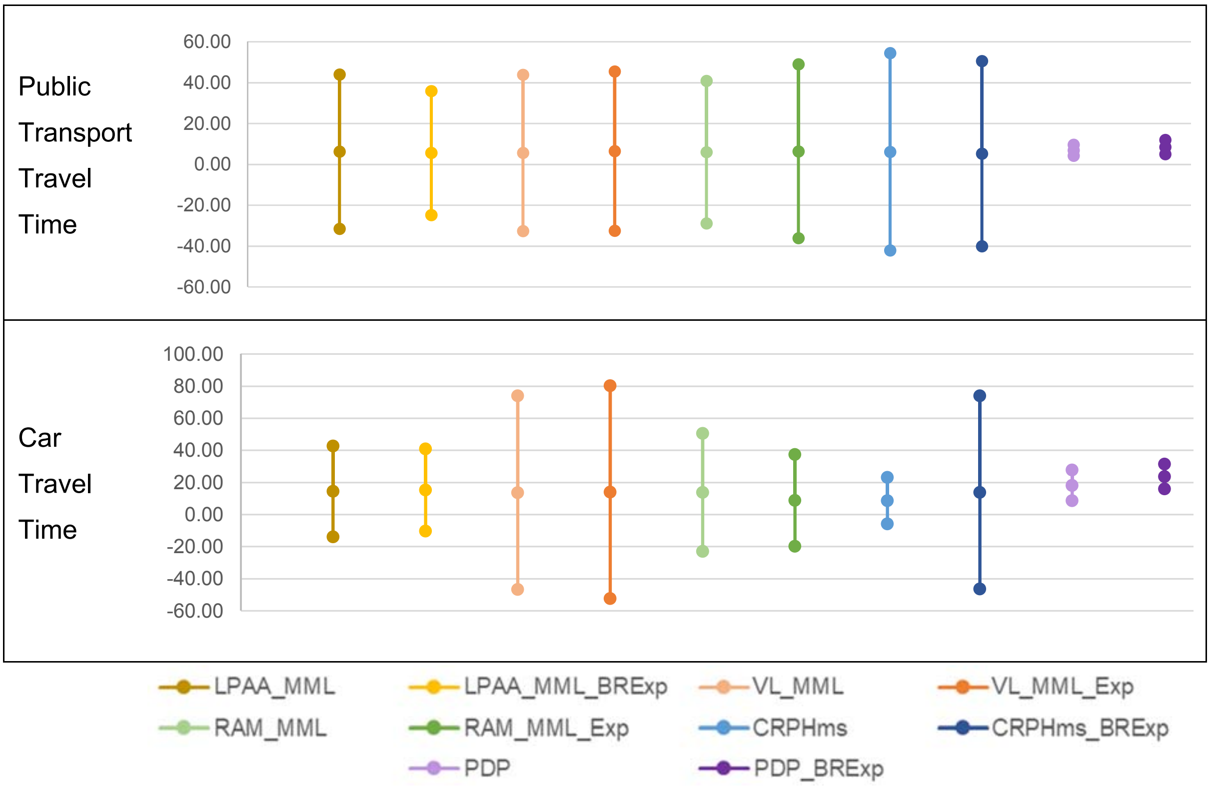


Figure 5-14: VTTS median and 95% confidence intervals for the models with and without behavioural refinements and/or experience

In summary, Table 5-21 presents the comparison of the attributes' median VTTS for models with and without behavioural refinements and/or experience using t-test. As can be seen, there is no clear pattern of increase or decrease when including behavioural refinements and/or experience. In the case of the LPAA model, the inclusion of behavioural refinements and experience reduces the median VTTS for all the modes, except for the car which has an increase. Moreover, this inclusion does not have a statistically significant influence on the median VTTS for train and metro (there is not enough evidence to suggest the estimates are statistically different). In the case of the VL heuristic, the inclusion of experience increases the median VTTS for all the modes. However, the difference is not statistically significant for the car or the bus considering a confidence level of 95% (the difference of the median VTTS for the bus is significant with a 90% confidence level). For the RAM process strategy, the only statistically significant difference is found for the car, and there is a decrease of the value when adding experience. For the process heterogeneity CRPH model, there is a statistically significant decrease in the median VTTS for the train and a statistically significant increase of the median VTTS for the car at a confidence level of 95%. There is also a statistically significant decrease of the median VTTS for the bus when considering behavioural refinements and experience at a 90% confidence level. When including process heterogeneity using the PDP approach, there is a statistically significant increase in every median VTTS when including behavioural refinements and experience.

Table 5-21: Comparison of attributes' median VTTS for models with and without behavioural refinements and/or experience using t-test

Travel Time	LPAA_MML_BRExp vs LPAA_MML	VL_MML_Exp vs VL_MML	RAM_MML_Exp vs RAM_MML	CRPHms_BRExp vs CRPHms	PDP_BRExp vs PDP
Bus	***-4.04***	1.66	0.83	-1.89	***38.94***
Train	-0.52	***2.33***	1.13	***-3.86***	***50.95***
Metro	-0.66	***2.95***	1.30	0.31	***65.56***
Car	***2.79***	0.37	***-12.51***	***9.79***	***52.29***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 5-15 shows graphically the stability of the VTTS estimate for the 25,000 different draws. There does not appear to be any pattern associated with the inclusion of behavioural refinements and/or experience. The stability seems to be similar for the models without and with behavioural refinements, except for CRPH. The CRPH preferred model with behavioural refinements and experience has superior stability for the public transport VTTS than the CRPH model that excludes these effects, but an inferior stability for the car VTTS.

These section shows that the inclusion of experience and behavioural refinements has a significant influence in the median, standard error and stability of the VTTS estimates. These differences are especially important in the PDP approach that considers all parameters as fixed. Therefore, the results suggest that when considering random parameters and alternative process strategies the inclusion of behavioural refinements and experience is not as important as when not considering them.

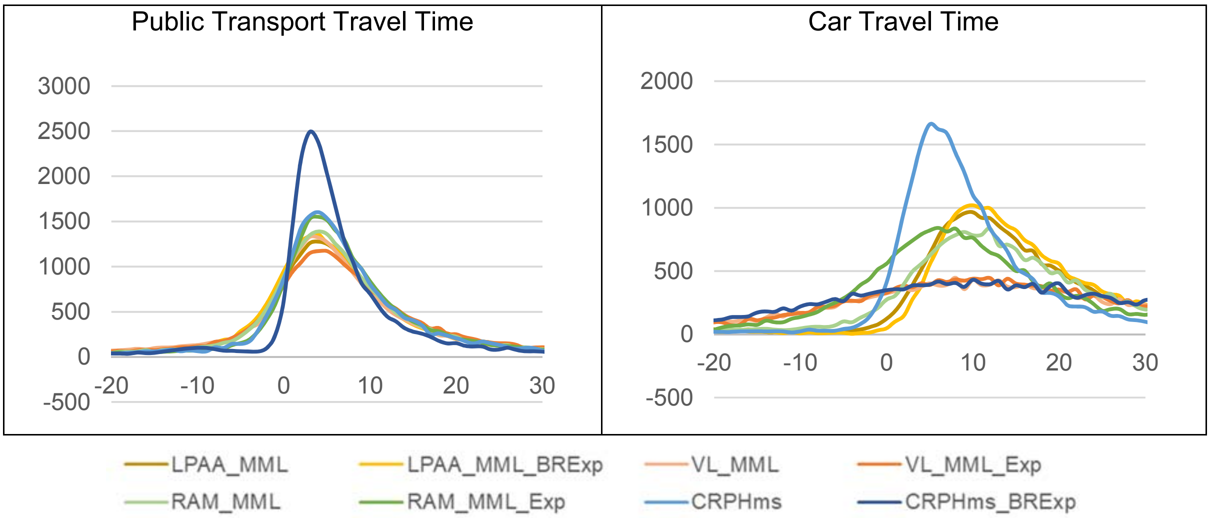


Figure 5-15: VTTS distribution for different draws using the models with and without behavioural refinements and/or experience

Summary: Comparison of the models with and without experience and behavioural refinements

Public transport VTTS

- median when adding experience and behavioural refinements
  - increases for VL, RAM and PDP
  - decreases for LPAA and CRPH
- standard error when adding experience and behavioural refinements
  - increases for VL, RAM and PDP
  - decreases for LPAA and CRPH

Car VTTS

- median when adding experience and behavioural refinements
  - increases for LPAA, VL, CRPH and PDP
  - decreases for RAM
- standard error when adding experience and behavioural refinements
  - increases for VL and CRPH
  - decreases for LPAA, RAM and PDP

Significant differences? Differences in the median VTTS are significant in the LPAA model for the bus and car only when including behavioural refinements/experience; in the VL model for the train and metro only; in the RAM model for the car only; in the CRPH model for the train and car; and in the PDP model for all.

Stability of the VTTS estimates? Similar stability with and without experience and behavioural refinements.

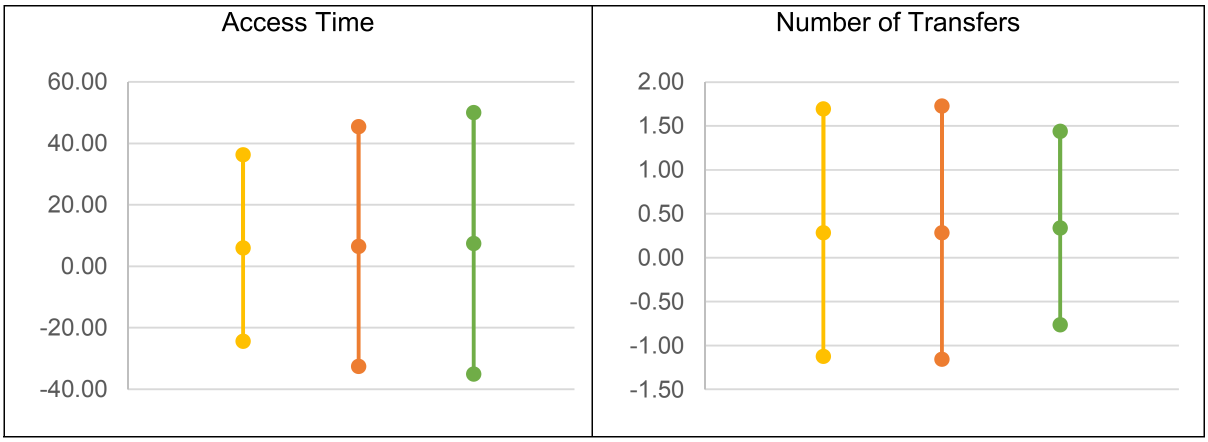
Conclusion: The inclusion of experience and behavioural refinements influences the estimates, even though the difference is not significant in every mode (except in the PDP model where all differences are significant).

5.5.3.2. Other Attributes: Egress time, access times, headway, transfers and crowding

5.5.3.2.1. Process Homogeneity using Different Heuristics

Figure 5-16 shows the WTP median and 95% confidence intervals of the other attributes for three models that represent process homogeneity: LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp. The LPAA heuristic obtains the lowest median WTP for the access time, egress time in public transport, headway, % of finding a seat available and standing density. The VL heuristic produces the lowest median WTP only for the number of transfers, and the RAM heuristic only for the egress time in the car. On the other hand, the LPAA heuristic does not obtain the highest value of the median WTP for any of the attributes. The VL heuristic is associated with the highest median WTP for the egress times (in public transport and car), and the RAM heuristic for all the others (access time, number of transfers, headway, % of finding a seat available and standing density).

Regarding the standard errors, the LPAA model produces the lowest standard errors for the WTP associated with access time, headway and the standing density. The RAM model has the lowest standard error of the WTP for the number of transfers, the egress times (public transport and car) and increase the % for finding a seat available. Contrarily, the RAM heuristic estimates the highest standard error of the WTP for access time, and the VL heuristic estimates the highest standard errors for all the other attributes (number of transfers, egress times (public transport and car), headway, % of seating and standing density).



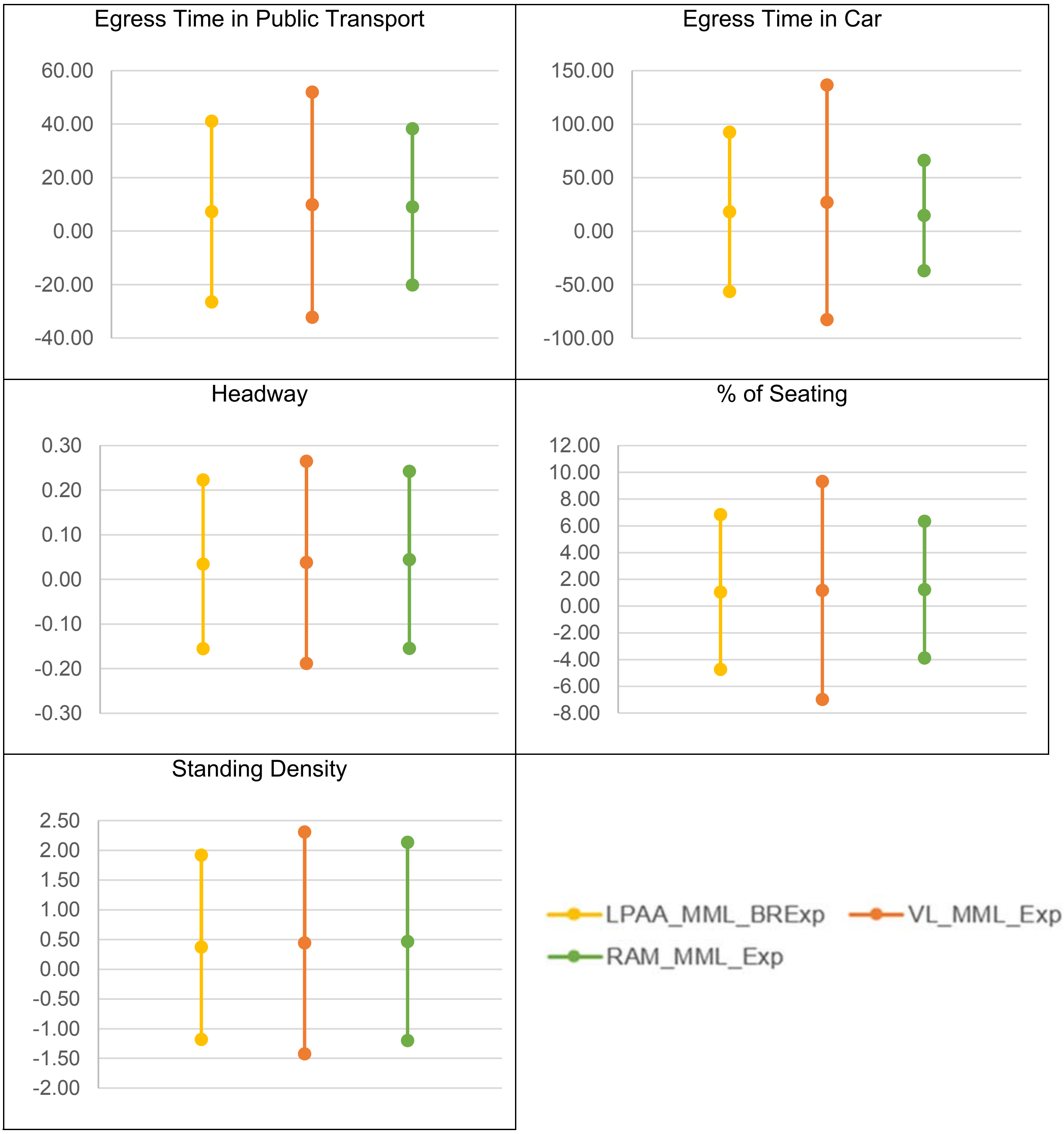


Figure 5-16: Attributes WTP median and 95% confidence intervals for the models with process homogeneity

For each mode of transport, the median WTP of the different models was compared using t-test to see if they were statistically different from each other, and the results are presented in Table 5-22. The indicators that are not statistically significant (shown with a **) represent the models that did not estimate significantly different median WTP. As can be seen, the median WTP estimates for the access time, headway and % of seating are not statistically different for the train and metro in the VL and LPAA models, at a 95% confidence level. The number of transfers is not statistically significant for the number of transfers with an 80% confidence level between the VL and LPAA model. The RAM and VL model also have several median WTP as statistically similar: the access time and egress time only for the bus, the % of finding a seat available for all the modes, and the standing density for the bus and train. The RAM and LPAA models have only two median WTP statistically similar attributes: the % of finding a seat available for the train and egress time for the car. The majority of the median WTP estimated using the VL and RAM heuristic are higher than those for the LPAA heuristic (represented with green arrows), and the ones estimated using the VL heuristic are higher than in the RAM model.

Table 5-22: Comparison of attributes' median WTP for models with process homogeneity using t-test

Attribute	Mode	VL_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs VL_MML_BRExp
Access Time	Bus	↑ ***3.68***	↑ ***5.66***	↑ 1.87
	Train	↓ -0.22	↑ ***2.63***	↑ ***2.65***
	Metro	↓ -0.27	↑ ***3.36***	↑ ***3.38***
Egress Time	Bus	↑ ***8.44***	↑ ***8.24***	↓ -1.83
	Train	↑ ***5.21***	↑ ***3.44***	↓ ***-2.53***
	Metro	↑ ***6.59***	↑ ***4.49***	↓ ***-3.06***
	Car	↑ ***7.87***	↓ ***-4.34***	↓ ***-11.85***
Headway	Bus	↑ ***4.06***	↑ ***7.02***	↑ ***2.17***
	Train	↑ 0.43	↑ ***3.67***	↑ ***3.05***
	Metro	↑ 0.54	↑ ***4.57***	↑ ***3.78***
Seat	Bus	↑ ***3.59***	↑ ***5.77***	↑ 0.66
	Train	↑ 0.35	↑ 1.57	↑ 0.95
	Metro	↑ 0.44	↑ ***2.04***	↑ 1.26
Stand	Bus	↑ ***6.69***	↑ ***8.52***	↑ 1.09
	Train	↑ ***2.27***	↑ ***4.04***	↑ 1.54
	Metro	↑ ***2.87***	↑ ***5.14***	↑ ***1.99***
Transfers	Bus	↑ ***3.03***	↑ ***7.13***	↑ ***3.17***
	Train	↓ -1.89	↑ ***2.20***	↑ ***4.50***
	Metro	↓ ***-2.39***	↑ ***2.82***	↑ ***5.72***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 5-17 shows the distribution of the WTP for the all attributes (except for the travel time that was analysed previously) using 25,000 different draws for the models that represent process homogeneity, which is required to analyse the stability of the WTP. The stability is superior in the LPAA model for the access time, headway and in the standing density WTP estimates. For all the other attributes, the RAM model has a superior stability.

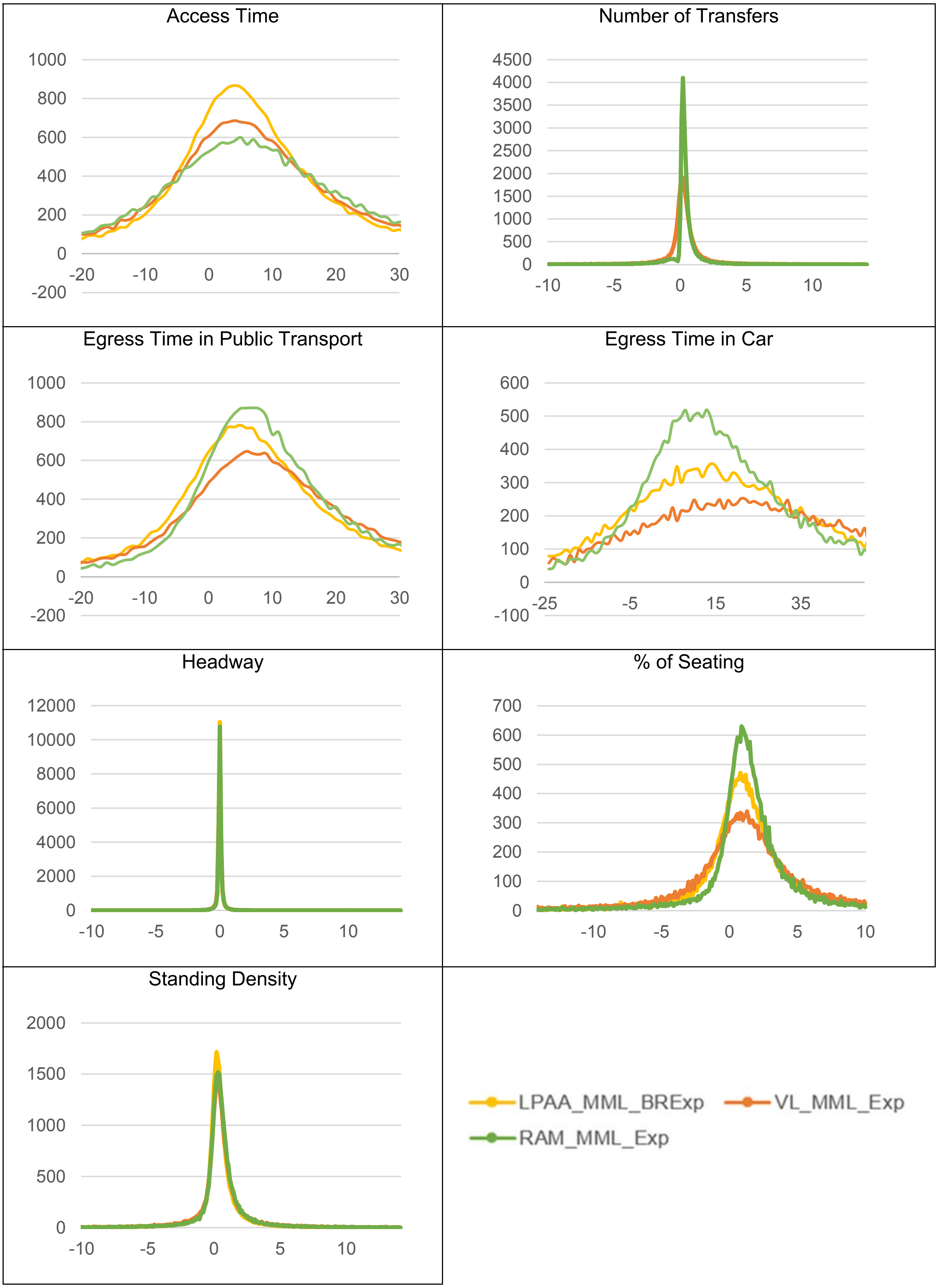


Figure 5-17: Attributes WTP distribution for different draws using the models with process homogeneity

In summary, the results of this section show that the WTP estimates when using the various process strategies are significantly different from each other for the majority of the attributes. Therefore, when considering process homogeneity, the selection of an appropriate process strategy is crucial since it has a significant influence in the outcome.

Summary: Comparison of the preferred process homogeneity models LPAA_MML_BRExp; RAM_MML_Exp and VL_MML_Exp

Access time WTP
- median (lower to higher): LPAA, VL, RAM
- standard error (lower to higher): LPAA, VL, RAM

Egress time public transport WTP
- median (lower to higher): LPAA, RAM, VL
- standard error (lower to higher): RAM, LPAA, VL

Egress time car WTP
- median (lower to higher): RAM, LPAA, VL
- standard error (lower to higher): RAM, LPAA, VL

Headway WTP
- median (lower to higher): LPAA, VL, RAM
- standard error (lower to higher): LPAA, RAM, VL

% of seats available and standing density WTP
- median (lower to higher): LPAA, VL, RAM
- standard error (lower to higher): RAM, LPAA, VL

Number of Transfers WTP
- median (lower to higher): VL, LPAA, RAM
- standard error (lower to higher): RAM, LPAA, VL

Significant differences? The majority of the differences are significant.

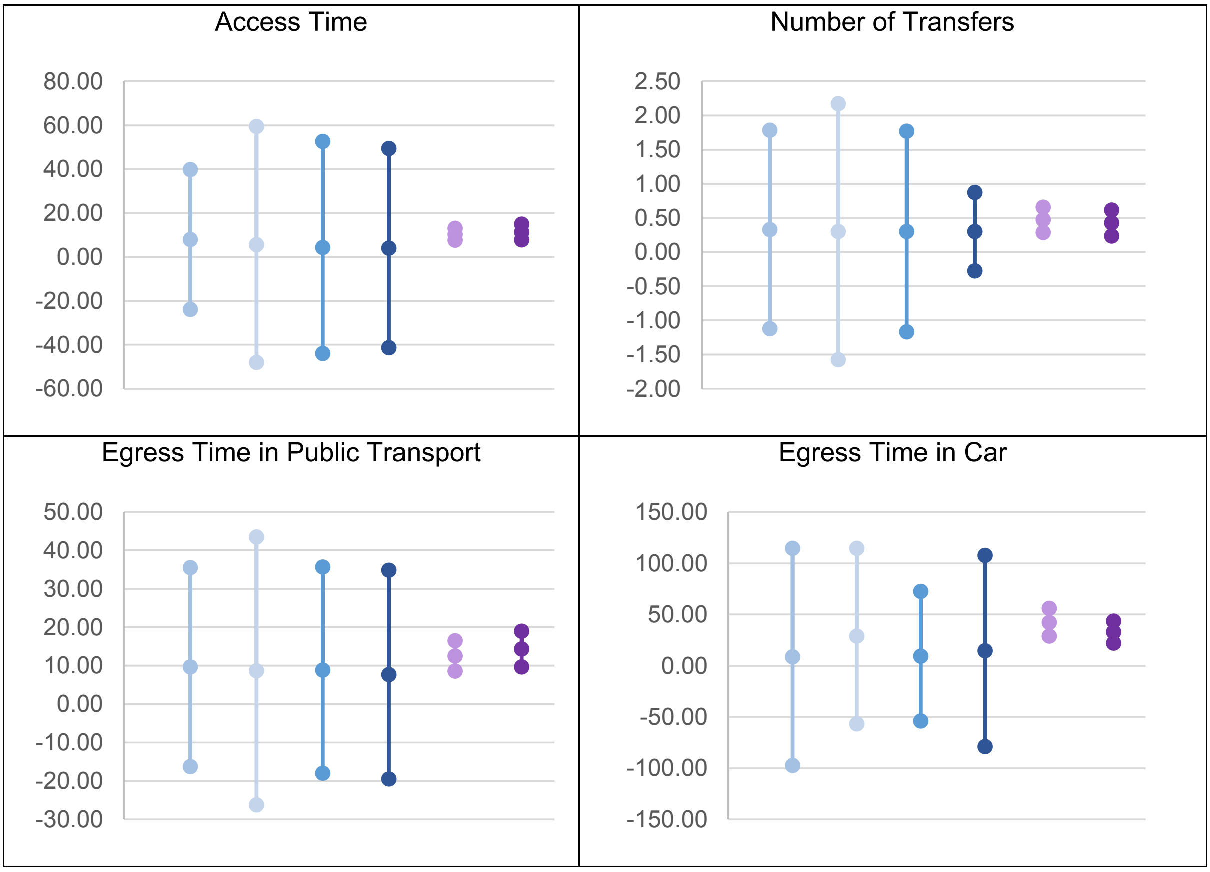
Stability of the WTP estimates? The stability is superior in the LPAA model for the access time, headway and in the standing density WTP estimates. For all the other attributes, the RAM model has a superior stability.

Conclusion: The use of LPAA, RAM or VL as the sole process strategy being used by individuals has a significant influence over the WTP estimates.

5.5.3.2.2. Probabilistic Decision Process versus Conditioning Random Process Heterogeneity

Figure 5-18 compares the PDP and CRPH approaches, both used to include process heterogeneity. For the majority of the attributes, the PDP or PDP_BRExp models estimate the highest median WTP values, except for the % of finding a seat available, which is highest for the CRPHms model followed by the PDP_BRExp model. As expected, the standard error is always larger in the CRPH model because it estimates every parameter as random, contrary to the PDP models.

Comparing the CRPH models, it can be seen that when including interactions with process strategies only in the mean estimate, the median WTP estimates are higher for all the attributes except for the egress time in the car and headway. In the case of the egress time for the car, the median WTP estimated is higher in the CRPHs model and it is significantly higher (more than double). For the headway attribute, the median WTP is higher when using the CRPHms_BRExp model followed by the CRPHm, but the difference between them is not statistically significant. The median WTP is lower in the model CRPHms_BRExp for all the attributes except for the headway, % of seats available and number of transfers. The % of seats available and number of transfers are lower in the CRPHs model. Analysing only the PDP models, the PDP_BRExp estimates higher median WTP for all the attributes except for the egress time in the car and number of transfers.



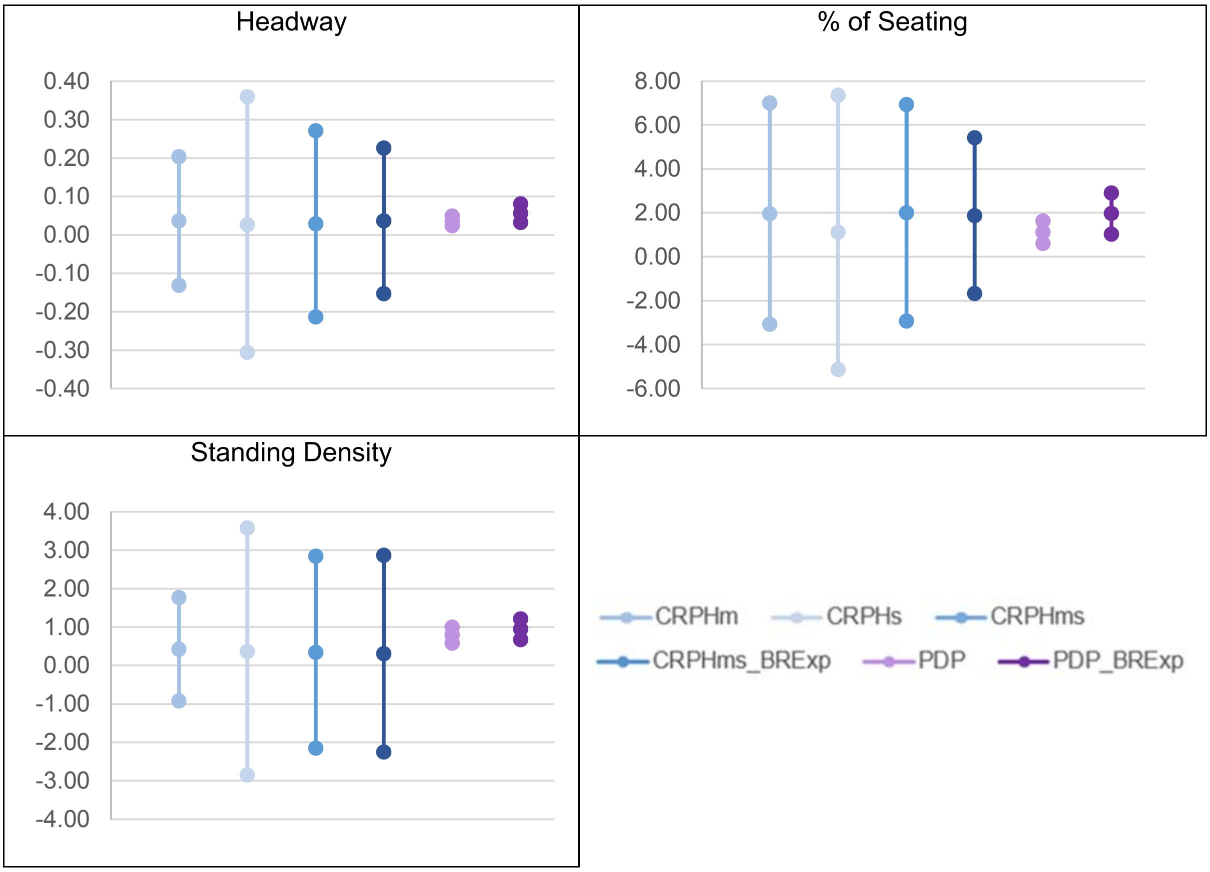


Figure 5-18: Attributes' WTP median and 95% confidence intervals for the models with process heterogeneity

Table 5-23 presents the results of the comparison between the median WTP estimates using t-test for each mode using the PDP and CRPH models (with and without behavioural refinements and experience). The majority of the median WTP estimates are higher and statistically different in the PDP approach relative to the CRPHms approach for the models without behavioural refinements, except for the % of seats available. This latter is higher and statistically different for the CRPHms model relative to the PDP model. In the case of the models with behavioural refinements, the same is true in the majority of the median WTP estimates except for the % of seats available. The median WTP for the % of seats available in the train is higher in the CRPHms_BRExp model than in the PDP_BRExp.

Table 5-23: Comparison of attributes' median WTP for models with process heterogeneity using t-test

Attribute	Mode	CRPHms vs PDP	CRPHms_BRExp vs PDP_BRExp
Access Time	Bus	***-12.68***	***-15.94***
	Train	***-17.19***	***-24.45***
	Metro	***-26.53***	***-33.33***
Egress Time	Bus	***-16.54***	***-27.49***
	Train	***-14.95***	***-32.79***
	Metro	***-28.80***	***-48.36***
	Car	***-60.07***	***-23.08***
Headway	Bus	***-6.25***	***-15.02***
	Train	***-3.66***	***-16.38***
	Metro	***-4.21***	***-15.08***
Seat	Bus	***15.80***	***-3.97***
	Train	***33.80***	0.83
	Metro	***30.62***	***-8.90***
Stand	Bus	***-20.40***	***-28.15***
	Train	***-25.71***	***-37.27***
	Metro	***-32.90***	***-44.54***
Transfers	Bus	***-13.68***	***-22.13***
	Train	***-13.65***	***-28.11***
	Metro	***-24.24***	***-45.82***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 5-19 summarises the stability of the WTP estimates for the CRPH models. The PDP is not included in these figures because, as explained above, they were not affected by the different draws, as they estimated all parameters as fixed. The graphs show relatively similar stability of WTP for all the models for access time, the number of transfers, egress time in public transport, and headway. For the car egress time and standing density WTP, the stability was superior for the CRPHms model. For the % of seats available WTP, the CRPHms_BRExp model had a superior stability.

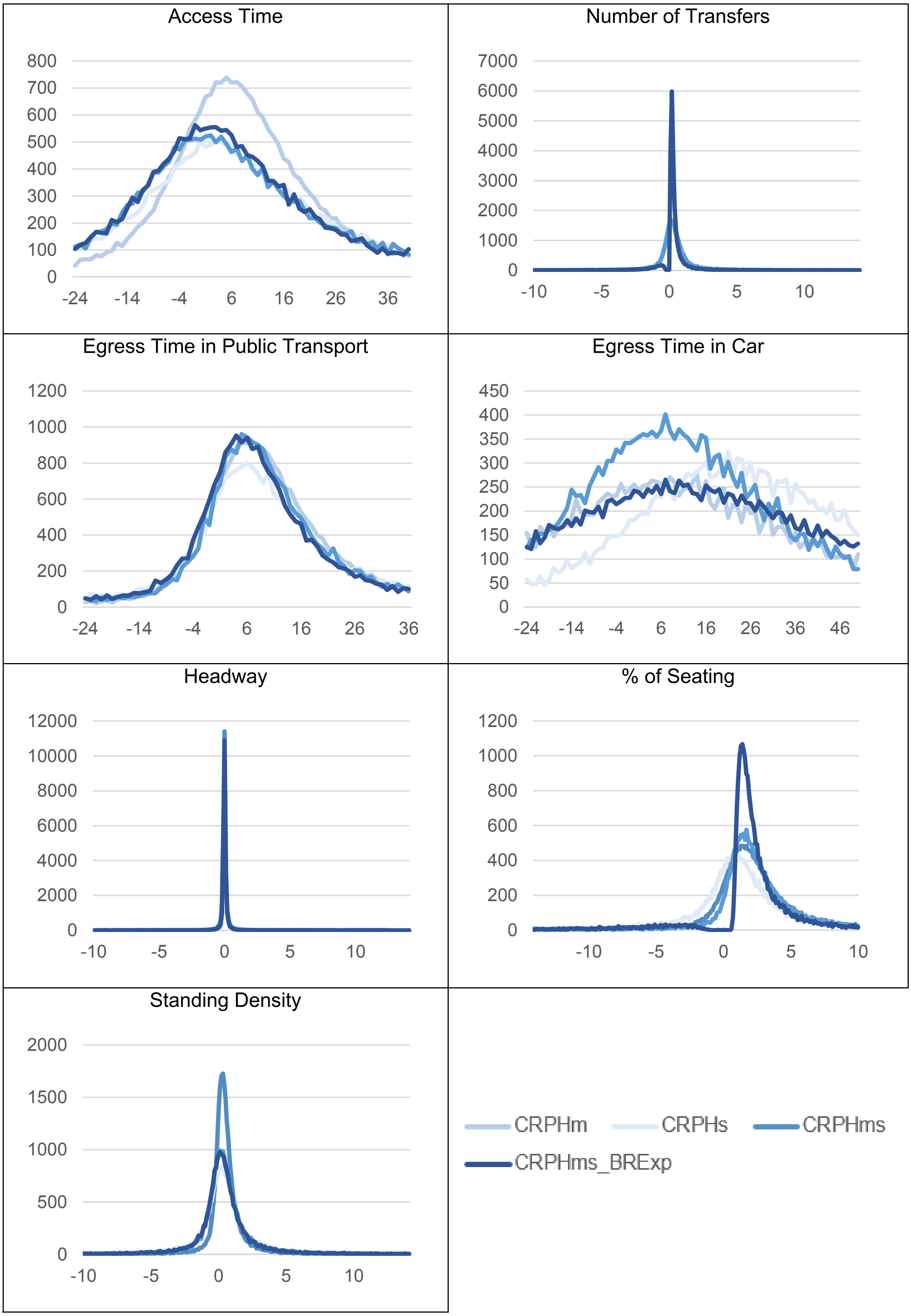


Figure 5-19: Attributes WTP distribution for different draws using the models with process heterogeneity

Summary: Comparison of the preferred models of two approaches to include process heterogeneity, PDP_BRExp and CRPHms_BRExp

All attributes' WTP

- median (lower to higher): CRPH, PDP
- standard error (lower to higher): PDP, CRPH

Significant differences? All the WTP estimates are significantly different from each other when using the two approaches: PDP or CRPH.

Stability of the WTP estimates? PDP is completely stable since every parameter is estimated as fixed.

Conclusion: The PDP approach produces significantly larger median WTP for all the attributes, but lower standard errors as it considers all fixed parameters.

5.5.3.2.3. Process Heterogeneity versus Process Homogeneity

Figure 5-20 presents the median WTP estimates and 95% confidence intervals for the three preferred models associated with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp, and RAM_MML_Exp) and two models for each of the process heterogeneity methods (CRPHms and CRPHms_BRExp; PDP and PDP_BRExp). Table 5-24 shows the comparison of the median estimates using t-test for the preferred process heterogeneity models versus the preferred process homogeneity models. As explained before, the green arrows represent an increase of median WTP for the process heterogeneity models, and the opposite is represented by the red arrows. The PDP_BRExp estimates a higher median WTP than the process homogeneity models for all the attributes, with all increases statistically significant. On the other hand, the CRPHms_BRExp median estimates are, in their majority, lower than the RAM model except for the % of seats available. They are also lower than in the VL model, except for the median WTP associated with train headway (although this difference is not statistically significant), the % of seats available, and the number of transfers (this difference is only significant for the train). Relative to the preferred LPAA model, half of the CRPHms_BRExp median WTP are higher and the other half are lower.

For the standard errors, they are lower for the CRPH_BRExp model compared to the process homogeneity models for the WTP associated with number of transfer, public transport egress time, and the % of seats available. It is greater than all the process homogeneity models for the access time and standing density. For the car egress time and public transport headway, the WTP standard error is in between the process homogeneity models ranges.

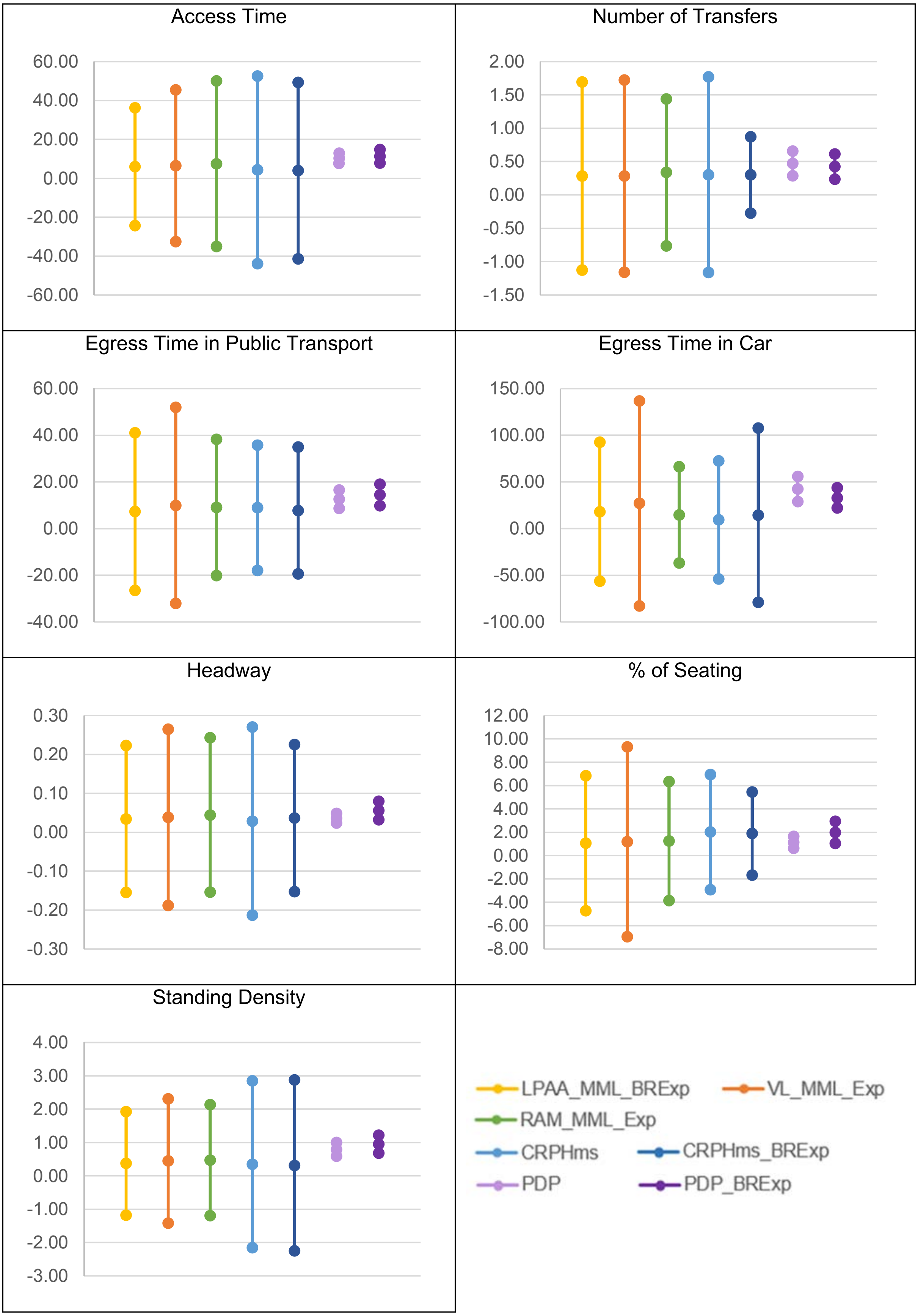


Figure 5-20: WTP median and 95% confidence intervals for the models with process homogeneity versus process heterogeneity

Table 5-24: Comparison of attributes' median WTP for models with process heterogeneity and process homogeneity using t-test

Attribute	Mode	CRPHms_BRExp vs			PDP_BRExp vs		
		LPAA_MML_BRExp	VL_MML_Exp	RAM_MML_Exp	LPAA_MML_BRExp	VL_MML_Exp	RAM_MML_Exp
Access Time	Bus	-1.69	***-4.21***	***-5.74***	***27.40***	***13.16***	***9.56***
	Train	***-6.62***	***-6.02***	***-8.27***	***21.78***	***18.79***	***13.66***
	Metro	***-9.53***	***-8.66***	***-11.53***	***27.58***	***23.78***	***17.20***
Egress Time	Bus	***4.27***	***-5.19***	***-4.17***	***32.84***	***11.67***	***19.92***
	Train	0.52	***-5.28***	***-3.39***	***25.87***	***15.84***	***27.09***
	Metro	-1.82	***-9.08***	***-7.46***	***32.61***	***19.94***	***33.76***
	Car	***-3.59***	***-10.21***	-0.33	***23.13***	***6.32***	***40.49***
Headway	Bus	***3.90***	-1.05	***-3.78***	***19.63***	***8.59***	***6.51***
	Train	0.73	0.21	***-3.42***	***13.30***	***11.54***	***8.53***
	Metro	-1.60	***-2.01***	***-5.56***	***16.78***	***14.56***	***10.92***
Seat	Bus	***17.56***	***7.90***	***10.25***	***24.52***	***10.34***	***15.08***
	Train	***17.56***	***13.96***	***18.37***	***19.60***	***14.83***	***21.64***
	Metro	***17.58***	***13.79***	***17.80***	***24.79***	***18.76***	***27.22***
Stand	Bus	0.05	***-4.81***	***-5.91***	***54.33***	***29.32***	***31.16***
	Train	***-3.93***	***-5.64***	***-7.16***	***47.28***	***39.93***	***42.31***
	Metro	***-6.09***	***-8.11***	***-9.94***	***59.77***	***50.48***	***53.29***
Transfers	Bus	***5.22***	0.72	***-3.73***	***17.24***	***9.98***	***7.83***
	Train	0.61	***3.22***	***-2.53***	***11.82***	***15.35***	***12.68***
	Metro	***-2.28***	0.86	***-7.33***	***14.99***	***19.45***	***16.00***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 5-21 summarises the stability of the attributes WTP for the different models. The results show that the process homogeneity models have a superior stability for the access time, car egress time and public transport standing density. For the number of transfers, egress time in public transport, and % of seats available, the CRPH model has a superior stability. There is no clear pattern regarding the stability of the WTP estimate in the process homogeneity or process heterogeneity models for all the attributes.

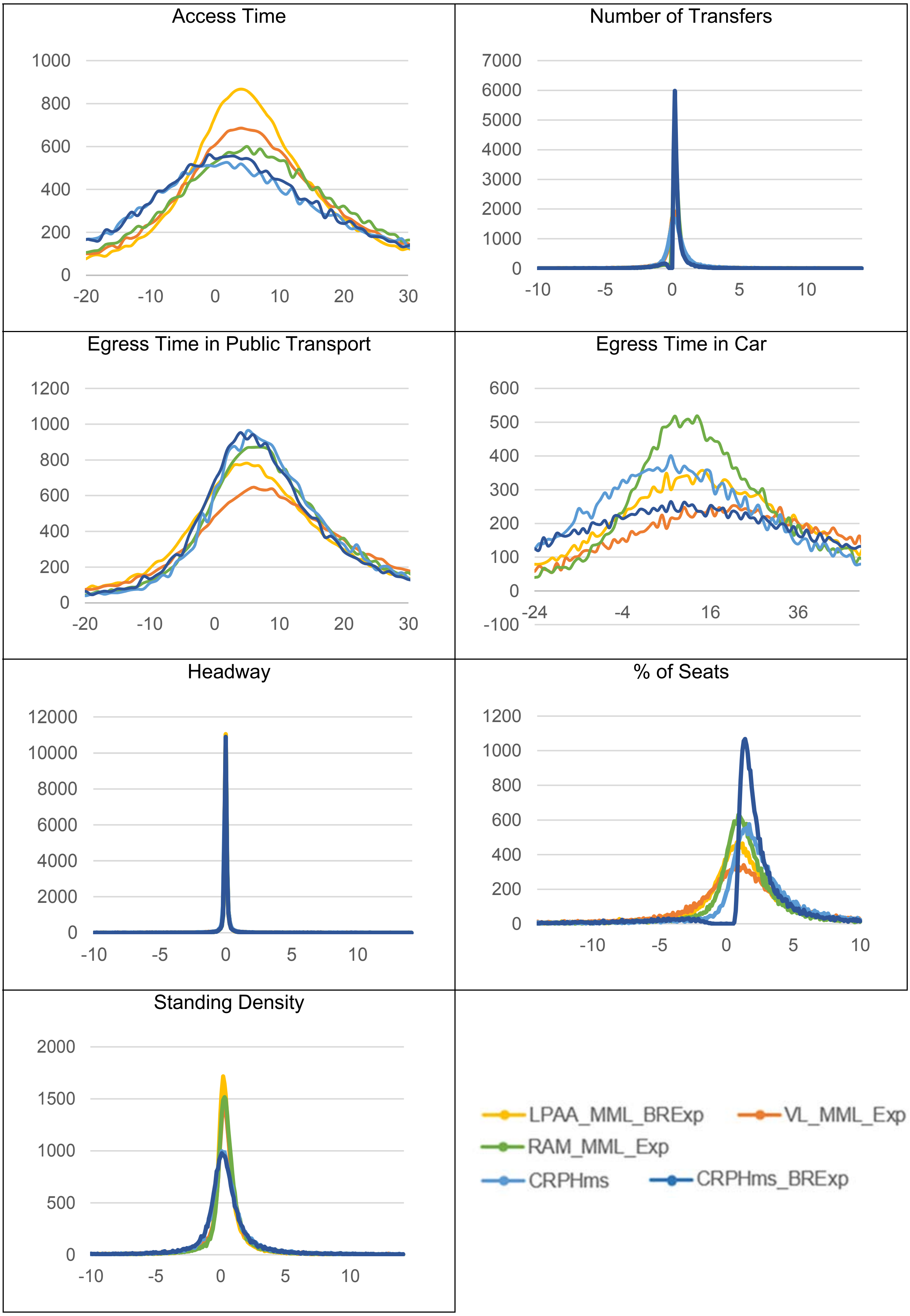


Figure 5-21: Attributes' WTP distribution for different draws using the models with process homogeneity versus process heterogeneity

Summary: Comparison of the preferred models with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp) with the preferred models with process heterogeneity (PDP_BRExp, CRPHms_BRExp)

Access time WTP
- median (lower to higher): CRPH, process homogeneity, PDP
- standard error (lower to higher): PDP, process homogeneity, CRPH

Egress time public transport WTP
- median (lower to higher): LPAA, CRPH, RAM, VL, PDP
- standard error (lower to higher): PDP, CRPH, process homogeneity

Egress time car WTP
- median (lower to higher): CRPH, process homogeneity, PDP
- standard error (lower to higher): PDP, RAM, LPAA, CRPH, VL

Headway WTP
- median (lower to higher): LPAA, CRPH, VL, RAM, PDP
- standard error (lower to higher): PDP, LPAA, CRPH, RAM, VL

% of seats available WTP
- median (lower to higher): process homogeneity, CRPH, PDP
- standard error (lower to higher): PDP, CRPH, process homogeneity

Standing density WTP
- median (lower to higher): CRPH, process homogeneity, PDP
- standard error (lower to higher): PDP, process homogeneity, CRPH

Number of Transfers WTP
- median (lower to higher): VL, LPAA, CRPH, RAM, PDP
- standard error (lower to higher): PDP, CRPH, process homogeneity

Significant differences? The majority of the median WTP are statistically different when considering process heterogeneity than when considering process homogeneity.

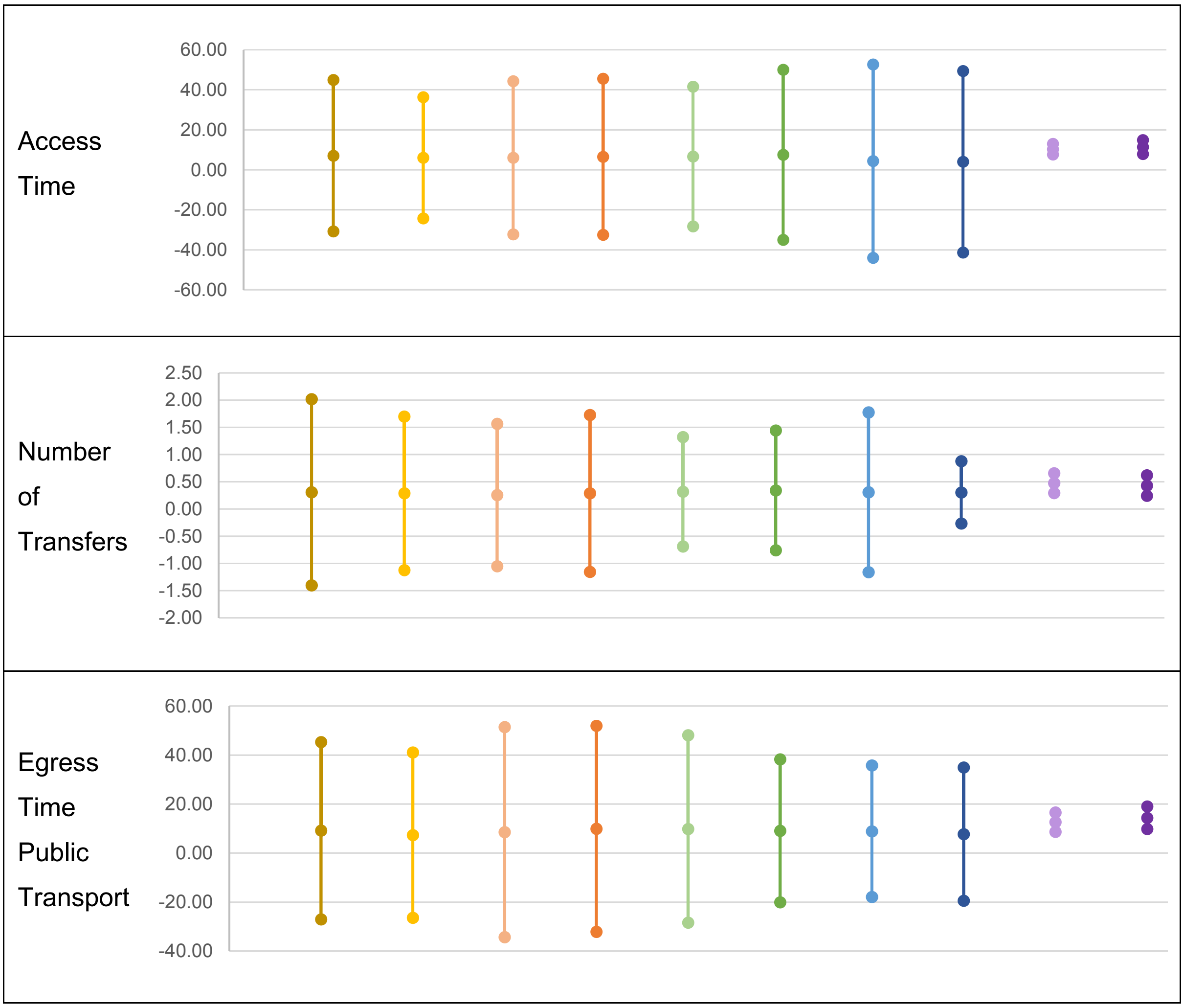
Stability of the WTP estimates? The process homogeneity models have a superior stability for the access time, car egress time and public transport standing density. For the number of transfers, egress time in public transport, and % of seats available, the CRPH model has a superior stability.

Conclusion: The process heterogeneity method directly influences the results. The PDP approach estimates the largest median and lower standard errors for the WTP relative to the process homogeneity models.

5.5.3.2.4. Behavioural Refinements and Experience

Figure 5-22 shows the WTP median and 95% confidence intervals estimates for the models with and without behavioural refinements and/or experience and Table 5-25 presents the comparison of the median estimates using t-test. For the LPAA model, there is a clear decrease in the median WTP estimates when allowing for behavioural refinements and experience, except in the WTP for the train and metro headway, which shows an increase. This is opposite for the VL models, where the inclusion of experience increases the median

WTP estimates for all attributes except for the car egress time. In the case of the RAM heuristic, when including experience there is an increase in the median WTP estimates for all the attributes except for the egress times (both in public transport and car). In the CRPH models, when including behavioural refinements and experience, many of the median WTP estimates are not statistically different from the evidence using the 'simpler' model. However, there is a significant decrease in the median WTP of most of the attributes in the CRPH_BRExp model relative to the CRPH model, except for the egress time in car and public transport headway, where there is a significant increase. Finally, in the case of the PDP models, there are significant differences in all the attributes' median WTP when including behavioural refinements and experience. Most of the attributes' median WTP increase in the PDP_BRExp relative to the PDP model, except for the egress time in car and the number of transfers.



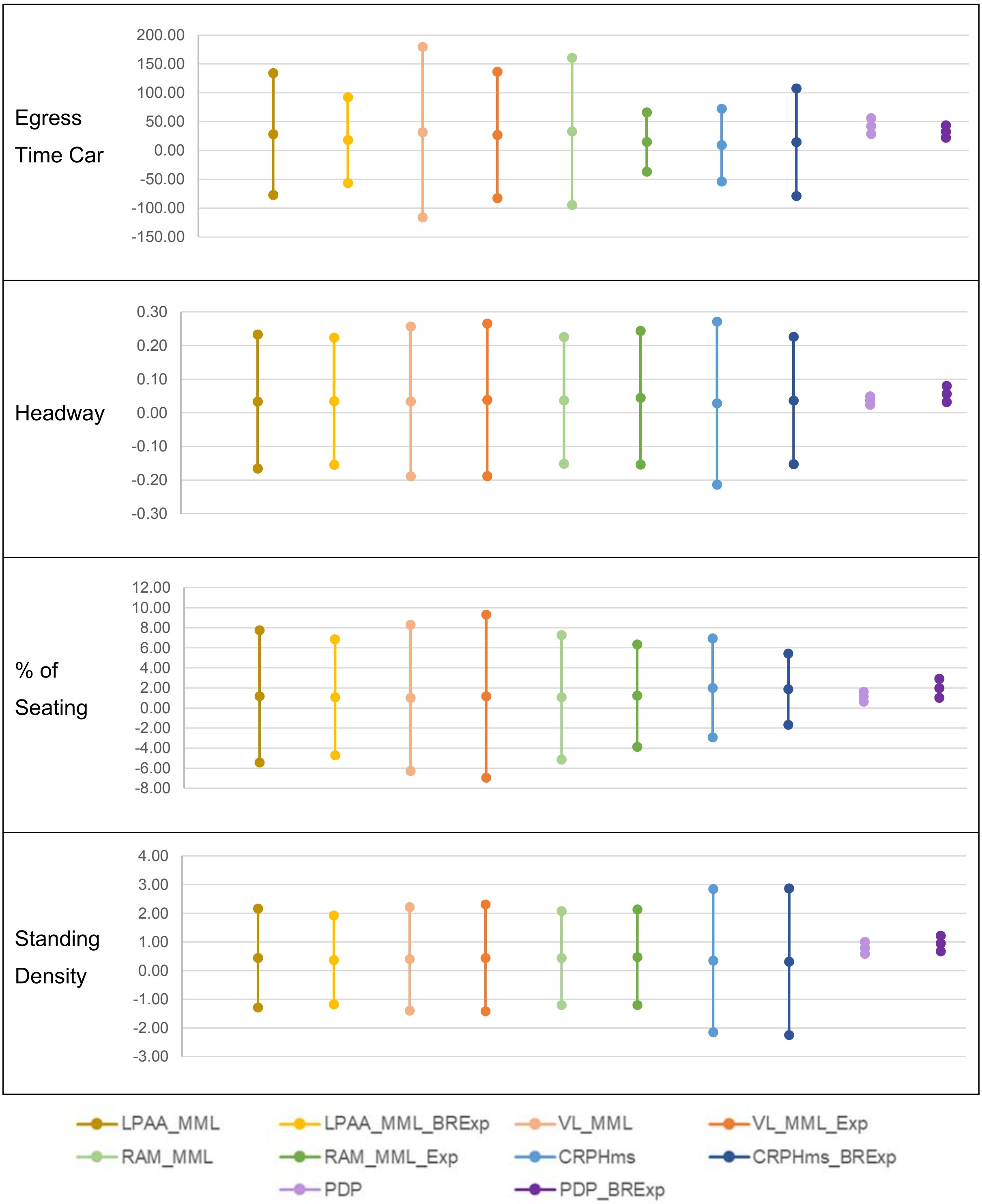


Figure 5-22: Attributes WTP median and 95% confidence intervals for the models with and without behavioural refinements and/or experience

Table 5-25: Comparison of attributes' median WTP for models with and without behavioural refinements and/or experience using t-test

Attribute	Mode	LPAA_MML_BRExp vs LPAA_MML	VL_MML_Exp vs VL_MML	RAM_MML_Exp vs RAM_MML	CRPHms_BRExp vs CRPHms	PDP_BRExp vs PDP
Access Time	Bus	↓ ***-5.14***	↑ 0.95	↑ 1.60	↓ -0.32	↑ ***24.14***
	Train	↓ -1.52	↑ 1.33	↑ ***2.29***	↓ ***-2.14***	↑ ***37.52***
	Metro	↓ -1.92	↑ 1.68	↑ ***2.92***	↑ 0.45	↑ ***43.34***
Egress Time	Bus	↓ ***-7.49***	↑ ***2.41***	↓ -1.74	↓ ***-2.18***	↑ ***34.75***
	Train	↓ ***-3.37***	↑ ***3.38***	↓ ***-2.44***	↓ ***-6.55***	↑ ***45.70***
	Metro	↓ ***-4.26***	↑ ***4.28***	↓ ***-3.09***	↓ ***-5.27***	↑ ***53.83***
	Car	↓ ***-9.49***	↓ ***-3.10***	↓ ***-15.65***	↑ ***5.33***	↓ ***-64.40***
Headway	Bus	↓ ***-2.22***	↑ 1.58	↑ ***2.96***	↑ ***4.69***	↑ ***79.83***
	Train	↑ ***2.25***	↑ ***2.22***	↑ ***4.20***	↑ ***4.59***	↑ ***112.45***
	Metro	↑ ***2.85***	↑ ***2.81***	↑ ***5.34***	↑ ***2.85***	↑ ***140.21***
Seat	Bus	↓ ***-3.94***	↑ 1.79	↑ ***2.27***	↓ -0.78	↑ ***82.79***
	Train	↓ -0.10	↑ ***2.52***	↑ ***3.20***	↓ ***-5.57***	↑ ***121.49***
	Metro	↓ -0.13	↑ ***3.18***	↑ ***4.17***	↓ ***-3.28***	↑ ***151.44***
Stand	Bus	↓ ***-6.86***	↑ 1.48	↑ 1.45	↓ -0.56	↑ ***51.93***
	Train	↓ ***-2.14***	↑ ***2.08***	↑ ***2.04***	↓ ***-2.78***	↑ ***68.20***
	Metro	↓ ***-2.70***	↑ ***2.63***	↑ ***2.67***	↓ -1.22	↑ ***81.83***
Transfers	Bus	↓ ***-3.69***	↑ 1.77	↑ 1.57	↑ 0.79	↓ ***-22.17***
	Train	↑ 0.51	↑ ***2.49***	↑ ***2.24***	↓ -1.84	↓ ***-22.66***
	Metro	↑ 0.64	↑ ***3.14***	↑ ***2.96***	↓ -0.04	↓ ***-31.10***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 5-23 graphically shows the stability of the attributes' WTP estimates (except for travel time). There does not appear to be any pattern associated with the inclusion of behavioural refinements and/or experience.

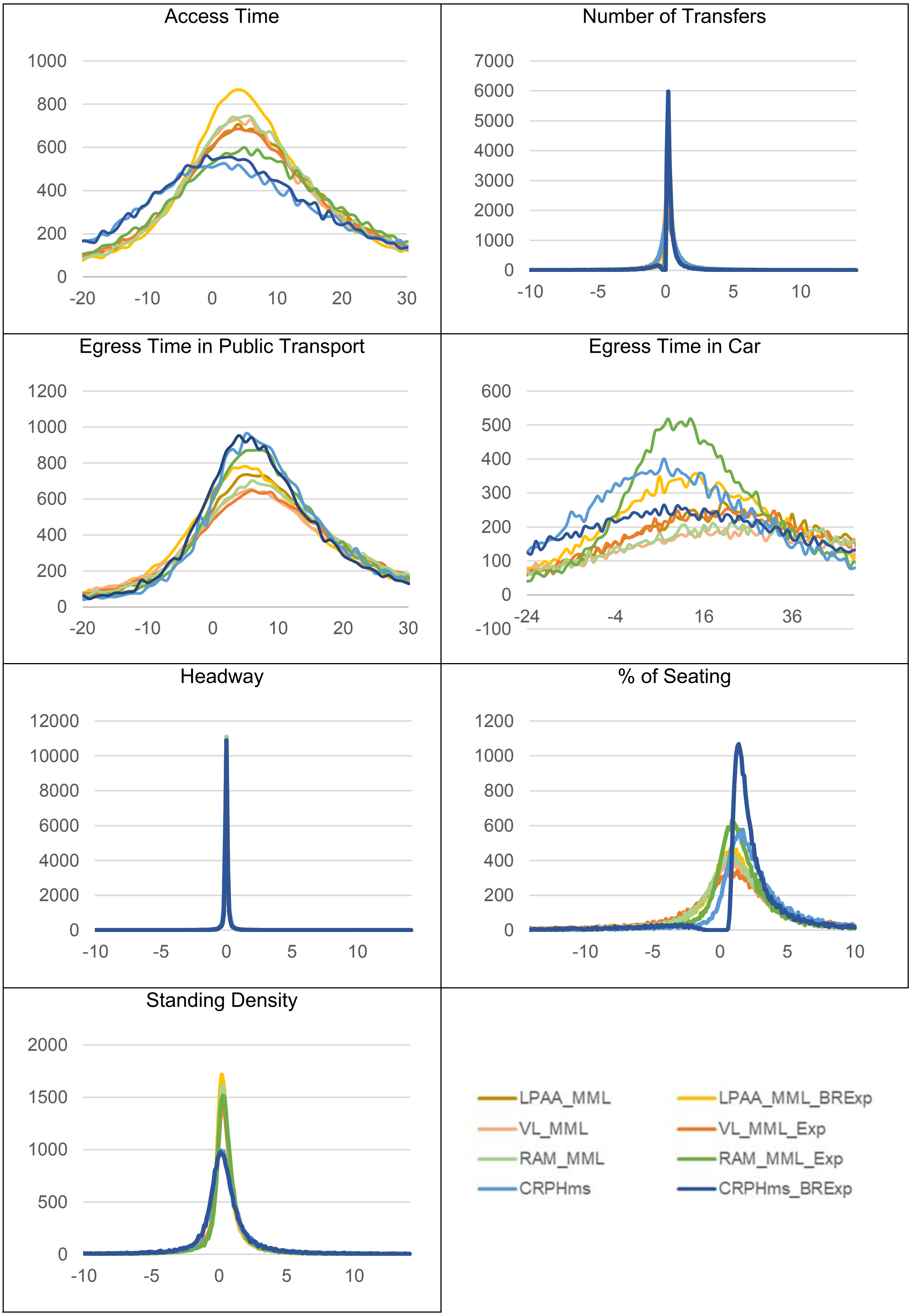


Figure 5-23: Attributes' WTP distributions for different draws using the models with and without behavioural refinements and/or experience

Summary: Comparison of the models with and without experience and behavioural refinements

Access time WTP

- median when adding experience and behavioural refinements
    - increases for VL, RAM and PDP
    - decreases for LPAA and CRPH
- standard error when adding experience and behavioural refinements
    - increases for VL, RAM and PDP
    - decreases for LPAA and CRPH

Egress time public transport WTP

- median when adding experience and behavioural refinements
    - increases for VL and PDP
    - decreases for LPAA, RAM and CRPH
- standard error when adding experience and behavioural refinements
    - increases for PDP and CRPH
    - decreases for LPAA, VL and RAM

Egress time car WTP

- median when adding experience and behavioural refinements
    - increases for CRPH
    - decreases for LPAA, VL, RAM and PDP
- standard error when adding experience and behavioural refinements
    - increases for CRPH
    - decreases for LPAA, VL, RAM and PDP

Headway WTP

- median when adding experience and behavioural refinements
    - increases for VL, RAM, PDP and CRPH
    - decreases for LPAA
- standard error when adding experience and behavioural refinements
    - increases for LPAA, VL, RAM and PDP
    - decreases for CRPH

% of seats available WTP

- median when adding experience and behavioural refinements
    - increases for VL, RAM and PDP
    - decreases for LPAA and CRPH
- standard error when adding experience and behavioural refinements
    - increases for VL and PDP
    - decreases for LPAA, RAM and CRPH

Standing density WTP

- median when adding experience and behavioural refinements
    - increases for VL, RAM and PDP
    - decreases for LPAA and CRPH
- standard error when adding experience and behavioural refinements
    - increases for VL, RAM, PDP and CRPH
    - decreases for LPAA

Number of transfers

- median when adding experience and behavioural refinements
    - increases for VL and RAM
    - decreases for LPAA, PDP and CRPH
- standard error when adding experience and behavioural refinements
    - increases for VL, RAM and PDP
    - decreases for LPAA and CRPH

Significant differences? There are significant differences in the median WTP estimates when considering behavioural refinements and/or experience in the LPAA model for all except for the access time, % of seats available and number of transfers for the train and metro. In the VL model the bus median WTP are not significantly different for any attribute except for the egress time, for the rest of the modes all of them are significantly different except for the access time. In the RAM model all the median WTP are significantly different except for the access time, egress time, standing density and number of transfers on the bus mode. In the CRPH model all the differences are significantly different, except for the bus and metro access times, bus % of seats available, bus and metro standing density, and the number of transfers in every mode. In the PDP model all the differences are significant.

Stability of the WTP estimates? Similar stability with and without experience and behavioural refinements.

Conclusion: The inclusion of experience and behavioural refinements significantly influences the WTP median and standard error for the majority of the attributes.

5.6. Conclusions

This chapter has reviewed different model specifications and what this might mean as a representation of individual preferences in the Metro Rail dataset. Three different process strategies, RAM, VL and LPAA were combined to estimate five types of models: (1) individuals use LPAA as their only process strategy; (2) individuals use VL as their only process strategy; (3) individuals use RAM as their only process strategy; (3) individuals may use any of the three process strategies with a certain probability (PDP); (4) individuals use a combination of the three process strategies by conditioning the parameters normally defined under LPAA with the alternative strategies (CRPH).

The first part of this chapter reviewed the best models for each type in terms of the overall model fit (using the log likelihood ratio test and the Vuong test). The results showed that the inclusion of experience and behavioural refinements has a statistically superior overall performance for all the model types. The preferred models also considered random parameters, as they provided a better overall model fit, except for the PDP approach that considers all parameters as fixed.

The models were first compared through their behavioural refinements and experience. The results showed significant differences in the behaviour of the experience parameter when considering different process strategies as the sole heuristic being used by individuals. Since the nature of the process strategies is different, it was anticipated that experience would have a different influence in each of them. Moreover, the results indicated that experience, risk attitudes and perceptual conditioning are less statistically significant when considering process heterogeneity than under process homogeneity. This is a crucial finding that suggests that the importance of including additional behavioural components is reduced when considering process heterogeneity; suggesting some amount of confoundment that can be captured through specific behavioural assumptions rather that a random distribution of parameter estimates. This raises the interesting question on what role random parameters play in capturing process heterogeneity when specific processing heuristics are accommodated.

General indicators, such as the log likelihood and the AIC were used to compare models. They showed that the CRPH preferred model (CRPHms_BRExp) was superior to the other ones. This was followed – not so closely – by the CRPHms, the LPAA_MML_BRExp and the RAM_MML_Exp models.

The WTP median and the 95% confidence intervals were compared to see if the different formulations produced different results, and if these differences were statistically significant. This section revealed significant differences in the WTP estimates for all the indicators when considering different process strategies as the only heuristic being used by respondents, and also differences when considering process homogeneity instead of process heterogeneity. What we find in this chapter is that when using the PDP approach to include process heterogeneity, the median WTP estimates are usually larger relative to the process homogeneity models such as LPAA, VL or RAM. Contrarily, when using the CRPH approach – considering our preferred model, CRPHms_BRExp - the median WTP estimates are usually lower in comparison to the process homogeneity models. This suggests that the inclusion of process heterogeneity and the approach used has a significant influence on the key behavioural outcomes such as WTP estimates; hence it is recommended that the findings in the chapter be used as a guide for researchers in considering the role of richer behavioural forms of models in representing choice making. The next chapter will use another dataset to establish the extent to which the findings in this chapter can be supported more generally.

CHAPTER 6

Results Northwest

6.1. Introduction

This chapter will present the model results using the NorthWest dataset described in Chapter 4 for the heuristics LPAA, VL and RAM, together with behavioural refinements and experience. For the utility expressions used in this chapter, the reader is referred to the 'Notational Glossary' at the beginning of the thesis. The chapter is structured in the exactly same way as chapter 5, presenting the models and their implications but excluding information already presented in chapter 5 (the reader is referred to chapter 5 for more details on each section). The next section presents the model results for the process homogeneity models. Section 3 will present the process heterogeneity model results using the PDP approach, while section 4 will present the process heterogeneity model results using the CRPH approach. Section 5 will compare the models using: (1) the results on behavioural refinements and experience; (2) the log likelihood and AIC indicators; and (3) the willingness to pay estimates (WTP) separated into the value of travel time savings (first part) and all the other attributes (second part). The last section discusses the main findings.

6.2. Simple MNL and MML Model Results

6.2.1. Linear Parameters and Additive Attributes

The simplest model, using the linear parameters and additive attributes (LPAA) model form, is the LPAA_MNL, which considers all the parameters as fixed. Model LPAA_MML is an equivalent model but estimates all the parameters as random. The LPAA_MNL_BRExp is a fixed parameter LPAA model with behavioural refinements and experience; and the LPAA_MML_BRExp model is a random parameter LPAA with behavioural refinements and experience. The results for these models are presented in Table 6-1. All the possible combinations of risk attitudes and experience were tested, but not all of them were found to be statistically significant.

Results show that experience was statistically significant for all the modes in the LPAA_MNL_BRExp model, but only for the train and car in the LPAA_MML_BRExp model. That is, when estimating random parameters, the influence of experience changes and it is not statistically significant for all the modes compared to the fixed parameters model. This suggests that maybe some of the experience effect is being captured by the heterogeneity of specific modal attributes. All the experience parameters were positive, suggesting that individuals are more likely to choose the mode they chose in their most recent trip. The results indicate that the utility conditioning form used to take into account respondents' experience (detailed in Section 3.5.1) was appropriate as a way of representing individual decision-making.

The risk attitudes associated with travel times and costs of the different modes was investigated. After several model revisions, results showed that it was appropriate to distinguish the risk attitudes common between the proposed new modal investments (i.e., new light rail, new heavy rail and new busway) and between the currently available modal facilities (i.e., bus, busway and train). Figure 6-1 presents the risk attitudes towards travel times. As can be seen, for the MNL models. The negative α and convexity of the function support risk taking attitude. Figure 6-2 presents the risk attitudes towards costs suggesting risk aversion towards the cost with all the parameters negative and concave shapes. The risk aversion attitude is similar for the parking cost, and the public transport fares in the MNL model. However, in the MML model, there are significant differences in the risk attitudes associated with public transport fares of the currently available modal facilities and towards the new modal investments. This highlights the different roles played by existing alternatives and a prospective mode.

Table 6-1: LPAA MNL and MML models (t-values in brackets)

			LPAA_MNL	LPAA_MML		LPAA_MNL_BRExp	LPAA_MML_BRExp	
Number of Parameters Estimated			**14**	**22**		**23**	**26**	
Log Likelihood at convergence			**-6,170.72**	**-4,701.06**		**-6,033.82**	**-4,611.45**	
Log likelihood at zero			**-7,838.25**					
AIC			**2.731**	**2.085**		**2.674**	**2.047**	
Parameters		**Alternatives**	**Mean**	**Mean**	**Std Dev**	**Mean**	**Mean**	**Std Dev**
Alternative Specific Constant New Light Rail	ASCLR	New Light Rail	1.76 (8.97)	3.63 (7.58)	-	3.35 (8.14)	5.48 (8.39)	-
Alternative Specific Constant New Heavy Rail	ASCNHR	New Heavy Rail	1.84 (10.83)	3.81 (8.34)	-	3.45 (8.64)	5.66 (8.89)	-
Alternative Specific Constant New Busway	ASCNBW	New Busway	0.63 (3.21)	2.58 (5.50)	-	2.32 (5.64)	4.47 (6.83)	-
Alternative Specific Constant Bus	ASCBS	Bus	1.51 (8.86)	3.21 (6.98)	-	3.02 (7.65)	5.13 (7.98)	-
Alternative Specific Constant Busway	ASCBW	Busway	1.23 (7.19)	3.02 (6.60)	-	2.87 (7.28)	4.85 (7.51)	-
Alternative Specific Constant Train	ASCTRAIN	Train	1.31 (7.60)	2.82 (6.23)	-	2.75 (7.10)	4.15 (6.53)	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Public Transport	-0.04 (13.76)	-0.07 (9.17)	0.11 (12.24)	-0.04 (13.09)	-0.07 (9.33)	0.11 (13.90)
Fare Public Transport	COSTPT	Public Transport	-0.25 (26.49)	-0.45 (20.28)	0.42 (17.07)	-0.76 (15.30)	-1.03 (11.62)	0.90 (15.05)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.17 (9.68)	-0.28 (5.34)	0.31 (9.1)	-0.20 (10.40)	-0.86 (12.40)	0.81 (19.05)
Parking Cost Car	COSTCRPC	Car	-0.03 (5.52)	-0.08 (4.32)	0.31 (11.31)	-0.11 (4.83)	-0.10 (5.88)	0.15 (8.43)
Travel Time Public Transport	TTPT	Public Transport	-0.04 (31.69)	-0.07 (22.87)	0.06 (18.9)	-0.02 (3.33)	-0.07 (22.46)	0.06 (19.39)
Travel Time Car	TTCR	Car	-0.03 (11.74)	-0.09 (10.06)	0.08 (15.06)	-0.03 (11.74)	-0.08 (11.02)	0.07 (15.58)
Egress Time	EGTIME	All Alternatives	-0.04 (12.89)	-0.06 (8.98)	0.07 (9.12)	-0.03 (11.86)	-0.06 (7.11)	0.07 (6.10)
Headway Public Transport	FREQPT	Public Transport	-0.02 (2.10)	-0.05 (2.98)	0.20 (12.91)	-0.02 (1.70)	-0.03 (1.93)	0.19 (13.83)
Experience Bus	EXPBS	Bus	-	-	-	0.20 (7.09)	-	-
Experience Busway	EXPBW	Busway	-	-	-	0.12 (4.22)	-	-
Experience Train	EXPTR	Train	-	-	-	0.18 (7.39)	0.25 (9.46)	-
Experience Car	EXPCR	Car	-	-	-	0.09 (3.00)	0.17 (2.68)	-

			LPAA_MNL	LPAA_MML		LPAA_MNL_BRExp	LPAA_MML_BRExp	
Risk Attitudes Travel Time Currently Available Modal Facilities	ALPHAEXTT	Bus, Busway, Train and Car	-	-	-	-0.19 (2.55)	-	-
Risk Attitudes Travel Time New Modal Investments	ALPHANEXTT	New Light Rail, New Heavy Rail, New Busway	-	-	-	-0.18 (2.36)	-	-
Risk Attitudes Travel Time Car	ALPHACRTT	Car	-	-	-	-	-	-
Risk Attitudes Cost Currently Available Modal Facilities	ALPHAEXCS	Bus, Busway, Train and Car	-	-	-	0.62 (16.89)	0.45 (9.54)	-
Risk Attitudes Cost New Modal Investments	ALPHANEXCS	New Light Rail, New Heavy Rail, New Busway	-	-	-	0.61 (14.46)	0.45 (9.24)	-
Risk Attitudes Fuel+Toll Car	ALPHACRTRCS	Car	-	-	-	-	-	-
Risk Attitudes Parking Car	ALPHACRPCS	Car	-	-	-	0.56 (5.02)	-	-

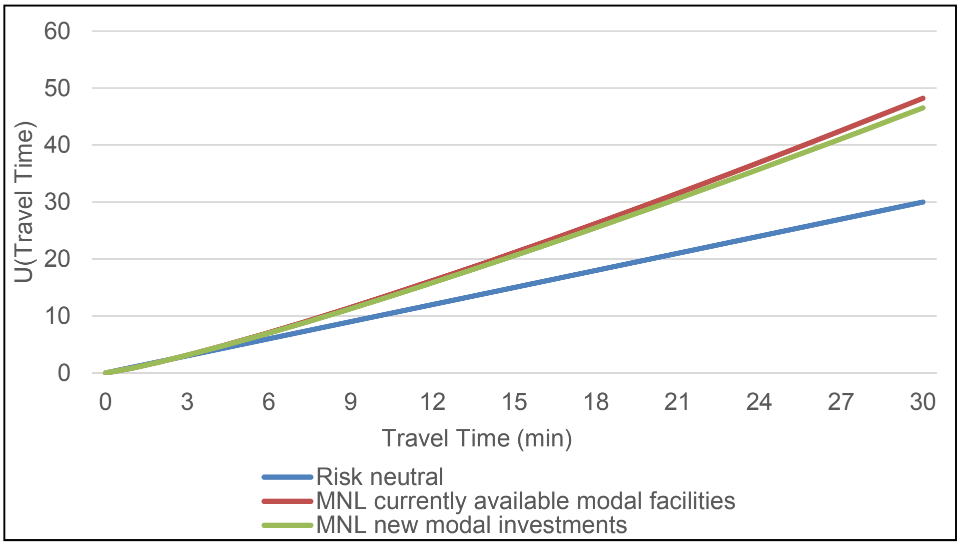


Figure 6-1: Risk attitudes towards the travel times in LPAA_MNL_BRExp

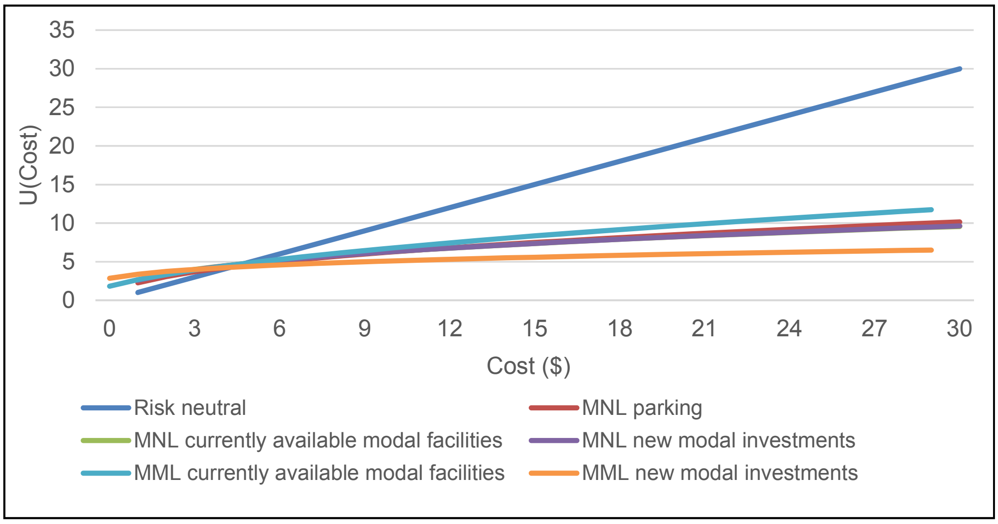


Figure 6-2: Risk attitudes towards costs in LPAA_MNL_BRExp and LPAA_MML_BRExp

The models can be compared through the log likelihood ratio test (explained in Section 3.8.2) with the results summarised in Table 6-2. The null hypotheses stating that the models are equivalent are rejected. This means that the inclusion of random parameters, behavioural refinements and experience significantly improve the models. The preferred model is the LPAA_MML_BRExp which has a significantly better overall fit and provides a behaviourally more appealing understanding of preferences through the influence of preference heterogeneity, risk attitudes, perceptual conditioning and overt experience.

Table 6-2: Log likelihood ratio test results for the LPAA models

	Random parameters		Behavioural refinements and experience	
	LPAA_MML vs. LPAA_MNL	LPAA_MML_BRExp vs. LPAA_MNL_BRExp	LPAA_MNL_BRExp vs. LPAA_MNL	LPAA_MML_BRExp vs. LPAA_MML
LR	2,939.336	2,844.736	273.808	179.208
Degrees of freedom	8	3	9	4
$\chi^2_{d.f.;0.001}$	26.124	16.266	27.877	18.467
Result	Reject null	Reject null	Reject null	Reject null

6.2.2. Value Learning

MNL and MML

The results of the models under value learning as the only process strategy are given in Table 6-3. The model, referred to as VL_MNL, considers all parameters as fixed across the sample; model VL_MML estimates all parameters as random; model VL_MNL_Exp considers all the parameters as fixed and includes experience conditioning the utility function; while model VL_MML_Exp considers experience but estimates all parameters as random.

Models VL_MNL_Exp and VL_MML_Exp find a statistically significant role for the experience in using bus, busway and car. The experience estimates were all positive suggesting that individuals are more likely to choose the mode they used in their most recent trip, after taking into consideration value learning effects.

Table 6-3: VL MNL and MML models (t-values in brackets)

			VL_MNL	VL_MML		VL_MNL_Exp	VL_MML_Exp	
Number of Parameters Estimated			**15**	**21**		**19**	**25**	
Log Likelihood at convergence			**-6,004.14**	**-4,729.70**		**-5,982.88**	**-4,702.90**	
Log likelihood at zero			**-7,838.25**					
AIC			**2.657**	**2.097**		**2.650**	**2.087**	
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Std Dev**	**Mean**	**Mean**	**Std Dev**
Alternative Specific Constant New Light Rail	ASCLR	New Light Rail	1.37 (11.12)	2.19 (10.99)	-	1.32 (10.98)	1.76 (8.44)	-
Alternative Specific Constant New Heavy Rail	ASCNHR	New Heavy Rail	1.16 (12.10)	2.29 (13.64)	-	1.07 (11.61)	1.83 (10.63)	-
Alternative Specific Constant New Busway	ASCNBW	New Busway	-0.07 (0.51)	1.00 (5.02)	-	-0.16 (1.17)	0.52 (2.53)	-
Alternative Specific Constant Bus	ASCBS	Bus	0.80 (8.03)	1.58 (9.30)	-	0.63 (6.67)	1.02 (5.83)	-
Alternative Specific Constant Busway	ASCBW	Busway	0.53 (5.43)	1.45 (8.68)	-	0.35 (3.52)	0.84 (4.72)	-
Alternative Specific Constant Train	ASCTR	Train	0.66 (6.78)	1.24 (7.35)	-	0.45 (4.70)	0.53 (2.94)	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Public Transport	-0.02 (8.90)	-0.07 (8.48)	0.11 (14.36)	-0.02 (8.93)	-0.06 (7.93)	0.11 (13.77)
Fare Public Transport	COSTPT	Public Transport	-0.16 (13.50)	-0.46 (17.02)	0.47 (16.81)	-0.17 (13.62)	-0.47 (17.35)	0.50 (16.17)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.10 (8.49)	-0.36 (7.75)	0.46 (10)	-0.10 (8.59)	-0.41 (8.38)	0.50 (10.02)
Parking Cost Car	COSTCRPC	Car	-0.02 (4.45)	-0.05 (4.28)	-	-0.02 (4.93)	-0.05 (4.76)	-
Travel Time Public Transport	TTPT	Public Transport	-0.02 (8.65)	-0.07 (21.24)	0.06 (17.89)	-0.02 (8.72)	-0.07 (21.56)	0.07 (19.17)
Travel Time Car	TTCR	Car	-0.01 (5.04)	-0.07 (7.61)	0.06 (10.22)	-0.01 (5.66)	-0.07 (9.29)	0.07 (11.56)
Egress Time	EGTIME	All Alternatives	-0.01 (7.28)	-0.04 (6.24)	0.08 (9.66)	-0.02 (7.43)	-0.04 (6.68)	0.08 (9.17)
Headway Public Transport	FREQPT	Public Transport	-0.04 (6.43)	-0.08 (4.84)	0.23 (15.31)	-0.04 (6.78)	-0.08 (4.59)	0.22 (14.93)
Experience Bus	EXPBS	Bus	-	-	-	0.29 (3.94)	0.18 (3.00)	-
Experience Busway	EXPBW	Busway	-	-	-	0.24 (3.29)	0.21 (3.63)	-
Experience Train	EXPTR	Train	-	-	-	0.32 (4.33)	0.30 (5.77)	-
Experience Car	EXPCR	Car	-	-	-	-0.32 (2.48)	-0.97 (4.59)	-
Concavity VL	CONC	All Alternatives	1.22 (7.57)	-	-	1.21 (7.25)	-	-

The concavity factor represents the way in which individuals weight the differences between the levels presented to them and their reference levels (VL). They are statistically different from zero when considering all parameters fixed, in the VL_MNL and VL_MNL_Exp models. The results of these weighting functions are presented graphically in Figure 6-3. Both models estimate a relatively similar concavity parameter (around 1.21), which shows that individuals tend to overweight higher differences as shown in Figure 6-3.

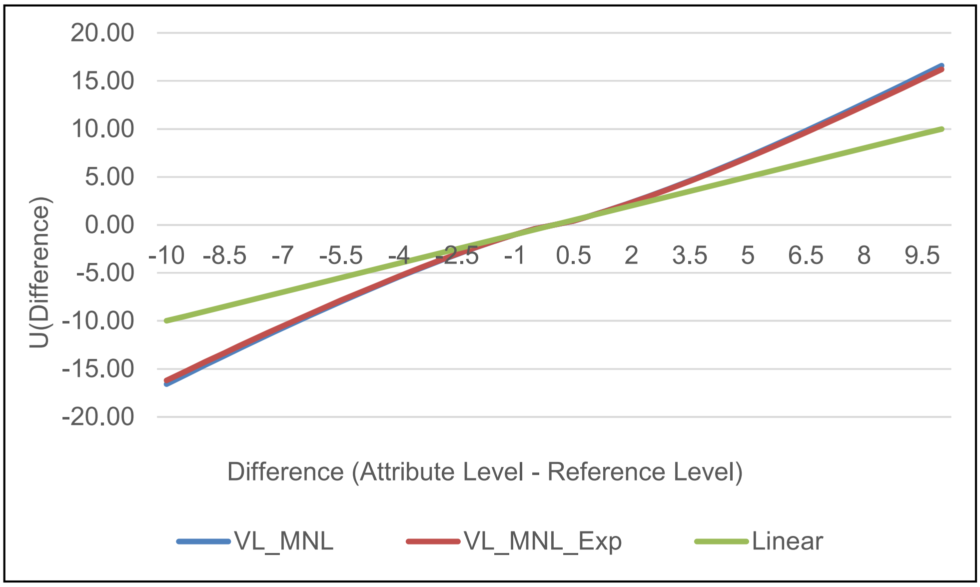


Figure 6-3: Concavity factor for VL models

As the models are nested, they can be compared using the log likelihood ratio test. The results are shown in Table 6-4. They suggest that both the inclusion of random parameters and experience significantly improve the overall statistical fit of the models. The preferred model on statistical and behavioural grounds is VL_MML_Exp.

Table 6-4: Log likelihood ratio test results for the VL models

	Random parameters		**Experience**	
	VL_MML vs. VL_MNL	VL_MML_Exp vs. VL_MNL_Exp	VL_MNL_Exp vs. VL_MNL	VL_MML_Exp vs. VL_MML
LR	2548.89	2559.952	42.522	53.584
Degrees of freedom	6	6	4	4
$\chi^2_{d.f.;0.001}$	22.458	22.458	18.467	18.467
Result	Reject null	Reject null	Reject null	Reject null

Choice set correlation

As explained in Section 3.4.2.1, it is important to consider if the model form proposed for VL induces any type of relationship between the choice sets. This section is equivalent to 5.2.2, where the four models are estimated. Models EC_LPAA and EC_VL represent a MNL model for the LPAA and VL heuristic, respectively, adding an error component common within individuals and different across individuals. Models EC_SeqLPAA and EC_SeqVL include the traditional error component and an additional error component that depends on the choice set sequence (included as a continuous variable).

Table 6-5 presents the results for the models. Models EC_LPAA and EC_VL show statistically significant estimates for the traditional error components for all the modes. That is, there is a significant part of the unobserved utility function that varies across - but not within - individuals. Model EC_SeqLPAA shows that the choice set sequence number significantly influences the unobserved part of the utility function for the new light rail mode, with 95% confidence level. This influence is statistically significant with 80% confidence level for the new heavy rail and car. The EC_SeqVL model has similar results, but the influence of the choice set sequence number in the unobserved part of the utility function has a higher significance level for the car (with a 95% confidence level) and a lower one in the new heavy rail (not significant).

In summary, the results show that VL does not induce any relationship between the error term and the choice set sequence relative to a traditional LPAA heuristic. There are some differences in the estimates for certain modes, as it is for the car and new heavy rail, but they are equivalent for all the rest. These findings suggests that the proposed model form for VL is appropriate.

Table 6-5: Choice set correlation model results (t-values in brackets)

			EC_LPAA	EC_SeqLPAA	EC_VL	EC_SeqVL
Number of Parameters Estimated			**21**	**28**	**22**	**29**
Log Likelihood at convergence			**-4,514.27**	**-4,522.32**	**-4,537.69**	**-4,488.28**
Log likelihood at zero			**-7,838.25**			
AIC			**2.002**	**2.009**	**2.013**	**1.994**
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Mean**	**Mean**
Alternative Specific Constant New Light Rail	ASCLR	New Light Rail	4.38 (6.35)	5.16 (9.86)	3.16 (8.09)	4.64 (8.10)
Alternative Specific Constant New Heavy Rail	ASCNHR	New Heavy Rail	4.57 (6.96)	5.56 (11.24)	3.56 (9.88)	4.72 (8.33)
Alternative Specific Constant New Busway	ASCNBW	New Busway	3.25 (4.82)	4.25 (8.09)	2.31 (5.79)	3.38 (5.66)
Alternative Specific Constant Bus	ASCBS	Bus	3.15 (4.78)	4.57 (9.01)	2.87 (7.69)	3.47 (5.83)
Alternative Specific Constant Busway	ASCBW	Busway	3.51 (5.27)	4.69 (9.41)	2.70 (7.43)	3.88 (6.86)
Alternative Specific Constant Train	ASCTR	Train	3.25 (4.79)	4.28 (8.54)	2.47 (6.58)	3.51 (6.10)
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-
Access Time	ACTIMEPT	Public Transport	-0.04 (6.94)	-0.04 (7.75)	-0.04 (6.28)	-0.05 (5.80)
Fare Public Transport	COSTPT	Public Transport	-0.36 (22.43)	-0.37 (22.54)	-0.35 (13.72)	-0.39 (13.69)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.17 (4.25)	-0.17 (4.51)	-0.09 (2.64)	-0.07 (2.08)
Parking Cost Car	COSTCRPC	Car	-0.05 (3.56)	-0.06 (4.14)	-0.04 (2.27)	-0.04 (2.30)
Travel Time Public Transport	TTPT	Public Transport	-0.06 (29.68)	-0.06 (29.47)	-0.07 (9.22)	-0.08 (9.17)
Travel Time Car	TTCR	Car	-0.05 (7.36)	-0.03 (5.91)	-0.05 (5.60)	-0.05 (5.27)
Egress Time	EGTIME	All Alternatives	-0.05 (8.43)	-0.05 (8.93)	-0.05 (6.64)	-0.05 (6.54)
Headway Public Transport	FREQPT	Public Transport	-0.01 (0.86)	-0.02 (1.38)	-0.02 (1.66)	-0.01 (0.55)
Concavity VL	CONC	All Alternatives	-	-	0.96 (1.64)	0.94 (2.32)
Error Components ASC New Light Rail	EC_LR	New Light Rail	1.44 (9.80)	-2.23 (9.35)	-1.96 (13.23)	1.82 (8.07)
Error Components Sequence New Light Rail	EC_LR_SEQ	New Light Rail	-	0.08 (2.34)	-	-0.11 (2.96)
Error Components ASC New Heavy Rail	EC_NHR	New Heavy Rail	1.34 (9.64)	1.51 (7.04)	1.91 (12.60)	-1.64 (6.88)
Error Components Sequence New Heavy Rail	EC_NHR_SEQ	New Heavy Rail	-	0.04 (1.32)	-	0.01 (0.21)
Error Components ASC New Busway	EC_NBW	New Busway	0.74 (3.18)	-0.58 (1.69)	-0.85 (3.75)	0.78 (2.38)
Error Components Sequence New Busway	EC_NBW_SEQ	New Busway	-	-0.04 (0.87)	-	0.05 (0.94)
Error Components ASC Bus	EC_BS	Bus	3.34 (15.55)	-2.81 (12.41)	2.69 (16.46)	3.01 (13.54)
Error Components Sequence Bus	EC_BS_SEQ	Bus	-	0.02 (0.75)	-	-0.02 (0.80)
Error Components ASC Busway	EC_BW	Busway	-1.83 (13.30)	-1.38 (7.25)	-1.36 (10.52)	-1.81 (8.52)
Error Components Sequence Busway	EC_BW_SEQ	Busway	-	0.02 (0.54)	-	0.03 (0.84)
Error Components ASC Train	EC_TR	Train	1.86 (14.27)	-1.74 (9.49)	-1.89 (13.52)	1.96 (8.67)
Error Components Sequence Train	EC_TR_SEQ	Train	-	-0.02 (0.61)	-	0.01 (0.24)
Error Components ASC Car	EC_CR	Car	3.94 (17.05)	4.54 (13.56)	3.70 (11.73)	-5.09 (9.51)
Error Components Sequence Car	EC_CR_SEQ	Car	-	-0.05 (1.50)	-	0.07 (2.10)

6.2.3. Relative Advantage Maximisation

This section presents the models that consider RAM as the only process strategy being used by individuals. The RAM_MNL model considers fixed parameters, and the RAM_MML model has random parameters. The RAM_MNL_Exp and RAM_MML_Exp models are equivalent to these ones but including experience. The results are shown in Table 6-6. Experience was statistically significant in decision making for every mode (bus, busway, train and car) with all estimates positive, showing that respondents are more likely to choose the same mode they used in their most recent trip.

The log likelihood ratio test used to compare the model results are shown in Table 6-7. All the null hypotheses stating that the models are equivalent are rejected. This shows that the consideration of all parameters as random and the inclusion of experience represents a significant improvement in the overall statistical performance of the models. The preferred model for the RAM heuristic is the one that considers random parameters and experience, RAM_MML_Exp.

Table 6-6: RAM MNL and MML models (t-values in brackets)

			RAM_MNL	RAM_MML		RAM_MNL_Exp	RAM_MML_Exp	
Number of Parameters Estimated			**14**	**22**		**18**	**25**	
Log Likelihood at convergence			**-6,169.89**	**-4,659.04**		**-6,053.10**	**-4,566.35**	
Log likelihood at zero			**-7,838.25**					
AIC			**2.730**	**2.067**		**2.680**	**2.027**	
Parameters	**Acronym**	**Alternatives**	**Mean**	**Mean**	**Std Dev**	**Mean**	**Mean**	**Std Dev**
Alternative Specific Constant New Light Rail	ASCLR	New Light Rail	2.10 (10.65)	2.59 (5.31)	-	2.76 (13.32)	4.74 (12.25)	-
Alternative Specific Constant New Heavy Rail	ASCNHR	New Heavy Rail	2.19 (12.48)	2.75 (5.85)	-	2.88 (15.26)	4.86 (13.19)	-
Alternative Specific Constant New Busway	ASCNBW	New Busway	0.99 (4.85)	1.58 (3.29)	-	1.62 (7.55)	3.64 (9.44)	-
Alternative Specific Constant Bus	ASCBS	Bus	1.85 (10.39)	2.16 (4.67)	-	2.20 (11.63)	3.94 (10.57)	-
Alternative Specific Constant Busway	ASCBW	Busway	1.58 (8.94)	1.97 (4.25)	-	2.16 (11.36)	3.91 (10.71)	-
Alternative Specific Constant Train	ASCTR	Train	1.68 (9.55)	1.72 (3.68)	-	2.09 (11.07)	3.38 (9.40)	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Public Transport	-0.03 (13.97)	-0.06 (10.70)	0.10 (14.86)	-0.03 (13.83)	-0.04 (8.09)	0.09 (11.13)
Fare Public Transport	COSTPT	Public Transport	-0.17 (25.08)	-0.45 (18.50)	0.49 (19.55)	-0.16 (23.62)	-0.37 (17.46)	0.45 (14.99)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.11 (9.74)	-0.12 (4.01)	0.05 (2.32)	-0.08 (9.20)	-0.06 (2.25)	-
Parking Cost Car	COSTCRPC	Car	-0.02 (5.33)	-0.08 (6.69)	0.26 (13.85)	-0.01 (3.47)	-0.12 (5.86)	0.15 (5.70)
Travel Time Public Transport	TTPT	Public Transport	-0.03 (30.07)	-0.05 (21.92)	0.05 (21.7)	-0.03 (28.76)	-0.05 (17.79)	0.06 (16.24)
Travel Time Car	TTCR	Car	-0.02 (12.08)	-0.10 (10.38)	0.17 (18.6)	-0.02 (11.26)	-0.08 (10.33)	0.14 (12.86)
Egress Time	EGTIME	All Alternatives	-0.02 (12.63)	-0.03 (7.83)	0.04 (5.24)	-0.02 (11.82)	-0.03 (7.06)	0.07 (8.37)
Headway Public Transport	FREQPT	Public Transport	-0.01 (2.01)	-0.04 (3.63)	0.19 (11.5)	-0.01 (1.85)	-0.05 (4.27)	0.19 (14.53)
Experience Bus	EXPBS	Bus	-	-	-	0.39 (7.80)	0.19 (4.27)	-
Experience Busway	EXPBW	Busway	-	-	-	0.14 (3.28)	0.11 (2.96)	-
Experience Train	EXPTR	Train	-	-	-	0.31 (7.40)	0.32 (8.98)	-
Experience Car	EXPCR	Car	-	-	-	0.32 (5.69)	0.66 (11.43)	-

Table 6-7: Log likelihood ratio test results for the RAM models

	Random parameters		Experience	
	RAM_MML vs. RAM_MNL	RAM_MML_Exp vs. RAM_MNL_Exp	RAM_MNL_Exp vs. RAM_MNL	RAM_MML_Exp vs. RAM_MML
LR	3,021.686	2,973.486	233.584	185.384
Degrees of freedom	8	7	4	3
$\chi^2_{d.f.;0.001}$	26.124	24.322	18.467	16.266
Result	Reject null	Reject null	Reject null	Reject null

6.3. Probabilistic Decision Process Model Results: LPAA, VL and RAM together with behavioural refinements and experience

The probabilistic decision process (PDP) models consider process heterogeneity using a latent class structure (although each class is not latent per se), where each class represents a different heuristic. The results of the models are presented in Table 6-8. The first class represents the RAM heuristic, the second class VL, and the third one the LPAA assumption. The PDP model considers a simple structure while the PDP_BRExp considers behavioural refinements in the LPAA heuristic (class 3) and experience in all the classes.

Model PDP suggests that that there is, on average, a 0.51 probability that respondents use the RAM heuristic, 0.32 probability of using the VL heuristic and a 0.17 probability of using the LPAA heuristic, which is almost equivalent to the results obtained in the Metro Rail data. Some of the parameters were found not to be statistically significant, especially for the LPAA class for fuel plus toll costs, car travel time, egress time and headway. In the VL class the access time, parking cost and car travel time were not statistically significant. In the RAM class the fuel plus toll costs and headway were not significant.

The PDP_BRExp model results show that there is, on average, a 0.15 probability that respondents use the RAM heuristic, with experience, a 0.34 probability of using the VL heuristic with experience and a 0.50 probability associated with the LPAA heuristic with behavioural refinements and experience. In the RAM heuristic the car travel time, egress time and headway were not statistically significant.

Table 6-8: PDP models (t-values in brackets)

			PDP			PDP_BRExp		
Number of Parameters Estimated			**35**			**44**		
Log Likelihood at convergence			**-4,953.14**			**-4,864.47**		
Log likelihood at zero			**-7,838.25**					
AIC			**2.202**			**2.167**		
Class Identification			**Class 1**	**Class 2**	**Class 3**	**Class 1**	**Class 2**	**Class 3**
Heuristic			RAM	VL	LPAA	RAM	VL	LPAA
Behavioural Refinements			N	N	N	Y	N	N
Experience			N	N	N	Y	Y	Y
Class Membership (%)			51%	32%	17%	15%	34%	50%
Parameters		**Alternatives**	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**
Alternative Specific Constant New Light Rail	ASCLR	New Light Rail	7.52 (13.81)	2.02 (7.40)	-2.63 (5.53)	-3.50 (3.90)	1.54 (6.49)	3.68 (4.76)
Alternative Specific Constant New Heavy Rail	ASCNHR	New Heavy Rail	7.19 (13.25)	2.12 (10.12)	-1.98 (4.94)	-2.34 (4.42)	1.38 (7.13)	3.41 (4.24)
Alternative Specific Constant New Busway	ASCNBW	New Busway	6.89 (12.25)	-0.12 (0.39)	-1.85 (3.67)	-2.12 (3.06)	-0.64 (2.14)	3.12 (3.92)
Alternative Specific Constant Bus	ASCBS	Bus	7.16 (13.04)	-0.08 (0.33)	1.09 (3.79)	1.36 (4.85)	-0.64 (2.23)	3.00 (3.74)
Alternative Specific Constant Busway	ASCBW	Busway	7.03 (13.03)	0.60 (2.77)	-0.38 (1.51)	-0.26 (0.88)	-0.03 (0.14)	3.01 (3.81)
Alternative Specific Constant Train	ASCTRAIN	Train	7.46 (13.78)	-0.91 (3.91)	-1.35 (3.61)	-1.06 (2.45)	-1.40 (5.58)	3.17 (3.8)
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-
Access Time	ACTIMEPT	Public Transport	-0.04 (7.25)	-	-0.11 (6.18)	-0.10 (6.95)	-	-0.05 (7.99)
Fare Public Transport	COSTPT	Public Transport	-0.38 (20.80)	-0.12 (5.90)	-0.12 (5.18)	-0.08 (5.85)	-0.06 (4.02)	-0.60 (14.89)
Fuel + Toll Cost Car	COSTCRTRC	Car	-	-0.23 (7.67)	-	-0.04 (2.90)	-0.11 (4.45)	-0.09 (2.38)
Parking Cost Car	COSTCRPC	Car	-0.08 (3.12)	-	-0.10 (11.15)	-0.03 (6.61)	-	-0.22 (4.77)
Travel Time Public Transport	TTPT	Public Transport	-0.06 (22.14)	-0.04 (12.08)	-0.02 (6.69)	-0.02 (5.55)	-0.01 (2.92)	-0.08 (23.5)
Travel Time Car	TTCR	Car	-0.04 (5.03)	-	-	-	-0.01 (2.93)	-0.09 (5.02)
Egress Time	EGTIME	All Alternatives	-0.03 (7.66)	-0.05 (7.40)	-	-	-0.01 (2.93)	-0.05 (7.14)
Headway Public Transport	FREQPT	Public Transport	-	-0.05 (2.76)	-	-	-0.03 (3.70)	-
Experience Bus	EXPBS	Bus	-	-	-	-	-	0.18 (3.57)
Experience Busway	EXPBW	Busway	-	-	-	-	-	-
Experience Train	EXPTR	Train	-	-	-	-	-	0.19 (6.02)
Experience Car	EXPCR	Car	-	-	-	0.72 (5.81)	-	-0.28 (2.52)
Risk Attitudes Travel Time Currently Available Modal Facilities	ALPHAEXTT	Bus, Busway, Train and Car	-	-	-	-	-	-
Risk Attitudes Travel Time New Modal Investments	ALPHANEXTT	New Light Rail, New Heavy Rail, New Busway	-	-	-	-	-	-
Risk Attitudes Travel Time Car	ALPHACRTT	Car	-	-	-	-	-	-
Risk Attitudes Cost Currently Available Modal Facilities	ALPHAEXCS	Bus, Busway, Train and Car	-	-	-	-	-	0.11 (3.52)
Risk Attitudes Cost New Modal Investments	ALPHANEXCS	New Light Rail, New Heavy Rail, New Busway	-	-	-	-	-	-
Risk Attitudes Fuel+Toll Car	ALPHACRTRCS	Car	-	-	-	-	-	-
Risk Attitudes Parking Car	ALPHACRPCS	Car	-	-	-	-	-	-
Concavity VL	CONC	All Alternatives	-	-	-	-	1.38 (4.56)	-

The parking cost was not significant for the VL heuristic, and the headway in the LPAA heuristic. The only behavioural refinement that appeared to be statistically significant in the LPAA class is a risk aversion attitude towards the cost of the currently available modal facilities (bus, busway and train). The graphical representation of this parameter is shown in Figure 6-4, exhibiting a concavity function which represents the risk aversion attitude. Experience in travelling by car was significant in the RAM heuristic with a positive sign, stating that individuals are more likely to choose the car if they used it in their recent trip. In the LPAA class, experience in bus, train and car was statistically significant. The parameter for experience in bus and train was positive, having an equivalent meaning as the car experience in the RAM heuristic. However, the car experience in the LPAA heuristic had a negative parameter, suggesting that individuals are less likely to choose the car if they used it in their recent trip. The PDP approach is different to the other types of models in the way that each heuristics is considered separately, and hence experience associated with each mode is included separately in each class. In this case, experience towards the car was statistically significant in the RAM heuristic – with a positive estimate – and in the LPAA heuristic – with a negative estimate. Even though the negative estimate is counterintuitive and it contradicts the results for the models that considered only an LPAA heuristic, the interpretation of this estimate should not be provided for each class separately but by considering all of them. If an individual used the car in his most recent trip, he will have a higher probability of choosing the car when using a RAM heuristic (class 1) and lower probability of choosing the car when using a LPAA heuristic (class 3).

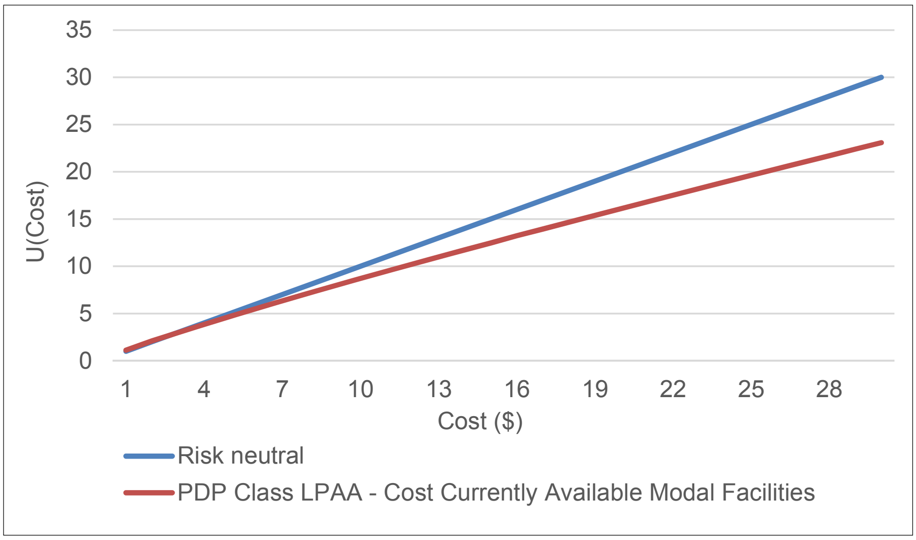


Figure 6-4: Risk attitudes towards parking cost in PDP_BRExp for the LPAA class

Similarly to what was found in the Metro Rail dataset, the inclusion of behavioural refinements and overt experience resulted in a major shift of class memberships, with a significant increase in the use of the LPAA heuristic (class 3). This finding emphasises the importance of including behavioural refinements which appear to have a statistically significant influence on preferences. Moreover, both models had some attributes that were found to not be statistically significant in a number of the heuristics. Figure 6-5 presents the concavity factor transformation for the VL heuristic (class 2). Since the concavity parameter is larger than 1, it suggests that individuals tend to over-weight larger differences.

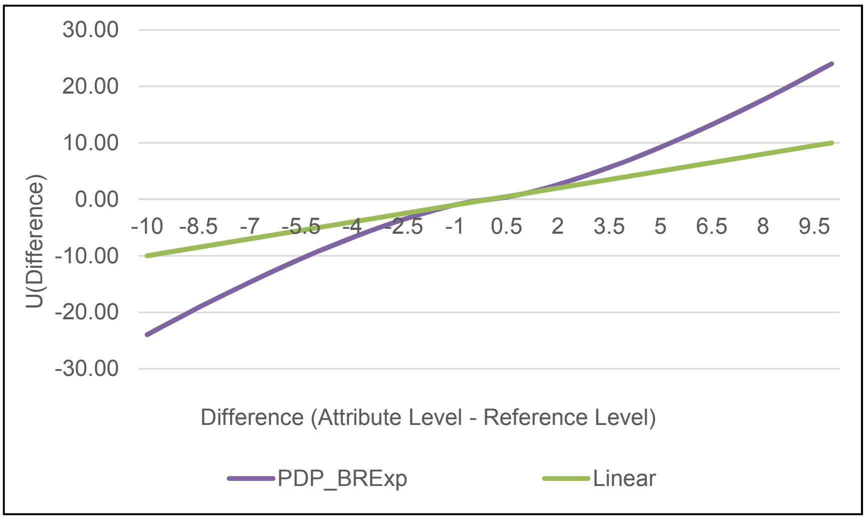


Figure 6-5: Concavity factor in PDP models

These models are not nested, since some of the parameters that are estimated in the PDP model are not considered in the PDP_BRExp model and vice versa. Such models can be compared using the Vuong test, with the results are shown in Table 6-9. The incorporation of experience and behavioural refinements significantly improves the overall performance of the models. Therefore, the preferred model is the PDP_BRExp.

Table 6-9: Vuong statistic test results for the PDP models

	PDP_BRExp vs. PDP
Mean	0.0123
Std Dev	0.0876
Sample Size	4530
Vuong Statistic	9.417
Result	Favours PDP_BRExp model

6.4. Conditioning Random Process Heterogeneity Model Results: LPAA, VL and RAM together with behavioural refinements and experience

Four models associated with conditioning random process heterogeneity (CRPH) models are presented in this section (for more details refer to Section 3.5.4) that allow for: (1) interactions between the process strategies and the mean estimates, referred to as CRPHm; (2) interactions between the process strategies and the standard deviation estimates, denoted as CRPHs; (3) interactions both with the mean and standard deviation estimates, referred to as CRPHms; and (4) and the same as CRPHms but with behavioural refinements (i.e., risk attitudes and perceptual conditioning) and experience (as defined in Section 3.5.1).

As explained in Section 3.5.4 and again in Section 5.4, the utility expression for this model can be written as follows:

$$U_i = \sum_n \left(\begin{bmatrix} \theta_{in} + \lambda^m_{VL,in} \cdot VL\left(x_{inqt}\right) + \lambda^m_{RAM,in} \cdot RAM\left(x_{inqt}\right) \\ + \left[\sigma_{in} + \lambda^s_{VL,in} \cdot VL\left(x_{inqt}\right) + \lambda^s_{RAM,in} \cdot RAM\left(x_{inqt}\right) \right] \cdot v \end{bmatrix} \cdot x_{inqt} \right) + \varepsilon_{iqt} \qquad (6.1)$$

Table 6-10 and 6-11 summarise the parameter estimates for the different types of models.

Table 6-10: CRPH models with interactions in the mean or standard deviation estimates (t-values in brackets)

			CRPHm				CRPHs			
Number of Parameters Estimated			**28**				**25**			
Log Likelihood at convergence			**-4,581.12**				**-4,619.24**			
Log likelihood at zero			**-7,838.25**							
AIC			**2.035**				**2.050**			
Parameters		**Alternatives**	θ	σ	λ_{VL}^{m}	λ_{RAM}^{m}	θ	σ	λ_{VL}^{s}	λ_{RAM}^{s}
Alternative Specific Constant New Light Rail	ASCLR	New Light Rail	2.35 (3.62)	-	-	-	3.28 (7.31)	-	-	-
Alternative Specific Constant New Heavy Rail	ASCNHR	New Heavy Rail	2.45 (3.82)	-	-	-	3.57 (8.34)	-	-	-
Alternative Specific Constant New Busway	ASCNBW	New Busway	1.22 (1.88)	-	-	-	2.34 (5.25)	-	-	-
Alternative Specific Constant Bus	ASCBS	Bus	1.92 (2.99)	-	-	-	3.01 (6.97)	-	-	-
Alternative Specific Constant Busway	ASCBW	Busway	1.71 (2.67)	-	-	-	2.78 (6.47)	-	-	-
Alternative Specific Constant Train	ASCTRAIN	Train	1.42 (2.24)	-	-	-	2.48 (5.80)	-	-	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-	-	-
Access Time	ACTIMEPT	Public Transport	-	-0.11 (12.20)	-0.003 (11.12)	-	-0.07 (11.44)	-0.18 (14.20)	0.003 (6.41)	-
Fare Public Transport	COSTPT	Public Transport	-0.62 (14.90)	0.39 (15.09)	0.01 (2.51)	-	-0.57 (20.16)	0.44 (20.33)	-	-
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.35 (6.26)	-0.32 (10.96)	-	-	-0.17 (4.10)	-	-0.01 (2.03)	-
Parking Cost Car	COSTCRPC	Car	-0.10 (4.41)	-0.32 (12.57)	-	-	-0.17 (9.79)	0.36 (17.46)	-	-
Travel Time Public Transport	TTPT	Public Transport	-0.03 (3.74)	-0.06 (18.13)	-0.001 (10.14)	-0.01 (5.27)	-0.07 (22.16)	-0.05 (20.65)	-	-
Travel Time Car	TTCR	Car	-0.13 (4.31)	0.08 (15.21)	-0.001 (5.73)	-0.01 (7.81)	-0.11 (12.42)	-0.14 (15.96)	-	-
Egress Time	EGTIME	All Alternatives	-0.08 (8.92)	0.08 (9.68)	-	-	-0.04 (7.70)	0.06 (8.20)	-	-
Headway Public Transport	FREQPT	Public Transport	-	-0.21 (10.91)	-0.01 (4.65)	-	-0.04 (2.75)	-0.38 (9.87)	0.01 (7.76)	0.03 (2.51)
Concavity VL	CONC	All Alternatives	-	-	-	-	-	-	-	-

Table 6-11: CRPH models with interactions in the mean and standard deviation estimates (t-values in brackets)

			CRPHms						**CRPHms_BRExp**					
Number of Parameters Estimated			**29**						**32**					
Log Likelihood at convergence			**-4,553.62**						**-4,494.74**					
Log likelihood at zero			**-7,838.25**											
AIC			**2.023**						**1.999**					
Parameters		**Alternatives**	θ	σ	λ_{VL}^{m}	λ_{RAM}^{m}	λ_{VL}^{s}	λ_{RAM}^{s}	θ	σ	λ_{VL}^{m}	λ_{RAM}^{m}	λ_{VL}^{s}	λ_{RAM}^{s}
Alternative Specific Constant New Light Rail	ASCLR	New Light Rail	2.64 (4.17)	-	-	-	-	-	3.89 (7.69)	-	-	-	-	-
Alternative Specific Constant New Heavy Rail	ASCNHR	New Heavy Rail	3.36 (5.42)	-	-	-	-	-	4.52 (9.35)	-	-	-	-	-
Alternative Specific Constant New Busway	ASCNBW	New Busway	1.79 (2.85)	-	-	-	-	-	2.96 (6.05)	-	-	-	-	-
Alternative Specific Constant Bus	ASCBS	Bus	2.47 (4.03)	-	-	-	-	-	3.62 (7.65)	-	-	-	-	-
Alternative Specific Constant Busway	ASCBW	Busway	2.22 (3.62)	-	-	-	-	-	3.46 (7.24)	-	-	-	-	-
Alternative Specific Constant Train	ASCTRAIN	Train	1.95 (3.17)	-	-	-	-	-	2.70 (5.59)	-	-	-	-	-
Alternative Specific Constant Car	ASCCAR	Car	-	-	-	-	-	-	-	-	-	-	-	-
Access Time	ACTIMEPT	Public Transport	-	-0.12 (12.84)	0.00 (11.93)	-	-	-	-0.03 (8.96)	-	0.00 (7.85)	-	-	0.05 (14.98)
Fare Public Transport	COSTPT	Public Transport	-0.74 (16.35)	0.59 (18.33)	-	-	-0.03 (17.39)	0.04 (17.20)	-0.64 (17.53)	0.47 (15.64)	-	-	-0.04 (21.39)	0.04 (18.63)
Fuel + Toll Cost Car	COSTCRTRC	Car	-0.30 (4.92)	-0.33 (8.40)	-	-	-	-	-0.14 (6.11)	-	-	-	-	0.14 (8.25)
Parking Cost Car	COSTCRPC	Car	-0.12 (6.08)	-0.09 (4.08)	-	-	-	-	-0.21 (10.02)	-0.24 (13.90)	-	-	-	-
Travel Time Public Transport	TTPT	Public Transport	-0.05 (5.69)	-0.06 (15.33)	0.00 (7.14)	-	-0.01 (3.15)	-	-0.07 (12.62)	-0.08 (16.84)	-	0.00 (2.65)	0.00 (2.28)	-
Travel Time Car	TTCR	Car	-0.10 (3.32)	0.06 (12.84)	0.00 (5.85)	-	-0.01 (6.06)	-	-0.12 (4.78)	0.08 (11.65)	0.00 (2.84)	-	-0.01 (5.04)	-
Egress Time	EGTIME	All Alternatives	-0.08 (8.94)	-	-	-	-	-	-0.06 (8.80)	0.06 (9.05)	-	-	-	-
Headway Public Transport	FREQPT	Public Transport	-0.04 (2.42)	-0.32 (9.83)	-	0.01 (5.21)	-	-	-0.04 (2.68)	-0.34 (13.38)	-	0.01 (6.52)	-	-
Experience Bus	EXPBS	Bus	-	-	-	-	-	-	-	-	-	-	-	-
Experience Busway	EXPBW	Busway	-	-	-	-	-	-	-	-	-	-	-	-
Experience Train	EXPTR	Train	-	-	-	-	-	-	0.29 (9.11)	-	-	-	-	-
Experience Car	EXPCR	Car	-	-	-	-	-	-	0.19 (3.08)	-	-	-	-	-
Concavity VL	CONC	All Alternatives	-	-	-	-	-	-	-	-	-	-	-	-

The results suggest that none of the behavioural refinements appear to be statistically significant when allowing for process heterogeneity in this form. Different combinations were tested, but none were statistically significant. Overt experience was statistically significant when conditioning the entire utility expression. Model CRPHms_BRExp included experience using the train and car with all of the estimates being positive. This suggests that, if an individual used any of these modes on his most recent trip, he will be more likely to choose the same mode again. The concavity factor was never statistically different in any of these models.

The interactions that were statistically significant for each model are presented in Table 6-12. A greater number of interactions were statistically significant in the CRPHms_BRExp model, followed by the CRPHms model. Model CRPHm considers five interactions between the mean and VL: access time, fare public transport, travel times (public transport and car) and headway, and two between the mean and RAM: travel times (public transport and car). The CRPHs model includes three interactions between the standard deviation and VL: access time, fuel plus toll costs and headway; and only one interaction between the standard deviation and the RAM heuristic in the headway.

Model CRPHms includes three interactions between the VL process strategy and the mean estimate: access time and travel time in public transport and car; and one interaction between the RAM heuristic and the mean estimate: headway. Three interactions are statistically significant between the VL heuristic and the standard deviation: public transport fare, travel time in public transport and car; and one interaction between the RAM heuristic and the standard deviation: public transport fare. Model CRPHms_BRExp includes the same interactions except for the VL with the mean public transport travel time but adds in: an interaction between the mean and RAM in public transport travel time, and an interaction between the RAM heuristic and standard deviation in the access time and fuel plus toll costs. The findings suggest that the travel times are mainly influenced by the VL heuristic, although the mean of the public transport travel time is also influenced by the RAM heuristic. The fares/costs are influenced in their standard deviation mostly by the RAM heuristic, although the public transport fare standard deviation is also influenced by VL.

Table 6-12: CRPH models interactions with process strategies

	CRPHm		CRPHs		CRPHms				CRPHms_BRExp			
	Mean		Std Dev		Mean		Std Dev		Mean		Std Dev	
	VL	RAM	VL	RAM	VL	RAM	VL	RAM	VL	RAM	VL	RAM
Access Time	X		X		X				X			X
Fare Public Transport	X						X	X			X	X
Fuel + Toll Cost Car			X									X
Parking Cost Car												
Travel Time Public Transport	X	X			X		X			X	X	
Travel Time Car	X	X			X		X		X		X	
Egress Time												
Headway Public Transport	X		X	X		X				X		
Total # of Interactions	**7**		**4**		**8**				**11**			

These models are not nested models given different interactions, and are compared using the Vuong statistic test. The results in Table 6-13 shows the results, which suggest that model CRPHms is not equivalent to model CRPHs at the 99% confidence level, and it is not equivalent to model CRPHm at the 80% confidence level. Hence, including interactions with process strategies both in the mean and standard deviation estimates significantly improves the overall model fit. Comparing model CRPHms_BRExp with CRPHms, the Vuong statistic favours model CRPHms at the 99% confidence level. This is interesting as it suggests that adding experience did not improve the overall model fit. However, the Vuong test only considers the mean parameters to estimate the individual contribution to the log-likelihood. Therefore, in this section two models will be the preferred ones: the CRPHms which the Vuong test favours, and the CRPHms_BRExp which has a better AIC indicator.

Table 6-13: Vuong statistic test results for CRPH models

	CRPHms vs. CRPHm	CRPHms vs. CRPHs	CRPHms_BRExp vs. CRPHms
Mean	0.021	0.102	-0.110
Std Dev	0.769	0.990	0.717
Sample Size	4530	4530	4530
Vuong Statistic	1.810	6.915	-10.357
Result	Favours CRPHms model with a 80% confidence level	Favours CRPHms model	Favours CRPHms model

6.5. Comparison of the Models

In this section the models will be compared through general indicators, such as the log likelihood and AIC, and more importantly, the willingness to pay (WTP) estimates. This section is equivalent to Section 5.5 but with the Northwest dataset. Models LPAA_MNL and LPAA_MML will be included for comparison even though they were not the preferred model of the LPAA section, but they are an excellent reference point given their dominant use in choice studies.

6.5.1. Behavioural Refinements, Experience and Concavity Factor

All of the preferred models included behavioural refinements and/or overt experience but the parameters and interpretations are different for each model. In the process homogeneity models, experience was a significant positive influence for all the modes (bus, busway, train and car) except for the LPAA_MML_BRExp that considered experience significant in favour of the train and car only. The majority of the experience parameters estimated were positive, suggesting individuals are more likely to choose the mode they used in their recent trip. The only one that was negative was experience towards the car when considering individuals use the VL heuristic as their only process strategy. As the VL heuristic takes into account reference levels with starting values equal to the characteristics of the mode individuals used on their most recent trip (i.e., experience), it could be expected that this heuristic had a different interaction with respondents' experience.

For the process heterogeneity models, when using the PDP approach, experience towards the bus, train and car are statistically significant in the LPAA class and towards the car in the RAM class. All of them were positive except for experience towards the car in the LPAA class. These results suggest that the statistically significant experience under process homogeneity is partly explaining process heterogeneity under a PDP approach. When using the CRPH approach to include process heterogeneity, experience towards the train and car were significant and both positive. These results align with the findings in the LPAA process homogeneity with random parameters (LPAA_MML_BRExp) model. As the CRPH method estimates the parameters as considered under an LPAA traditional process strategy adding interactions with alternative heuristics, it might be expected that experience influences them in a similar way.

Regarding behavioural refinements, in the LPAA process homogeneity model (LPAA_MML_BRExp) results show a risk aversion towards all public transport fares (for the currently available modes and the new modal investments). In the PDP approach, the LPAA class showed a significant risk averse attitude towards the currently available modal facilities. When using the CRPH approach, none of the risk attitudes seem to be statistically significant, which shows that this approach is taking into consideration other things not considered in the PDP approach.

The results of the influence of experience and risk attitudes in the different models suggest that they are statistically more significant when considering process homogeneity than when considering process heterogeneity. This is a crucial finding suggesting that the importance of including additional behavioural components is reduced when considering process heterogeneity.

The concavity factor was significant in only three estimated models: VL_MNL, VL_MNL_Exp, and PDP_BRExp. All the other models consider a linear evaluation of the difference between the attribute level and the reference level. The estimated concavity factors are shown in Figure 6-6. The results suggest that individuals tend to overweight larger differences.

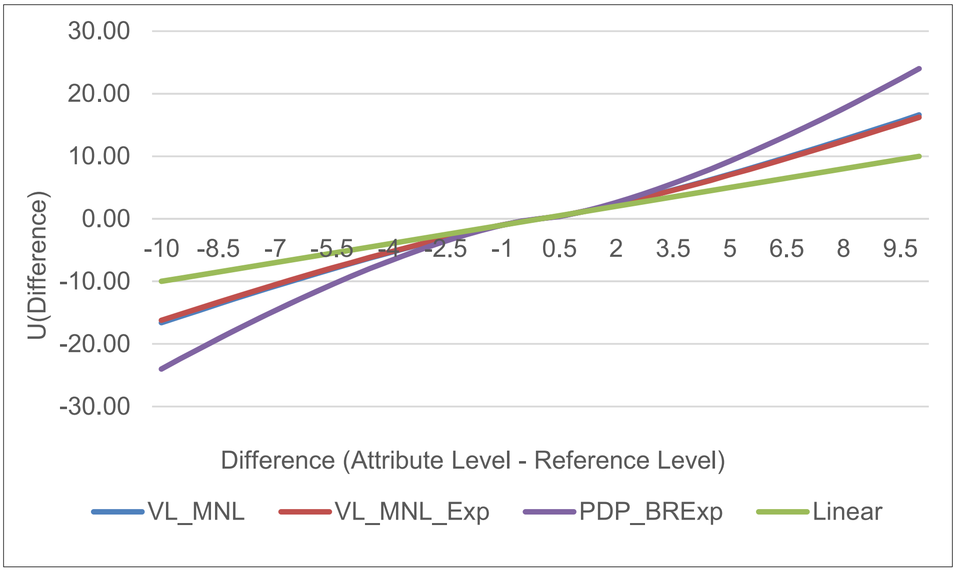


Figure 6-6: Concavity factors' comparison

6.5.2. Log likelihood and AIC

The models presented above will be compared in terms of general indicators, such as their log likelihood and AIC indicators as summarised in Table 6-14. The AIC indicator values the log likelihood while penalising the number of parameters estimated, and it is presented using a colour scale where the darker the green the better the AIC indicator and the darker the red the worse it is The models that have a superior AIC are CRPHms_BRExp and the CRPHms. These are followed by the RAM_MML_Exp and LPAA_MML_BRExp models. Model LPAA_MNL is the most inferior in terms of the AIC, followed by the RAM_MNL and then by the VL_MNL. When considering preference heterogeneity but not process heterogeneity, the RAM_MML is superior, followed by the LPAA_MML and then by the VL_MML. The fact that the relative performance of the process strategies changes when considering preference heterogeneity might be suggesting some sort of confounding between them.

Table 6-14: General indicators

	Number of Parameters Estimated	Log Likelihood at convergence	Log likelihood at zero	AIC
LPAA_MNL	14	-6,170.72	-7,838.25	2.731
LPAA_MML	22	-4,701.06		2.085
LPAA_MML_BRExp	26	-4,611.45		2.047
VL_MNL	15	-6,004.14		2.657
VL_MML	21	-4,729.70		2.097
VL_MML_Exp	25	-4,702.90		2.087
RAM_MNL	14	-6,169.89		2.730
RAM_MML	22	-4,659.04		2.067
RAM_MML_Exp	25	-4,566.35		2.027
PDP	35	-4,953.14		2.202
PDP_BRExp	44	-4,864.47		2.167
CRPHm	28	-4,619.14		2.052
CRPHs	25	-4,619.24		2.050
CRPHms	29	-4,553.62		2.023
CRPHms_BRExp	32	-4,494.74		1.999

6.5.3. Willingness to Pay Estimates

Willingness to pay estimates (WTP) are often the most important behavioural policy outcome in a choice study. The purpose of this subsection is equivalent to Section 5.5.3; the reader is referred to that section for more details on the structure. Table 6-15 and 6-16 present the WTP estimates for the models that consider process homogeneity and process heterogeneity, respectively.

Table 6-15: Willingness to pay estimates for the models with process homogeneity

		LPAA_MNL		LPAA_MML		LPAA_MML_BRExp		VL_MML		VL_MML_Exp		RAM_MML		RAM_MML_Exp	
		Median	Std Dev	Median	Std Dev	Median	Std Dev	Median	Std Dev	Median	Std Dev	Median	Std Dev	Median	Std Dev
Travel Time $/person hour	**Light Rail**	9.34	0.49	6.54	11.08	5.94	9.74	5.54	11.70	5.35	11.95	4.18	10.02	4.30	13.08
	New Heavy Rail	9.34	0.49	6.54	11.08	5.96	9.77	5.54	11.70	5.35	11.95	4.18	10.01	4.30	13.04
	New Busway	9.34	0.49	6.54	11.08	7.13	11.71	5.54	11.70	5.35	11.95	4.18	9.92	4.31	12.96
	Bus	9.34	0.49	6.54	11.08	7.10	11.66	5.54	11.70	5.35	11.95	4.17	9.94	4.30	12.98
	Busway	9.34	0.49	6.54	11.08	6.95	11.43	5.54	11.70	5.35	11.95	4.18	9.98	4.31	13.04
	Train	9.34	0.49	6.54	11.08	5.86	9.61	5.54	11.70	5.35	11.95	4.16	9.97	4.29	13.01
	Car	21.88	2.99	13.78	45.52	9.35	15.13	16.81	35.18	15.17	31.56	17.51	103.20	28.63	116.99
Access Time $/person hour	**Light Rail**	10.24	0.83	6.15	18.80	5.89	16.32	5.18	18.45	4.38	17.82	4.18	16.77	3.40	16.60
	New Heavy Rail	10.24	0.83	6.15	18.80	5.91	16.37	5.18	18.45	4.38	17.82	4.20	16.79	3.43	16.61
	New Busway	10.24	0.83	6.15	18.80	7.07	19.64	5.18	18.45	4.38	17.82	4.25	16.84	3.47	16.65
	Bus	10.24	0.83	6.15	18.80	7.04	19.54	5.18	18.45	4.38	17.82	4.23	16.77	3.46	16.60
	Busway	10.24	0.83	6.15	18.80	6.90	19.14	5.18	18.45	4.38	17.82	4.20	16.73	3.42	16.56
	Train	10.24	0.83	6.15	18.80	5.81	16.10	5.18	18.45	4.38	17.82	4.23	16.60	3.45	16.43
Egress Time $/person hour	**Light Rail**	8.58	0.78	5.32	13.00	4.71	11.24	3.27	13.74	3.08	12.96	2.42	6.64	2.67	13.00
	New Heavy Rail	8.58	0.78	5.32	13.00	4.73	11.28	3.27	13.74	3.08	12.96	2.43	6.66	2.69	13.03
	New Busway	8.58	0.78	5.32	13.00	5.66	13.52	3.27	13.74	3.08	12.96	2.45	6.65	2.70	13.02
	Bus	8.58	0.78	5.32	13.00	5.63	13.45	3.27	13.74	3.08	12.96	2.43	6.64	2.69	13.01
	Busway	8.58	0.78	5.32	13.00	5.52	13.18	3.27	13.74	3.08	12.96	2.42	6.62	2.67	12.99
	Train	8.58	0.78	5.32	13.00	4.65	11.10	3.27	13.74	3.08	12.96	2.43	6.65	2.69	13.03
	Car	27.10	2.30	8.97	36.62	6.60	14.41	9.57	39.22	8.80	33.58	6.33	27.39	13.03	56.95
Headway $/person minute	**Light Rail**	0.08	0.04	0.07	0.54	0.04	0.46	0.10	0.60	0.09	0.58	0.05	0.47	0.07	0.56
	New Heavy Rail	0.08	0.04	0.07	0.54	0.04	0.46	0.10	0.60	0.09	0.58	0.05	0.48	0.07	0.58
	New Busway	0.08	0.04	0.07	0.54	0.05	0.55	0.10	0.60	0.09	0.58	0.05	0.49	0.07	0.58
	Bus	0.08	0.04	0.07	0.54	0.05	0.55	0.10	0.60	0.09	0.58	0.05	0.48	0.07	0.58
	Busway	0.08	0.04	0.07	0.54	0.05	0.53	0.10	0.60	0.09	0.58	0.05	0.48	0.07	0.58
	Train	0.08	0.04	0.07	0.54	0.04	0.45	0.10	0.60	0.09	0.58	0.05	0.48	0.06	0.58

Table 6-16: Willingness to pay estimates for the models with process heterogeneity

		PDP		PDP_BRExp		CRPHm		CRPHs		CRPHms		CRPHms_BRExp	
		Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error	Median	Std Error
Travel Time $/person hour	**Light Rail**	9.79	0.84	9.57	1.26	8.14	12.52	5.82	8.19	7.77	11.18	7.87	10.26
	New Heavy Rail	9.81	0.84	9.36	1.17	7.84	11.85	5.84	8.19	7.39	10.60	7.45	10.33
	New Busway	9.84	0.84	9.46	1.23	7.88	14.91	5.83	8.19	14.52	15.65	13.66	13.08
	Bus	9.83	0.84	11.41	1.52	7.98	14.95	5.84	8.18	13.77	15.51	12.98	12.60
	Busway	9.88	0.84	10.86	1.29	7.28	12.88	5.84	8.19	10.91	13.31	11.03	12.71
	Train	9.80	0.84	10.75	1.34	7.61	11.08	5.84	8.18	6.76	9.82	6.65	10.23
	Car	22.39	11.06	26.23	9.57	16.09	44.05	15.55	60.45	21.26	44.95	19.39	41.88
Access Time $/person hour	**Light Rail**	7.60	1.81	7.57	1.72	3.66	18.64	5.90	18.15	4.06	17.88	4.14	19.68
	New Heavy Rail	7.59	1.81	7.56	1.72	3.70	18.02	5.90	18.08	4.01	17.19	3.93	21.26
	New Busway	7.84	1.87	7.25	1.73	3.64	21.96	5.90	18.11	7.95	26.66	8.38	20.82
	Bus	7.80	1.86	8.82	2.22	2.41	21.64	5.92	19.72	4.98	25.87	6.79	21.34
	Busway	7.69	1.83	8.74	2.14	3.83	19.99	5.91	17.92	6.74	22.90	7.07	23.01
	Train	7.59	1.81	8.95	2.07	8.95	18.61	6.00	11.83	9.02	17.35	5.83	19.88
Egress Time $/person hour	**Light Rail**	7.05	1.15	5.94	1.58	5.21	13.73	3.55	7.96	5.53	13.15	5.99	9.01
	New Heavy Rail	7.06	1.15	5.88	1.55	5.38	13.31	3.54	7.95	5.50	12.65	5.82	8.97
	New Busway	7.22	1.18	5.61	1.38	4.86	16.11	3.53	7.95	10.15	19.23	9.69	12.07
	Bus	7.19	1.18	6.53	1.50	4.87	16.07	3.53	7.95	9.52	18.85	9.29	11.59
	Busway	7.11	1.16	6.78	1.63	5.38	14.78	3.53	7.95	8.67	16.67	8.41	11.13
	Train	7.05	1.15	6.94	1.81	5.55	12.70	3.53	7.94	5.28	11.86	5.40	8.79
	Car	33.91	11.89	13.23	4.31	8.71	35.92	6.31	26.13	14.38	49.37	11.71	31.65
Headway $/person minute	**Light Rail**	0.05	0.03	0.10	0.04	0.15	0.60	0.06	0.17	0.06	0.25	0.07	0.22
	New Heavy Rail	0.05	0.03	0.05	0.02	0.04	0.56	0.06	0.47	0.05	0.60	0.07	0.57
	New Busway	0.05	0.03	0.05	0.02	0.04	0.68	0.06	0.53	0.10	0.93	0.11	0.81
	Bus	0.05	0.03	0.02	0.01	0.02	0.68	0.06	0.58	0.10	1.01	0.11	0.86
	Busway	0.05	0.03	0.05	0.02	0.04	0.62	0.06	0.52	0.09	0.83	0.10	0.77
	Train	0.05	0.03	0.04	0.02	0.03	0.54	0.06	0.50	0.05	0.60	0.06	0.59

The following subsections will analyse the median, standard error, and stability[13] of the WTP estimates. We begin by focussing on the value of travel time savings (VTTS) for public transport and car, given it is considered one of the most important metrics in user time benefit calculations, with the following section summarising WTP estimates for other attributes.

6.5.3.1. Value of Travel Time Savings

6.5.3.1.1. Process Homogeneity using Different Heuristics

Figure 6-7 graphically presents the VTTS median and the 95% confidence intervals for public transport and car for the three models that represent process homogeneity: LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp.

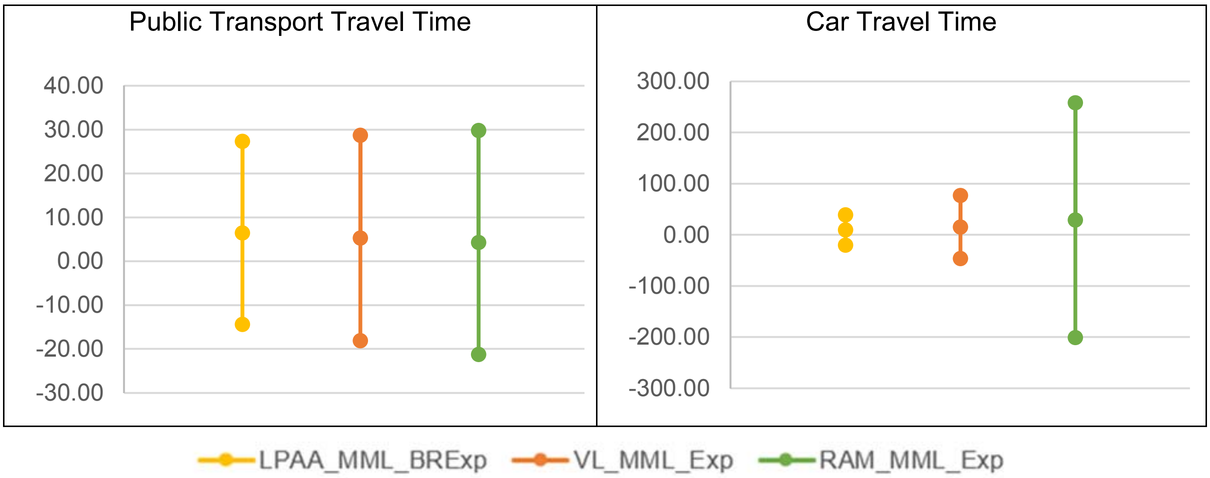


Figure 6-7: VTTS median and the 95% confidence intervals for the models with process homogeneity

The median VTTS for public transport is larger for the LPAA model ($6.54 per person hour), followed by the VL model ($5.35) and the lowest value is for the RAM model ($4.30). The standard error is largest for the RAM heuristic ($13.02) with the LPAA heuristic having the smallest standard error ($10.65). The median VTTS when using the car is significantly smaller for the LPAA model ($9.35 per person hour), followed by the VL model ($16.81) and the largest value if for the RAM model ($28.63). The standard error follows the same order as the median,

[13] The reader is referred to the introduction of Section 5.5.3 for the definition of stability

which is lower for the LPAA model ($15.13) and significantly larger for the RAM model ($116.99).

For each mode of transportation, the VTTS were compared across the process homogeneity models to see if they were statistically different from each other using the median and standard errors with a t-test (Section 3.8.1), and the results are presented in Table 6-17. An absolute value larger than 1.96 represents statistically significant different estimates at the 95% confidence level. The results show that all the estimates are significantly different from each other, revealing the impact of considering different process strategies in the VTTS estimates.

Table 6-17: Comparison of median VTTS for models with process homogeneity using t-test

Travel Time	VL_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs VL_MML_BRExp
New Light Rail	***-2.31***	***-6.07***	***-3.57***
New Heavy Rail	***-2.25***	***-5.80***	***-3.38***
New Busway	***-4.96***	***-7.51***	***-2.74***
Bus	***-7.05***	***-10.79***	***-4.00***
Busway	***-6.27***	***-9.84***	***-3.79***
Train	***-2.14***	***-6.26***	***-3.87***
Car	***10.39***	***10.21***	***6.94***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-8 shows the distribution of the VTTS obtained from 25,000 different draws representing their stability. The results show that the stability of the VTTS for public transport is similar when using different process strategies, but it is slightly worse in the RAM model and better in the LPAA model. The car VTTS' stability is significantly better in the LPAA model followed by the VL model, and it is significantly worse in the RAM model.

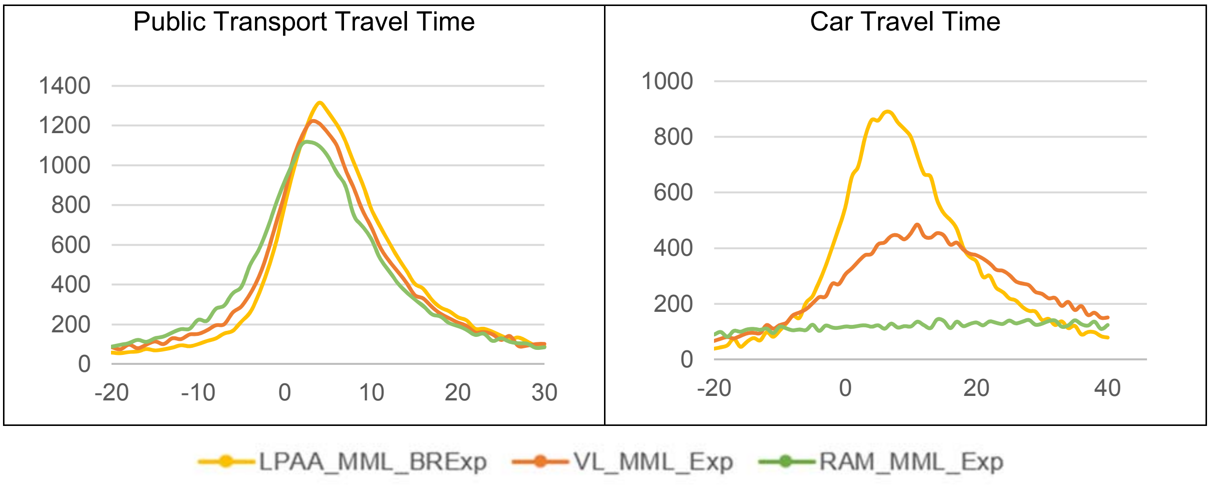


Figure 6-8: VTTS distribution for different draws using the models with process homogeneity

The results of this section suggest that the estimates for the car and public transport VTTS for models that use different process strategies are considerably different in terms of their median and standard errors. The stability of the VTTS in public transport is relatively similar for the different models but significantly different for the car VTTS. The results suggest that the process strategy considered has an important influence on the estimates and conclusions.

Summary: Comparison of the preferred process homogeneity models LPAA_MML_BRExp; RAM_MML_Exp and VL_MML_Exp

Public transport VTTS

- median (lower to higher): RAM, VL, LPAA
- standard error (lower to higher): LPAA, VL, RAM

Car VTTS

- median (lower to higher): LPAA, VL, RAM
- standard error (lower to higher): LPAA, VL, RAM

Significant differences? All the estimates (public transport and car) for the different models are significantly different from each other.

Stability of the VTTS estimates? Similar stability for the public transport VTTS. The car VTTS stability is better in the LPAA model and worse in the RAM model.

Conclusion: The use of LPAA, RAM or VL as the sole process strategy being used by individuals has a significant influence over the VTTS estimates.

6.5.3.1.2. Probabilistic Decision Process versus Conditioning Random Parameter Heterogeneity

Figure 6-9 compares the process heterogeneity models using the PDP and CRPH approaches. The median VTTS for public transport is higher for the PPD_BRExp model

($10.24 per person hour), followed by the CRPHms model ($10.19), the CRPHms_BRExp model ($9.84), the PDP model ($9.82), the CRPHm model ($7.79) and finally the CRPHs model ($5.83 per person hour). There are significant changes in the estimate when including process heterogeneity using the different models. For the car VTTS, the lowest value is for CRPHs ($15.55 per person hour), followed by: the CRPHm ($16.09 per person hour), the CRPHms_BRExp ($19.39), CRPHms ($21.26), PDP ($22.39) and finally PDP_BRExp ($26.23 per person hour). It is important to note that the car VTTS are higher for all the PDP models relative to the CRPH models. Regarding the standard errors for the VTTS, it is always lower in the PDP approach as it considers every parameter fixed while the CRPH approach estimates random parameters.

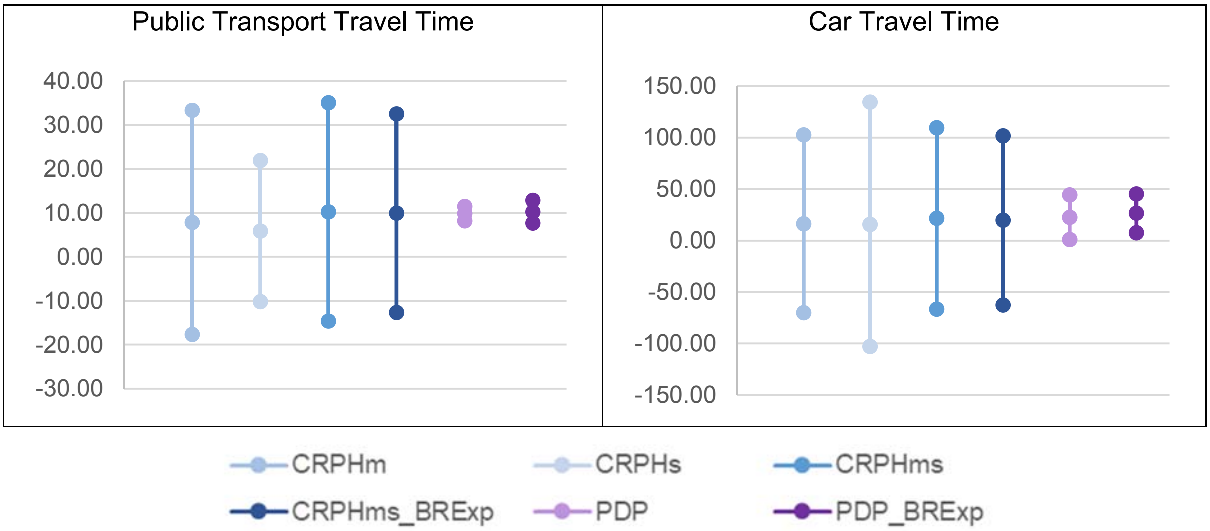


Figure 6-9: VTTS median and 95% confidence intervals for the models with process heterogeneity

Table 6-18 presents the results when comparing the VTTS for each mode using the PDP and CRPH model (without and with behavioural refinements and experience). The results show that all the estimates are statistically different (absolute value larger than 1.96 with a 95% confidence level) when considering the CRPH method relative to the PDP method, except for the car VTTS in the models without behavioural refinements and experience, and the busway VTTS for the models with behavioural refinements and experience. For the signs, the VTTS are lower under the CRPH heuristic relative to the PDP for the new light rail, new heavy rail, train and car; and are higher for the new busway, bus and busway modes. In conclusion, the results show that the VTTS for public transport and car are significantly different when considering the CRPH and PDP approaches.

Table 6-18: Comparison of median VTTS for models with process heterogeneity using t-test

Travel Time	CRPHms vs PDP	CRPHms_BRExp vs PDP_BRExp
New Light Rail	***-10.84***	***-9.88***
New Heavy Rail	***-12.99***	***-10.49***
New Busway	***13.88***	***14.83***
Bus	***17.08***	***8.34***
Busway	***4.96***	0.85
Train	***-19.91***	***-25.60***
Car	-1.52	***-9.94***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-10 shows the stability of the VTTS for the CRPH models (the PDP is not considered as it estimates fixed parameters). The stability for public transport VTTS is higher in the CRPHs model, and the other ones have a relatively similar stability. The stability for the car VTTS is similar across all the CRPH models.

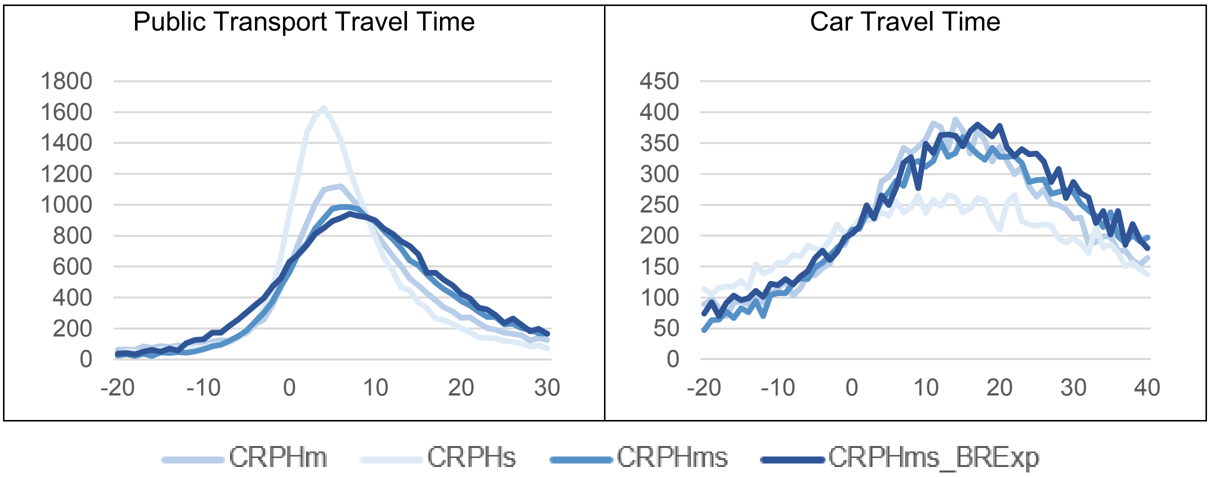


Figure 6-10: VTTS distribution for different draws using the models with process heterogeneity

Summary: Comparison of the preferred models for the two approaches to include process heterogeneity, PDP_BRExp and CRPHms_BRExp

Public transport and car VTTS

- median (lower to higher): CRPHms_BRExp, PDP_BRExp
- standard error (lower to higher): PDP_BRExp, CRPH_BRExp

Significant differences? The majority of the VTTS estimates are significantly different from each other when using the two approaches.

Stability of the VTTS estimates? PDP is completely stable since every parameter is estimated as fixed.

Conclusion: The PDP_BRExp model produces significantly larger median VTTS than the CRPH_BRExp for the public transport and car VTTS, but lower standard errors as it considers all fixed parameters.

6.5.3.1.3. Process Heterogeneity versus Process Homogeneity

Figure 6-11 presents the VTTS for the three preferred models with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp, and RAM_MML_Exp) and two models for each of the process heterogeneity methods (CRPHms and CRPHms_BRExp; PDP and PDP_BRExp). The process homogeneity models have a lower VTTS median for public transport relative to the process heterogeneity models. The lowest is for the RAM_MML_Exp model ($4.30 per person hour) and the highest for the PDP_BRExp model ($10.24). The standard error for the VTTS under the PDP approach is very low as it estimates fixed parameters. For the other models, the standard error for the public transport VTTS is lower in the LPAA model ($10.65) followed by the CRPHms_BRExp model ($11.54), and larger in the RAM model ($13.02).

The median for the car VTTS is lower in the LPAA preferred model ($9.35 per person hour), followed by the VL ($15.17) and then followed by the CRPHms_BRExp model ($19.39). The car VTTS median is larger under the PDP approach (considering both models). The standard error is lower in the PDP models, followed by the LPAA model ($10.65 per person hour), the CRPHms_BRExp ($11.54), VL model ($11.95), CRPHms model ($12.68) and RAM model ($13.02).

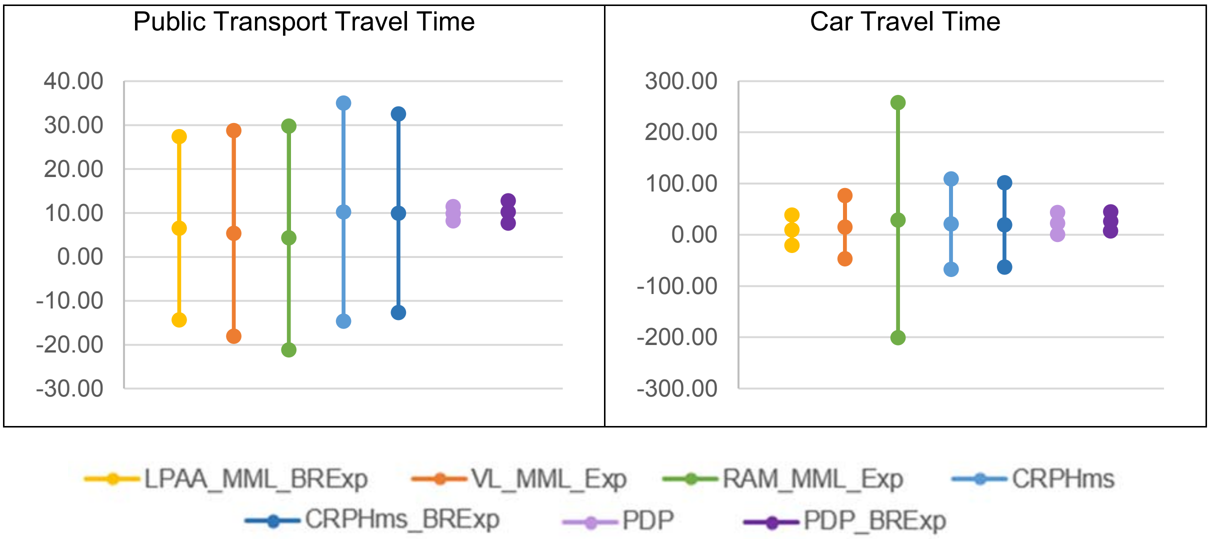


Figure 6-11: VTTS median and 95% confidence intervals for the models with process homogeneity versus process heterogeneity

Table 6-19 presents the comparison of VTTS for the preferred models with process heterogeneity and homogeneity. The first three columns compare the model CRPHms_BRExp with the three preferred models of process homogeneity and the last three columns the PDP_BRExp model with the same three models. The results show that all the VTTS are significantly different and higher in the CRPH and PDP model than in the process homogeneity models, except for the car VTTS that is lower in the CRPH and PDP models than in the RAM model. However, the car VTTS difference between the CRPH and RAM model is significant, while it is not significant between the PDP and RAM model (with 95% confidence level). In conclusion, the results show that there are significant differences in the VTTS when considering process homogeneity versus considering process heterogeneity.

Table 6-19: Comparison of median VTTS for models with process heterogeneity and process homogeneity using t-test

Travel Time	CRPHms_BRExp vs			PDP_BRExp vs		
	LPAA_MML_BRExp	VL_MML_Exp	RAM_MML_Exp	LPAA_MML_BRExp	VL_MML_Exp	RAM_MML_Exp
New Light Rail	***8.25***	***9.67***	***12.97***	***22.28***	***21.18***	***24.18***
New Heavy Rail	***6.01***	***7.61***	***10.81***	***19.77***	***19.10***	***22.06***
New Busway	***17.26***	***21.80***	***23.59***	***9.19***	***15.92***	***18.40***
Bus	***23.08***	***29.60***	***32.31***	***24.70***	***33.90***	***36.63***
Busway	***15.35***	***20.98***	***23.78***	***21.88***	***29.56***	***32.22***
Train	***3.66***	***5.35***	***9.22***	***32.51***	***28.98***	***31.87***
Car	***14.08***	***5.02***	***-4.64***	***58.88***	***20.94***	-1.28

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-12 summarises the stability findings of the VTTS for the different models. The VTTS for public transport has a relatively similar level of stability, however the stability is slightly superior in the process homogeneity models. The level of stability for the car VTTS is superior in the LPAA_MML_BRExp model and inferior in the RAM_MML_Exp model. For the rest of the models the stability is relatively similar.

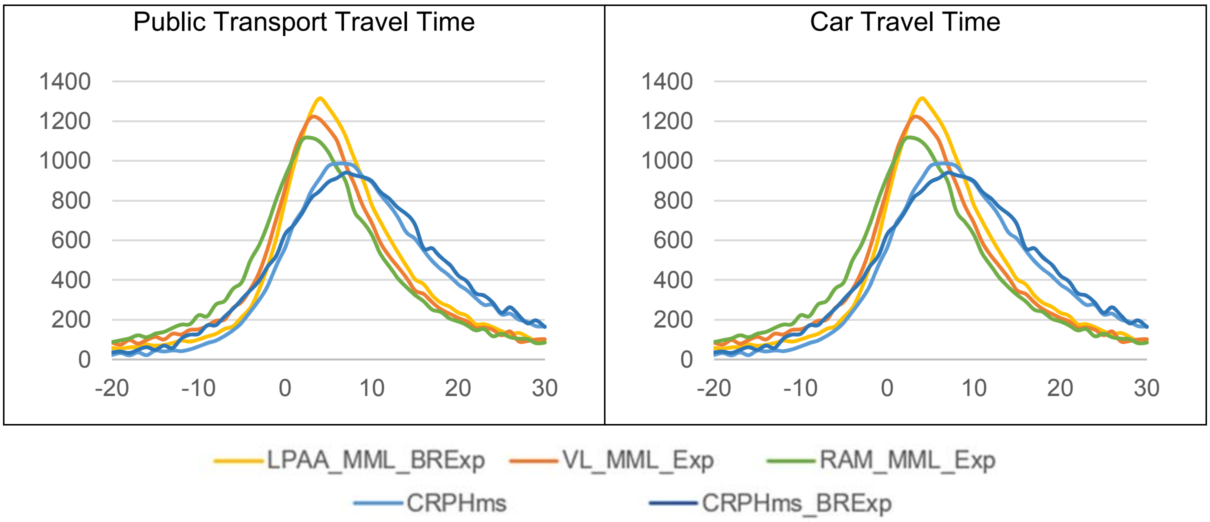


Figure 6-12: VTTS distribution for different draws using the models with process homogeneity versus process heterogeneity

Summary: Comparison of the preferred models with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp) with the preferred models with process heterogeneity (PDP_BRExp, CRPHms_BRExp)

Public transport VTTS

- median (lower to higher): process homogeneity models, process heterogeneity models
- standard error (lower to higher): PDP, LPAA, CRPH, VL, RAM

Car VTTS

- median (lower to higher): LPAA, VL, CRPH, PDP, RAM
- standard error (lower to higher): PDP, LPAA, CRPH, VL, RAM

Significant differences? The majority of the differences between the VTTS are statistically different when considering process heterogeneity instead of process homogeneity.

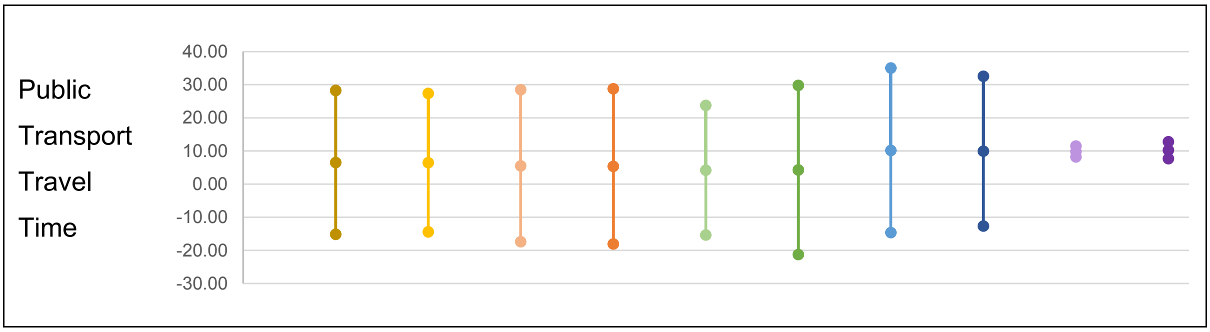
Stability of the VTTS estimates? For the public transport VTTS the process homogeneity models have a slightly superior stability. For the car VTTS the preferred LPAA has a superior stability and the preferred RAM has an inferior stability, the rest of them have a similar stability.

Conclusion: The process heterogeneity method directly influences the results. The median estimates are, in their majority, higher in the process heterogeneity models than in the process homogeneity models.

6.5.3.1.4. Behavioural Refinements and Experience

Figure 6-13 shows the median VTTS with their 95% confidence intervals for the preferred models with and without behavioural refinements and/or experience. When including behavioural refinements and/or experience for the VTTS in public transport, there is a decrease in the median value for the LPAA MML model (from $6.54 to $6.49 per person hour), in the VL MML model (from $5.54 to $5.35 per person hour), and in the CRPHms MML model (from $10.19 to $9.94 per person hour). Oppositely, there is an increase in the median value for the RAM MML model (from $4.18 to $4.30 per person hour) and the PDP MML (from $9.82 to $10.24 per person hour). The standard error VTTS increases when adding behavioural refinements and/or experience in the VL MML, RAM MML and PDP models, and decreases in the LPAA MML and CRPHms models.

When adding behavioural refinements and/or experience there is a decrease in the median car VTTS for the LPAA MML model (from $13.78 to $9.35 per person hour), the VL MML model (from $16.81 to $15.17 per person hour), and the CRPHms model (from $21.26 to $19.39 per person hour), and an increase for the PDP model (from $22.39 to $26.23 per person hour),. There is also an increase for the car VTTS median in the RAM MML model (from $17.51 to $28.63 per person hour). The standard error VTTS increases when adding behavioural refinements and/or experience in the RAM MML model, and decreases for the rest of them. What this suggests for decreases is that the marginal (dis)utility in the numerator is lower and/or the marginal dis(utility) in the denominator is higher relative to the models without these additional observed effects. The inverse is the case for an increase. The best we can claim is that the relationship between travel time and travel cost effects is noticeably changed when such behavioural refinements and overt experience are accounted for, and given the statistical significance of such effects, the implications on choosing practical estimates of VTTS is profound.



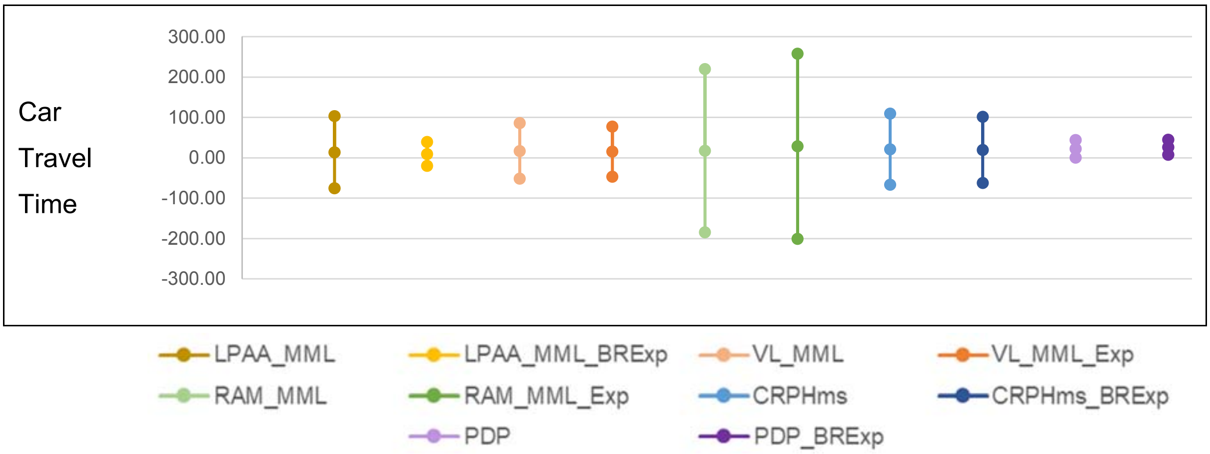


Figure 6-13: VTTS median and 95% confidence intervals for the models with and without behavioural refinements and/or experience

Table 6-20 presents the comparison of the VTTS for models with and without behavioural refinements and/or experience. As can be seen, there is no clear pattern of increase or decrease when including behavioural refinements and/or experience. In the case of the LPAA model, the inclusion of behavioural refinements and experience produces a significant reduction (estimates are significantly different) in the VTTS in the new light rail, new heavy rail, train and car and a significant increase (estimates are significantly different) for the bus. This inclusion does not have a statistically significant influence on the VTTS in the new busway and existing busway (there is not enough evidence to suggest the estimates are statistically different). In the case of the VL heuristic, the inclusion of experience produces a significantly different and lower VTTS in the car only, the other differences are not significant. In the RAM heuristic, experience produces a significantly different and higher VTTS in the car; for all the other modes the differences are not significant. In the CRPH model, the inclusion of behavioural refinements and experience produces a significantly different and lower VTTS for new busway and bus only; the differences for the rest of the modes are not statistically different from zero. In the PDP model, the results are significantly different and lower for the VTTS in the new modal investments and a significant increase in the currently available modal facilities.

Table 6-20: Comparison of median VTTS for models with and without behavioural refinements and/or experience using t-test

Travel Time	LPAA_MML_BRExp vs LPAA_MML	VL_MML_Exp vs VL_MML	RAM_MML_Exp vs RAM_MML	CRPHms_BRExp vs CRPHms	PDP_BRExp vs PDP
New Light Rail	***-2.44***	-0.70	0.43	0.40	***-8.77***
New Heavy Rail	***-2.25***	-0.66	0.42	0.22	***-17.85***
New Busway	1.72	-0.54	0.36	***-1.97***	***-11.78***
Bus	***2.35***	-0.78	0.53	***-2.64***	***61.55***
Busway	1.70	-0.75	0.48	0.41	***40.92***
Train	***-2.98***	-0.75	0.48	-0.48	***38.83***
Car	***-5.77***	***-2.17***	***4.45***	-1.90	***16.40***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-14 shows graphically the stability of the VTTS estimate for the 25,000 different draws. There does not appear to be any consistent pattern associated with the inclusion of behavioural refinements and/or experience. The stability seems to be similar for the models with and without behavioural refinements.

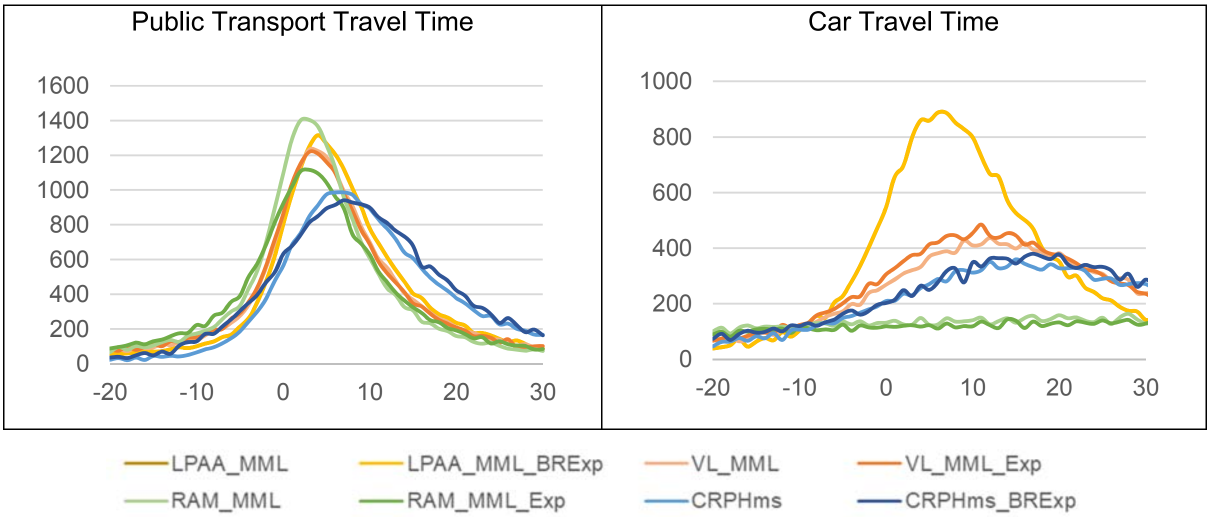


Figure 6-14: VTTS distribution for different draws using the models with and without behavioural refinements and/or experience

These section shows that the inclusion of experience and behavioural refinements has a statistically significant influence on the VTTS estimates. These differences are especially important in the PDP approach that considers all parameters as fixed, and in the LPAA model.

6.5.3.2. Other Attributes: Egress time, access times and headway

Summary: Comparison of the models with and without experience and behavioural refinements

Public transport VTTS

- median when adding experience and behavioural refinements
  - increases for RAM and PDP
  - decreases for LPAA, VL and CRPH
- standard error when adding experience and behavioural refinements
  - increases for VL, RAM and PDP
  - decreases for LPAA and CRPH

Car VTTS

- median when adding experience and behavioural refinements
  - increases for RAM
  - decreases for LPAA, VL, PDP and CRPH
- standard error when adding experience and behavioural refinements
  - increases for RAM
  - decreases for LPAA, VL, PDP and CRPH

Significant differences? The majority of the differences between the VTTS for the models with and without behavioural refinements and experience are significant in the LPAA and PDP model; in the VL and RAM model only the differences in the car are significant; and in the CRPH only for the new busway and bus.

Stability of the VTTS estimates? Similar stability with and without experience and behavioural refinements.

Conclusion: The inclusion of experience and behavioural refinements significantly influences the VTTS estimates although there is no clear patter if they produce an increase or decrease in the median and standard errors.

6.5.3.2.1. Process Homogeneity using Different Heuristics

Figure 6-15 shows the WTP median and 95% confidence intervals of the other attributes for three models that represent process homogeneity: LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp. The median WTP for access time and egress time in public transport are lower in the RAM model ($3.44 and $2.68 per person hour, respectively), followed by the VL model ($4.38 and $3.08, respectively), and higher for the LPAA model ($6.43 and $5.15, respectively). The median WTP for car egress time is lower in the LPAA model ($6.60 per person hour) and higher in the RAM model ($13.03). In the case of the headway, the LPAA model estimates the lowest median WTP ($0.05 per person minute) and the VL model estimates the largest ($0.09 per person minute).

For the standard errors, the LPAA model produces the lowest standard errors for the WTP associated with public transport and car egress times and headway. The RAM model has the lowest standard error of the WTP for the access time. Contrarily, the RAM heuristic estimates the highest standard error of the WTP for the public transport and car egress times and headway, and the LPAA heuristic for the access time.

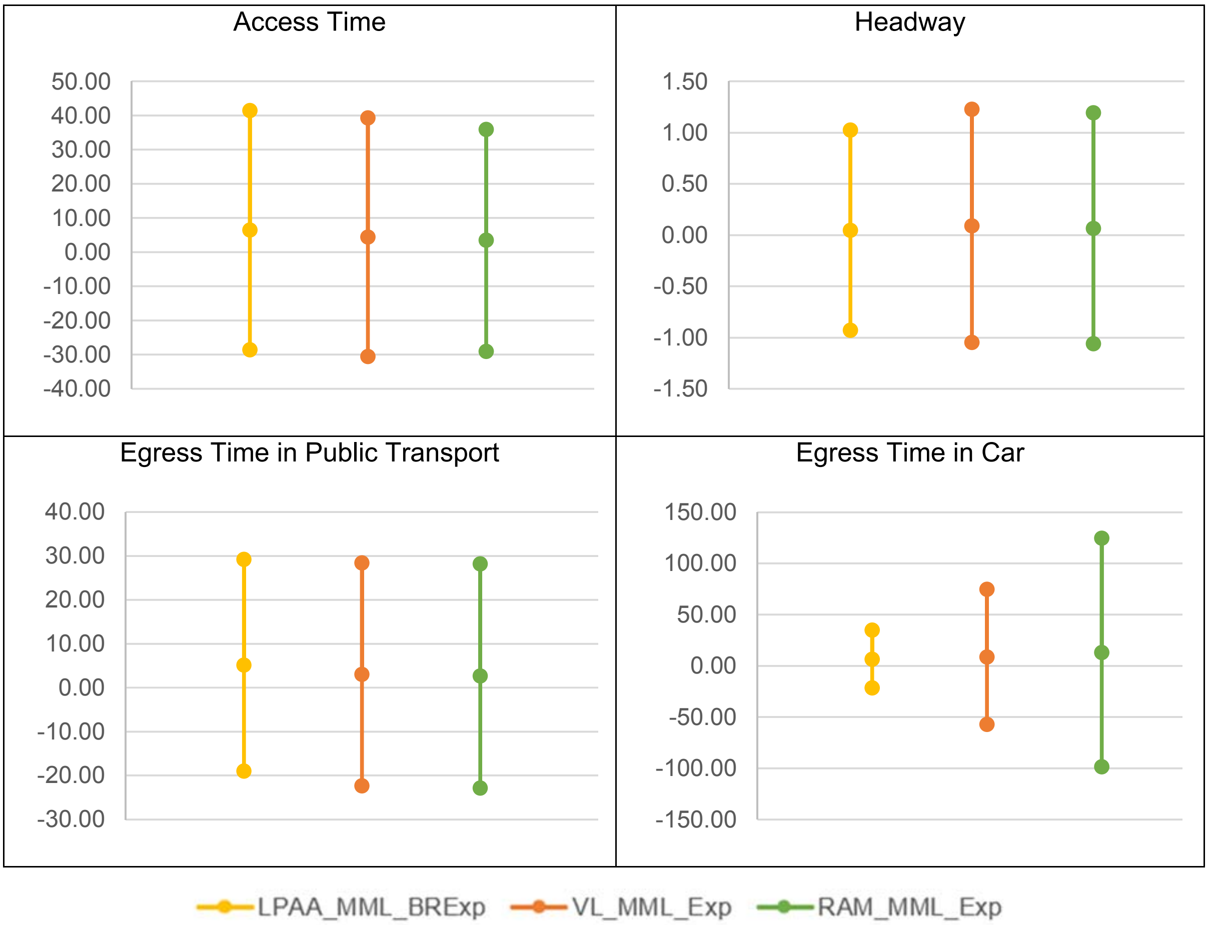


Figure 6-15: Attributes WTP median and 95% confidence intervals for the models with process homogeneity

For each mode of transport, the WTP of the different models was compared to see if they were statistically different from each other, and the results are presented in Table 6-21. The WTP estimates for the access time and public transport egress times are significantly different and lower in the LPAA preferred model than in the VL and RAM models. The WTP for the car egress time is significantly different and higher in the LPAA model than in the other two (VL and RAM). The WTP for the headway is significantly different and higher in the LPAA model than in the VL, and it is significantly different and higher than the RAM model only for the train (the others are not statistically different). The results show that the RAM and VL models estimate significantly different WTP for the access time (for all the modes except the new busway), car egress time, and headway for the bus and train with 95% confidence level. All of them are lower in the RAM model except for the car egress time which is lower in the VL model.

Table 6-21: Comparison of attributes' median WTP estimates for models with process homogeneity using t-test

Attribute	Mode	VL_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs LPAA_MML_BRExp	RAM_MML_Exp vs VL_MML_BRExp
Access Time	New Light Rail	⇩ ***-3.76***	⇩ ***-6.44***	⇩ ***-2.42***
	New Heavy Rail	⇩ ***-3.61***	⇩ ***-6.08***	⇩ ***-2.24***
	New Busway	⇩ ***-4.72***	⇩ ***-6.51***	⇩ -1.74
	Bus	⇩ ***-6.76***	⇩ ***-9.40***	⇩ ***-2.55***
	Busway	⇩ ***-6.20***	⇩ ***-8.84***	⇩ ***-2.53***
	Train	⇩ ***-3.83***	⇩ ***-6.60***	⇩ ***-2.47***
Egress Time	New Light Rail	⇩ ***-5.74***	⇩ ***-7.15***	⇩ -1.34
	New Heavy Rail	⇩ ***-5.48***	⇩ ***-6.76***	⇩ -1.22
	New Busway	⇩ ***-6.41***	⇩ ***-7.33***	⇩ -0.96
	Bus	⇩ ***-9.20***	⇩ ***-10.59***	⇩ -1.43
	Busway	⇩ ***-8.51***	⇩ ***-9.91***	⇩ -1.43
	Train	⇩ ***-5.93***	⇩ ***-7.37***	⇩ -1.36
	Car	⇧ ***3.76***	⇧ ***6.83***	⇧ ***4.00***
Headway	New Light Rail	⇧ ***3.84***	⇧ 1.86	⇩ -1.85
	New Heavy Rail	⇧ ***3.64***	⇧ 1.73	⇩ -1.72
	New Busway	⇧ ***2.24***	⇧ 0.85	⇩ -1.36
	Bus	⇧ ***3.27***	⇧ 1.22	⇩ ***-2.01***
	Busway	⇧ ***3.25***	⇧ 1.24	⇩ -1.94
	Train	⇧ ***4.18***	⇧ ***2.00***	⇩ ***-1.96***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-16 shows the distribution of the WTP for the all attributes (except for the travel time that was analysed previously) using 25,000 different draws for the models that represent process homogeneity, which is required to analyse the stability of the WTP. The stability for all the attributes is relatively similar across the models, except for the headway where the LPAA model has a superior stability and the RAM model an inferior stability.

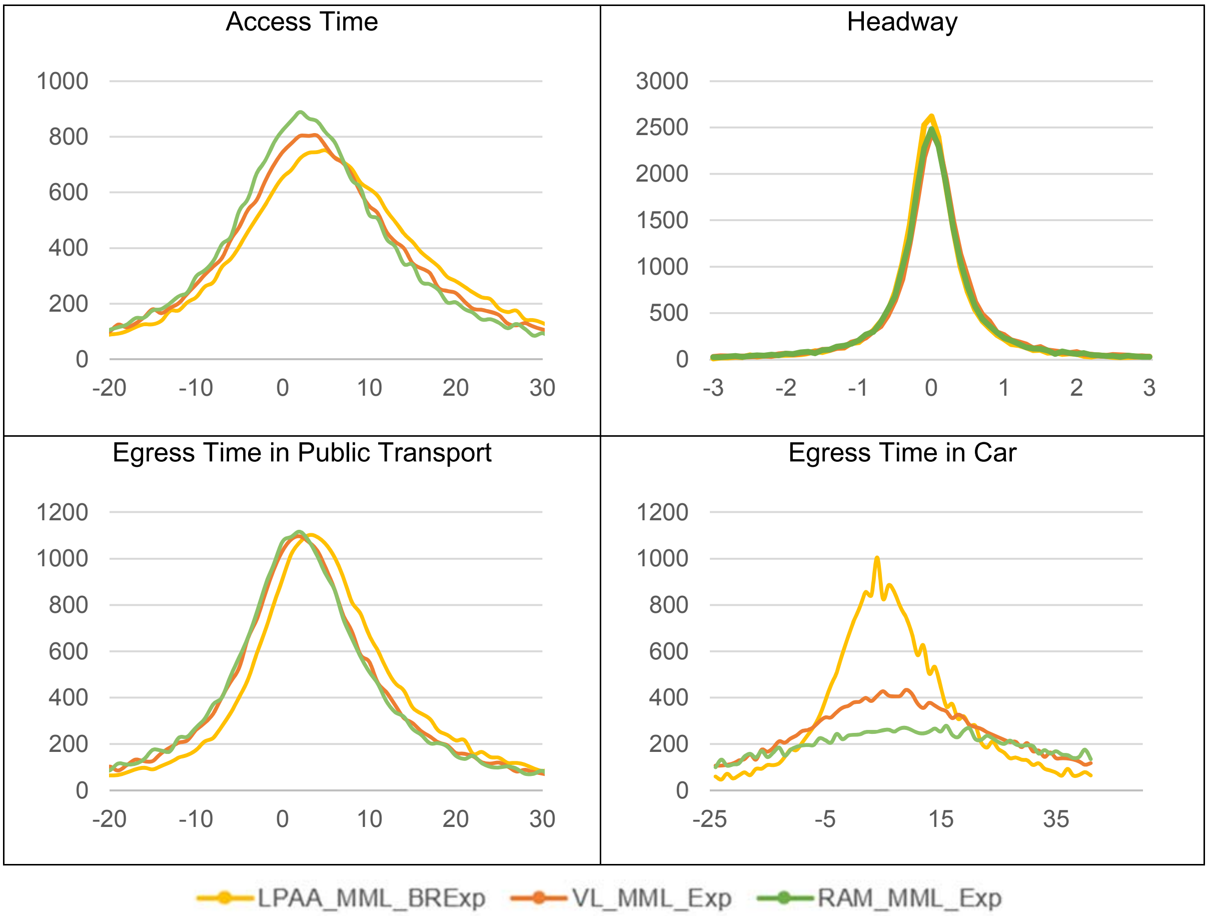


Figure 6-16: Attributes WTP distribution for different draws using the models with process homogeneity

The results of this section show that the WTP estimates are significantly different from each other using the t-test to compare two parameters (Section 3.8.1) for the majority of the attributes when using the various process strategies. Therefore, when considering process homogeneity, the selection of an appropriate process strategy is crucial since it has a significant influence on the outcome.

Summary: Comparison of the preferred process homogeneity models LPAA_MML_BRExp; RAM_MML_Exp and VL_MML_Exp

Access time WTP
- median (lower to higher): RAM, VL, LPAA
- standard error (lower to higher): RAM, VL, LPAA

Egress time public transport WTP
- median (lower to higher): RAM, VL, LPAA
- standard error (lower to higher): LPAA, VL, RAM

Egress time car WTP
- median (lower to higher): LPAA, VL, RAM
- standard error (lower to higher): LPAA, VL, RAM

Headway WTP
- median (lower to higher): LPAA, RAM, VL
- standard error (lower to higher): LPAA, VL, RAM

Significant differences? The majority of the differences are statistically significant between the LPAA model and the RAM or the VL. Models RAM and VL seem to be more similar between them.

Stability of the WTP estimates? The stability is similar for the models, except for the headway which has a superior stability in the LPAA model.

Conclusion: The use of LPAA, RAM or VL as the sole process strategy being used by individuals has a significant influence on the WTP estimates.

6.5.3.2.2. Probabilistic Decision Process versus Conditioning Random Process Heterogeneity

Figure 6-17 compares the PDP and CRPH approaches, both used to include process heterogeneity. For the access time and car egress times, the PDP models obtain the highest median WTP values compared to the CRPH models. In the case of the public transport egress time, the median WTP it is larger in the CRPHms ($7.44 per person hour) followed by the CRPHms_BRExp ($7.43) model and then for the PDP model ($7.11), the lowest median WTP is in CRPHs model ($3.53). The headway median WTP is smaller in the PDP, PDP_BRExp and CRPHm model ($0.05 per person minute) and larger in the CRPHms_BRExp model ($0.09 per person minute). As expected, the standard error is always larger in the CRPH model because it estimates every parameter as random, contrary to the PDP models.

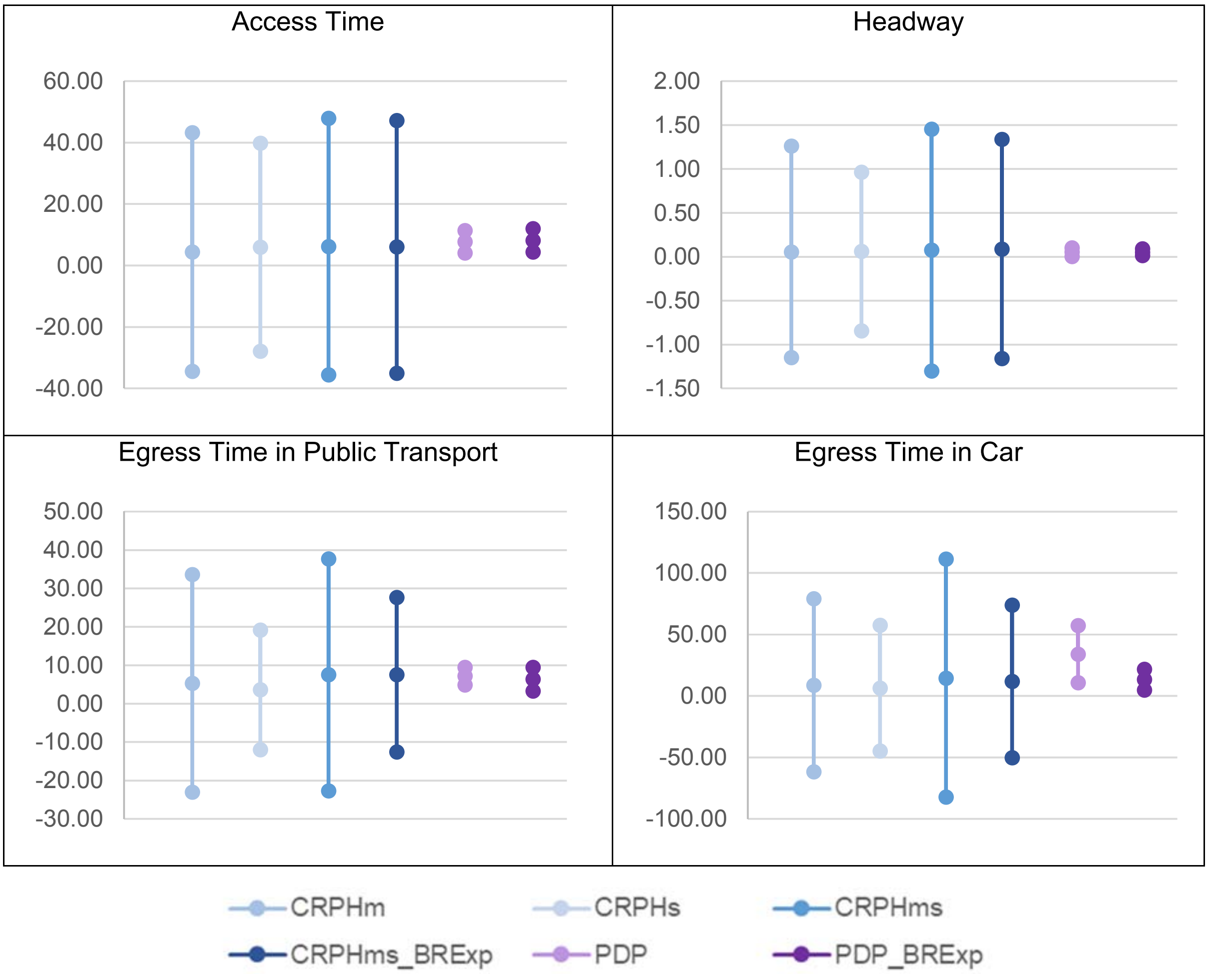


Figure 6-17: Attributes WTP median and 95% confidence intervals standard error for the models with process heterogeneity

Table 6-22 presents the results of the comparison between the WTP estimates for each mode in the PDP and CRPH models (with and without behavioural refinements and experience). Both median and standard error are used to compare the WTP estimates using the t-test (Section 3.8.1). In the models without behavioural refinements or experience, the access time WTP is significantly different and lower in the CRPH model for new light rail, new heavy rail, bus and busway, and significantly higher for the train. For the egress times WTP, it is significantly different and lower in the CRPH model for the new light rail, new heavy rail, train and car, and significantly different and higher for the other modes (i.e., new busway, bus and busway). In the case of the models with behavioural refinements and experience, the access time WTP is always significantly different and lower in the CRPH model except for the new busway which is significantly different and higher in the CRPH model. The egress times WTP are significantly different and higher in the CRPH model for the new busway, bus, and busway, and significantly different and lower for the train and car. The headway WTP estimate is significantly different and lower in the CRPH for the new light rail, and significantly different and higher for the rest of the modes (for the new heavy rail it is significantly higher with an 80% confidence level, the rest all with 95% confidence level).

Table 6-22: Comparison of attributes median WTP estimates for models with process heterogeneity using t-test

Attribute	Mode	CRPHms vs PDP	CRPHms_BRExp vs PDP_BRExp
Access Time	New Light Rail	***-11.86***	***-10.43***
	New Heavy Rail	***-11.86***	***-9.72***
	New Busway	0.20	***2.51***
	Bus	***-7.34***	***-6.36***
	Busway	***-2.66***	***-4.67***
	Train	***5.26***	***-10.06***
Egress Time	New Light Rail	***-6.97***	0.34
	New Heavy Rail	***-7.00***	-0.37
	New Busway	***7.07***	***15.62***
	Bus	***8.28***	***15.90***
	Busway	***6.02***	***9.32***
	Train	***-9.56***	***-11.09***
	Car	***-24.01***	***-2.96***
Headway	New Light Rail	1.08	***-8.59***
	New Heavy Rail	0.28	1.80
	New Busway	***2.45***	***3.79***
	Bus	***2.92***	***7.25***
	Busway	***2.65***	***4.30***
	Train	0.03	***2.00***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-18 summarises the stability of the WTP estimates for the CRPH models. The PDP is not included in these figures because it estimated all parameters as fixed. The graphs show a superior stability for the WTP of all the attributes for model CRPHs. The worst stability in the public transport and car egress times is for model CRPHms and in the headway for model CRPHms_BRExp. In the case of the access time, there is a similar stability between the CRPHm, CRPHms and CRPHms_BRExp models.

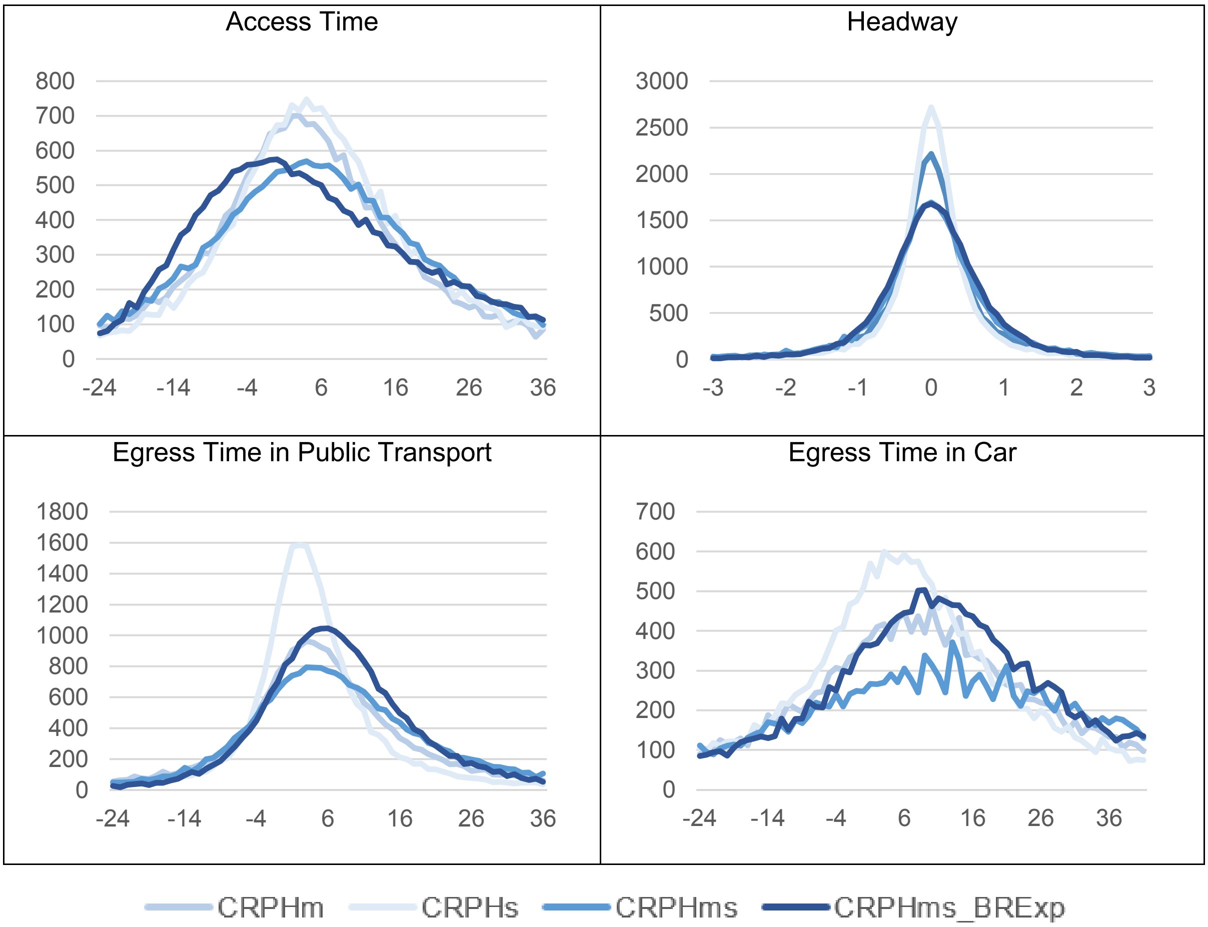


Figure 6-18: Attributes WTP distribution for different draws using the models with process heterogeneity

Summary: Comparison of the preferred models of two approaches to include process heterogeneity, PDP_BRExp and CRPHms_BRExp

Access time and car egress time WTP

- median (lower to higher): CRPH, PDP
- standard error (lower to higher): PDP, CRPH

Public transport egress time and headway WTP

- median (lower to higher): PDP, CRPH
- standard error (lower to higher): PDP, CRPH

Significant differences? Almost all the WTP estimates are significantly different from each other when using the two approaches.

Stability of the WTP estimates? PDP is completely stable since every parameter is estimated as fixed.

Conclusion: The PDP approach produces significantly different WTP for all the attributes relative to the CRPH approach.

6.5.3.2.3. Process Heterogeneity versus Process Homogeneity

Figure 6-19 presents the median and standard error WTP estimates for the three preferred models associated with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp, and RAM_MML_Exp) and two models for each of the process heterogeneity methods (CRPHms and CRPHms_BRExp; PDP and PDP_BRExp), and Table 6-23 shows the comparison of the estimates using the median and standard error (Section 3.8.1). The results suggest that the majority of the attributes' median WTP estimates are larger in the process heterogeneity models than in the process homogeneity models, and there are significant differences between them. Relative to the CRPH preferred model, the egress times median WTP are all significantly different and lower in the LPAA and VL model for all the modes and in the RAM for all the modes except the car (which are not statistically different). The access time median WTP are significantly different and higher in the CRPH models than in the LPAA model only for the new busway and significantly different and lower for the new light rail and new heavy rail. The CRPH also estimates significantly different and higher access time WTP than the RAM and VL models for the new busway, bus, busway and train (for the new light rail and new heavy rail the differences are not significant). For the headway median WTP, the CRPH models estimates significantly different and higher values for all the modes considering 90% confidence level. The headway WTP estimates seem relatively similar between the CRPH and VL model, except for the new light rail and train which are significantly different and lower in the CRPH model (90% confidence level).

Comparing the PDP preferred model with the process homogeneity models, the PDP model estimates significantly different and higher median WTP estimates for the access times and egress times for all the modes (with a few exceptions that are not statistically different and presented in Table 6-23 with **). The PDP approach estimates significantly different and lower headway median WTP than the VL model for all the modes except for the new light rail; significantly different and higher median WTP than the LPAA model for the new light rail and significantly different and lower for the bus; and significantly different and higher than the RAM model for the new light rail, and significantly different and lower for the bus, busway and train.

The standard errors are always lower in the PDP as it considers only fixed parameters. The second lowest standard error is in the LPAA model for the headway and car egress time, for the CRPH model in the public transport egress time and for RAM in the access time. The largest standard error is in the RAM model for the egress times, and in the CRPH model for the access time and headway.

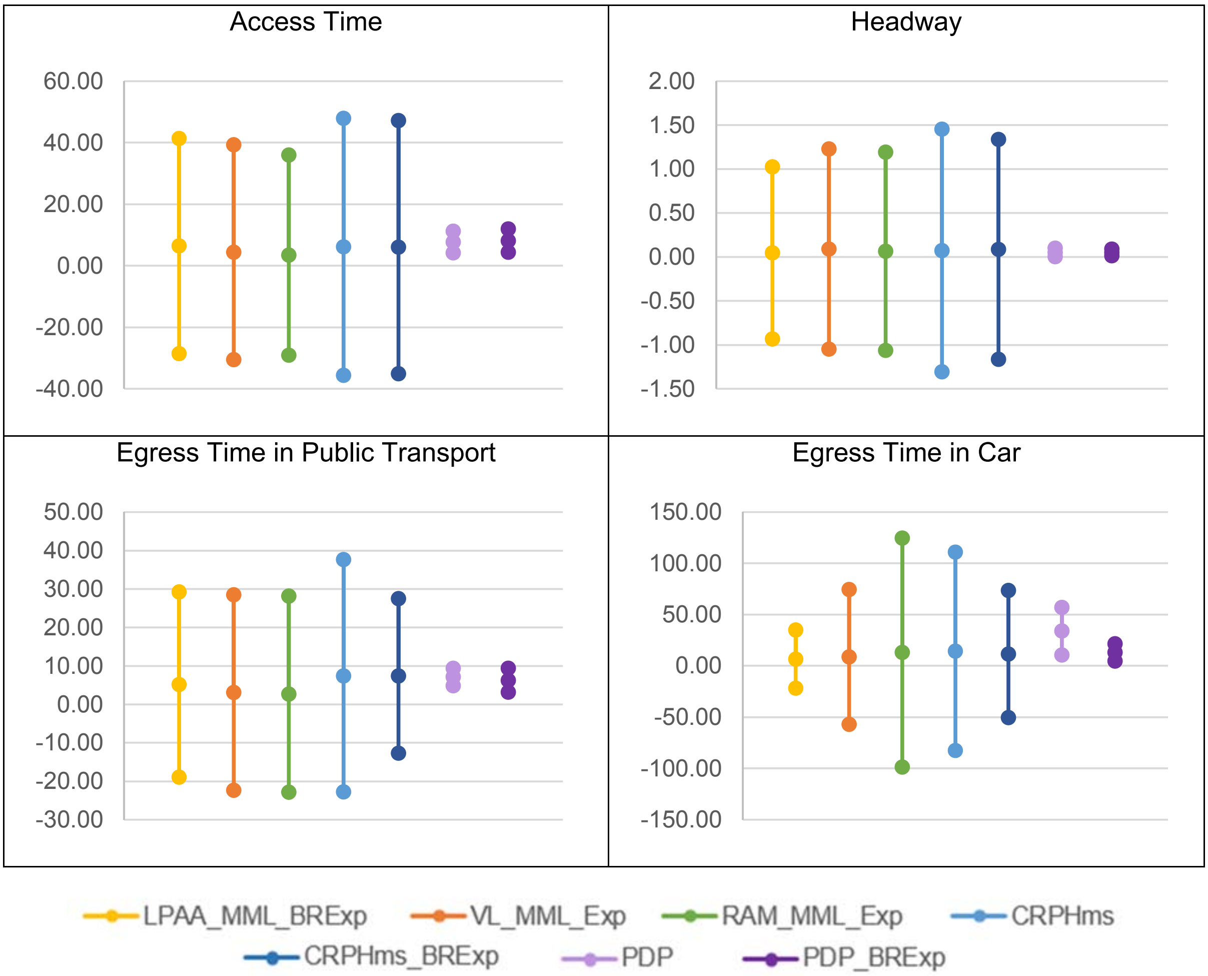


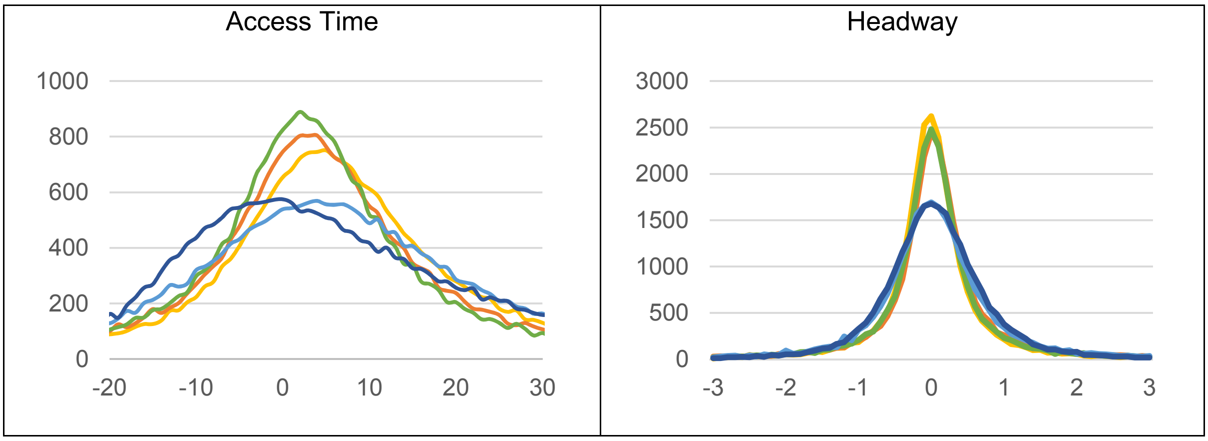
Figure 6-19: WTP median and 95% confidence intervals for the models with process homogeneity versus process heterogeneity

Table 6-23: Comparison of attributes median WTP estimates for models with process heterogeneity and process homogeneity using t-test

Attribute	Mode	CRPHms_BRExp vs LPAA_MML_BRExp	CRPHms_BRExp vs VL_MML_Exp	CRPHms_BRExp vs RAM_MML_Exp	PDP_BRExp vs LPAA_MML_BRExp	PDP_BRExp vs VL_MML_Exp	PDP_BRExp vs RAM_MML_Exp
Access Time	New Light Rail	↓ ***-4.11***	↓ -0.53	↑ 1.74	↑ ***6.16***	↑ ***10.72***	↑ ***15.04***
	New Heavy Rail	↓ ***-4.20***	↓ -0.92	↑ 1.07	↑ ***5.75***	↑ ***10.16***	↑ ***14.15***
	New Busway	↑ ***2.12***	↑ ***6.78***	↑ ***8.56***	↑ 0.42	↑ ***7.45***	↑ ***10.50***
	Bus	↓ -0.58	↑ ***5.83***	↑ ***8.29***	↑ ***6.09***	↑ ***16.63***	↑ ***21.53***
	Busway	↑ 0.36	↑ ***5.95***	↑ ***8.27***	↑ ***6.16***	↑ ***15.65***	↑ ***20.51***
	Train	↑ 0.05	↑ ***3.50***	↑ ***5.94***	↑ ***12.47***	↑ ***16.41***	↑ ***21.39***
Egress Time	New Light Rail	↑ ***5.34***	↑ ***11.11***	↑ ***12.64***	↑ ***6.50***	↑ ***13.20***	↑ ***15.03***
	New Heavy Rail	↑ ***4.35***	↑ ***9.96***	↑ ***11.33***	↑ ***5.80***	↑ ***12.28***	↑ ***13.91***
	New Busway	↑ ***10.33***	↑ ***17.35***	↑ ***18.30***	↓ -0.19	↑ ***9.02***	↑ ***10.32***
	Bus	↑ ***13.86***	↑ ***24.04***	↑ ***25.49***	↑ ***4.45***	↑ ***17.80***	↑ ***19.73***
	Busway	↑ ***10.78***	↑ ***20.10***	↑ ***21.60***	↑ ***6.11***	↑ ***18.26***	↑ ***20.21***
	Train	↑ ***3.42***	↑ ***9.55***	↑ ***11.10***	↑ ***13.15***	↑ ***19.03***	↑ ***20.83***
	Car	↑ ***9.18***	↑ ***3.95***	↓ -1.26	↑ ***27.53***	↑ ***8.18***	↑ 0.22
Headway	New Light Rail	↑ ***3.24***	↓ -1.94	↑ 0.48	↑ ***7.71***	↑ 1.19	↑ ***3.88***
	New Heavy Rail	↑ 1.86	↓ -1.62	↑ 0.11	↑ 0.72	↓ ***-4.06***	↓ -1.64
	New Busway	↑ ***2.92***	↑ 1.08	↑ ***2.19***	↓ -0.40	↓ ***-3.46***	↓ -1.55
	Bus	↑ ***3.79***	↑ 1.21	↑ ***2.80***	↓ ***-4.34***	↓ ***-8.58***	↓ ***-5.79***
	Busway	↑ ***3.29***	↑ 0.54	↑ ***2.20***	↓ -0.41	↓ ***-4.79***	↓ ***-2.07***
	Train	↑ 1.66	↓ ***-2.21***	↓ -0.28	↑ 0.12	↓ ***-5.20***	↓ ***-2.45***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-20 summarises the stability of the attributes WTP for the different models. The stability in the public transport egress time is relatively similar for all the models. For the access time and headway WTP the stability is superior in the process homogeneity models than in the CRPH. The car egress time WTP stability is superior only in the LPAA model, the rest of the models have a similar stability.



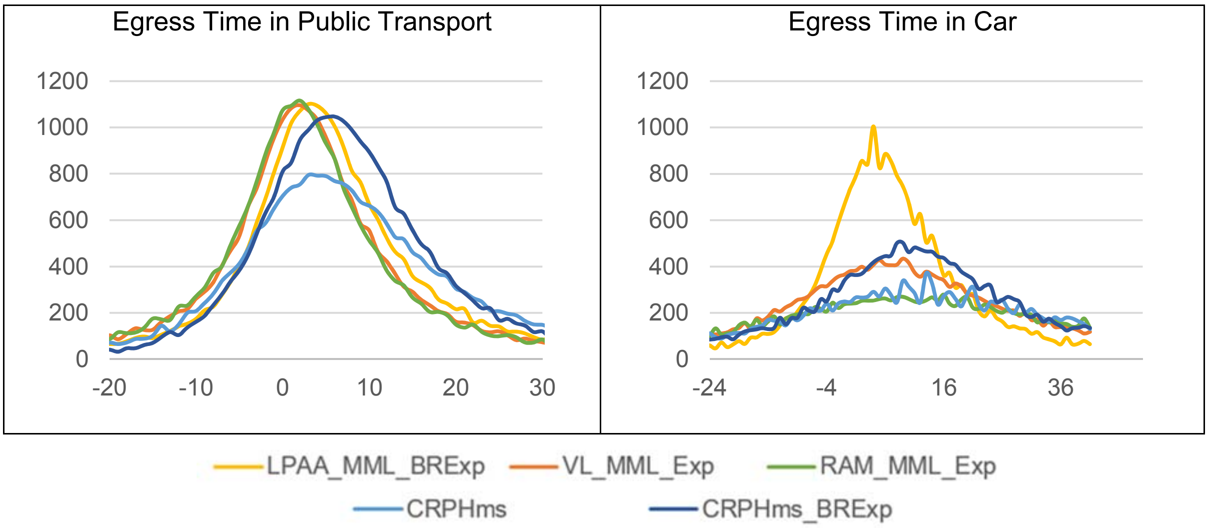


Figure 6-20: Attributes WTP distribution for different draws using the models with process homogeneity versus process heterogeneity

Summary: Comparison of the preferred models with process homogeneity (LPAA_MML_BRExp, VL_MML_Exp and RAM_MML_Exp) with the preferred models with process heterogeneity (PDP_BRExp, CRPHms_BRExp)

Access time WTP

- median (lower to higher): RAM, VL, CRPH, LPAA, PDP
- standard error (lower to higher): PDP, RAM, VL, LPAA, CRPH

Egress time public transport WTP

- median (lower to higher): RAM, VL, LPAA, PDP, CRPH
- standard error (lower to higher): PDP, CRPH, LPAA, VL, RAM

Egress time car WTP

- median (lower to higher): LPAA, VL, CRPH, RAM, PDP
- standard error (lower to higher): PDP, LPAA, CRPH, VL, RAM

Headway WTP

- median (lower to higher): LPAA, PDP, RAM, CRPH, VL
- standard error (lower to higher): PDP, LPAA, RAM, VL, CRPH

Significant differences? The majority of the differences between the VTTS are statistically different when considering process heterogeneity (i.e., PDP or CRPH) instead of process homogeneity (i.e., LPAA, RAM or VL) for all modes.

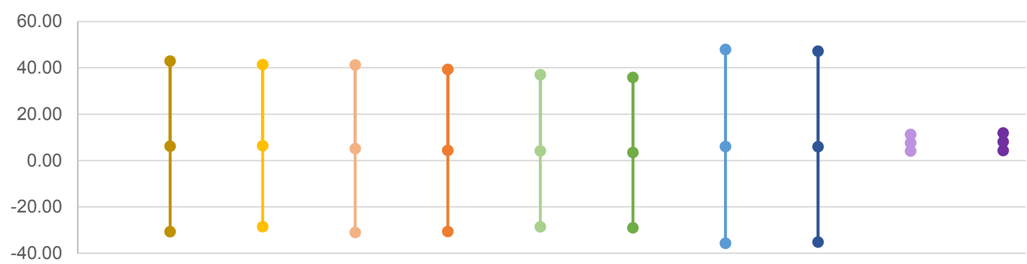
Stability of the WTP estimates? The stability of the WTP for the access time, car egress time and headway is superior in the process homogeneity models. For the public transport egress time the stability of the models is relatively similar.

Conclusion: The process heterogeneity method directly influences the results and it produces significantly different median WTP estimates relative to the process homogeneity models.

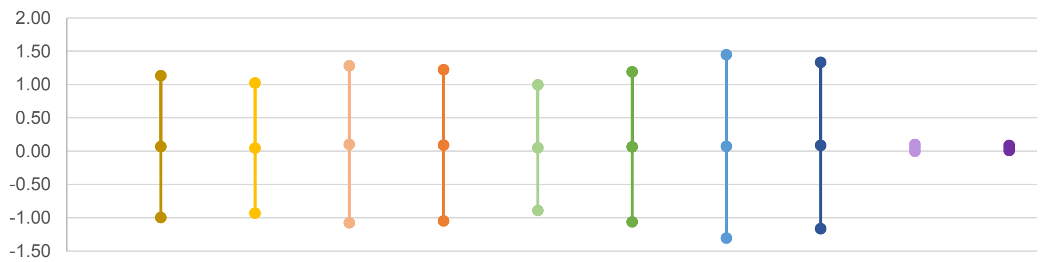
6.5.3.2.4. Behavioural Refinements and Experience

Figure 6-21 shows the WTP median and 95% confidence interval estimates for the models with and without behavioural refinements and/or experience, and Table 6-24 presents the comparison. When adding behavioural refinements and experience in the LPAA model: the median access time WTP for the bus significantly increases; the median egress time WTP for the new light rail, new heavy rail, train and car significantly decreases; and the median headway WTP for the new light rail, new heavy rail and train significantly decreases. In the VL model, there are significant increases in the median access time WTP for the bus, busway and train. In the RAM model, there are significant decreases in the median access time WTP for the new light rail, bus, busway and train; and a significant increase in the median car egress time WTP. In the CRPH process heterogeneity model, there is a significant increase in the median access time WTP for the bus and a significant decrease for the train; there is a significant decrease of the median car egress time WTP; and a significant increase in the median headway WTP for the new light rail. Finally, when adding behavioural refinements and experience in the PDP model there are significant decreases in the median egress time WTP for all the modes, in the headway WTP for all the modes except new light rail, and in the access time WTP for the new busway. There are also significant increases in the median access times WTP for the bus, busway and train, and for the car egress time WTP. In conclusion, there is a significantly higher influence on the median WTP estimates for the PDP model, followed by the LPAA model. This is an important finding as it is suggesting that when considering process heterogeneity and random parameters the importance of behavioural refinements and experience becomes less influential.

Access Time



Headway



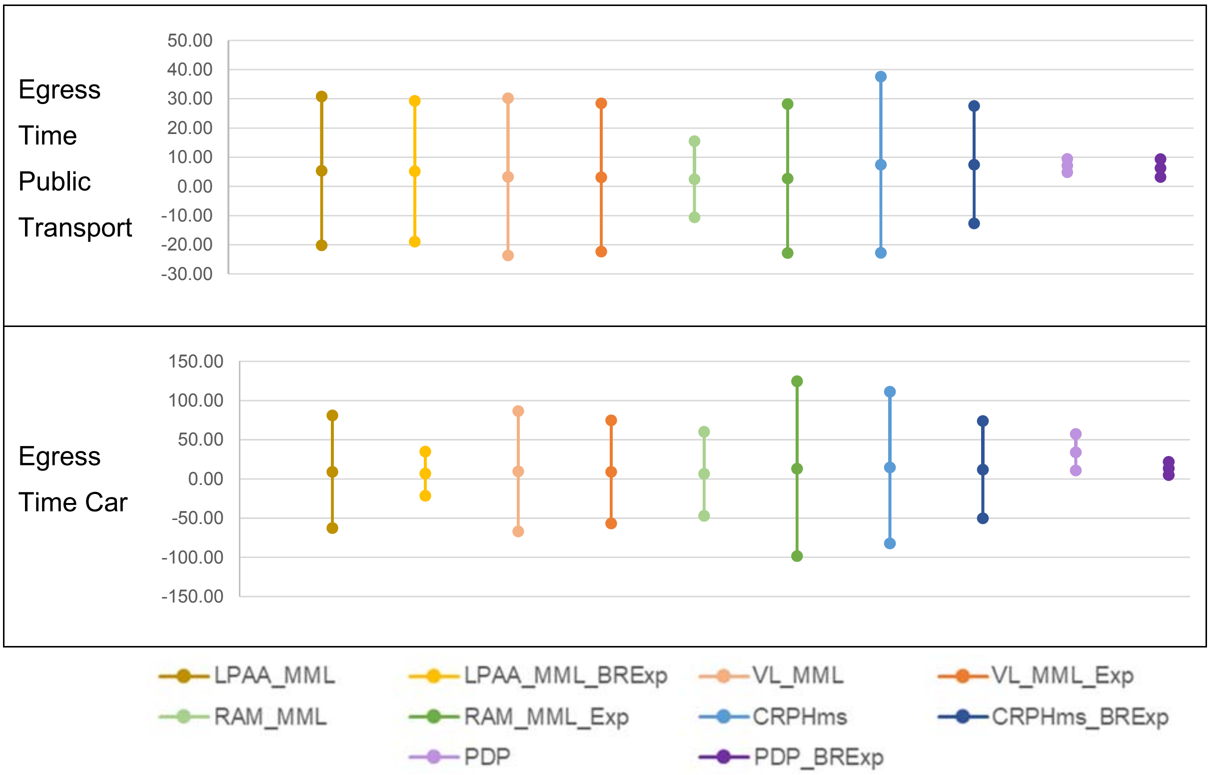


Figure 6-21: Attributes WTP median and 95% confidence intervals for the models with and without behavioural refinements and/or experience

Table 6-24: Comparison of attributes' median WTP estimates for models with and without behavioural refinements and/or experience using t-test

Attribute	Mode	LPAA_MML_BRExp vs LPAA_MML	VL_MML_Exp vs VL_MML	RAM_MML_Exp vs RAM_MML	CRPHms_BRExp vs CRPHms	PDP_BRExp vs PDP
Access Time	New Light Rail	↓ -0.64	↓ -1.89	↓ ***-1.99***	↑ 0.18	↓ -0.85
	New Heavy Rail	↓ -0.56	↓ -1.79	↓ -1.88	↓ -0.15	↓ -0.79
	New Busway	↑ 1.58	↓ -1.46	↓ -1.53	↑ 0.58	↓ ***-10.77***
	Bus	↑ ***2.20***	↓ ***-2.11***	↓ ***-2.21***	↑ ***3.64***	↑ ***23.46***
	Busway	↑ 1.79	↓ ***-2.02***	↓ ***-2.12***	↑ 0.65	↑ ***24.03***
	Train	↓ -0.89	↓ ***-2.02***	↓ ***-2.15***	↓ ***-7.79***	↑ ***31.79***
Egress Time	New Light Rail	↓ ***-2.13***	↓ -0.63	↑ 1.06	↑ 1.75	↓ ***-34.51***
	New Heavy Rail	↓ ***-1.98***	↓ -0.60	↑ 1.00	↑ 1.17	↓ ***-34.98***
	New Busway	↑ 0.84	↓ -0.48	↑ 0.81	↓ -0.94	↓ ***-41.23***
	Bus	↑ 1.12	↓ -0.70	↑ 1.18	↓ -0.69	↓ ***-23.48***
	Busway	↑ 0.69	↓ -0.67	↑ 1.11	↓ -0.86	↓ ***-10.74***
	Train	↓ ***-2.54***	↓ -0.67	↑ 1.16	↑ 0.50	↓ ***-3.24***
	Car	↓ ***-3.77***	↓ -0.93	↑ ***6.61***	↓ ***-2.85***	↓ ***-102.14***
Headway	New Light Rail	↓ ***-2.24***	↓ -0.99	↑ 1.03	↑ ***2.68***	↑ ***64.57***
	New Heavy Rail	↓ ***-2.11***	↓ -0.94	↑ 0.96	↑ 0.92	↓ ***-3.33***
	New Busway	↓ -1.07	↓ -0.76	↑ 0.78	↑ 0.43	↓ ***-8.13***
	Bus	↓ -1.58	↓ -1.10	↑ 1.13	↑ 0.66	↓ ***-91.08***
	Busway	↓ -1.61	↓ -1.05	↑ 1.09	↑ 0.72	↓ ***-8.99***
	Train	↓ ***-2.46***	↓ -1.05	↑ 1.08	↑ 0.82	↓ ***-15.88***

Note: Bold and italic estimates are the ones significant at a 95% confidence level

Figure 6-22 graphically shows the stability of the attributes' WTP estimates (except for travel time). There does not appear to be any pattern associated with the inclusion of behavioural refinements and/or experience.

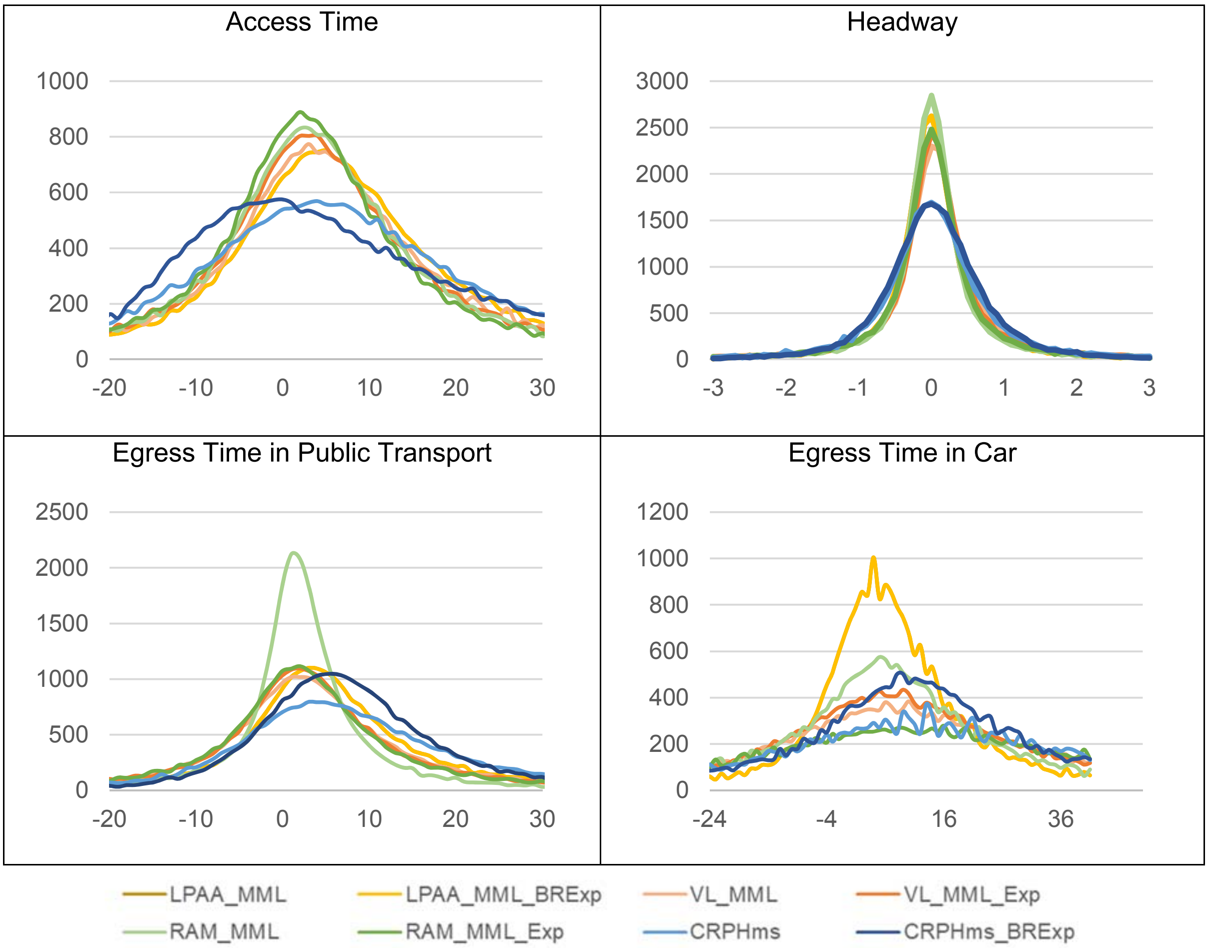


Figure 6-22: Attributes WTP distribution for different draws using the models with and without behavioural refinements and/or experience

Summary: Comparison of the models with and without experience and behavioural refinements
Access time WTP

- median when adding experience and behavioural refinements
  - increases for LPAA, PDP
  - decreases for VL, RAM and CRPH
- standard error when adding experience and behavioural refinements
  - increases for PDP
  - decreases for LPAA, VL, RAM and CRPH

Egress times (public transport and car) WTP

- median when adding experience and behavioural refinements
  - increases for RAM
  - decreases for LPAA, VL, PDP and CRPH
- standard error when adding experience and behavioural refinements
  - increases for RAM
  - decreases for LPAA, VL, PDP and CRPH

Headway WTP

- median when adding experience and behavioural refinements
  - increases for RAM and CRPH
  - decreases for LPAA, VL, PDP
- standard error when adding experience and behavioural refinements
  - increases for RAM
  - decreases for LPAA, VL, PDP and CRPH

Significant differences? The majority of the WTP estimates are significantly different when considering behavioural refinements and experience for the PDP model. The LPAA model also has several significant differences.

Stability of the WTP estimates? Similar stability with and without experience and behavioural refinements.

Conclusion: The inclusion of experience and behavioural refinements significantly influences the WTP median and standard error especially in the PDP model and LPAA.

6.6. Conclusions

This chapter has used a new dataset, the Northwest dataset, to estimate the exact same models as estimated in the previous chapter on the Metro Rail dataset. For the three different process strategies, RAM, VL and LPAA five types of models were presented: (1) individuals use LPAA as their only process strategy; (2) individuals use VL as their only process strategy; (3) individuals use RAM as their only process strategy; (3) individuals may use any of the three process strategies with a certain probability (PDP); (4) individuals use a combination of the three process strategies by conditioning the parameters normally defined under LPAA with the alternative strategies (CRPH). Assessing the models for each process heuristic and for process heterogeneity, in terms of the overall model fit (using the log likelihood ratio test and the Vuong test), offers encouraging evidence that the inclusion of experience and behavioural refinements has a statistically superior overall performance for all the model types. The preferred models support random parameters, providing a better overall model fit, except for the PDP approach where all parameters are fixed.

The behavioural refinements and experience of the models was first compared. The results showed statistically significant differences in the behaviour of the experience parameter when considering different process strategies as the sole heuristic being used by individuals. Since the nature of the process strategies is different, it was anticipated that overt experience would have a different influence in each of them. Moreover, the results suggest that experience and risk attitudes are less statistically significant when considering process heterogeneity than under process homogeneity. This is an important finding suggesting that the importance of including additional behavioural components is reduced when considering process heterogeneity; hinting some amount of possible confoundment of specific behavioural assumptions and a random distribution of parameter estimates. Consistent with the MetroRail dataset from Chapter 5, this raises the interesting question of what role random parameters play in capturing process heterogeneity when specific processing heuristics are accommodated. In general, we would tend to prefer explicit systematic sources of explanation and not unobserved non-systematic sources of preference heterogeneity, unless there are problems in applying such systematic sources in practice (such as forecasting future levels).

Statistical fit indicators, such as the log likelihood ratio and the AIC were used to compare models. They showed that the CRPH preferred model (CRPHms_BRExp) was superior to the other ones. This was followed – not so closely – by the CRPHms, the RAM_MML_Exp, and the LPAA_MML_BRExp models.

The WTP median estimates and the 95% confidence intervals were compared to see if the different formulations produced similar or different results. The evidence suggest significant differences in the WTP estimates for all the indicators when considering different process strategies where each is the only heuristic being used by respondents, as well as differences when assuming process homogeneity in contrast to process heterogeneity. The most important finding is that the inclusion of process heterogeneity together with the process heuristic selected has a significant influence on the key behavioural outcomes such as WTP estimates.

The key question that needs to answered is whether we have identified, from all the models and process heuristics assumptions, pointers to support one (or maybe more) behaviourally appealing ways of representing the preferences of a sample of travellers, and what this mean for willingness to pay estimates in contrast (at least) to the very standard and traditional simple

LPAA form from which WTP estimates are commonly obtained. What we find in this chapter is that the preferred model CRPHms_BRExp results in WTP estimates for key attributes that are typically lower when considering process homogeneity under LPAA, VL or RAM. The CRPHms_BRExp model supports the idea that there is significant taste heterogeneity across the sample, where the mean and standard deviation of the parameters - normally defined under and LPAA model - interact with alternative process strategies, such as VL and RAM. The next chapter will compare the results in Chapter 5 and the current chapter, using the evidence to offer a number of more portable findings based on two datasets.

CHAPTER 7

Conclusions and Future Research

7.1. Summary

The objective of this thesis was to integrate multiple decision process strategies, Value Learning (VL) and Relative Advantage Maximisation (RAM) in particular, alongside the traditional LPAA 'process rule' with behavioural refinements (i.e., risk attitudes, perceptual conditioning and experience), to take into account process endogeneity in choice responses. A novel approach was proposed to include process heterogeneity, referred to as *conditioning of random process heterogeneity (CRPH)*, where the mean and standard deviation of the parameters normally defined under an LPAA heuristic are conditioned by process strategies. This approach takes into account the relationship between process heterogeneity and taste heterogeneity. In the following sections we synthesise the main contributions, re-visiting the research questions proposed in Chapter 1. We also suggest areas of future research that build on the contributions herein and offer some concluding remarks.

7.2. Main Contributions of this Thesis

A number of research questions were presented in previous chapters to position the focus of the research, recognising the greater emphasis that should be placed on the duality of choosing a process rule and the outcome choice. This interest aligns with McFadden's (2001, p. 374) call to include information, experience and decision processes in the traditionally used random utility maximization (RUM) framework, referred to in Chapter 1:

"*What lies ahead? I believe that the basic RUM theory of decision-making, with a much larger role for experience and information in the formation of perceptions and expression of preferences, and allowance for the use of rules as agents for preferences, can describe most economic choice behavior in markets, surveys, and the laboratory. If so, then this framework* *can continue for the foreseeable future to form a basis for microeconometric analysis of consumer behavior and the consequences of economic policy."*

We revisit each research question and assess the new evidence obtained from the myriad of models presented herein that investigate the role of process homogeneity, process heterogeneity, behavioural refinements and experience, using two datasets in the Sydney context - the Metro Rail dataset and the Northwest dataset. The use of two datasets offered an opportunity to establish the extent to which the proposed methods can be supported more generally. A number of assessment criteria were used to assess how the various process strategies and behavioural refinements (i.e. taste heterogeneity, risk attitudes, perceptual conditioning) contribute to the relative 'performance' of each model form. These are goodness of statistical fit, class probability allocations (when using latent class models), statistical significance of each attribute (parameter) of interest, and the profile of median and standard deviation estimates of willingness to pay (WTP). The last part of this Section revisits the conceptual framework for decision making proposed in Chapter 1, and how the findings support and reinforce the value of the behavioural extensions developed in this thesis.

Research Question 1: Are preferences better represented when considering multiple decision process strategies, risk attitudes, perceptual conditioning and experience? How might they work together?

Table 7-1 presents the AIC indicator (considering the log likelihood while penalising the number of parameters estimated) as a representation of the overall statistical performance of the models with and without behavioural refinements and/experience. The results suggest that behavioural refinements and experience improve all the models. A colour scale is used to show the greatest improvements (darker green) and lesser improvements (lighter green), in terms of the AIC difference value. In both datasets, the largest AIC improvements, when including these additional parameters, are associated with the LPAA MNL and RAM MNL models.

Table 7-1: Summary and comparison of AIC indicators for models with and without behavioural refinements and/or experience (BRExp)

	AIC Metro Rail			AIC Northwest		
	Without BRExp	With BRExp	Difference AIC (Without BRExp - With BRExp)	Without BRExp	With BRExp	Difference AIC (Without BRExp - With BRExp)
LPAA MNL	1.339	1.301	0.039	2.731	2.674	0.056
LPAA MML	1.067	1.054	0.013	2.085	2.047	0.038
VL MNL	1.265	1.259	0.006	2.657	2.650	0.008
VL MML	1.075	1.071	0.004	2.097	2.087	0.010
RAM MNL	1.342	1.301	0.042	2.730	2.680	0.050
RAM MML	1.065	1.054	0.011	2.067	2.027	0.040
PDP	1.080	1.068	0.012	2.202	2.167	0.035
CRPH	1.052	1.047	0.005	2.023	1.999	0.025

*Colour scale represents a detriment (red) or improvement (green) in the AIC indicator when including behavioural refinements and/or experience. A darker colour tone represents a higher detriment or improvement.

When including behavioural refinements and experience in the PDP model, there was a major shift in the class memberships in both datasets, presented in Figure 7-1. In the PDP_MNL models the probability of individuals using the RAM heuristic was the highest (0.50 in the Metro Rail dataset and 0.51 in the Northwest dataset) and the lowest was for the LPAA heuristic (0.18 in the Metro Rail dataset and 0.17 in the Northwest dataset). When adding behavioural refinements and experience, models PDP_MNL_BRExp, the probability of individuals using the LPAA heuristic was the highest (0.45 in the Metro Rail dataset and 0.50 in the Northwest dataset), and the lowest in the Metro Rail dataset was for the VL heuristic (0.18) and in the Northwest dataset was for the RAM heuristic (0.15). These findings emphasise that inclusion of behavioural refinements and experience change the distribution of preferences across the assessed heuristics and benchmark reference form (LPAA) and matter in a statistical sense as well as an intuitively appealing behavioural sense. There is a case for including these phenomena in choice studies.

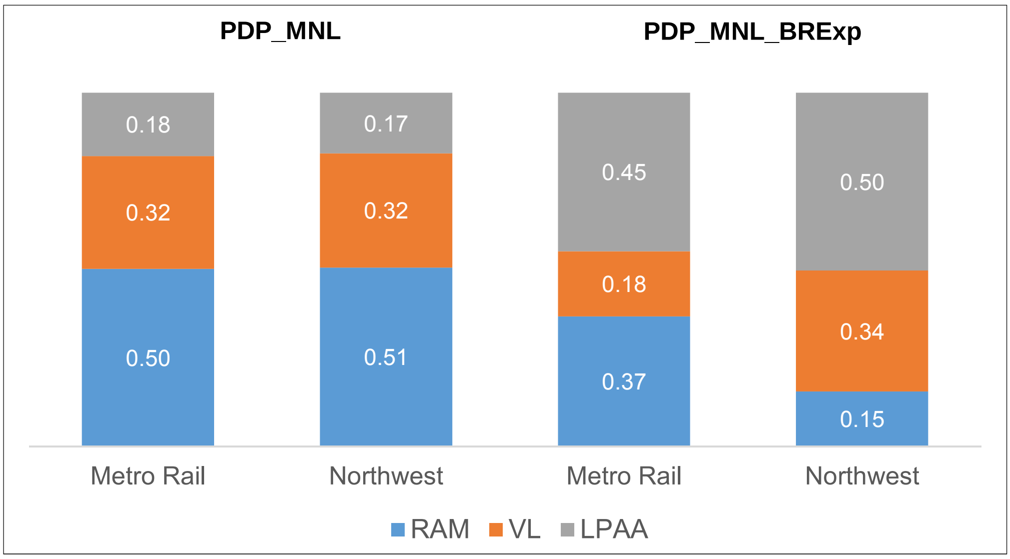


Figure 7-1: Class probability distribution for PDP models

The second part of this research question refers to whether preferences are better represented by multiple decision process strategies. Table 7-2 presents the difference between the AIC indicators for the models with process homogeneity and process heterogeneity. The red colour scale represents a detriment in the AIC indicator when considering process heterogeneity relative to process homogeneity, and a darker (lighter) red represents a larger (lower) detriment. The green colour scale represents an improvement in the AIC indicator, where a larger (lower) improvement is represented by darker (lighter) green. As can be seen, the PDP approach used to integrate multiple process strategies has a worse AIC indicator than most of the process homogeneity models (with the exception of the VL MML with behavioural refinements and experience, BRExp, in the Metro Rail dataset). Contrarily, the CRPH models used to consider process heterogeneity have a much improved AIC indicator than all the process homogeneity models, for both datasets. This reveals the importance and impact of random parameters when understanding preferences and raises the question of what role random parameters play in capturing process heterogeneity when specific processing heuristics are accommodated.

Table 7-2: Comparison of AIC indicators for models with process homogeneity and process heterogeneity

Dataset	Models		Difference AIC (Process homegeneity model - Process heterogeneity model)	
	Process Homogeneity Model	Process Heterogeneity Model	Models without BRExp	Models with BRExp
Metro Rail	LPAA MML	PDP	-0.013	-0.014
	LPAA MML	CRPH	0.015	0.007
	VL MML	PDP	-0.005	0.003
	VL MML	CRPH	0.023	0.024
	RAM MML	PDP	-0.015	-0.014
	RAM MML	CRPH	0.013	0.007
Northwest	LPAA MML	PDP	-0.117	-0.12
	LPAA MML	CRPH	0.062	0.048
	VL MML	PDP	-0.105	-0.08
	VL MML	CRPH	0.074	0.088
	RAM MML	PDP	-0.135	-0.14
	RAM MML	CRPH	0.044	0.028

*Colour scale represents a detriment (red) or improvement (green) in the AIC indicator when considering process heterogeneity relative to process homogeneity. A darker colour tone represents a higher detriment or improvement.

When allowing for multiple process strategies, experience and behavioural refinements to capture preference variations, the overall performance of the models is statistically better (and behaviourally more appealing). However, it is important to understand how these components work together. Table 7-3 presents a summary of the parameters that were shown to be statistically significant in the models that included behavioural refinements and/or experience for each dataset. The last column in the table presents the percentage of the parameters that were statistically significant out of all that were tested. In the LPAA MNL model, which represents a relatively simple model with fixed parameters and including one – and the most commonly used - process strategy, 50% of the parameters tested were statistically significant in the Metro Rail dataset, and 82% in the Northwest dataset. The notable difference between these percentages is in perceptual conditioning, which was only tested in the Metro Rail dataset but was not found statistically significant in Northwest data because no attribute was presented with levels of variation. However, both data sources found the experience parameter for all the modes and more than half of the possible risk attitudes to be statistically significant. When adding random parameters to the LPAA (LPAA MML), there is a decrease on the number of parameters found to be statistically significant: in the Metro Rail dataset from 50% to 43% and in Northwest from 82% to 36%. The larger decrease for the Northwest data relative to the Metro Rail is mainly due to the experience parameters where half were not statistically significant, whereas in the Metro Rail dataset all remained statistically significant.

The models that considered VL or RAM as the only process strategy being used only included experience and not behavioural refinements. In the Metro Rail dataset 2/3 of the experience parameters were statistically significant in the VL model, and all were in the RAM model. In the Northwest dataset, all the experience parameters were statistically significant in the VL and RAM models. There was no difference in the number of significant behavioural refinements and/or experience parameters when adding random parameters in the VL or RAM models for both datasets.

The last two rows in each dataset in Table 7-3 summarise the evidence for the process heterogeneity models. There is a large decrease in the percentage of behavioural refinements and experience parameters found to be statistically significant for these models. For the PDP model in the Metro Rail dataset, only a 25% of these parameters were statistically significant and 22% in the Northwest dataset. Even though the PDP model tested many more experience parameters (3 times more representing each class), approximately only 1/4 of them were statistically significant. For the CRPH approach, these percentages decrease even more: in the Metro Rail dataset, to 21%, and in the Northwest dataset, to 18%. The finding that the process heterogeneity models identified statistically significantly less behavioural refinements and experience parameters is an important finding. This suggests that the importance of including additional behavioural components is reduced in the presence of process heterogeneity (i.e., the PDP and CRPH models) relative to process homogeneity (i.e., LPAA, VL or RAM models), and that the risk of confoundment may increase with more complex model forms. This statement is very important as it is saying that by including process heterogeneity other behavioural refinements are not as important. It could certainly be because process heterogeneity is somehow confounded with the behavioural refinements, but we cannot be certain of this. We can only argue that less attention is required on other behavioural refinements when including process heterogeneity.

Table 7-3: Summary of the behavioural refinements and experience significant in the preferred models

Dataset	Preferred Models	Model Characteristics		Behavioural Refinements and Experience Parameters				
		Random Parameters	Multiple decision process strategies	Experience	Risk attitudes	Perceptual conditioning	Number of Parameters	% of Total Possibilities
Metro Rail	LPAA MNL	No	No	Yes (3/3)	Yes (3/4 travel time + 1/5 cost)	No (0/2)	7	50%
	LPAA MML	Yes	No	Yes (3/3)	Yes (0/4 travel time + 2/5 cost)	Yes (1/2)	6	43%
	VL MNL	No	No	Yes (2/3)	-	-	2	67%
	VL MML	Yes	No	Yes (2/3)	-	-	2	67%
	RAM MNL	No	No	Yes (3/3)	-	-	3	100%
	RAM MML	Yes	No	Yes (3/3)	-	-	3	100%
	PDP	No	Yes	Yes (4/9)	Yes (0/4 travel time + 1/5 cost)	No (0/2)	5	25%
	CRPHms	Yes	Yes	Yes (3/3)	No (0/4 travel time + 0/5 cost)	No (0/2)	3	21%
Northwest	LPAA MNL	No	No	Yes (4/4)	Yes (2/3 travel time + 3/4 cost)	-	9	82%
	LPAA MML	Yes	No	Yes (2/4)	Yes (0/3 travel time + 2/4 cost)	-	4	36%
	VL MNL	No	No	Yes (4/4)	-	-	4	100%
	VL MML	Yes	No	Yes (4/4)	-	-	4	100%
	RAM MNL	No	No	Yes (4/4)	-	-	4	100%
	RAM MML	Yes	No	Yes (4/4)	-	-	4	100%
	PDP	No	Yes	Yes (4/16)	Yes (0/3 travel time + 1/4 cost)	-	5	22%
	CRPHms	Yes	Yes	Yes (2/4)	No (0/3 travel time + 0/4 cost)	-	2	18%

Research Question 2: How do decision process strategies interact with each other?

Two approaches were used to integrate multiple decision processing strategies. The PDP approach suggests, from the perspective of the decision maker, that there is no statistically significant *interaction* between two or three process strategies but an individual respondent will select an independent process strategy, in the presence of other process heuristics, up to a probability. In contrast, the CRPH approach hypothesises that the assessed decision process strategies interact with each other with individuals using more than one rule to evaluate a specific attribute. This interaction between process strategies is included in the mean parameter estimate across the sample and in the standard deviation (capturing preference heterogeneity). The CRPH approach is an innovative approach to testing for preference heterogeneity at the attribute level in a setting of process heuristics compared to the PDP approach, the most common approach used in the transport literature (and other literatures) to integrate multiple heuristics in choice making models.

Table 7-4 summarises the AIC indicators for the preferred PDP and CRPH models for each dataset. There is a significant improvement in the AIC when considering the CRPH approach: in the Metro Rail dataset this improvement is 0.021 and in the Northwest dataset it is 0.169. These differences are significant and show that the CRPH provides an improved representation of preferences in both datasets. This supports the position that process strategies can and often do interact with each other in preference revelation.

Table 7-4: AIC indicator of preferred PDP and CRPH models

Dataset	Model	Number of Parameters Estimated	Log Likelihood at convergence	Log likelihood at zero	AIC
Metro Rail	PDP_MNL_BRExp	47	-5,007.53	-13,125.44	1.068
	CRPHms_BRExp	35	-4,922.41		1.047
Northwest	PDP_MNL_BRExp	44	-4,864.47	-7,838.25	2.167
	CRPHms_BRExp	32	-4,494.74		1.999

The interactions that are statistically significant for the preferred CRPH models are shown in Table 7-5 with different colours for each dataset, where yellow represents statistically significant interactions in the preferred CRPH model for both datasets. Four interactions between the standard deviation and the process strategies were present in both datasets: access time with RAM; fare public transport with both VL and RAM; and travel time public transport with VL. There was only one interaction between the mean and process strategies present in both datasets: headway public transport and VL. The attributes transfer in public transport, % seating probability and density were only available in the Metro Rail dataset.

There are both similarities and differences between the models' interactions, with the evidence suggesting dataset specific effects. In both datasets several interactions were found significant between the mean and standard deviation estimates – traditionally defined under an LPAA heuristic - with both the VL and RAM heuristic. This is not surprising as there is a sense that more complex model forms typically reveal differences in evidence between datasets that is not observed in simpler models because of the absence of additional sources of behavioural variance. What this suggests is that if the behaviourally richer models are an improved explanation of choice making then it becomes more likely that data must be collected in the setting in which a study is focussed, limiting the ability to make inferential statements about the portability of evidence. This may not be such good news for practitioners who are looking for evidence of WTP estimates that can be taken from one context and used into another context. Intuitively we have uncovered further support for the view that accounting for more sources of variability in preferences of a sample of choice reduces the ability to transfer evidence.

Table 7-5: Process strategies' interactions in CRPH preferred models for the datasets

	Mean		**Standard deviation**	
	VL	RAM	VL	RAM
Access Time	Northwest			Both
Fare Public Transport			Both	Both
Fuel + Toll Cost Car			Metro Rail	Northwest
Parking Cost Car	Metro Rail		Metro Rail	
Travel Time Public Transport		Northwest	Both	
Travel Time Car	Northwest		Northwest	
Egress Time				
Transfer Public Transport				
Headway Public Transport		Both		
% Seat Public Transport	Metro Rail			
Density Public Transport				Metro Rail

Northwest
Metro Rail
Both

Research Question 3: Is there any relationship between process heterogeneity and taste heterogeneity?

The CRPH method proposes that there is a relationship between the process heterogeneity and taste heterogeneity by including interactions between the parameters' mean and standard deviation as well as different process strategies. The analysis of research question 2 compared the interactions that were statistically significant in the preferred CRPH models, supporting a relationship between process heterogeneity and taste heterogeneity. To establish the extent to which this finding might still apply under a larger set of process heuristics, we estimated a series of CRPH models with alternative process rules and interactions for each dataset as summarised in Table 7-6.

Table 7-6: Process strategies' interactions in the CRPH models for the two datasets

	CRPHm		**CRPHs**		**CRPHms**				**CRPHms_BRExp**			
	Mean		**Std Dev**		**Mean**		**Std Dev**		**Mean**		**Std Dev**	
	VL	RAM	VL	RAM	VL	RAM	VL	RAM	VL	RAM	VL	RAM
Access Time												
Fare Public Transport												
Fuel + Toll Cost Car												
Parking Cost Car												
Travel Time Public Transport												
Travel Time Car												
Egress Time												
Transfer Public Transport												
Headway Public Transport												
% Seat Public Transport												
Density Public Transport												

Northwest Metro Rail Both

For each CRPH model forms there are numerous statistically significant interactions between the process strategies and taste heterogeneity. When including interactions between the mean and process strategies, the model is embedding taste heterogeneity through the process strategies. For interactions between the standard deviation and process strategies, the model also accounts for taste heterogeneity through process strategies. The results show that these relationships are statistically significant and contribute behaviourally appealing evidence in representing choice making in both datasets. Moreover, the findings suggest that the relationship between taste heterogeneity and process heterogeneity is attribute-specific and should not be considered common between all the attributes. This is shown in Table 7-6 where each attribute is represented by a different combination of interactions.

Research Question 4: How do the various behavioural elements of the integrated choice process influence key behavioural outputs such as willingness to pay estimates and confidence intervals?

The median willingness to pay (WTP) estimates and confidence intervals for the preferred models were presented and compared in Section 5 and Section 6 for each dataset. The median WTP is the main focus (in contrast to mean estimates) of the analysis because when estimating all parameters as random, the mean WTP is highly dependent on the draws, and thus is not very stable; in fixed parameters models the mean and median WTP are equivalent (more details in Section 3.7). We compare the main findings for the attributes present in both datasets: travel times, access time, egress times, and headway. The first part of the question refers to the influence of behavioural refinements and/or overt experience on these estimates, and the second part to the influence of process heterogeneity. Table 7-7 presents the median WTP estimates for the attributes on each dataset relative to year 2009[14], where the last column represents the percentage difference between the datasets. There are statistically significant differences between the datasets, especially in the access times, egress times and headway. The egress times were presented very differently in the surveys, where the Metro Rail survey presented more detailed egress time information (i.e., walk time, public transport time, car pick up or taxi time), and the Northwest survey presented it only as the total time (i.e., time getting from main mode to destination). This we believe is contributing to difference in evidence between the datasets. The travel times in bus and car were also presented differently in the Metro Rail survey, with different levels of variation. The remaining attributes were relatively similar in both surveys. The results show that there are statistically significant differences in the median WTP estimates between the datasets.

---

[14] Calculated using the annual inflation rate provided by the Reserve Bank of Australia. Respondents were asked for the interval that best represented their income, so a uniform distribution was assumed within each intervals to include inflation.

Table 7-7: Median WTP estimates for the different models in each dataset relative to year 2009

Attribute	Model	WTP median estimates		
		Metro Rail	Northwest	% change (Metro Rail - Northwest)
Public Transport Travel Time ($ per person hour)	LPAA_MML_BRExp	5.52	7.68	-39%
	VL_MML_Exp	6.43	6.33	2%
	RAM_MML_Exp	6.36	5.09	20%
	PDP_BRExp	8.41	12.11	-44%
	CRPHms_BRExp	5.20	11.76	-126%
Car Travel Time ($ per person hour)	LPAA_MML_BRExp	15.26	11.06	28%
	VL_MML_Exp	13.94	17.95	-29%
	RAM_MML_Exp	8.80	33.87	-285%
	PDP_BRExp	23.60	31.03	-31%
	CRPHms_BRExp	13.74	22.94	-67%
Public Transport Access Time ($ per person hour)	LPAA_MML_BRExp	5.98	7.61	-27%
	VL_MML_Exp	6.46	5.18	20%
	RAM_MML_Exp	7.47	4.07	46%
	PDP_BRExp	11.31	9.64	15%
	CRPHms_BRExp	3.98	7.13	-79%
Public Transport Egress Time ($ per person hour)	LPAA_MML_BRExp	7.28	6.09	16%
	VL_MML_Exp	9.89	3.64	63%
	RAM_MML_Exp	9.03	3.18	65%
	PDP_BRExp	14.37	7.43	48%
	CRPHms_BRExp	7.70	8.79	-14%
Car Egress Time ($ per person hour)	LPAA_MML_BRExp	18.08	7.81	57%
	VL_MML_Exp	26.96	10.41	61%
	RAM_MML_Exp	14.74	15.41	-5%
	PDP_BRExp	32.89	15.65	52%
	CRPHms_BRExp	14.44	13.86	4%
Headway ($ per person minute)	LPAA_MML_BRExp	0.03	0.06	-61%
	VL_MML_Exp	0.04	0.11	-178%
	RAM_MML_Exp	0.04	0.08	-74%
	PDP_BRExp	0.06	0.06	-6%
	CRPHms_BRExp	0.04	0.10	-181%

*Colour scale represents a decrease (red) or increase (green) in the WTP estimates in the Metro Rail data models relative to the Northwest data models. A darker colour tone represents a larger increase or decrease.

a) *Behavioural refinements and/or experience*

Table 7-8 presents the percentage variation in the WTP median and standard error when including behavioural refinements and/or overt experience in the different models. There is no clear pattern whether the estimates or standard errors are higher or lower when adding behavioural refinements and/or experience. However, the results in both datasets show that there are significant changes in the estimates when considering behavioural refinements and/or experience.

Table 7-8: Influence of behavioural refinements and/or experience in WTP median and standard error estimates

Model	Attribute	Metro Rail		Northwest	
		% change in median	% change in standard error	% change in median	% change in standard error
LPAA MML	Public Transport Travel Time	-11%	-20%	-1%	-4%
	Car Travel Time	6%	-10%	-32%	-67%
	Public Transport Access Time	-15%	-20%	5%	-5%
	Public Transport Egress Time	-20%	-7%	-3%	-5%
	Car Egress Time	-37%	-29%	-26%	-61%
	Headway	3%	-5%	-32%	-8%
VL MML	Public Transport Travel Time	15%	2%	-3%	2%
	Car Travel Time	2%	10%	-10%	-10%
	Public Transport Access Time	8%	2%	-15%	-3%
	Public Transport Egress Time	16%	-2%	-6%	-6%
	Car Egress Time	-15%	-26%	-8%	-14%
	Headway	14%	2%	-13%	-4%
RAM MML	Public Transport Travel Time	7%	22%	3%	31%
	Car Travel Time	-36%	-22%	63%	13%
	Public Transport Access Time	13%	22%	-18%	-1%
	Public Transport Egress Time	-8%	-24%	11%	96%
	Car Egress Time	-55%	-60%	106%	108%
	Headway	21%	6%	24%	19%
PDP	Public Transport Travel Time	22%	33%	4%	55%
	Car Travel Time	30%	-19%	17%	-13%
	Public Transport Access Time	10%	33%	6%	6%
	Public Transport Egress Time	15%	18%	-12%	35%
	Car Egress Time	-22%	-21%	-61%	-64%
	Headway	56%	91%	-1%	-23%
CRPH	Public Transport Travel Time	-15%	-6%	-2%	-9%
	Car Travel Time	60%	317%	-9%	-7%
	Public Transport Access Time	-7%	-6%	-2%	-1%
	Public Transport Egress Time	-13%	1%	0%	-33%
	Car Egress Time	55%	48%	-19%	-36%
	Headway	28%	-22%	17%	-9%

*Colour scale represents a decrease (red) or increase (green) in the median WTP and on its standard error when including behavioural refinements and/or experience. A darker colour tone represents a larger increase or decrease.

b) Process homogeneity versus process heterogeneity

Table 7-9 presents the percentage change in the median WTP estimates when considering process heterogeneity relative to the different process homogeneity models. In the Metro Rail dataset, there is a clear increase in the median WTP when using the PDP approach to allow for process heterogeneity versus all the process homogeneity models, and a decrease when using the CRPH approach (except for the car travel time relative to the VL model and the public transport egress time and headway relative to the LPAA model, where there are minor increases).

In the Northwest dataset there is an increase in the majority of the median WTP estimates when considering process heterogeneity relative to all the process homogeneity models, with a few exceptions. There appears to be no common pattern between the datasets in the median WTP estimates when assessing process heterogeneity versus process homogeneity. This suggests, given the proviso of data differences, that the proposed CRPH method will not always lead to higher or lower WTP estimates; however they provide a statistically improved and behaviourally more appealing way to represent preferences that underlie choice making.

Table 7-9: Influence on the median WTP estimates when considering process heterogeneity versus process homogeneity

Attribute	Process Homogeneity Model	% change in median WTP estimates when considering process heterogeneity			
		Metro Rail		Northwest	
		PDP	CRPH	PDP	CRPH
Public Transport Travel Time ($ per person hour)	LPAA_MML_BRExp	52%	-6%	58%	53%
	VL_MML_Exp	31%	-19%	91%	86%
	RAM_MML_Exp	32%	-18%	138%	131%
Car Travel Time ($ per person hour)	LPAA_MML_BRExp	55%	-10%	181%	107%
	VL_MML_Exp	69%	-1%	73%	28%
	RAM_MML_Exp	168%	56%	-8%	-32%
Public Transport Access Time ($ per person hour)	LPAA_MML_BRExp	89%	-33%	27%	-6%
	VL_MML_Exp	75%	-38%	86%	38%
	RAM_MML_Exp	51%	-47%	137%	75%
Public Transport Egress Time ($ per person hour)	LPAA_MML_BRExp	97%	6%	22%	44%
	VL_MML_Exp	45%	-22%	104%	141%
	RAM_MML_Exp	59%	-15%	134%	177%
Car Egress Time ($ per person hour)	LPAA_MML_BRExp	82%	-20%	101%	78%
	VL_MML_Exp	22%	-46%	50%	33%
	RAM_MML_Exp	123%	-2%	2%	-10%
Headway ($ per person minute)	LPAA_MML_BRExp	63%	6%	8%	85%
	VL_MML_Exp	46%	-5%	-44%	-4%
	RAM_MML_Exp	26%	-18%	-23%	32%

*Colour scale represents a decrease (red) or increase (green) in the median WTP when considering process heterogeneity relative to process homogeneity. A darker colour tone represents a larger increase or decrease.

The Conceptual Framework

Given the evidence emanating for the responses to the research questions, we can now assess the contribution of the initially proposed conceptual framework (presented in Chapter 1), to improve the way in which preference revelation in choice models can be enhanced. Drawing on the relationships summarised in Figure 7-2, our findings align well with what this framework proposes, with evidence supporting behaviourally relevant roles for multiple process strategies, behavioural refinements and experience, each in turn offering additional bases of understanding preferences.

The results using the Northwest dataset support a position where risk attitudes are present even where there is limited or no variability in the levels of the attributes, encouraging the assessment of behavioural refinements despite the characteristics of the design of the available choice experiments. Overt experience also has an important role to play in decision-making. The empirical evidence supports McFadden's (2001) call for more effort in building in process rules and experience in choice modelling. Specifically we offer new evidence on how experience, multiple decision process strategies and behavioural refinements all interact. When adding more behavioural relevance (often seen a modelling complexity) into discrete choice models through random parameters and process heterogeneity, the inclusion of additional behavioural refinements may not be necessary.

The preferred preference revelation model form, *conditioning of random process heterogeneity* (CRPH), supports a behavioural paradigm in which individuals use more than one process heuristic in decision making, supporting heterogeneity in processing information related to alternatives on offer. The impact on important behavioural outputs such as willingness to pay is profound and has important policy relevance in project appraisal.

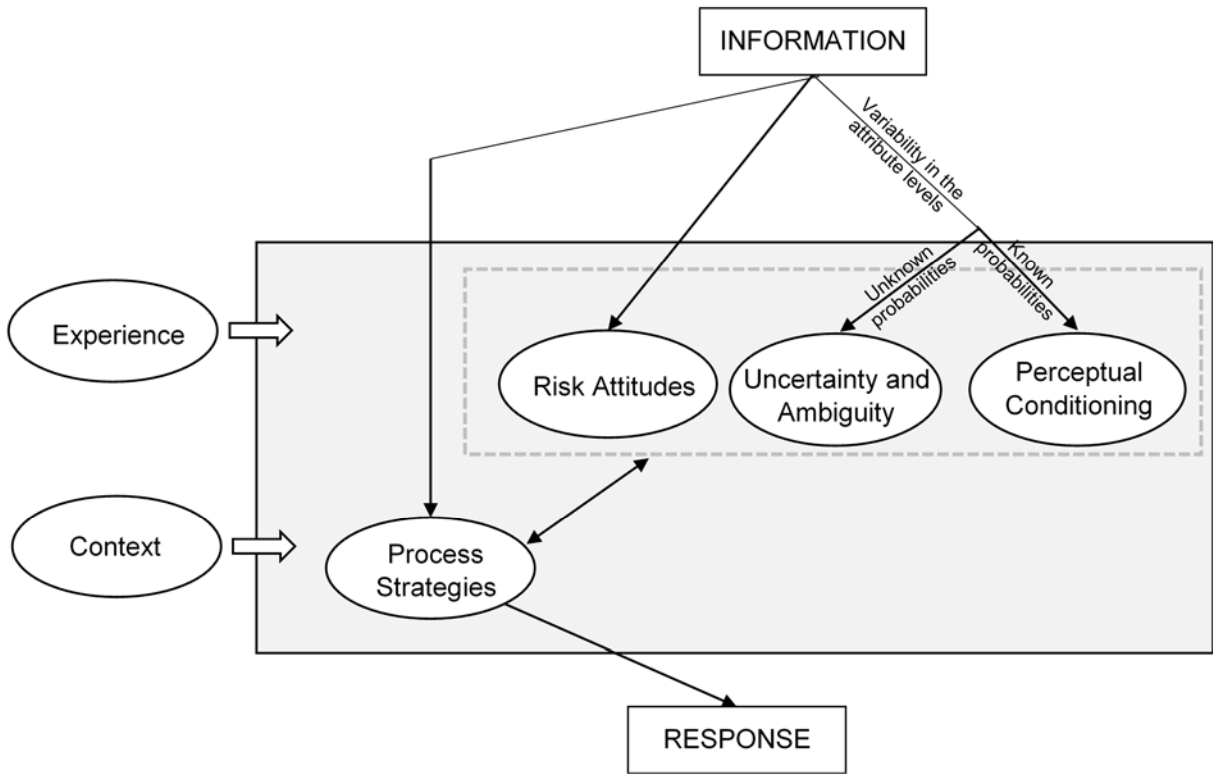


Figure 7-2: Proposed conceptual framework for decision making

In summary, we recommend that a choice study should allow for taste heterogeneity through the standard random parameter specification as well as an overlay (or interaction) with one or more process heuristics, overt experience and behavioural refinements, especially with risk attitudes (we suggest to test for the inclusion of perceptual conditioning). What this tends to do is to provide a better representation of preferences which is translated into significantly different (higher or lower) median estimates of key attributes such as travel time savings. The dataset requirements to be able to include process heterogeneity are more demanding since greater variance in attribute levels is required to identify the various contributions. The two datasets used in this thesis were constructed through a Bayesian-efficient design using D-error as the optimality criteria. Therefore, the dataset required for further inquiry and practical application only need to have enough variability in the attribute levels. If any of the attributes is presented with more than one levels of occurrence, then perceptual conditioning can and should be tested in those attributes. Otherwise, only risk attitudes can and should be tested. In terms of data collection, it is strongly advised to include questions regarding individual experience in the alternatives presented, since this is the only way the experience can be included in the modelling. No questions regarding the process strategies used by respondents are necessary to estimate this type of models, although they could be asked to help guide the selection of process strategies but this was not studied in this research.

7.3. Areas for Future Research

The discussion in this thesis on the contribution of multiple process strategies, behavioural refinements and overt experience in choice modelling, as well as the confoundment between process heterogeneity and taste heterogeneity, suggests several lines of future research.

One useful extension of this research is the assessment of other process strategies such as attribute non-attendance, extremeness aversion, majority of confirming dimensions, and random regret minimisation. This is particularly relevant when considering the CRPH method proposed in this research, where the interaction between process strategies was found to be a key feature of improved behavioural models designed to reveal preferences in choice making. Integrating other process strategies would enrich the use of this method.

Another extension would be the comparison of the proposed model form to other techniques that have been used in literature to combine process strategies, such as multiple heuristics weighting models, or to other simpler functions that argue that a process strategy acts upon a subset of the attributes.

Another area of research relates to data collection. It would be interesting to analyse if respondents are aware of the process strategies they use by asking them different questions. These responses can then be included or excluded in the modelling, similar to what has been conducted using attribute non-attendance (stated or inferred ANA). It will be interesting to investigate how different surveys in terms of the attributes, their levels and the questions influence the use of process strategies being used.

The model proposed in this research considers overt experience conditioning the utility function. Possible endogeneity issues are acknowledged and addressed by the panel data consideration in the form of random parameters. However, this could be investigated in more detail using different model forms along the lines of causal inference, which has rarely been used in travel demand modelling (Brathwaite and Walker, 2018).

The application of the method investigated herein in different choice studies would aid our understanding of their appeal (given we have tested these models on two datasets only) in different contexts, not only in terms of the geographical location and the socio-economic characteristics of respondents, but also in the design of the choice experiment. This will allow the approach to be assessed in studies where the data can be pooled to control for design differences, which we know do to varying degrees impact on WTP estimates (see Hensher et al. 2015). The methods should also be investigated beyond transport studies.

7.4. Concluding remarks

This thesis has proposed and tested the role that different choice model forms (from relatively simple to complex forms) might play in improving our understanding of sources of influence on preference revelation in choice making. For the first time, we have simultaneously integrated into a discrete choice model multiple decision process strategies, risk attitudes, perceptual conditioning and overt experience, and investigated the relationship between the richer behavioural paradigms embedded in process heterogeneity and taste heterogeneity (commonly defined through random parameters). Two datasets were used to provide a more generalised overview of the outcomes under these different behavioural assumptions.

One of the most important findings is that when process heterogeneity is accounted for through specific heuristics such as value learning, less behavioural refinements and overt experience may not been needed to be incorporated as explicitly influencing sources. This helps in identifying appealing parsimonious preference expressions in choice models. When taste heterogeneity is overlayed in these more parsimonious models through random parameters, we find that the interaction between random parameters and processing heuristics adds new insights into the relationship between preference heterogeneity and process heterogeneity. These two phenomena are correlated and hence behaviourally condition each other in important ways. This resulted in a new functional specification of the preference expression, referred to as *conditioning of random process heterogeneity (CRPH)*. The evidence herein is strong enough to suggest empirically that there exists (in two datasets at least) a significant attribute-specific relationship between process strategies and random parameters.

References

Abdellaoui, M., L'Haridon, O., Zank, H., 2010. Separating curvature and elevation: A parametric probability weighting function. J. Risk Uncertain. 41, 39–65. https://doi.org/10.1007/s11166-010-9097-6

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Alberini, A., Kanninen, B., Richard, T., 2017. Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data Author (s): Anna Alberini , Barbara Kanninen and Richard T . Carson Published by : University of Wisconsin Press Stable URL : http://www.jstor.org/stable/3147170 Modelin 73, 309–324.

Andersen, S., Harrison, G.W., Hole, A.R., Lau, M., Rutström, E.E., 2012. Non-linear mixed logit. Theory Decis. 73, 77–96. https://doi.org/10.1007/s11238-011-9277-0

Ariely, D., Loewenstein, G., Prelec, D., 2003. "Coherent Arbitrariness": stable demand curves without stable preferences. Q. J. Econ. 1996–1997. https://doi.org/http://www.mitpressjournals.org/loi/qjec

Arrow, K.J., 1965. The theory of risk aversion, in: In: Aspects of the Theory of Risk Bearing. Yrjo Jahnssonin Saatio, Helsinki. Reprinted in: Essays in the Theory of Risk Bearing. Markham, Chicago 90–109 (1971). ISBN: 0841020019.

Arrow, K.J., 1963. Liquidity preference. Lecture VI in Lecture Notes for Economics 285. Econ. Uncertainty, pp. 33-53, undated, Stanford Univ.

Balbontin, C., Hensher, D.A., Collins, A.T., 2017a. Do familiarity and awareness influence voting intention: The case of road pricing reform? J. Choice Model. 1–16. https://doi.org/10.1016/j.jocm.2017.01.005

Balbontin, C., Hensher, D. a., Collins, A.T., 2017b. Integrating attribute non-attendance and value learning with risk attitudes and perceptual conditioning. Transp. Res. Part E Logist. Transp. Rev. 97, 172–191. https://doi.org/10.1016/j.tre.2016.11.002

Bateman, I.J., Burgess, D., Hutchinson, W.G., Matthews, D.I., 2006. Preference learning versus coherent arbitrariness: NOAA guidelines or a Learning Design Contingent Valuation (LDCV). Work. Pap. - Cent. Soc. Econ. Res. Glob. Environ. 55, 1–25. https://doi.org/10.1016/j.jeem.2007.08.003

Bateman, I.J., Carson, R.T., Day, B., Dupont, D., Louviere, J.J., Morimoto, S., Scarpa, R., Wang, P., 2008. Choice set awareness and ordering effects in discrete choice experiments. CSERGE Work. Pap. EDM 08-01.

Bates, J., 1954. A Model for the Science of Decision. Philos. Sci. 21, 326–339.

Bell, D.R., Lattin, J.M., 2000. Looking for Loss Aversion in Scanner Panel Data: The Confounding Effect of Price Response Heterogeneity. Mark. Sci. 19, 185–200. https://doi.org/10.1287/mksc.19.2.185.11802

Ben-Akiva, M., Bolduc, D., 1996. Multinomial Probit with a Logit Kernel and a General Parametric Specification of the Covariance Structure.

Ben-Akiva, M., Morikawa, T., 1990. Estimation of switching models from revealed preferences and stated intentions. Transp. Res. Part A Gen. 24, 485–495. https://doi.org/10.1016/0191-2607(90)90037-7

Bernoulli, D., 1738. Specimen Theoriae Novae de Mensure Sortis. Cem. Acad. Sci. Imperalis Petropolitanae V, 175–192.

Bierlaire, M., 2016. PythonBiogeme : a short introduction. Rep. TRANSP-OR 160706, Ser. Biogeme. Transp. Mobil. Lab. Sch. Archit. Civ. Environ. Eng. Ec. Polytech. Fédérale Lausanne, Switzerland.

Blanchard, O.J., Fischer, S., 1989. Lectures on macroeconomics. MIT press.

Bliemer, M.C.J., Rose, J.M., 2013. Confidence intervals of willingness-to-pay for random coefficient logit models. Transp. Res. Part B Methodol. 58, 199–214. https://doi.org/10.1016/j.trb.2013.09.010

Brathwaite, T., Walker, J.L., 2018. Causal inference in travel demand modeling (and the lack thereof). J. Choice Model. 26, 1–18. https://doi.org/10.1016/j.jocm.2017.12.001

Camerer, C.F., Ho, T.H., 1994. Violations of the betweenness axiom and nonlinearity in probability. J. Risk Uncertain. 8, 167–196. https://doi.org/10.1007/BF01065371

Cameron, T.A., Quiggin, J., 1994. Estimation Using Contingent Valuation Data from a “Dichotomous Choice with Follow-Up” Questionnaire: Reply. J. Environ. Econ. Manage. 27, 218–234. https://doi.org/10.1006/jeem.1998.1026

Campbell, D., Hensher, D.A., Scarpa, R., 2014. Bounding WTP distributions to reflect the “actual” consideration set. J. Choice Model. 11, 4–15. https://doi.org/10.1016/j.jocm.2014.02.004

Campbell, D., Hensher, D.A., Scarpa, R., 2012. Cost thresholds, cut-offs and sensitivities in stated choice analysis: Identification and implications. Resour. Energy Econ. 34, 396–411. https://doi.org/10.1016/j.reseneeco.2012.04.001

Campbell, D., Hutchinson, W.G., Scarpa, R., 2008. Incorporating discontinuous preferences into the analysis of discrete choice experiments. Environ. Resour. Econ. 41, 401–417. https://doi.org/10.1007/s10640-008-9198-8

Cantillo, V., Heydecker, B., Ortúzar, J. de D., 2004. A discrete choice model incorporating thresholds forperception in attribute values.

Cantillo, V., Ortuzar, J.D., Williams, H., 2007. Modeling discrete choices in the presence of inertia and serial correlation. Transp. Sci. 41, 195–205. https://doi.org/10.1287/trsc.1060.0178

Carlsson, F., Kataria, M., Lampi, E., 2008. Ignoring attributes in choice experiments. Proc. AEARE Conf. Gothenbg. 2008.

Carlsson, F., Martinsson, P., 2006. How much is too much? An investigation of the effect of the number of choice sets, starting point and the choice of bid vectors in choice experiments. Work. Pap. 1–21.

Carson, K.S., Chilton, S.M., Hutchinson, W.G., 2009. Necessary conditions for demand revelation in double referenda. J. Environ. Econ. Manage. 57, 219–225. https://doi.org/10.1016/j.jeem.2008.07.005

Carson, R., Groves, T., 2011. Incentive and information properties of preference questions: commentary and extensions, in: The International Handbook on Non-Market Environmental Valuation. Edward Elgar Northampton, MA., pp. 300–321.

Carson, R., Groves, T., List, J., Machina, M., 2004. Probabilistic influence and supplemental benefits: A field test of the two key assumptions underlying stated preferences. Pap. Present. NBER Public Econ. Work. Palo Alto.

Carson, R.T., Groves, T., 2007. Incentive and informational properties of preference questions. Environ. Resour. Econ. 37, 181–210. https://doi.org/10.1007/s10640-007-9124-5

Carson, R.T., Wilks, L., Imber, D., 1994. Valuing the Preservation of Australia’s Kakadu Conservation Zone. Oxf. Econ. Pap. 46, 727–749.

Chen, Y., Sun, Y., 2003. Age Differences in Financial Decision-Making: Using Simple Heuristics. Educ. Gerontol. https://doi.org/10.1080/713844418

Chernev, A., 2004. Extremeness Aversion and Attribute-Balance Effects in Choice. J. Consum. Res. 31, 249–263. https://doi.org/10.1086/422105

Chorus, C., 2012. Random Regret Minimization: An Overview of Model Properties and

Empirical Evidence. Transp. Rev. 32, 75–92. https://doi.org/10.1080/01441647.2011.609947

Chorus, C.G., 2010. A new model of random regret minimization. Eur. J. Transp. Infrastruct. Res. 10, 181–196.

Chorus, C.G., Arentze, T. a., Timmermans, H.J.P., 2008. A Random Regret-Minimization model of travel choice. Transp. Res. Part B Methodol. 42, 1–18. https://doi.org/10.1016/j.trb.2007.05.004

Chorus, C.G., Rose, J.M., Hensher, D. a., 2013. Regret minimization or utility maximization: It depends on the attribute. Environ. Plan. B Plan. Des. 40, 154–169. https://doi.org/10.1068/b38092

Collins, A.T., 2012. Attribute nonattendance in discrete choice models: measurement of bias, and a model for the inference of both nonattendance and taste heterogeneity. PhD Thesis, ITLS, Univ. Sydney.

Collins, A.T., Hensher, D.A., 2015. The influence of varying information load on inferred attribute non-attendance, in: Rasouli, S., Timmermans, H.J.P. (Eds.), Bounded Rational Choice Behavior: Applications in Transport. Emerald Group Publishing, UK, pp. 73–94.

Collins, A.T., Rose, J.M., Hensher, D.A., 2013a. A model for the simultaneous inference of attribute nonattendance and taste heterogeneity. Work. Pap.

Collins, A.T., Rose, J.M., Hensher, D. a., 2013b. Specification issues in a generalised random parameters attribute nonattendance model. Transp. Res. Part B Methodol. 56, 234–253. https://doi.org/10.1016/j.trb.2013.08.001

Couch, A., Keniston, K., 1960. Yeasayers and naysayers: Agreeing response set as a personality variable. J. Abnorm. Soc. Psychol. 60, 151–174. https://doi.org/10.1037/h0040372

Daly, A., Hess, S., Train, K., 2012. Assuring finite moments for willingness to pay in random coefficient models. Transportation (Amst). 39, 19–31. https://doi.org/10.1007/s11116-011-9331-3

Day, B., Pinto, J.L., 2010. Ordering anomalies in choice experiments. J. Environ. Econ. Manage. 59, 271–285. https://doi.org/10.1016/j.jeem.2010.03.001

DeShazo, J.R., 2002. Designing Transactions without Framing Effects in Iterative Question Formats. J. Environ. Econ. Manage. 43, 360–385. https://doi.org/10.1006/jeem.2000.1185

Domenecich, T.A., McFadden, D., 1975. Urban Travel Demands: A Behavioural Analysis. North-Holland, Amsterdam.

Edwards, W., 1961. Behavioral decision theory1,2. Annu. Rev. Psychol. 12, 473–498.

Eeckhoudt, L., Gollier, C., Schlesinger, H., 2005. Economic and financial decisions under risk, in: Economic and Financial Decisions under Risk. Princeton University Press, pp. 3–25.

Fasolo, B., McClelland, G.H., Todd, P.M., 2007. Escaping the tyranny of choice: When fewer attributes make choice easier 7, 13–26. https://doi.org/10.1177/1470593107073842

Fiebig, D.G., Keane, M.P., Louviere, J., Wasi, N., 2010. The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. Mark. Sci. 29, 393–421. https://doi.org/10.1287/mksc.1090.0508

George, E., McCulloch, R., 1993. Variable selection via Gibbs sampling. J. Am. Stat. Assoc. 88, 881–889. https://doi.org/10.1080/01621459.1993.10476353

Gigerenzer, G., Todd, P.M., 2008. Chapter 105 Rationality the Fast and Frugal Way: Introduction. Handb. Exp. Econ. Results. https://doi.org/10.1016/S1574-0722(07)00105-9

Gigerenzer, G., Todd, P.M., Group, A.R., 1999. Simple Heuristics that Make Us Smart, Oxford Uni. ed.

Gilboa, I., Pazgal, A., 2001. Cumulative Discrete Choice. Mark. Lett. 12, 119–130. https://doi.org/10.1023/A:1011134718403

Gilboa, I., Schmeidler, D., 1995. Case-Based Decision Theory. Q. J. Econ. 110, 605–639.

Gilbride, T.J., Allenby, G.M., Brazell, J.D., 2006. Models for Heterogeneous Variable Selection. Am. Mark. Assoc. 43, 420–430.

Goldstein, W.M., Einhorn, H.K., 1987. Expression theory and the preference reversal phenomena. Psychol. Rev. 94, 236.

Goodwin, P.B., 1977. Habit and Hysteresis in Mode Choice. Urban Stud. 14, 95–98. https://doi.org/10.1080/00420987720080101

Gourville, J.T., Soman, D., 2007. Extremeness Seeking : When and Why Consumers Prefer the Extremes. HBS Mark. Res. Pap. 1–2.

Gourville, J.T., Soman, D., 2005. Overchoice and Assortment Type: When and Why Variety Backfires. Mark. Sci. 24, 382–395. https://doi.org/10.1287/mksc.1040.0109

Greene, W., 2001. Fixed and Random Effects in Nonlinear Models. WP Stern Sch. Bus. 1–

45.

Greene, W.H., Hensher, D.A., 2003. A latent class model for discrete choice analysis: Contrasts with mixed logit. Transp. Res. Part B Methodol. 37, 681–698. https://doi.org/10.1016/S0191-2615(02)00046-2

Grether, D., Wilde, L., 1984. An Analysis of Conjunctive Choice: Theory and Experiments. J. Consum. Res. https://doi.org/10.1086/208976

Hanemann, W., 1991. Reviewer's comments, in D. Imber, G. Stevenson and L. Wilks, "'A Contingent Valuation Survey of the Kakadu Conservation Zone,'" Resource Assessment Commission, RAC Research Paper No. 3, pp. 187-192, Canberra, Australia.

Harrison, G.W., Rutström, E.E., 2009. Expected utility theory and prospect theory: One wedding and a decent funeral. Exp. Econ. 12, 133–158. https://doi.org/10.1007/s10683-008-9203-7

Heath, C., Tversky, A., 1991. Preference and belief: Ambiguity and competence in choice under uncertainty. J. Risk Uncertain. 4, 5–28. https://doi.org/10.1007/BF00057884

Hensher, D.A., 2010a. Hypothetical bias, choice experiments and willingness to pay. Transp. Res. Part B Methodol. 44, 735–752. https://doi.org/10.1016/j.trb.2009.12.012

Hensher, D.A., 2010b. Attribute Processing, Heuristics, and Preference Construction in Choice Analysis, in: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and the State-of-Practice. pp. 35–69.

Hensher, D.A., 2006a. How do respondents process stated choice experiments? Attribute consideration under varying information load. J. Appl. Econom. 21, 861–878. https://doi.org/10.1002/jae.877

Hensher, D.A., 2006b. Revealing differences in willingness to pay due to the dimensionality of stated choice designs: An initial assessment. Environ. Resour. Econ. 34, 7–44. https://doi.org/10.1007/s10640-005-3782-y

Hensher, D.A., 1976. Valuation of commuter travel time savings: an alternative procedure, in: Heggie, I.G. (Ed.), Modal Choice and Value of Travel Time. Clarendon Press, Oxford, pp. 108–131.

Hensher, D.A., 1975. Perception and Commuter Modal Choice - An Hypothesis. Urban Stud. 101–104. https://doi.org/10.1080/00420987520080091

Hensher, D.A., Balbontin, C., Collins, A.T., 2017. Heterogeneity in decision processes: embedding extremeness aversion, risk attitude and perceptual conditioning in multiple process rules choice making, in: Paper Presented in The Fifth International Choice Modelling Conference 3 – 5 April 2017, Cape Town, South Africa.

Hensher, D.A., Collins, A.T., 2011. Interrogation of Responses to Stated Choice Experiments: Is there sense in what respondents tell us? J. Choice Model. 4, 62–89. https://doi.org/10.1016/S1755-5345(13)70019-8

Hensher, D.A., Collins, A.T., Greene, W.H., 2013a. Accounting for attribute non-attendance and common-metric aggregation in a probabilistic decision process mixed multinomial logit model: a warning on potential confounding. Transportation (Amst). 40, 1003–1020. https://doi.org/10.1007/s11116-012-9447-0

Hensher, D.A., Greene, W.H., 2010. Non-attendance and dual processing of common-metric attributes in choice analysis: A latent class specification. Empir. Econ. 39, 413–426. https://doi.org/10.1007/s00181-009-0310-x

Hensher, D.A., Greene, W.H., 2003. The mixed logit model: the state of practice. Transportation (Amst). 30, 133–176. https://doi.org/10.1023/A:1022558715350

Hensher, D.A., Greene, W.H., Ho, C., 2015a. Random regret minimisation and random utility maximisation in the presence of preference heterogeneity: An empirical contrast. Work. Pap. 1–11.

Hensher, D.A., Greene, W.H., Li, Z., 2011a. Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel time savings. Transp. Res. Part B Methodol. 45, 954–972. https://doi.org/10.1016/j.trb.2011.05.023

Hensher, D.A., Ho, C.Q., 2016. Experience conditioning in commuter modal choice modelling – Does it make a difference? Transp. Res. Part E Logist. Transp. Rev. 95, 164–176. https://doi.org/10.1016/j.tre.2016.09.010

Hensher, D.A., Li, Z., 2012. Valuing travel time variability within a rank-dependent utility framework and an investigation of unobserved taste heterogeneity. J. Transp. Econ. Policy 46, 293–312.

Hensher, D.A., Li, Z., Rose, J.M., 2013b. Accommodating risk in the valuation of expected travel time savings. J. Adv. Transp. https://doi.org/10.1002/atr.160

Hensher, D.A., Rose, J., Greene, W.H., 2005. The implications on willingness to pay of respondents ignoring specific attributes. Transportation (Amst). 32, 203–222. https://doi.org/10.1007/s11116-004-7613-8

Hensher, D.A., Rose, J.M., 2012. The influence of alternative acceptability, Attribute thresholds and choice response certainty on automobile purchase preferences. J. Transp. Econ. Policy 46, 451–468.

Hensher, D.A., Rose, J.M., 2009. Valuation of Travel Time Reliability in an Extended Expected Utility Theory Framework, in: In Keynote Paper Presented at the 12th International Conference on Travel Behavior Research (Vol. 18).

Hensher, D.A., Rose, J.M., 2007. Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study. Transp. Res. Part A Policy Pract. 41, 428–443. https://doi.org/10.1016/j.tra.2006.09.006

Hensher, D.A., Rose, J.M., Collins, A.T., 2013. Understanding buy in for risky prospects : Incorporating degree of belief into the ex ante assessment of support for alternative road pricing schemes. J. Transp. Econ. Policy 47, 453–473.

Hensher, D.A., Rose, J.M., Collins, A.T., 2011b. Identifying commuter preferences for existing modes and a proposed Metro in Sydney, Australia with special reference to crowding. Public Transp. 3, 109–147. https://doi.org/10.1007/s12469-010-0035-4

Hensher, D.A., Rose, J.M., Greene, W.H., 2015b. Applied Choice Analysis - Second Edition. Cambridge University Press. https://doi.org/10.1017/CBO9780511610356

Hensher, D.A., Rose, J.M., Greene, W.H., 2012. Inferring attribute non-attendance from stated choice data: Implications for willingness to pay estimates and a warning for stated choice experiment design. Transportation (Amst). 39, 235–245. https://doi.org/10.1007/s11116-011-9347-8

Hensher, D.A., Rose, J.M., Leong, W., Tirachini, A., Li, Z., 2013c. Choosing Public Transport—Incorporating Richer Behavioural Elements in Modal Choice Models. Transp. Rev. 33, 92–106. https://doi.org/10.1080/01441647.2012.760671

Herriges, J.A., Shogren, J.F., 1996. Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning. J. Environ. Econ. Manage. 30, 112–131. https://doi.org/10.1006/jeem.1996.0008

Hess, S., Rose, J.M., 2007. A latent class approach to modelling heterogeneous information processing strategies in SP studies. Oslo Work. Valuat. Methods Transp. Plan. 1–26.

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., Caussade, S., 2013. It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity. Transportation (Amst). 40, 583–607. https://doi.org/10.1007/s11116-012-9438-1

Hess, S., Stathopoulos, A., Daly, A., 2012. Allowing for heterogeneous decision rules in discrete choice models: An approach and four case studies. Transportation (Amst). 39, 565–591. https://doi.org/10.1007/s11116-011-9365-6

Hess, S., Train, K., 2017. Correlation and scale in mixed logit models. J. Choice Model. 23, 1–8. https://doi.org/10.1016/j.jocm.2017.03.001

Holmes, T.P., Boyle, K.J., 2005. Dynamic Learning and Context-Dependence in Valuation Questions. Land Econ. 81, 114–126. https://doi.org/10.3368/le.81.1.114

Holt, C., Laury, S., 2002. Risk aversion and incentive effects. Am. Econ. Rev. 92, 1644–1655.

Huber, J., Payne, J.W., Puto, C., 1982. Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. J. Consum. Res. 9, 90–98.

Hurwicz, L., 1972. On informationally decentralized systems. In:McGuire CB, Radner R (eds) Decision and organisation. University of Minnesota Press, North-Holland.

Ingersoll, J., 1987. Theory of financial decision making. J. Finance 43, 1–259. https://doi.org/10.2307/2328335

Islam, T., Louviere, J.J., Burke, P.F., 2007. Modeling the effects of including/excluding attributes in choice experiments on systematic and random components. Int. J. Res. Mark. 24, 289–300. https://doi.org/10.1016/j.ijresmar.2007.04.002

Johnston, R.J., Boyle, K.J., Adamowicz, W. (Vic), Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., Vossler, C.A., 2017. Contemporary Guidance for Stated Preference Studies. J. Assoc. Environ. Resour. Econ. 4, 319–405. https://doi.org/10.1086/691697

Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–92. https://doi.org/10.2307/1914185

Kivetz, R., Netzer, O., Srinivasan, V., 2004. Alternative Models for Capturing the Compromise Effect. J. Mark. Res. 41, 237–257. https://doi.org/10.1509/jmkr.41.3.237.35990

Ladenburg, J., Olsen, S.B., 2008. Gender-specific starting point bias in choice experiments: Evidence from an empirical study. J. Environ. Econ. Manage. 56, 275–285. https://doi.org/10.1016/j.jeem.2008.01.004

Lancsar, E., Louviere, J., 2006. Deleting “irrational” responses from discrete choice experiments: A case of investigating or imposing preferences? Health Econ. 15, 797–811. https://doi.org/10.1002/hec.1104

Lawrence, C., Zhou, J., Tits, A., 1994. User’s guide for CFSQP version 2.0: ser’s guide for CFSQP Version 2.0: AC code for solving (large scale) constrained nonlinear (minimax) optimization problems, generating iterates satisfying all inequality constraints. Univ. Maryl.

Leong, W., 2014. PhD Thesis - Embedding Decision Heuristics in Discrete Choice Models : Assessing the Merits of Majority of Confirming Dimensions , Extremeness Aversion , and Reference.

Leong, W., Hensher, D.A., 2015. Contrasts of Relative Advantage Maximisation with Random Utility Maximisation and Regret Minimisation. J. Transp. Econ. Policy 49, 167–186(20).

Leong, W., Hensher, D.A., 2012a. Embedding Decision Heuristics in Discrete Choice Models: A Review. Transp. Rev. 32, 313–331. https://doi.org/10.1080/01441647.2012.671195

Leong, W., Hensher, D.A., 2012b. Embedding multiple heuristics into choice models: An exploratory analysis. J. Choice Model. 5, 131–144. https://doi.org/10.1016/j.jocm.2013.03.001

Leong, W., Hensher, D. a., 2014. Relative advantage maximisation as a model of context dependence for binary choice data. J. Choice Model. 11, 30–42. https://doi.org/10.1016/j.jocm.2014.05.002

Li, B., Hensher, D.A., 2017. Risky weighting in discrete choice. Transp. Res. Part B Methodol. 102, 1–21. https://doi.org/10.1016/j.trb.2017.04.014

Li, Z., Hensher, D., 2011. Prospect Theoretic Contributions in Understanding Traveller Behaviour: A Review and Some Comments. Transp. Rev. 31, 97–115. https://doi.org/10.1080/01441647.2010.498589

Loomes, G., Moffatt, P.G., Sugden, R., 2002. A microeconometric test of alternative stochastic theories of risky choice. J. Risk Uncertain. 24, 103–130. https://doi.org/10.1023/A:1014094209265

Luce, R.D., 1959. Individual choice behaviour. John Wiley & Sons, Inc, New York.

McFadden, D., 2001. Economic Choices. Am. Econ. Assoc. 91, 351–378.

McFadden, D., 1974. Conditional Logit Analysis of Qualititative Choice Behaviour, P. Zarembk. ed. Frontiers in Econometrics. Academic Press, New York.

McFadden, D., 1973. Conditional Logit Analysis of Qualitative Choice Behavior, in: Frontiers in Econometrics, Ed. by P. Zarembka, Wiley, New York. New York: Academic Press, pp. 105–142.

McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. J. Appl. Econom. 15, 447–470. https://doi.org/10.1002/1099-1255(200009/10)15:5<447::AID-

JAE570>3.0.CO;2-1

McIntosh, E., Ryan, M., 2002. Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences. J. Econ. Psychol. 23, 367–382. https://doi.org/10.1016/S0167-4870(02)00081-8

McNair, B.J., Bennett, J., Hensher, D.A., 2011. A comparison of responses to single and repeated discrete choice questions. Resour. Energy Econ. 33, 554–571. https://doi.org/10.1016/j.reseneeco.2010.12.003

McNair, B.J., Hensher, D.A., Bennett, J., 2012. Modelling Heterogeneity in Response Behaviour Towards a Sequence of Discrete Choice Questions: A Probabilistic Decision Process Model. Environ. Resour. Econ. 51, 599–616. https://doi.org/10.1007/s10640-011-9514-6

Mirrlees, J. a, 1971. An Exploration in the Theory of Optimal Taxation. Rev. Econ. Stud. 38, 175–208.

Mitchell, R.C., Carson, R.T., 1989. Using Surveys to Value Public Goods: The Contingent Valuation Method. Resources for the Future, Washington, DC.

Moser, R., Raffaelli, R., 2014. Does attribute cut-off elicitation affect choice consistency? Contrasting hypothetical and real-money choice experiments. J. Choice Model. 11, 16–29. https://doi.org/10.1016/j.jocm.2014.02.003

Nicolau, J.L., 2011. Testing prospect theory in airline demand. J. Air Transp. Manag. 17, 241–243. https://doi.org/10.1016/j.jairtraman.2010.11.001

Oehlert, G.W., 1992. A Note on the Delta Method. Am. Stat. 46, 27. https://doi.org/10.2307/2684406

Ortúzar, J.D.D., Willumsen, L., 2011. Modelling Transport - Fourth Edition. John Wiley and Sons, Chichester.

Payne, J.W., 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. Organ. Behav. Hum. Perform. 16, 366–387. https://doi.org/10.1016/0030-5073(76)90022-2

Plott, C.R., 1996. Rational individual behavior in markets and social choice processes: the discovered preference hypothesis, in: Arrow, K.J., Colombatto, E., Perlman, M., Schmidt, C. (Eds.), Rational Foundations of Economic Behavior. Palgrave Macmillan. https://doi.org/10.1057/9780230389724

Polak, J.W., Hess, S., Liu, X., 2008. Characterising Heterogeneity in Attitudes to Risk in

Expected Utility Models of Mode and Departure Time Choice, in: Presented at the 87th Transportation Research Board Annual Meeting, January 13-17, Washington D.C.

Pratt, J., 1964. Risk aversion in the small and in the large. Econometrica 32, 122–136.

Prelec, D., 1998. The Probability Weighting Function Author. Econometrica 66, 497–527.

Puckett, S.M., Hensher, D.A., 2009. Revealing the extent of process heterogeneity in choice analysis: An empirical assessment. Transp. Res. Part A Policy Pract. 43, 117–126. https://doi.org/10.1016/j.tra.2008.07.003

Quiggin, J., 1982. A theory of anticipated utility. J. Econ. Behav. Organ. https://doi.org/10.1016/0167-2681(82)90008-7

Rasouli, S., Timmermans, H., 2014. Applications of theories and models of choice and decision-making under conditions of uncertainty in travel behavior research. Travel Behav. Soc. 1, 79–90. https://doi.org/10.1016/j.tbs.2013.12.001

Rigby, D., Burton, M., 2006. Modeling disinterest and dislike: A bounded Bayesian mixed logit model of the UK market for GM food. Environ. Resour. Econ. 33, 485–509. https://doi.org/10.1007/s10640-005-4995-9

Rossi, P.E., Allenby, G.M., McCulloch, R., 2006. Bayesian Statistics and Marketing. John Wiley & Sons, Inc, UK.

Russo, J.E., Dosher, B.A., 1983. Strategies for multiattribute binary choice. J. Exp. Psychol. Learn. Mem. Cogn. 9, 676–696. https://doi.org/10.1037/0278-7393.9.4.676

Samuelson, P. a, 1954. the Pure Theory of Public Expenditure. Rev. Econ. Stat. 36, 387–389.

Scarpa, R., Gilbride, T.J., Campbell, D., Hensher, D.A., 2009. Modelling attribute non-attendance in choice experiments for rural landscape valuation. Eur. Rev. Agric. Econ. 36, 151–174. https://doi.org/10.1093/erae/jbp012

Scarpa, R., Rose, J.M., 2008. Design efficiency for non-market valuation with choice modelling: How to measure it, what to report and why. Aust. J. Agric. Resour. Econ. 52, 253–282. https://doi.org/10.1111/j.1467-8489.2007.00436.x

Scarpa, R., Thiene, M., Hensher, D.A., 2010. Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services: Does It Matter? Land Econ. 86, 817.

Senbil, M., Kitamura, R., 2004. Reference Points in Commuter Departure Time Choice: A Prospect Theoretic Test of Alternative Decision Frames. J. Intell. Transp. Syst. 8, 19–31. https://doi.org/10.1080/15472450490437726

Senna, L., 1994. The influence of travel time variability on the value of time. Transportation (Amst). 21, 203–228.

Sharpe, K.M., Staelin, R., Huber, J., 2008. Using Extremeness Aversion to Fight Obesity: Policy Implications of Context Dependent Demand. J. Consum. Res. 35, 406–422. https://doi.org/10.1086/587631

Simon, H.A., 1955. A Behavioral Model of Rational Choice. Q. J. Econ. 69, 99–118. https://doi.org/10.2307/1884852

Simonson, I., Tversky, A., 1992. Choice in Context : Tradeoff Contrast and Extremeness Aversion. J. Mark. XXIX, 281–296.

Slovic, P., Fischhaff, B., Lichtenstein, S., 1977. Behavioural Decision Theory. Annu. Rev. Psychol. 28, 1–39.

Stott, H.P., 2006. Cumulative prospect theory's functional menagerie. J. Risk Uncertain. 32, 101–130. https://doi.org/10.1007/s11166-006-8289-6

Swait, J., 2001. A non-compensatory choice model incorporating attribute cutoffs. Transp. Res. Part B Methodol. 35, 903–928. https://doi.org/10.1016/S0191-2615(00)00030-8

Swait, J., Adamowicz, W., 2001. The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. J. Consum. Res. 28, 135–148. https://doi.org/10.1086/321952

Thayer, M.A., 1981. Contingent Valuation Techniques for Assesing Environmental Impacts: Further Evidence. J. Environ. Econ. Manage. 8, 27–44.

Todd, P.M., Gigerenzer, G., 2007. Environments That Make Us Smart. Curr. Dir. Psychol. Sci. 16, 167–171.

Train, K., 2009. Discrete Choice Methods with Simulation, Books.Google.Com.

Train, K., 1999. Halton Sequences for Mixed Logit. Work. Pap. 1–18.

Tversky, A., 1972. Elimination by Aspects: A Theory of Choice. Psychol. Rev. 79, 281–299.

Tversky, A., Fox, C.R., 1995. Weighing risk and uncertainty. Psychol. Rev. https://doi.org/10.1037/0033-295X.102.2.269

Tversky, A., Kahneman, D., 1992. Advances in Prospect Theory : Cumulative Representation of Uncertainty. J. Risk Uncertain. 5, 297–323.

Tversky, A., Kahneman, D., 1991. Loss Aversion in Riskless Choice: A Reference-Dependent Model. Q. J. Sci. 106, 1039–1061.

Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. Science 185, 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Tversky, A., Simonson, I., 1993. Context- dependent Preferences. Manage. Sci. 39, 1179–1189.

van Cranenburgh, S., Guevara, C.A., Chorus, C.G., 2015. New insights on random regret minimization models. Transp. Res. Part A Policy Pract. 74, 91–109. https://doi.org/10.1016/j.tra.2015.01.008

Von Neumann, J., Morgenstern, O., 1947. Theory of Games and Economic Behavior. Princeton University Press.

Vuong, Q., 1989. Likelihood ratio tests for model selection and non nested hypotheses. Econom. J. Econom. Soc. 57, 307–333.

Weller, P., Oehlmann, M., Mariel, P., Meyerhoff, J., 2014. Stated and inferred attribute non-attendance in a design of designs approach. J. Choice Model. 11, 43–56. https://doi.org/10.1016/j.jocm.2014.04.002

Williams, H.C.W.L., Ortúzar, J. de D., 1982. Behavioural theories of dispersion and the mis-specification of travel demand models. Transp. Res. Part B Methodol. 16, 167–219. https://doi.org/10.1016/0191-2615(82)90024-8

Wu, G., Gonzalez, R., 1999. Nonlinear Decision Weights in Choice Under Uncertainty. Manage. Sci. 45, 74–85. https://doi.org/10.1287/mnsc.45.1.74

Xu, H., Zhou, J., Xu, W., 2011. A decision-making rule for modeling travelers' route choice behavior based on cumulative prospect theory. Transp. Res. Part C Emerg. Technol. 19, 218–228. https://doi.org/10.1016/j.trc.2010.05.009

Young, W., 1984. A non-tradeoff decision making model of residential location choice. Transp. Res. Part A Policy Pract. I, 1–11.

Zellman, E., Kaye-Blake, W., Abell, W., 2010. Identifying consumer decision-making strategies using alternative methods. Qual. Mark. Res. An Int. J. https://doi.org/10.1108/13522751011053635

Appendix

Appendix A1: Metro Rail attribute and attribute levels

Table A-1: Car attribute and attribute levels for Metro Rail study

	Attribute	Unit	# Levels	Levels	Pivot/Rule
Main mode	Quickest trip time	minutes	5	-25%, -20%, -15%, -10%, -5%	Value less than average pivot time
	Travel time on average	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported trip time
	Slowest trip time	minutes	5	10%, 15%, 20%, 25%, 30%	Value more than average pivot time
	Quickest trip time	%	5	10%, 15%, 20%, 25%, 30%	-
	Travel time on average	%	5	40%, 45%, 50%, 55%, 60%	-
	Slowest trip time	%		-	100% - (Travel time on average % + Quickest trip time %)
	Fuel cost	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Base calculated using the average speed and assuming an average fuel efficiency of 8 litres for every 100 km travelled, and fuel of $1 per litre
	Toll cost	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25% or $0, $2, $4, $6, $8	Pivot from reported toll, or if non-toll road, use monetary levels
	Parking cost	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25% or $0, $5, $10, $15, $20	Pivot from reported parking cost, or if no cost reported, use monetary levels
Egress mode	Walk time to final destination	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot off respondent provided level if alternative is indicated as eligible
	Public transport time (including waiting time)	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot off respondent provided level if alternative is indicated as eligible
	Fare (one way)	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot off respondent provided level if alternative is indicated as eligible

Table A-2: Bus attribute and attribute levels for Metro Rail study

	Attribute	Unit	# Levels	Levels	Pivot/Rule
Access mode	Walk time	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported walk time
	Car travel time	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported drive time
	Park and ride cost	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported drive to station and park cost
Main mode	Quickest trip time	minutes	5	-25%, -20%, -15%, -10%, -5%	Value less than average pivot time
	Travel time on average	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported trip time
	Slowest trip time	minutes	5	10%, 15%, 20%, 25%, 30%	Value more than average pivot time
	Quickest trip time	%	5	10%, 15%, 20%, 25%, 30%	-
	Travel time on average	%	5	40%, 45%, 50%, 55%, 60%	-
	Slowest trip time	%		-	100% - (Travel time on average % + Quickest trip time %)
	Fare (one way)	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivoted from respondent provided level
	Number of transfers	number	3	0, 1, 2	-
	Headway	Minutes	6	10, 20, 30, 40, 50, 60	-
	Seating	% to find a seat	7	0%, 10%, 20%, 30%, 40%, 50%, 75%	-
	Standing	Standees per square metre	9	0, 0.49, 0.81, 1.14, 1.79, 2.44, 3.10, 3.75, 4.40	Calculated using capacity of buses with % of occupation
Egress mode	Walk time to final destination	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported time
	Car pick up from stop or station	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported time
	Taxi Fare	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported time

Table A-3: Metro and train attribute and attribute levels for Metro Rail study

	Attribute	Unit	# Levels	Levels	Pivot/Rule
Access mode	Walk time	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported walk time
	Public transport time (including waiting time)	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported public transport time
	Fare (one way)	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported
	Car travel time	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported drive time
	Park and ride cost	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported drive to station and park cost
Main mode	Travel time on average	minutes	5	-25%, -20%, -15%, -10%, -5%	Pivot from reported trip time
	Fare (one way)	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivoted from respondent provided level
	Number of transfers	number	2	0, 1	-
	Headway	minutes	6	2, 4, 6, 8, 10, 12	-
	Seating	% to find a seat	7	0%, 10%, 20%, 30%, 40%, 50%, 75%	-
	Standing	Standees per square metre	10	0, 0.18, 0.56, 1.13, 1.65, 2.22, 2.75, 3.31, 3.84, 4.40	Calculated using capacity of metro with % of occupation
Egress mode	Walk time to final destination	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported time
	Public transport time (including waiting time)	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported time
	Fare (one way)	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported fare
	Train time (including waiting time)	minutes	5	2, 4, 6, 8, 10	-
	Train fare (one way)	AUD$	5	$2.60, $2.80, $3.00, $3.20, $3.60	-
	Car pick up from station	minutes	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported time
	Taxi Fare	AUD$	5	-25%, -12.5%, 0%, 12.5%, 25%	Pivot from reported time

Appendix A2: Northwest attribute and attribute levels

Table A-4: New light rail (NLR), new heavy rail (NHR), new busway (NBW) attribute and attribute levels for inter-regional trips in Northwest study

	Attribute	Unit	# Levels	Levels	Pivot/Rule
Access mode	Walk time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported walk time
	Car time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported car time
	Bus time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported bus time
	Bus fare	AUD$	4	-25%, 0%, 25%, 50% or $1, $2, $3, $4	Pivot from reported bus fare as access mode
Main mode	Fare (one-way) / running cost (for car)	AUD$	4	-25%, 0%, 25%, 50%	Pivot from estimated fares according to origin and destination
	In-vehicle travel time	minutes	4	-25%, 0%, 25%, 50%	Pivot from estimated travel times according to origin and destination
	Frequency of service				
	Time spent transferring at Beecroft station	minutes	4	2,4,6,8 for LR 0,2,4,6 for HR - for BW	-
Egress mode	Time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported egress time

Table A-5: New light rail (NLR), new heavy rail (NHR), new busway (NBW) attribute and attribute levels for intra-regional trips in Northwest study

	Attribute	Unit	# Levels	Levels	Pivot/Rule
Access mode	Walk time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported walk time
	Car time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported car time
	Bus time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported bus time
	Bus fare	AUD$	4	-25%, 0%, 25%, 50% or $1, $2, $3, $4	Pivot from reported bus fare as access mode
Main mode	Fare (one-way) / running cost (for car)	AUD$	4	Up to 4 stations: -25%, 0%, 25%, 50% of $2.20 else, % of $2.80	-
	In-vehicle travel time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported travel time
	Waiting time	minutes	4	8,10,13,15 for NLR 3,4,5,6 for NHR and NBW	-
Egress mode	Time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported egress time

Table A-6: Bus, busway and train attribute and attribute levels for inter and intra-regional trips in Northwest study

	Attribute	Unit	# Levels	Levels	Pivot/Rule
Access mode	Walk time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported walk time
	Car time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported car time
	Bus time	minutes	4	-25%, 0%, 25%, 50% - for bus	Pivot from reported bus time
	Bus fare	AUD$	4	-25%, 0%, 25%, 50% or $1, $2, $3, $4 - for bus	Pivot from reported bus fare as access mode
Main mode	Fare (one-way) / running cost (for car)	AUD$	4	-25%, 0%, 25%, 50%	Pivot from reported main mode fare
	In-vehicle travel time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported main mode travel time
	Waiting time	minutes	4	If reported level ≤ 2: 1, 2, 3, 4 If 3: 2, 3, 4, 5 If 4: 3, 4, 5, 6 If 5: 4, 5, 6, 8 If 6: 5, 6, 8, 9 If 7: 5, 7, 9, 11 If 8: 6, 8, 10, 12 If 9: 7, 9, 11, 14 If 10: 8, 10, 13, 15	Pivot from reported waiting time
Egress mode	Time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported egress time

Table A-7: Car attribute and attribute levels for inter and intra-regional trips in Northwest study

	Attribute	Unit	# Levels	Levels	Pivot/Rule
Main mode	Fare (one-way) / running cost (for car)	AUD$	4	-25%, 0%, 25%, 50%	Pivot from estimated fares according to origin and destination
	Toll cost (one way)	AUD$	4	-25%, 0%, 25%, 50% or $0, $2, $4, $6	Pivot from reported toll, or if non-toll road, use monetary levels
	Parking cost (one day)	AUD$	4	-25%, 0%, 25%, 50% or $0, $2, $4, $6	Pivot from reported parking cost, or if no cost reported, use monetary levels
	In-vehicle travel time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported travel time
Egress mode	Time	minutes	4	-25%, 0%, 25%, 50%	Pivot from reported egress time