School of Earth and Planetary Sciences

Modelling Train Station Choice under Uncertainty for Park and Ride Users

Chunmei Chen

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

May 2019

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Human Ethics (For projects involving human participants/tissue, etc.) The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Research Ethics and Biosafety Office of the University of Western Australia, Approval Number **RA/4/1/5370.**

Chunmei Chen

Signature……………………………………………

Date………17/05/2019 …………………..………

ABSTRACT

Park and Ride (P&R) is an important component of many train systems as it is recognised to be an efficient way to reduce traffic congestion and increase train usage. However, high demand for P&R can result in some less than desirable side-affects, including pressure on the parking at and around the station and congestion on the approach roads, which may result in unreliable travel times to the train station and even over-crowding on the trains. When the capacities of the P&R car parks are well below the levels required to cater for the demand, the unwanted side-effects or problems include uncertainty in finding a space or an uncertain parking search time and illegal or inconvenient, (e.g. to local residents), parking on the surrounding streets. Research to date related to these issues is limited. The aim of this research is to conduct a systematic study to understand the decision-making process of P&R users choosing their departure train station under uncertainty, such as travel time, parking search time and crowding on the train. The research identifies key factors influencing P&R users' choice, measures P&R users' risk attitude for the variation of each uncertain situation and evaluates individuals' preference heterogeneity for station choice.

Four key tasks, to meet the key objectives, were undertaken: 1) develop a novel framework for estimating P&R users' station choice under uncertainty; 2) investigate and identify the key factors, (both certain and uncertain), affecting train station choice for P&R users; 3) develop station choice models under uncertainty for P&R users; and 4) implement and validate the station choice models.

The data used to model station choice were obtained by a stated choice (SC) experiment designed using a D-efficiency approach and validated with an eye tracking experiment, (consisting of a 60Hz Remote Eye Tracking Device (RED) and a laptop). The questionnaire produced from the SP experiment had two alternative stations with 18 attributes and their corresponding attributes levels. In total, more than 600 respondents were involved in the experiment and about 2400 usable questionnaires were obtained.

The station choice models in the research, similar to those in previous literature, were developed within a discrete choice theory framework but different models were applied. For example, mixed logit (ML) was used to develop the sub-model of station choice under parking search time uncertainty and the overall station choice model, and a latent class (LC) model was used to examine the effect of crowding uncertainty on station choice. The station choice models developed with the ML or LC approach can not only be used to calculate choice probabilities but also to reveal the effect of individuals' preference heterogeneity on station choice from different viewpoints. Moreover, the utility functions in the station choice models were established with a combination of cumulative prospect theory (CPT) and extended expected utility theory (EEUT) with the mean-variance approach, which not only successfully explained the effect of uncertainty on station choice for P&R users but also made the station choice models more realistic by incorporating both subjective and objective information.

Six factors, (travel time to the departure station, parking search time, crowding on trains, safety, train frequency and ticket fare), were identified as key factors influencing station choice, the first three being uncertain factors. Based on this, four models of station choice were developed including three sub-models of station choice under uncertainty related to travel time, parking search time and crowding, and an overall station choice model under uncertainty combining the above three separate factors. The results showed that P&R users' risk attitude did have an effect on station choice and the attitude towards the variability of travel time to station, parking search time and crowding is risk averse. Moreover, P&R users who have personally experienced higher travel time variations and greater differences between perceived and estimated travel times tend to be more risk averse towards their station choice under travel time variability than those who have experienced or perceived less travel time variability. Furthermore, the higher the risk aversion that commuters accessing a train station displayed, the fewer boarding at that station and correspondingly, the less crowded the trains stopping at that station.

This research has developed a systematic methodology for modelling station choice under uncertainty for P&R users. The results could be used to support planning decisions on the location, price and capacity of P&R facilities, and provide evidence for evaluating upcoming investment decisions.

ACKNOWLEDGEMENTS

While studying for my PhD, I have received a considerable amount of support, encouragement and suggestions from a number of people, to whom I would like to express my sincere gratitude.

Firstly, I would like to express my foremost appreciation to my main supervisor, Associate Professor Jianhong (Cecilia) Xia, for her continuous support of my PhD study and for her patience, passion and immense knowledge. Her guidance has helped me throughout my research and in the writing of this thesis. I cannot imagine having a better supervisor and mentor for my PhD study. It is because of her encouragement and support that I have been able to overcome many difficulties and complete my study. Thank you very much, Cecilia, for your patience, passion and generosity.

As well as my main supervisor, I would like to thank my associate supervisor, Assistant Professor Brett Smith, from the University of Western Australia and co-supervisor, Senior Lecturer Todd Robinson, for taking the time from their busy schedules to meet and discuss my study and for offering so much insightful guidance and advice on experimental design, data collection and model development.

My sincere thanks also goes to all team members in the Arc Linkage project, (LP110201150), including Emeritus Professor John Taplin, Associate Professor Doina Olaru, Assistant Professor Brett Smith, Dr Ying Huang from the Business School, University of Western Australia, Dr Renlong Han and Gary McCarney from the Departments of Transport and Planning and Dr Ting (Grace) Lin from Curtin University, for contributing their valuable time in providing suggestions and discussing the development and implementation of my experimental design and data collection. Thank you all very much.

Special thanks go to Professor Robert Lewis from Fordham University and Professor Paláncz Béla from Budapest University of Technology and Economics for offering me advice and suggestions on how to develop the station choice models.

I would like to express my deepest appreciation to the Arc Linkage Gant (LP110201150) and the APA (I) scholarship, which has financially supported my PhD study.

A special thanks also to Gary McCarney for helping me proofread and edit my thesis and for providing suggestions for my research.

I give my most sincere thanks to all the staff and postgraduate students in the School of Earth Sciences, Curtin University for providing help and encouragement in times of need. Thank you to Professor Bert Veenendaal, Associate Professor Jon Kirby, Lori Patterson and Caroline Rockliff for their assistance during the life of my study. Special thanks also go to Dr Saud Aboshiqah, Dr Xiaoying Wu, Dr Xin Liu, Dr Ting(Grace) Lin, Dr Qian Sun, Feiyan Yu, Zhiwen Song, Qian Yu, Heng zhou, Yuchen Liu and Muhammad Thaqif Muhammed Najmuddin, and especially the members of lab 301A, for providing a tolerant, friendly and very collaborative environment.

Lastly, but not the least, I would like to thank my family and friends. Words cannot express how grateful I am to you, my parents, my brothers and my mother-in law, for all of the sacrifices that you have made on my behalf and for supporting me spiritually throughout PhD study. I would also like to thank my beloved husband, Zhonghui. Thank you for supporting me in all my endeavours. To my beloved kids, Yiming and Leyao, I would like to express my thanks for always cheering me up.

RELATED PUBLICATIONS

Refereed Conference Papers:

Chen, C.; Xia, J.; Smith, B.; Han, R. (2014). "Development of a conceptual framework for modelling train station choice under uncertainty for park and ride users". Paper presented at the TRB 93rd Annual meeting, USA

Chen, C.; Xia, J.; Smith, B.; Olaru, D.; Taplin, J.; Han, R. (2015). "Influence of parking on train station choice under uncertainty for park and ride users". Paper presented at the 6th International Conference on Applied Human Factors and Ergonomics (AHFE2015) and the Affiliated Conferences, AHFE 2015, USA

Chen, C.; Xia, J.; Smith, B.; Olaru, D.; Taplin, J.; Han, R. (2017). "Influence of travel time variability on train station choice for park-and-rider users". Paper presented at World Conference on Transport Research - WCTR 2016 Shanghai, China.

Olaru, D., Smith, B., Xia,J., **Chen, C.,** Lin,T., Huang Y., Han, R. (2013), "Ticket to ride": factors affecting park-and-ride travel in Perth, WA, 2013, Paper presented at European Transport Conference, Germany.

Journal papers:

Shao, C., Xia, J. C., Lin, T., Goulias, K. G., **Chen, C**., Han, R. (2015). Logistic regression models for the nearest train station choice: A comparison of captive and non-captive stations. Case Studies on Transport Policy, 3(4), 382-391.

PUBLICATIONS AND PRESENTATIONS BASED ON THIS THESIS

Some of the work presented in this thesis has already been published by the thesis author and co-authors, they are listed as follows:

Conference papers (peer-reviewing and fully refereed)

(1) **Chen**, C., Xia, J., Smith, B., Olaru, D., Taplin, J., & Han, R. (2015). Influence of Parking on Train Station Choice under Uncertainty for park-and-ride Users. Procedia Manufacturing, 3, 5126-5133. http://dx.doi.org/https://doi.org/10.1016/j.promfg.2015.07.537

(2) **Chen**, C., Xia, J., Smith, B., Olaru, D., Taplin, J., & Han, R. (2017). Influence of travel time variability on train station choice for park-and-rider usersWorld Conference on Transport Research - WCTR 2016 Shanghai, held in Shanghai, China, (pp. 2473-2489). http://dx.doi.org:10.1016/j.trpro.2017.05.274

Conference poster (peer-reviewing and fully refereed)

(3) **Chen, C**.; Xia, J.; Smith, B.; Han, R. (2014). “Development of a conceptual framework for modelling train station choice under uncertainty for park and ride users”. Paper presented at the TRB 93rd Annual meeting, USA

Thesis contribution of papers published

Author: Chunmei Chen

Paper	(1)	(2)	(3)
Conception and design	✓	✓	✓
Methodology	✓	✓	✓
Data collection and analysis	✓	✓	✓
Modelling	✓	✓	✓
Interpretation and discussion	✓	✓	✓
Critically revising	✓	✓	✓

I acknowledge that these represent my contribution to the above research output

Signed: Chunmei Chen

Author: Jianhong (Cecilia) Xia

Paper	1	2	3
Conception and design	✓	✓	✓
Methodology	✓	✓	✓
Data collection and analysis			
Modelling			
Interpretation and discussion			
Critically revising	✓	✓	✓

I acknowledge that these represent my contribution to the above research output

Signed:

Author: Brett Smith

Paper	1	2	3
Conception and design			
Methodology			
Data collection and analysis	✓	✓	
Modelling			
Interpretation and discussion			
Critically revising	✓	✓	✓

I acknowledge that these represent my contribution to the above research output

Signed:

Author: Doina Olaru

Paper	1	2
Conception and design		
Data collection and analysis		
Methodology		
Modelling		
Interpretation and discussion		
Critically revising	✓	✓

I acknowledge that these represent my contribution to the above research output

Signed: Olaru

Author: John Taplin (who has passed away)

Paper	1	2
Conception and design		
Data collection and analysis		
Methodology		
Modelling		
Interpretation and discussion		
Critically revising	✓	✓

I acknowledge that these represent my contribution to the above research output

Signed:

Author: Renlong Han

Paper	1	2	3
Conception and design			
Data collection and analysis			✓
Methodology			
Modelling			
Interpretation and discussion			
Critically revising	✓	✓	✓

I acknowledge that these represent my contribution to the above research output

Signed: Renlong Han

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1. 1 Introduction

In many, if not all, cities around the world, the level of private car ownership has been rising, resulting in increased congestion on the roads and higher levels of traffic pollution (Phang & Toh, 2004). A counter-measure, often adopted by responsible agencies, has been to reduce parking supply and/or increase the price of parking within city centres to discourage car trips, primarily by commuters. Hence, more and more commuters need to depend on the public transport (PT) system to get to work, with the train, as a key PT mode, becoming a more and more important alternative for the daily commute (Ginn, 2009).

In many cities, park and ride (P&R) is a key part of the train system and P&R use has been increasing in line with increasing train usage. P&R combines the efficiency of the private car in serving short suburban trips, i.e., from home to the station, with the speed, high capacity and efficiency of the train in providing line haul mass transit to major employment hubs. It is accepted as an efficient way to reduce traffic congestion on roads and our reliance on fossil fuel for transportation. It can also provide the basis for creating competitive economies and liveable inclusive communities. Data from the Department of infrastructure (2005) reveal that P&R use in Australia is higher than in many other countries around the world. Perth, the fourth largest city in Australia, has a relatively high proportion of P&R users about 23,000 per day (Department of Transport, 2010; Public Transport Authority, 2012-2017). As most of the P&R stations fill up within the morning peak, latent P&R demand is likely to be higher than this. Based on a projection by the Department of Transport, Western Australia, the demand for P&R in 2021 could reach nearly 30,000. Moreover, a survey conducted jointly by Curtin University, the University of Western Australia, the Department of Transport and the Public Transport Authority on 2 July 2012 revealed that the distribution of demand for P&R was very uneven. It also showed that about fifty percent of commuters did not choose their nearest station as their departure point, which increased the pressure on the limited supply at some stations and resulted in the under-use of the available P&R supply at others (Shao, Xia, Lin, Goulias, & Chen, 2015). However, P&R users' choice of departure train station is relatively complex and involves many factors including travel time, parking search time, crowding on trains and dropping children at school on the way to work etc. It can vary day-to-day due to uncertainty in the traffic networks, parking situation, the situation on trains etc. Therefore, a systematic study is required to understand the decision making process of P&R users in choosing their departure train station under uncertainty and to provide a basis for matching P&R demand with supply.

Moreover, P&R should, ideally, improve traffic conditions on the freeways running parallel to the train lines with P&R facilities. However, this is often not the case. In Perth, as predicted by Bureau of Transport and Regional Economics [BTRE] (2007), the cost of congestion in Perth could reach $2.1 billion by 2020. Therefore, an adequate choice behavioural model when the road network is congested is essential for evaluating upcoming transport infrastructure investment decisions.

Research in the field is very limited and conducted primarily in North America, the UK and the Netherlands. Moreover, most studies focused on access mode choice and station choice combined. The earliest station choice study can be traced back to the 1970s in North America. Liou and Talvitie (1974) identified the process of station choice as having two sequential stages, i.e., mode choice first then station choice. Based on this choice process, they developed a multinomial logit (MNL) model of station choice. Subsequently, this choice process was recognised by a few authors who studied and tested the effects of different factors, (such as the location of station, access cost, parking supply, etc.), on station choice with nested logit models, cross-nested logit models or MNL models (Adcock, 1997; Davidson & Yang, 1999; Debrezion, Pels, & Rietveld, 2007; Desfor, 1975; Fan, Miller, & Badoe, 1993; Fox, 2005; Harata & Ohta, 1986; Mahmoud, Habib, & Shalaby, 2014). Other researchers studied the prediction of rail passenger ridership with station choice models developed within simple MNL models (Kastrenakes, 1988; Lythgoe & Wardman, 2004; Lythgoe, Wardman, & Toner, 2004).

Factors in the previous literature influencing station choice included the location of a station, the cost of access and parking attributes, (such as parking capacity, parking cost, etc.), which are known (i.e., certain). Uncertain or variable factors, (such as travel time, parking search time, crowding on trains, etc.), were not considered in these models. However, the effects of some uncertain factors on travel choice have been identified. For example, Li, Hensher, and Rose (2010) identified travel time as an uncertain factory and studied the effects of its variability on route choice. Therefore, it is time to explore the influence of the variability of these uncertain factors on station choice to produce more accurate and reliable predictive models.

In order to address this problem, a station choice model focusing on the effects of the attributes only related to P&R users will be developed to better understand P&R users' choice of departure train station under uncertainty.

In summary, this research focuses on developing an adequate station choice model specifically for P&R users in an uncertain environment, which can not only predict P&R demand, (and thereby provide evidence for evaluating upcoming P&R investment decisions), but also update the strategic transport evaluation model (STEM) used for assessing land use and transport policy by the Department of Transport in Perth, Western Australia.

1. 2 Research objectives

This research mainly aims to model the choice of departure train station for P&R users based on the effect of the three uncertain factors identified from data on P&R users' trips between home and departure train station, i.e. travel time to the station, parking search time and the crowding on trains. This study could be used to support planning decisions on the location and capacity of P&R facilities and provide evidence for evaluating upcoming transport infrastructure investment decisions. Moreover, it could assist the Department of Transport to update the STEM model. To achieve these objectives, the key tasks include:

- Identify and attempt to quantify/ measure the key uncertain factors affecting railway station choice.
    - ❖ Analyse P&R users' travel characteristics to explore the key factors affecting their decision for departure train station. Identify and qualify uncertain factors affecting train station choice.
    - ❖ Design a stated preference experiment to collect abound sample..
- Develop a novel framework for estimating P&R users' station choice under uncertainty.
- Develop station choice sub-models for P&R users based on the effects of the uncertain factors.
    - ❖ Measure the effects of variability of the uncertain factors on P&R users' and their risk attitude towards this variability
    - ❖ Explore the relationship between the P&R users' risk attitude and boarding numbers.

- Develop an overall station choice model under uncertainty and evaluate the effects of each uncertain factor on station choice for P&R users.
- Evaluate and validate the whole station choice model.

1. 3 Research significance

The earlier research related to station choice aimed to explore how rail passengers choose their departure train station under certain situations. Train station choice model development, in the past decades, progressed from standard logit models to more complicated cross-nested models within a discrete choice model framework and the inclusion of additional relevant certain variables. In practice, the choice situations are full of uncertainty, which is in contrast to the “theoretical” situations the traditional models simulated. For example, travel time to the station can vary significantly by time of day (Li et al., 2010). In addition, traditional models didn’t consider P&R users’ risk attitude towards these uncertain circumstances or variables. Therefore, more advanced models, such as the latent class model, mixed logit model, etc., could to be used to explore and better understand station choice behaviour.

The effects of uncertainty on choice have been tested in the transportation field, but have tended to focus on travel mode choice, departure time choice and route choice. There appears to be little or no research specific to choice of departure station under uncertainty. In the current research of travel choice under uncertainty, the effect of the variability of uncertain situations on travel choice is evaluated using the mean-variance method and the variation of uncertain stations is measured using expect utility theory (EUT), extended expect utility theory (EEUT) and cumulative prospect theory (CPT). To my knowledge, this is the first time that this modelling approach, (i.e. EEUT, CPT, mixed logit models and latent class models), has been applied to the exploration of station choice behaviour for P&R users. The significance of this research is in developing robust models to understand the effects of uncertain factors on the choice of departure train station for P&R users and measuring their risk attitude towards the variability of the uncertain factors. The contributions of this research include:

- Developing a new framework to explore P&R users’ choice of departure train station under uncertainty;
- Establishing new station choice models to predict P&R users’ choice of departure train station under uncertainty;
- Evaluating the effects of uncertain factors on station choice;

- Estimating the variation of uncertain factors within the EEUT and CPT framework; and
- Measuring P&R users' risk attitude towards these variations.

1.4 Research methodology

This study develops a rigorous, realistic and easily computed station choice model that clearly explains P&R users' choice of departure train station under uncertain situations and analyses their risk attitude towards the variations of the uncertain factors. The detailed steps are listed as follows:

- **Step 1**. Data collection
  - ❖ Clarifying the choice sets based on the decision-making tree of P&R users, literature review and real preference data via intercept surveys.
  - ❖ Designing a stated choice (SC) experiment and an eye tracking experiment, and collecting SP data to explore the mechanism of choice behaviour of P&R users under uncertainty.
- **Step 2.** Development of station choice models

All station choice models will be developed within a discrete choice theory framework, in which variations in the uncertain factors will be measured using cumulative prospect theory (CPT) or extended expected utility theory (EEUT) and their effect on station choice evaluated with the mean-variance method. All parameters in the models will be estimated with the Nlogit 5 software package.

- **Step 3.** Analysis of results
  - ❖ Based on the estimates and shape of the value and weighting functions, we will obtain the commuters' risk attitude towards the variations in the uncertain factors;
  - ❖ Comparing the parameters indicating respondents' risk attitude with rail passengers' ridership for each station, we will identify the effects of respondents' risk attitude on rail ridership; and
  - ❖ The latent class model and mixed logit model will be used to determine P&R users' preference heterogeneity.
- **Step 4.** Evaluation of the station choice model

Eye tracking techniques will be applied to assess the SC experiment and, from that, to evaluate the overall station choice model.

1.5 Thesis outline

This thesis comprises nine chapters as shown in Figure 1.1 and described below:
Chapter 1 introduces the area of research. It sets out the background, specifies the problems addressed, identifies the study's scope, and lists its objectives and the research's significance in improving the strategic transport evaluation model and sustainable development of the transport system in Perth.
Chapter 2 starts with a brief description of the concept of P&R as a travel mode, then reviews the literature related to P&R, train station choice and travel choice under uncertainty, which provides an overview of the gaps and limitations in the previous literature associated with departure train station choice.
Chapter 3 sets out an efficient and effective methodology to develop station choice models under uncertainty. The study area, the decision making process of train station choice for P&R users, the identification of choice sets, sample stations and surveys are included in the chapter.
Chapter 4 discusses the design of the two experiments used in the research, one for developing the station choice model and the other for validating the model. Data collection and collation, experimental design, determination of sample size and the survey implementation are included in the chapter.
Chapter 5 develops the station choice sub-model for exploring station choice behaviour under travel time uncertainty. This chapter includes the development of the travel time sub–model, the measurement of respondents' risk attitude towards uncertainty of travel time and the analysis of the impact of respondents' real travel time experiences on their risk attitude towards station choice.
Chapter 6 develops the station choice sub-model based on the effect of crowding on trains. Crowding indices affecting station choice for P&R users are identified first in the chapter, then the crowding sub-model is developed. The measurement of P&R users' risk attitude towards crowding on trains, the exploration of the relationship between individuals' risk attitude and boarding numbers, and an understanding of the effects of personal preference heterogeneity on station choice are also included in the chapter.

Chapter 7 develops the station choice model based on the uncertainty/variability of the parking search time. The chapter investigates the effects of parking attributes on station preference and measures P&R users' risk attitudes towards the variability of parking search time.

Chapter 8 develops the overall station choice model under uncertainty. This chapter includes identification of key factors for P&R users' choice of departure train station, the establishment of the overall model, investigation for individuals' preference heterogeneity on station choice and elasticity analysis for the key factors.

The thesis concludes with Chapter 9. It summarises the major findings, discusses the limitations of the research with respect to the objectives and provides recommendations for future research.

It is worth noting that the similarity of parts of chapters in the thesis is high as this information is from my PhD candidacy report approved by Curtin University on August 12, 2012 and three published papers (i.e. (Chen et al., 2015; Chen, Xia, Smith, & Han, 2014; Chen et al., 2017)), in which my contribution is about 80%. These papers are contained within the thesis as follows.

- Chapter 2 used part of the candidacy report;
- Chapter 3 used part of candidacy report and papers I, II and III;
- Chapter 5 used paper II;
- Chapter 7 used paper III; and
- Chapter 9 used papers II and III

1.6 Chapter summary

This chapter has established the objectives of and rationale for the research on modelling train station choice for P&R users under uncertainty and sets out the key tasks and thesis structure. The next chapter will review the literature relevant to P&R station choice and travel choice under uncertainty so as to identify the research gaps in the research and methods required to fill these gaps.

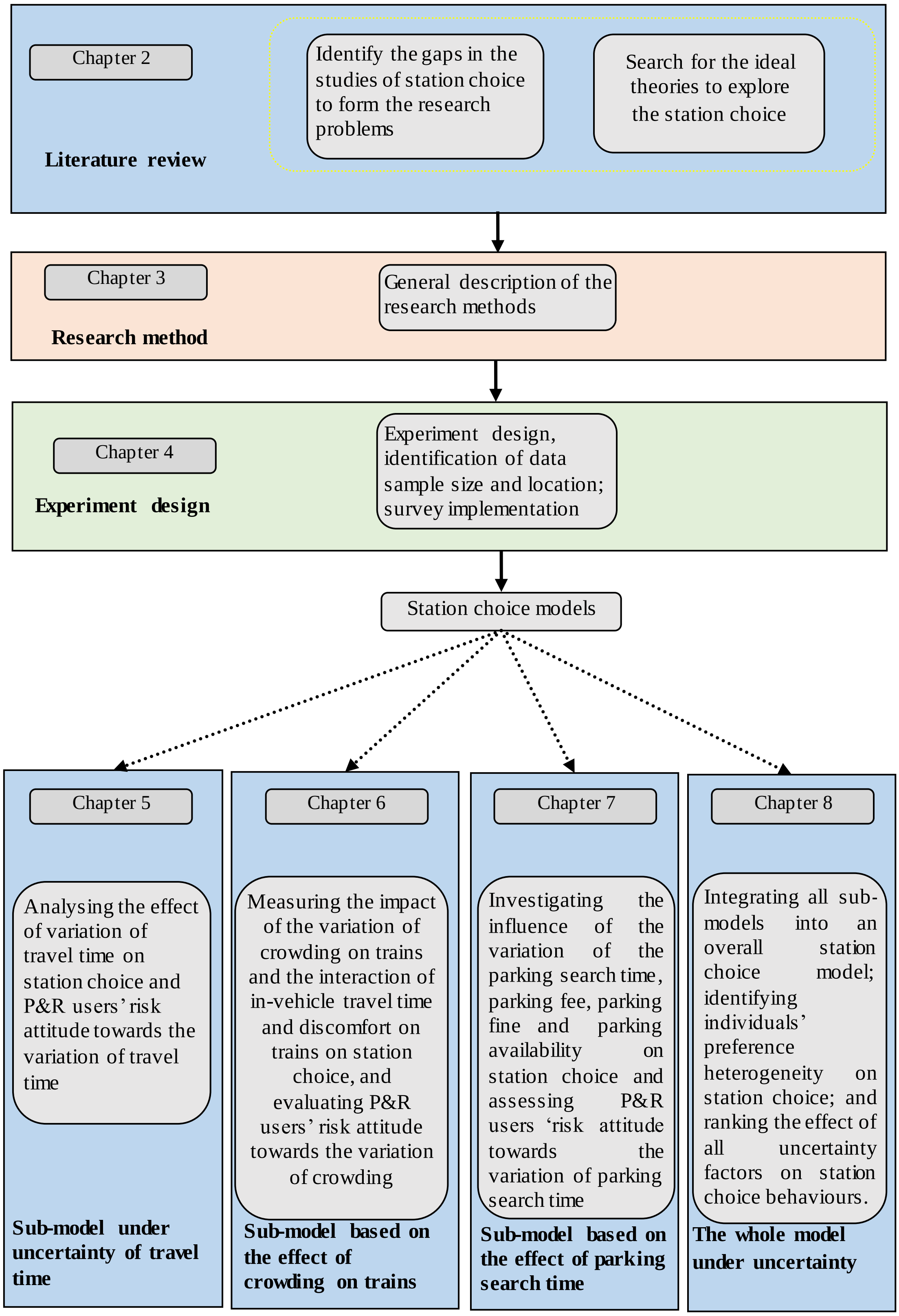


Figure 1.1 Research structure and relationship to thesis chapters

CHAPTER 2 LITERATURE REVIEW

This chapter presents a review of the literature associated with the study of P&R, station choice and travel choice under uncertainty. It starts with a brief description of the concept of P&R as a travel mode, including its advantages and disadvantages. Sections 2.2, 2.3 and 2.4 then provide, respectively, a comprehensive review of the research literature related to P&R, train station choice and travel choice under uncertainty. The review summarises the current evidence in this field and highlights the research gaps to be addressed in this thesis.

It is worth emphasising that some of the content in the chapter is from my candidacy report submitted to and approved by Curtin University.

2. 1 The development of P&R

Park and Ride (P&R) is a form of travel involving multimodal transport that encourages car users to transfer to Mass Public Transport (MPT), such as buses, trains, etc., in order to reduce road congestion, increase PT patronage, decrease air pollution and achieve cost and time savings for P&R users. In other words, a trip by the P&R mode involves the use two or more travel modes and requires a transfer between private car and public transport (Ison & Rye, 2008). The concept of P&R can be disaggregated into four constituent phases, as shown in Figure 21: private transport access; a parking service, public transport departure, and destination.

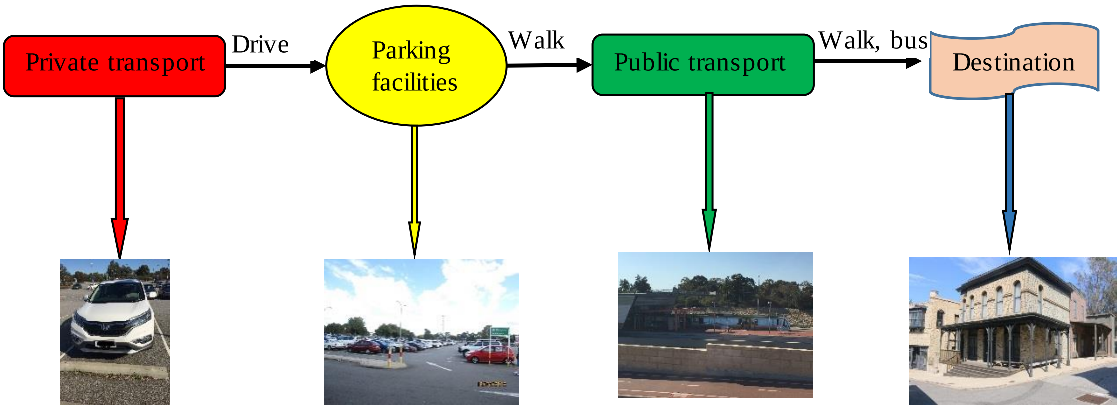


Figure 2.1 P&R phases

2.1.1 Advantages

P&R has gained enormous popularity since it was first introduced in the United States in the 1930s as a measure to increase transit ridership and manage travel demand (Noel, 1988). Many countries, including UK, USA, Australia and China, have witnessed its benefits. It has four key roles: ① providing an alternative travel mode to drive all the way, which improves accessibility (Bolger, Colguhoun and Morrall 1992; Flint 1992; Niblette and Palmer 1993; Noel 1988); ② alleviating traffic congestion, reducing energy consumption and decreasing pollution related to the automobile by encouraging car users to use public transport (PT) for part of their trip. In other words, it is an efficient way to manage travel demand (Hong Kong Transport Department (HKTD), 1995); ③ reducing costs and saving time for its users, and improving users' comfort (Noel, 1988); and ④ controlling parking demand in central areas (Williams, 1999).

2.1.2 Disadvantages

Even though much of the literature, (such as Bolger, Colguhoun and Morrall 1992; Flint 1992; Niblette and Palmer 1993; Noel 1988), proved that P&R can improve a transportation system's efficiency, some unexpected problems emerged as it was implemented. Not all P&R users were found to be transferring from car to PT but included those that previously caught the bus or walked to the PT (Arne, 2004; Harris, Cooper, & Whitfield, 1998; Mingardo, 2013; Noel, 1988; Parkhurst, 1996; Parkhurst, 2000). Some P&R schemes, therefore, increased traffic on some roads and the vehicle kilometres travelled rather than reducing them, such that traffic conditions became worse on some highways and the pollution by vehicles increased (Cousins, 1977; Meek, Ison, & Enoch, 2009; Parkhurst, 1995; Topp, 1995). In addition, many P&R facilities were found to be built on greenfield, (or green belt) sites, which may have adverse environmental impacts (Pearson Education Ltd, 2014).

Although there are some issues related to P&R, overall, P&R was found to play a very important role in benefitting commuters, reducing traffic congestion and decreasing environmental pollution. More importantly, it has been identified as an efficient strategy to manage travel demand and can improve the sustainability of a transport system(Olaru, Smith, Xia, & Lin, 2014). Therefore, P&R has been accepted and implemented worldwide.

2. 2 Research related to P&R

The research related to the studies of P&R is substantial and has therefore been divided into five key aspects, as discussed below.

2.2.1 Guidelines to design and plan P&R facilities

In order to develop sound P&R facilities, many agencies have developed guidelines based on specific local conditions and requirements. In the USA, the State of Florida published the "State Park and Ride Lot Program Planning Manual" in 1989. It was used by the Florida Department of Transport (FDOT) and other state agencies to plan, design and implement P&R facilities. It has been updated and revised a number of times since then including in 1996, 2001 and 2012, the latter being a major revision to incorporate the latest strategies, initiatives and legislation, with the name changed from "manual" to "guide" (Frederick R. Harris, 2012).

After that, similar documents were published one after another. For example, Bolger, Colguhoun, and Morrall (1992) prepared planning guidelines for the P&R facilities of a Light Rail Transit System in Calgary, a city in the Canadian province of Alberta. The guidelines included the criteria for locating P&R facilities, access and egress requirements and the size and design of the parking facilities. Robert and Spillar (1997) developed P&R Planning and Design Guidelines for New York, USA that provided advice on selecting the optimum locations to maximise use and best serve community needs, and on the design of the P&R facilities. The American Associaton of State Highway and Transportation Officials (2004) prepared a guide for P&R facilities for Washington DC, USA Compared to the guidelines for New York, it is more detailed and comprehensive. It considered the operation and maintenance of P&R facilities, and the integration of architecture, landscape, art and P&R facilities. Cherrington et al. (2017) produced a "Decision–Making Toolbox to Plan and Manage Park-and-Ride Facilities for Public Transport", which is a guide book on planning and managing P&R in America and covers everything from the concept to the operation of P&R facilities. In the Australian context, Austroads and Departments of Transport in various states have developed guidelines. For example, Austroads (2017) published the "Guide to Traffic Management Part11: Parking", which included guidance for planners and engineers to plan, design and implement efficient P&R facilities. In Queensland, The Department of Transport and Main Roads (2015) prepared the "Public Transport Infrastructure Manual (PTIM) 2015", which provided good practice guidelines for the planning and design of P&R facilities.

2.2.2 Locating and pricing P&R facilities

Given the consistent recognition of the importance of locating P&R facilities to effectively intercept vehicles on their journey to their destinations (Dickins, 1991; Faghri, Lang, Hamad, & Henck, 2002; Farhan & Murray, 2005; Robert & Spillar, 1997), many studies associated with locating P&R facilities have been conducted. Diverse approaches have been proposed. For example Horner and Grubesic (2001) and Faghri et al. (2002) used geographic information system (GIS) tools to identify optimal P&R locations. Pickett (2005) suggested that the decision on locating P&R facilities should consider a range of factors including the key stakeholders involved, what has worked and not worked in the past, (i.e. reviewing successful and less successful facilities), new technology, pricing and market strategies, and targeting key user groups. Wang, Yang, and Lindsey (2004) used an optimal approach to locate and price P&R facilities based on profit maximisation and social cost minimisation. Horner and Groves (2007) proposed a network flow-based approach while Krasić and Lanović (2013) used an Analytic Hierarchy Process (AHP) method to evaluate the potential locations of P&R facilities. Fan, Khan, Ma, and Jiang (2013) developed a bi-level programming model to locate P&R facilities with the capacity to capture the supply-demand interactions of commuters. Generally, all these approaches have contributed to efficiently locating P&R facilities and increasing P&R attractiveness.

2.2.3 P&R modelling

Two types of research related to P&R modelling were found. The first is associated with the prediction of P&R demand and patronage (Fernandez, Cea, Florian, & Cabrera, 1994; Islam, Liu, Sarvi, & Zhu, 2015). In this context, many P&R demand models were developed with the consideration of road network equilibrium. The detailed specifications of these models are diverse, including nested logit models (Fernandez et al., 1994; Fox, 2005; García & Marín, 2005), multinomial logit models (Islam et al., 2015; Li, Lam, Wong, Zhu, & Huang, 2007; Xiong & Yang, 2008), deterministic continuum equilibrium models (Liu, Huang, Yang, & Zhang, 2009), and disaggregate demand models (Florian & Los, 1979). The second type of research, although limited compared to the first, is the modelling of P&R access station choice. The objectives are usually to investigate the factors influencing the choice of P&R access station and to estimate P&R train station demand. For example, Vijayakumar, El-Geneidy, and Patterson (2011), based on the data from an Origin to Destination (OD) survey and boarding data, identified the factors affecting the driving distance to, and demand at, suburban rail stations in the Montreal region as well as personality, (e.g. gender and age), trips characteristics, (e.g. in-vehicle travel time), and station characteristics, (e.g. parking bays). Mahmoud et al. (2014) used a multinomial logit approach to develop three access station choice models for different regional P&R users to estimate the access distance by car and the number of P&R users boarding at each station.

In summary, most of the previous studies focused on modelling P&R demand based on road network equilibrium not the choices of individuals. Moreover, the research specific to P&R access station choice is very limited and only identified a few factors influencing P&R station choice. In practice, more factors, (such as parking search time, crowding on trains, etc.), may affect P&R station choice. In order to accurately predict P&R demand, the systematic investigation of P&R station choice has been identified as research gaps that should be filled.

2.2.4 Factors influencing travellers' choice for P&R

The factors affecting P&R schemes, based on previous literature, can be divided into four groups. The first is related to the service quality of public transportation, such as safety (Sgurgaikar & Deakin, 1927), convenience, seats availability, reliability, frequency of public transit system, and difficulty in parking (Arne, 2004; Bos, Heijden, Molin, & Timmermans, 2004; Harris et al., 1998; Li, Lam, Wong, Zhu, & Huang, 1994; Robinson, 1994), and in-vehicle travel time in public transit and transfer time at P&R station (Islam et al., 2015). The second is related to the journey cost, such as travel time and cost (Bos et al., 2004; Harris et al., 1998; Lam, Holyoad, & Lo, 2001; Li et al., 1994; Seik, 1997), and the severity of road congestion (Mogridge, 1990; Qin, Guan, & Zhang, 2012). The third group is associated with travellers' characteristics, such as income and car ownership (Arne, 2004) and individual preference (Bos et al., 2004). The last group is P&R facility location (Mingardo, 2013) and parking fare (Islam et al., 2015; Kono, Uchida, & Andrade, 2014). Even though diverse factors were identified in different papers, nobody considered all of them in analysing P&R choice behaviour. Therefore, research is needed to systematically study the effect of a variety of factors on travellers' choice.

2.2.5 Assessment of performance of the P&R

In the previous literature, assessing the performance of P&R facilities mainly focused on their contribution in generating a mode shift from car-only modes to more sustainable transport modes. The methods to assess their performance chiefly included surveys and modelling, with most researchers using the survey method as the results can provide direct evidence of the P&R facility's performance. Generally, the surveys to assess a P&R facility's performance were intercept surveys, in which the questionnaires used included some specific questions, such as *how did you get to the train station or bus stop?, are you satisfied with the P&R facilities?, how much will you save by using P&R mode?*, etc. Researchers can easily count the number of arrivals at the station by car or bus and assess the P&R facility's performance based on this and the answers to the above questions. For example, in the UK, Parkhurst (1995) surveyed 600 users in York and 1750 users in Oxford and the results confirmed the success of P&R in attracting users. Mingardo (2013) surveyed 700 travellers at nine rail-based P&R stations around Rotterdam, The Netherlands, and further confirmed the good performance of some P&R facilities in outer suburbs. Dijk and Montalvo (2011) conducted surveys in 45 cities in Europe and assessed the current levels of usage of the P&R facilities, mapped out the range of P&R strategies adopted by these European cities, and investigated the policies and objectives underpinning these strategies Meek, Ison, and Enoch (2011), Parkhurst and Meek (2014) examined P&R facilities in the UK and confirmed the attractiveness and effectiveness of P&R.

Another way to assess the performance of a P&R system is modelling. For example, Karamychev and Reeven (2011) developed a discrete choice model to test the effect of the introduction of P&R on social welfare and distribution of car traffic. The results indicate that P&R did efficiently intercept, (capture), cars travelling to the city centre at the outskirts of the centre.

In general, the studies adopted two indicators of P&R usage. The first, and by far the most common, is the proportion of public transport passengers using P&R, i.e. the number P&R users on the PT service divided by the total using the PT service. It is commonly used in UK and Australian cities. For example, in Oxford it is 17% (Huntley, 1993), in York 12% and in London 21%, with a P&R market share in some Australia cities of up to 15% (Vincent & Hamilton, 2007). A second index, although seldom used, is P&R usage as a proportion of commuter trips to the centre business districts (Vincent & Hamilton, 2007).

The effects of P&R are relative complex. Even though the success of P&R facilities was generally confirmed by the usage surveys, issues were identified in some cases, such as lower than expected patronage levels (Arne, 2004; Heggie & Papoulias, 1975, 1976); the misuse of the P&R carparks (e.g. non-P&R users parking there), insufficient (too infrequent) public transit services, personal safety concerns, low service quality of the public transit (Abbas & Sergany, 2008), and low mode shift rate to public transit (Olaru et al., 2013; Olaru et al., 2014; Wiseman, Bonham, Mackintosh, Straschko, & Xu, 2012). Therefore, further studies are needed to address more sustainable P&R.

2.2.6 Summary

As P&R has become recognised as an efficient and economic travel mode option, particularly for low-density cities, more and more authors have devoted themselves to this research. Currently, the research related to P&R covers many aspects, including planning and design, location optimisation, implementation and operation, and demand prediction. Generally, the underlying objective of the research is to provide a basis for improving P&R's attractiveness and increasing rail patronage. Given that few previous P&R demand models took P&R access station choice behaviour into account, and there isn't a thorough study on P&R access station choice either, to date we cannot accurately predict P&R demand. Correspondingly, there are issues related to the balance of supply and demand of P&R. Therefore, to develop sustainable P&R, the systematic study of P&R users' choice for departure train station and its introduction into P&R demand models should be addressed.

2. 3 Research on train station choice

The literature related to train station choice modelling is also very limited. It mainly focuses on the discussion of the several aspects listed below.

2.3.1 Objectives for studying station choice

The objectives of station choice studies are generally limited to three aspects. The first is improving the transport model packages by incorporating station choice sub-models. For example, Fox (2005) improved the Policy Responsive Integrated Strategy Model (PRISM) used in the West Midlands region of the UK and Fox, Andrew Daly, Bhanu

Patruni, and Milthorpe (2011) extended the Sydney Strategic Model (STM). The second is providing evidence for planning the rail network. For example, Mahmoud et al. (2014) located a new train station based on the results of a station choice study. Givoni and Rietveld (2014) helped transport policy makers plan train stations in an urban region and design the required road network around the stations. Debrezion, Pels, and Rietveld (2009) predicted travel demand at the station level to allow transportation managers to select the station locations when extending the existing lines and planning new lines, design the P&R facilities and plan the connecting (feeder) bus services. The third is improving the demand models. For example, Davidson and Yang (1999) produced an accurate station-level ridership forecast model for MIS studies and facilities planning. Lythgoe and Wardman (2004) established a new parkway of forecasting model that had improved features and was easier to apply compared to previous demand models. Blainey and Evens (2011) improved rail demand models by analysing the catchment areas that P&R stations should serve and informed and enhanced the Passenger Demand Forecasting Handbook (PDFH) guidance on catchment definition for demand modelling.

In general, the research related to station choice mainly aims to improve demand models so as to more accurately predict rail patronage and provide planners and policy makers with useful information to plan all kinds of transportation facilities. However, the behaviour of individual station choice is very complex and cannot be predicted accurately based on the current limited literature. Moreover, the evidence for improving railway service quality, as obtained by exploring station choice behaviour, has not been investigated so far. Therefore, separately evaluating the effect of factors influencing station choice, as a basis for improving rail service quality, should be considered as one of the tasks in this study.

2.3.2 Station choice models

According to our knowledge, most station choice models were developed within discrete choice theory, with five types of logit models found in the previous literature. They are multinomial logit (MNL) models, nested logit (NL) models, probit models, cross-nested logit models and latent class (LC) models.

The MNL model is the simplest with a closed form and predominated in the earlier history of station choice. The earliest MNL model of station choice was developed by Liou and Talvitie (1974) for access mode and access station selection. With this model, the sequence of a traveller's decision-making process for the access mode and station choices and the attributes with greatest effects were identified. After that, a number of authors developed MNL models of station choice based on different considerations. For example, Kastrenakes (1988), developed an MNL model for New Jersey Transit to predict rail ridership based on four variables (i.e. station location with respect to home, access time to station, train frequency and the generalised (time & distance) cost of travel between the access station and the final destination). Adcock (1997) investigated the factors determining a rail passenger's station choice and used them to develop an MNL model of station choice. Moreover, he estimated the relative importance of each factor with the model. Debrezion et al. (2007) established an MNL model for Dutch railway passengers to choose their departure train station with consideration of accessibility to the train station and service quality provided at the station. The MNL models were developed based on the axiom of the "independence of irrelevant alternatives (IIA)", but, in reality, some correlation between the choices exists. A number of authors have therefore raised concern over whether the MNL models adequately explained station choice behaviour. Hence, some researchers began to try other discrete choice models.

The essence of the nested logit (NL) model is that it can take into account the correlations between those features of an alternative that cannot be readily observed by dividing (partitioning) the full set of alternatives (choice set) into sub-sets containing similar alternatives, (usually referred to as nests). Given that a traveller's decision-making process is a function that includes the interaction of both station choice and access mode choice, some authors began to use the NL models to explore jointly station choice and access mode behaviour. For example, Harata and Ohta (1986) developed a three level NL model to analyse access mode and station choice behaviour, and tested its prediction accuracy and temporal stability. Fan et al. (1993) developed an NL model and an MNL model for Greater Toronto to analyse rail access mode and station choices and subway automobile access station choice for commuter trips by rail and subway during the morning peak. Davidson and Yang (1999) developed a two level NL model in which station choice was taken as the lower level and mode of access the upper level. Debrezion et al. (2009) developed a two level NL model to explore Dutch railway passengers' choice of departure station and access mode in two decision structures, (i.e. access mode upper level and station choice lower level, and vice versa). Givoni and Rietveld (2014) used a two level NL model to determine the optimum number of train stations that should be provided for the urban area under investigation. Generally, within the NL model, the traveller's decision making process is made of access mode choice and station choice, and most models adopted two-level structures per Figure 2.2.

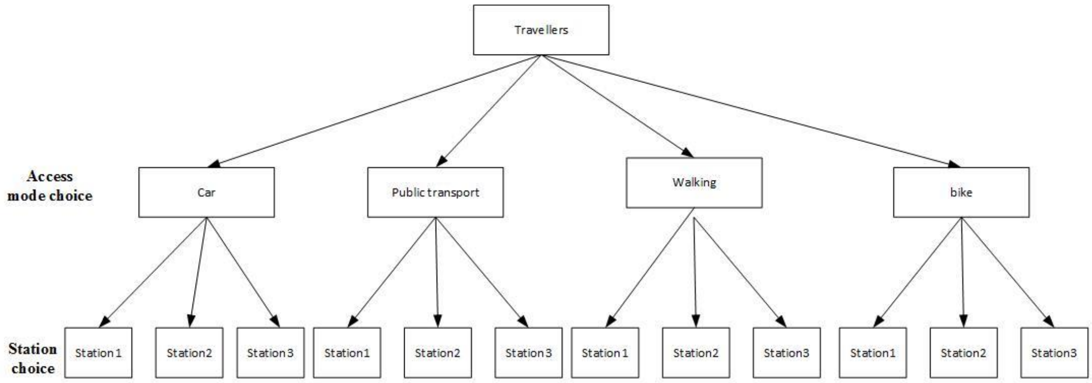


Figure 2.2 The two-level structure of NL models

Similar to the NL model, the cross-nested logit (CNL) model also belongs to the set of general extreme value (GEV) models and, thus, also allows correlations of attributes over alternatives and is easy to manipulate. Not only that, the CNL model can contain multiple nests. Therefore, Lythgoe et al. (2004) used it to explore access mode choice and station choice behaviour and forecast demand for journeys from new stations.

Even though the GEV model can relax the IIA property of the MNL model, it is unable to take into account random variations in taste, i.e. individual differences, or incorporate longitudinal data if the unobserved factors for each individual are correlated over time. Therefore, Desfor (1975) developed a probit station choice model, which could overcome all the limitations of MNL models, to test the effects of station location, access cost to competing stations and parking availability on a commuter's station choice, and to estimate the catchment area of each station on the line, which made a basis for the demand of the entire line.

In summary, access mode and station choice behaviour were modelled in the earlier research as a hierarchical decision process with access mode as the upper level. However, a traveller's decision-making process is very complex and cannot be accurately predicted. Therefore, Chakour and Eluru (2014) simultaneously considered two segments of access mode and station choice behaviour: Segment 1—station choice then access mode choice, and Segment 2—access mode choice then station choice. They then applied a latent segmentation approach to allocate individuals to the two segments based on probabilities determined by the individuals socio-demographic status, the level of service provided, land uses and urban form around the station, the characteristics of the trip and the facilities at the station

In general, the previous station choice models were developed in discrete choice theory and the trend of their forms was from simple to more complex, i.e. MNL models, NL models, CNL models and probit models. The MNL model is the easiest and the most widely used discrete choice model due to its function for choice probability having a closed form and being readily interpretable. It was derived based on the assumption that the distribution of unobserved factors across alternatives is independent and identically distributed (IID) with a Type I extreme-value distribution, which means that there is no correlation between the unobserved factors over the alternatives, individual's responses are the same, (homogeneity), and all alternatives have the same variance. This assumption has been violated by many studies (Ben-Akiva & Francois, 1983; Daganzo, 1979; McFadden, 1978; McFadden & Train, 2000). NL models, CNL models, and Probit models partially relaxed the assumption, i.e. the independent requirement, by allowing the unobserved utility to be correlated across alternatives but retained the assumption that the random components of the alternatives were identically distributed. Specific to both NL models and CNL models, the restriction in the alternatives in different nests was relaxed, while still exhibiting the independence of irrelevant alternatives (IIA) property within each nest. However, they belong to the General Extreme Value (GEV) class of models, so they assumed that the random components of alternatives followed the general extreme value distribution, rather than the Type I extreme-value distribution. For Probit models, they completely relaxed the independence assumption but only allowed the unobserved utility of alternatives to follow the normal distribution. Actually, we cannot assume that the distribution of the unobserved factors across alternatives is identical and that an individual's preference to attributes of alternatives is homogenous, which may lead to the results of the analysis of station choice using GEV models and Probit models deviating from reality, i.e. be less accurate and reliable than desirable. Therefore, more flexible discrete choice models that can fully relax the IID assumption should be applied when exploring station choice behaviour.

2.3.3 Attributes affecting station choice

Based on Ortúzar and Willumsen (2011), a traveller's choice of travel mode is made based on three types of factors: the traveller's characteristics, the travel mode's characteristics and the service quality of the transport facilities. The factors influencing a P&R user's choice of departure train station, based on different trips between origin and departure train station are classified as follows:

(a) Depart from home

The departure time needs to be considered as it is very sensitive to changes in travel time in peak hours and costs (Jong et al., 2003), traffic congestion on the road network during the peak travel time (Yang & Huang, 1997) and arrival time at the destination (Bajwa, Bekhor, Kuwahara, & Chung, 2008). Moreover, the survey for Bay Area Rapid Transit (BART) rail station parking lots revealed 56 percent of commuters leave home between 4:00 am and 6:00 am to ensure a space is available at their preferred train station along on the I-80 corridor (Shirgaokar & Deakin, 2005), which confirmed that the departure time can be very important for a P&R user's choice. Martinovich (2008) also found that many stations in Perth started to fill before 6:00 am and were completely full by 7:30 am. Additionally, Olaru et al. (2014) revealed P&R station choice is linked to departure time.

(b) Drive to train station

During this stage, accessibility to the train station, location of the station and fuel cost were found to affect station choice. Accessibility has two measurements, namely, distance to the station from last stop and travel time (also called access time).

Distance to the station from last stop can be measured in a number of ways. The easiest is straight-line distance, i.e. the distance in a straight line between the last stop and the station, and was used in the station choice models developed by Desfor (1975), Adcock (1997), Mahmoud et al. (2014), etc. It is the least accurate as it rarely reflects the distance travellers actually travel. Another index is real journey distance determined by the shortest route to the station from the last stop. Debrezion et al. (2007), Blainey and Evens (2011), Givoni and Rietveld (2014) used it to develop their station choice models.

The second measurement of accessibility is travel or access time, and, within the context of P&R, refers to the time spent travelling from the last stop to the departure train station. It was used by Kastrenakes (1988), Fan et al. (1993), Lythgoe and Wardman (2004), Lythgoe et al. (2004), etc. Whichever measurement was used, accessibility always had a negative effect on station choice behaviour in the past studies.

Fuel cost was included, as part of general travel cost, in the station choice model developed by Kastrenakes (1988), Fan et al. (1993), Lythgoe and Wardman (2004), Fox (2005), etc.

(c) Parking at the station

During this stage, parking capacity and parking cost were the first factors considered in station choice models. Miller and Cheah (1991) asserted that the parking charge could be one of the key factors in car access station choice because they found that car commuters may compare the P&R parking charge, (plus rail fare), with the parking fee at the workplace for the drive all the way mode. Later, Fan et al. (1993) and Davidson and Yang (1997) proved this assertion and revealed that the effect of parking supply (or capacity) on station choice is positive.

After that, more parking attributes were identified. Debrezion et al. (2007), Vijayakumar et al. (2011), Mahmoud et al. (2014) found that station choice was related to parking attractiveness, measured by parking availability, and showed a positive effect on station choice. In addition, Mahmoud et al. (2014) investigated the effect of parking cost on station choice and a negative sign was found.

Kastrenakes (1988) also examined parking availability and parking fee in his station choice model but they were dropped from the model due to their counterintuitive signs.

(d) Rail station service quality

The quality of service at a rail station is usually described by the quality of the train services from that station and the range of services and facilities at the train station itself (Debrezion et al., 2007). The quality of rail service is measured by train frequency, network connectivity and coverage. The supplementary service quality is evaluated based on the availability of other facilities at the station, such as parking bays, safety facilities, shops, etc. (Debrezion et al., 2007, 2009; Kastrenakes, 1988; Wardman & Whelan, 1999). Additionally, Fox (2005) introduced the number transferring to rail into his station choice model. Usually, these factors were taken as independent variables in the station choice models but Debrezion et al. (2009) incorporated them into a new index (called service quality index) and introduced it into an NL model combining station choice and access mode choice.

(e) Travel on trains

The time spent travelling on the train, (also called in-vehicle travel time), was identified as a factor influencing station choice by a number of authors. Desfor (1975) was the first to introduce travel time on train, as a trip cost, into a station choice model and identified its contribution to the disutility of the train alternative. Fan et al. (1993) and Davidson and Yang (1997) put it into their study of modelling rail access mode and station choice and revealed its negative effect on access station choice. Fox (2005) concluded that P&R users with car access usually minimise, as much as possible, the travel time to the station relative to the travel time spent on the train.

(f) Additional factors and uncertainty

Over time, more and more attributes have been included in the station choice models but these are still not sufficient to fully explain, and therefore predict, a P&R user's choice of departure train station. There are two key reasons for this. The first is that some important factors, including crowding on trains, comfortability of trains, parking search time, parking availability and safety, are missing from the station choice models. The second is that the factors used in the current station choice models are assumed to be certain, i.e. the factors have fixed values known by the P&R user when making his/her station choice. In practice, a number of factors can vary by time of day and/or from day to day. These "uncertain" factors include travel time to the station, (influenced by traffic congestion levels, weather and incidents such as crashes or breakdowns), and parking availability at the station and, linked with this, the time taken searching for a free space. It has been found that these uncertain factors, and their level of variability, do have an effect on a traveller's choice of station. Therefore, testing the effects of additional attributes and of the uncertain factors, (and their variability), on station choice should be further work.

2.3.4 Study area summary

To date, station choice studies have focused mainly on the USA (Davidson & Yang, 1999; Desfor, 1975; Kastrenakes, 1988; Liou & Talvitie, 1974), UK (Blainey & Evens, 2011; Fox, 2005; Lythgoe & Wardman, 2004; Lythgoe et al., 2004; Wardman & Whelan, 1999), Canada (Chakour & Eluru, 2014; Fan et al., 1993; Mahmoud et al., 2014), Netherlands (Debrezion et al., 2007, 2009; Givoni & Rietveld, 2014), and Japan (Desfor, 1975). There is a research gap in this field in Australia, especially in Western

Australia, even though P&R makes up a sizable share of total public transport patronage.

In order to fill the gaps in modelling station choice behaviour, a general station choice model based on data from Perth, Western Australia, (and including both certain and uncertain factors), will be developed.

2. 4 Research on travel choice under uncertainty

The study of travel choice behaviour has a long history and the topics covered span a wide area, including route choice, travel mode choice, departure time choice, location choice and station choice. Most of the early research assumed that complete and perfect information was available to all travellers when making their station choice decision. However, this assumption of decision making under certainty is not realistic. For example, when choosing public transport, travellers may not be able to predict whether or not a seat would be available as this can vary by time of day and/or from day-to-day, even at the same time of day and for the same train. When choosing a park bay, drivers cannot be sure that they will get a parking bay and even if they can, where it may be located and how long it may take to find it, as parking availability and the location of free spaces can also vary by time of day and from day to day. When choosing a route, drivers cannot forecast which route will have the least congestion (Rasouli & Timmermans, 2014). Based on this, more and more researchers have recognised that travel choice may be made in uncertain environments and have applied decision making theories under uncertainty (or risk) to explore travel choice behaviours.

2.4.1 Decision making theories under uncertainty (or risk)

There are many decision making theories that can be used to understand choice behaviour under uncertainty (or risk): discrete choice models (Bates, 1987; Ben-Akiva & Steven, 1985; Greene & Hensher, 2003; Small, 1987; Truong & Hensher, 1985), von Neumann-Morgenstern expected utility theory (Savage, 1972) and non-expected utility theory, such as weighted expected utility (Chew & MacCrimmon, 1979), rank dependent utility (Quiggin, 1993), extended expected utility theory (Li, Hensher, & Rose, 2009), prospect theory (Kahneman & Tversky, 1979), and cumulative prospect theory (Tversky & Kahneman, 1992). To date, the main theories used to explore travel choice are expected utility theory (or extended expected utility theory), and unexpected utility, (such as prospect theory). For example, Sikka (2012) and Tian and Gao (2013) applied expected utility theory to study the route choice based on the effect of travel time uncertainty. Avineri and Bovy (2014) integrated prospect theory into a travel choice model to analyse travellers' response to risk and uncertainty. Li et al. (2009), Li, Tirachini, and Hensher (2012), and Li et al. (2010) used extended expected utility theory to analyse route choice under travel time uncertainty.

2.4.2 Methods to measure uncertainty

Travel time uncertainty has been recognised in many choice situations, such as route choice, departure time choice, travel mode choice, etc. The three most popular methods to evaluate travel time uncertainty on travel choice are mean-variance, scheduled delay method and mean-lateness. They are described below.

(a) Mean-variance

The approach, proposed by Jackson and Jucker (1982) defines the utility of a choice option as a function of expected travel time, (mean travel time), and variability in travel time (standard-deviation). A commuter minimises the sum of these two terms when choosing a departure station.

(b) Scheduling delays

This approach is based on a traditional utility maximisation framework. It focuses on the time constraints, (e.g. train departure time), a commuter may face and their associated costs due to early or late arrival. There is a desire to minimise the frequency of late arrivals but maximise the time spent at home relative to the train waiting time. This approach is mainly used for departure time choice studies (Carrion & Levinson, 2012; Gaver, 1968; Knight, 1974).

(c) Mean-Lateness

This approach is proposed by the Association of Trains Operating Companies, UK and used widely in passenger rail in UK. It suggests that a traveller's expected utility consists of the scheduled journey time and the mean lateness at destination (Assiociation of Train Operating Companies, 2002). Travellers aim to minimise the lateness of arrival at their final destination, such as work or educational institution.

2.4.3 Application of travel choice under uncertainty

To date, the research related to travel choice under uncertainty is limited and only focused on three choice behaviours, departure time choice, and route choice and travel mode. In general, the decision-making process under uncertainty is made up of two components. The first considers variations in the transport network that affect travel times by time of day and from day to day including the more predictable variations due to congestion and unpredictable and irregular variations due to road works and incidents such as crashes and breakdowns (Circella, Dell'Orco, & Sassanelli, 2005). The second is the degree of confidence the decision makers have in their assessment of the network conditions, i.e. their level of uncertainty in how much the uncertain factors may vary. The understanding of travel choice under the both types of uncertainty is summarised as follows:

(a) Departure time choice under uncertainty

Originally, Gaver (1968) discovered that travellers determined their time of departure based on their required time of arrival and the likely travel time. After that, a number of authors investigated the effect of travel time uncertainty on travellers' departure time choice using different models. For example, McFadden et al. (1977) and Small (1978) separately developed departure time choice models based on a multinomial logit and schedule delay approach, and used them to evaluated the effect of the uncertainty of late arrival, social-demographic factors and mode choice on commuters' departure decisions. Abkowitz (1981) extended their studies. More factors, such as the level of flexible work practices, available travel modes, individual characteristics, (such as income, age, social status and occupation), and quality of transportation service, were included in the logit model. Moreover, transit commuters were studied separately from all commuters. In all these studies, scheduling disutility is traded off against the possible advantages due to variations in congestion over the rush hour. Noland and small (1995) developed a generalised cost model where commuters balanced the costs (disabilities) of less convenient travel times and the penalties for late arrivals against the desire to reduce the time spent in congestion to a minimum. With the model, they reinforced the conclusion from Gaver (1968). Ettema, Tamminga, Timmermans, and Arentze (2005) developed a micro-simulation model to account for departure time choice under travel time uncertainty for routine trips, in which individuals' decisions about departure time for routine trips were made based on a mental model, where the mean travel times and their variances were specified for a range of departure times. Lam (2000) developed a theoretical model to analyse commuters' joint decisions of route and departure time under travel time uncertainty and found that travellers may use toll roads, where available, to reduce travel time uncertainty and therefore depart later. Jou, Kitamura, Weng, and Chen (2008) used prospect theory to test commuters' asymmetric responses to gains and losses by comparing their actual arrival times against their expected arrival times, the latter also being referred to as the reference times. They found that the reactions of commuters were consistent with expectations. Li et al. (2012) developed a non-linear schedule model and revealed that the more uncertain travel time is, the earlier commuters depart. Siu and Lo (2014) extended the schedule cost approach into probabilities, by applying weightings to the expected and scheduled travel time costs, and developed a punctuality-based travel choice model. It revealed that more risk-averse travellers choose to depart from home earlier than less risk-averse travellers.

Generally, the previous studies mainly focused on identifying uncertain factors influencing departure time choice and developing choice model under uncertainty. Based on this, commuters' responses to travel uncertainty were tested and commuters' risk attitudes towards uncertainty were measured. Moreover, most of the departure time choice models under uncertainty were developed using the scheduled delay approach, in which only travel time and congestion were considered as uncertain.

(b) Route choice under uncertainty

Route choice is one of main decisions made daily by travellers in an uncertain environment. The uncertain environment results from a number of factors including variations in traffic levels, variations in the capacity of road network due to road works, lane or road closures etc., incidents such as breakdowns or crashes, weather conditions or faulty traffic operation, e.g. faulty or poorly set traffic signals increasing congestion at an intersection (Avineri & Prashker, 2003). The uncertainty presented in route choice models is usually only travel time uncertainty or travel time variability (Bekhor, Ben-Akiva, & Ramming, 2001; Outram & Thompson, 1977). For example, Shao, Lam, and Tam (2006) used travel time reliability to measure the magnitude of unexpected delay, introduced it into a route choice sub-model and formulated travel time reliably based on a traffic assignment model. This model may assist transportation planners to better understand how travellers behave in a congested road network environment and, hence, plan and operate their strategic road network. Most of the existing route choice models under uncertainty are developed within random utility theory with the underlying the assumption that travellers are rational, homogeneous and have perfect knowledge. Discrete choice models, with mean-variance expected utility functions formulated by von Neumann and Morgenstern (1947), were mainly used. In these models, travellers' decisions are made in terms of expected travel time and its variation. By applying different decision rules, different route choice models were produced, including shortest path model (Dial, 1969; Drrksrra, 1959), logit route choice (Dial, 1971), probit route choice (Burrell, 1976), C-logit (also a multinomial logit model capturing the correlation among alternatives in a deterministic way) (Cascetta, Nuzzolo, Russo, & Vitetta, 1996; de Palma & Picard, 2005), and mixed logit model (Bogers, Viti, & Hoogendoorn, 2005). There is also literature recording the application of other decision making theories under uncertainty (or risk) or utility functions on route choice. For example, Yang and Jiang (2014) applied cumulative prospect theory to model route choice behaviour and analysed travellers' risk attitudes. Wang, Liao, Gao, and Timmermans (2018) used a schedule delay approach to develop utility functions of route choice and examine the extent that travel delays and risk attitudes affect route choice.

In summary, the studies of route choice under uncertainty mainly focused on identifying uncertain environments for route choice, measuring the effect of travel time uncertainty on route choice, developing route choice models and evaluating travellers' risk attitudes towards travel time uncertainty. Generally, the route choice models under uncertainty were developed within discrete choice theory based on stated preference data, in which the effect of travel time uncertainty on route choice was estimated using the mean-variance approach or the schedule delay approach within an expected utility theory or cumulative prospect theory framework.

(c) Travel mode choice under uncertainty

Similar to route choice, travel mode choice can also be a daily decision for some travellers. It has attracted more attention due to its close relationship to the policies and strategies used to develop the overall transportation system, manage travel demand and mitigate traffic congestion. Traditionally, travel mode choice models were developed within random utility maximisation (RUM) theory and only focused on work-trips, in which the factors considered included parking fee, travel time, transfer time, social-demographic data, etc. and their values were fixed. As the fact that travel time uncertainty affects travel choice became more widely recognised, a few researchers began to explore travel mode choice under uncertainty. In the literature, only departure time (Pan & Zuo, 2013; Shukla, Ma, Wickramasuriya Denagamage, Huynh, & Perez, 2015), and travel time (Matthieu & Ben-Akiva, 2014) were taken as uncertain factors influencing mode choice. The travel mode choice models were developed within discrete choice theory, such as MNL models (Matthieu & Ben-Akiva, 2014), probit models and nested logit models. These models only showed linear relationships between the factors and mode choice, which is not consistent with reality. Some authors found that the mode choice decision was usually combined with other choice decisions, such as departure time choice (Bhat, 1998) or route choice (Eluru, Chakour, & El-Geneidy, 2012). For this reason, new methods, such as machine learning (Rasmidatta, 2006), artificial neural networks (ANN) (Cantarella & de Luca, 2003; Shmueli, Salomon, & Shefer, 1996), decision tree (Xie, Lu, & Parkany, 2003), rank-dependent utility theory (Matthieu & Ben-Akiva, 2014), and cumulative prospect theory (Ben-Elia, Erev, & Shiftan, 2008; Ben-Elia & Shiftan, 2010) were applied to explore travel mode choice behaviour.

(d) Sub-summary

Focused on this literature, the studies of travel choice behaviours under uncertainty mainly included identifying the uncertain factors, developing travel choice models under these uncertainties, analysing the effect of these uncertainties on travel choice, respondents' risk attitudes towards the uncertainty and the data used in choice models. In general, most of choice models under travel time uncertainty were developed within discrete choice theory and the trend is from simple models (i.e. MNL) to more advanced models (e.g. probit logit and nested logit). The utility of choice based on the effect of travel time uncertainty is evaluated using the mean-variance approach and the schedule delay approach within expected utility theory and cumulative prospect theory frameworks. However, more factors were also identified as uncertain in other research, such as parking search time (Avineri & Prashker, 2003; Hunt & Teply, 1993), overcrowding on trains (Hunt & Teply, 1993), but have not yet been considered in the above literature. Therefore, the research will apply the decision-making theories under uncertainty (or risk) and develop methods to measure the effect of uncertainty on station choice.

2. 5 Chapter summary

This chapter has reviewed previous research related to P&R, station choice and travel decision making under uncertainty (or risk). The history of the development of P&R schemes was introduced first. Then, the research relating to P&R was reviewed, including designing, planning, locating, and pricing of P&R facilities, the models and methods used to understand P&R choice behaviour, the factors influencing commuters' choices for P&R mode, assessing P&R facilities and measuring patronage of P&R. The research related to station choice was discussed afterwards, including discussion of the objectives for exploring station choice behaviour, the methods and approaches used to explore station choice behaviour, attributes affecting station choice, the methods to collect data, data type and the countries where station choice behaviour has been studied. Lastly, the literature associated with travel choice under uncertainty (or risk) was reviewed, including the decision making theories under uncertainty (or risk) and the methods to measure uncertainty, and their application for travel choice behaviour.

Generally, literature on station choice is limited, with that on P&R access station choice even more so, and that on station choice under uncertainty for P&R users almost non-existent.

Therefore, to fill the research gap, modelling station choice for P&R users in uncertain environments is required. Based on the exploration of previous literature mentioned in the chapter, a methodology to understand a P&R user's choice of departure train station facing uncertain situations is presented in the next chapter.

CHAPTER 3 METHODOLOGY

The previous chapter identified a number of gaps in the research related to modelling station choice. These include the need to understand the demand for P&R services in Perth and P&R users' station choice under uncertainty using a consistent and robust methodology. This chapter presents an overview of the methodology adopted for this research project to meet the above requirements.

It is worth reiterating that some of content in the chapter is from my published papers and my candidacy report approved by Curtin University.

3. 1 Study area

Perth, the capital of Western Australia (see the figure 3.1), was selected as the case study area for three main reasons. Firstly, Perth is a city with a high level of car ownership and car usage, resulting in significant peak-hour traffic congestion, (particularly on the freeways), and difficulties in parking in the central business district (CBD), which has the highest concentration of employment. Car ownership data indicate that Perth had about 700 motor vehicles per 1000 people in 2011, and increasing by about 2% per annum since then. Over a similar period, the total length of freeways has not increased in line with car ownership, (approximately 125 km in 2006 and unchanged in 2015) (Bureau of Infrastructure Transport and Regional Economics (BITRE), 2017). Demand for road space is therefore increasing at a greater rate than supply, resulting in increasing congestion on Perth's freeways. In order to address this and promote sustainable mobility, the P&R mode has been developed over the last 20 years or so. P&R marries private car use, (for the home to station leg), with efficient, fast and high capacity public transport to major employment nodes, and is used extensively by commuters all over the world (Cairns, 1998; Ginn, 2009).

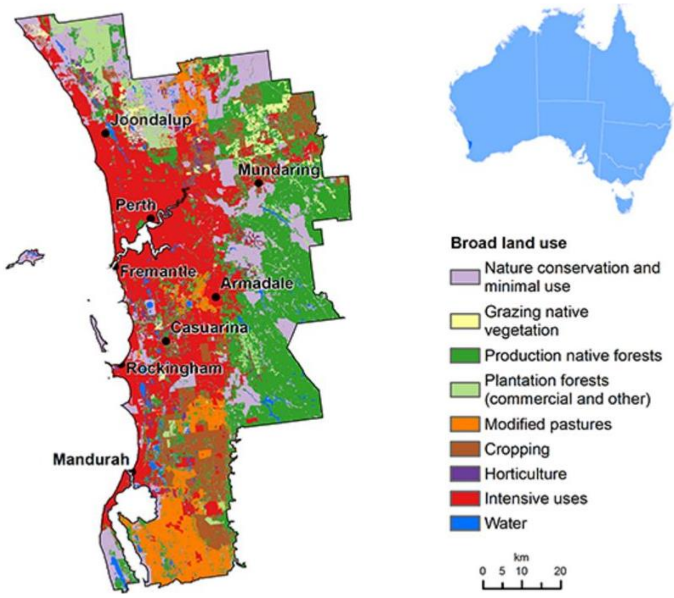


Figure 3.1 Perth and its land use map

(Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES), 2017)

The second reason for selecting Perth is that it has a very low-density, (around 318 persons per square kilometre in 2017 (Population 2018, 2018)) (see Figure 3.1), which cannot be efficiently served by the public transport system. Therefore, residents living in more remote outer suburbs have to depend on the private car for most, if not all, of their daily travel needs. However, the shortage and cost of parking in the CBD and congestion on the freeways limit their ability to drive directly to the inner city area, where the bulk of the employment is located. Hence, more and more commuters living in the outer suburbs, and indeed many in suburbs closer in, choose P&R as their travel mode, allowing them to drive to the station on the less congested local roads then use to train to reach the CBD quickly and without parking hassles.

The third reason is that there are a number of problems related to the P&R facilities in Perth. Perth currently has five rail lines and seventy stations, with a total track length of about 173 kilometres. In 2017, around 21,000 parking bays were provided at these train stations (Department of Transport, 2010; Public Transport Authority, 2011-2017). However, most of the station car parks, especially those on the newer north/south lines, fill up quickly during the morning peak, indicating that supply is inadequate to meet the overall P&R demand. In addition, a survey jointly conducted in 2 July 2012 by the University of Western Australia, Curtin University, the Department of Planning (DoP), and the Public Transport Authority (PTA) revealed that the demand for P&R facilities was distributed unevenly and did not fully match the distribution of supply. To date, it is unclear why some P&R facilities are more likely to be chosen by commuters than others, but a better understanding of this would allow planners to better match the future provision of station parking with the latent demand for P&R.

3. 2 Research workflow

There are three main objectives for modelling departure station choice under uncertainty for P&R users, namely, evaluation of the effects of uncertain factors on station choice for P&R users, assessment of variability of uncertain factors and measurement of the effect of respondents' risk attitude towards station choice. In order to achieve these objectives, the following work was conducted.

- Clarification of research hypothesis;
- Identification of decision making process of departure train station for P&R users;
- Determination of choice set;
- Data collection and analysis;
- Development of train station choice models;
- Analysis of results; and
- Implementation of the method

These steps are shown in Figure 3.2 and described below.

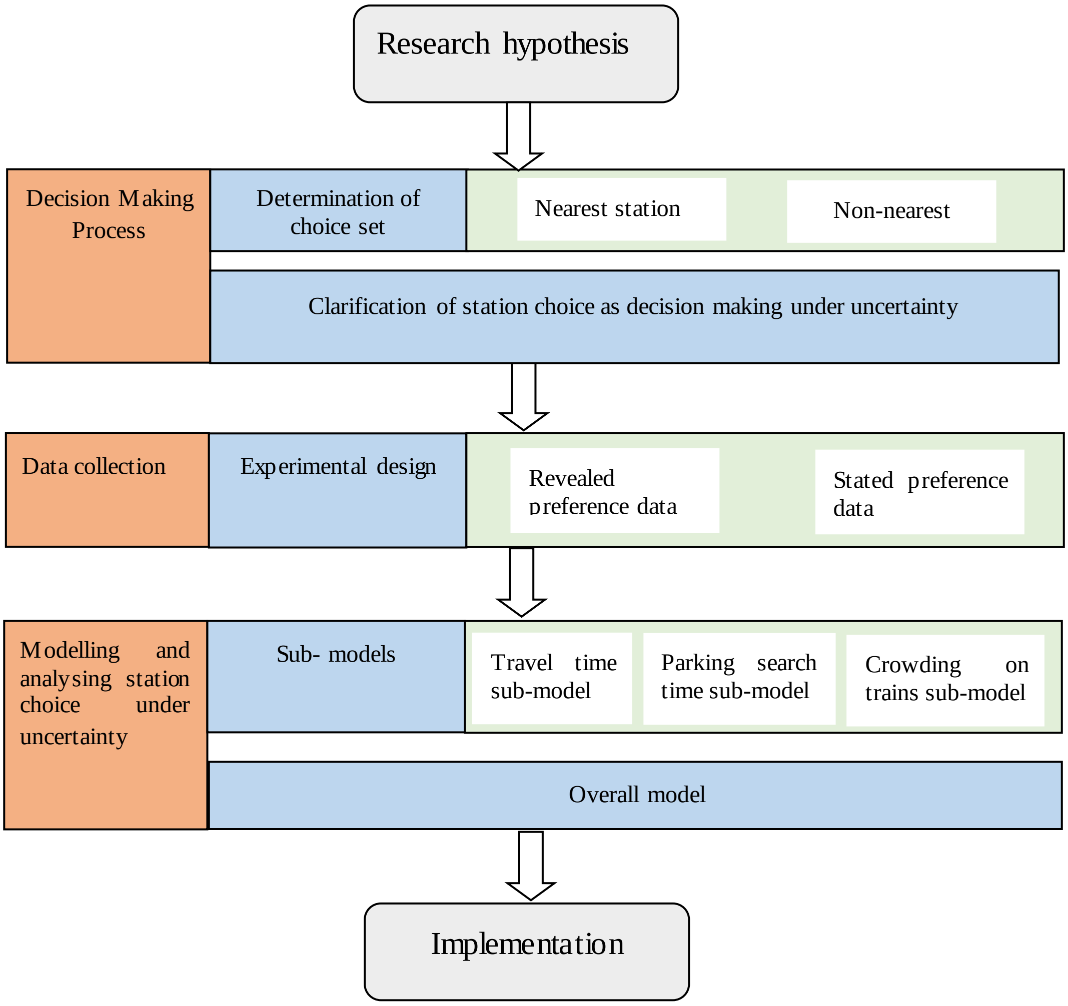


Figure 3.2 Workflow for modelling station choice of P&R users under uncertainty

3. 3 Decision making process of train station choice for P&R users

According to the responses from the train station survey, the decision making process of P&R users for departure train station usually starts with the selection of the most convenient (routine) train station based on their departure time, the train frequency and the expected arrival time. Then, they may modify this decision based on parking availability around the station. If parking is likely to be available around the most convenient station, they would choose that station. If parking is likely to be difficult to find, based on past experience, they may decide to transfer to another station where past experience has indicated spaces are more likely to be available. The choice may also depend upon the time of departure, compared to the normal departure time. For an earlier than usual departure, the most convenient station may be chosen as there would be more chance of a parking space. For a later departure, the alternative station would be more likely to be selected.

Alternatively, they may decide to use another travel mode, e.g. drive all the way, the decision being based on a number of other factors including total travel time, level of crowding on trains, overall cost, service quality, etc. In making this decision, uncertainty in some of the variables would be taken into consideration. The key one is probably the likelihood of finding a parking space and, linked with this, the parking search time. When the P&R station car park is full or close to full, a P&R user may spend a lot of time driving round trying to find a free space. They may have to park in a street around the station, (if such parking is available). Hence, the parking search time would vary with the parking location, time of day, (increasing as car parks fill), and the ratio of demand to supply. The P&R user cannot fully predict the number of available spaces, if any, in advance, so the parking search time is uncertain. Travel time to the station, (subject to localised congestion), and crowding on the trains, (at Perth's peak hour 5 minute frequency, a 1 minute delay in the train could add up to 20% more passengers on that train), are also variable and therefore uncertain.

The process for determining station choice for P&R users under uncertainty is shown in Figure 3.3.

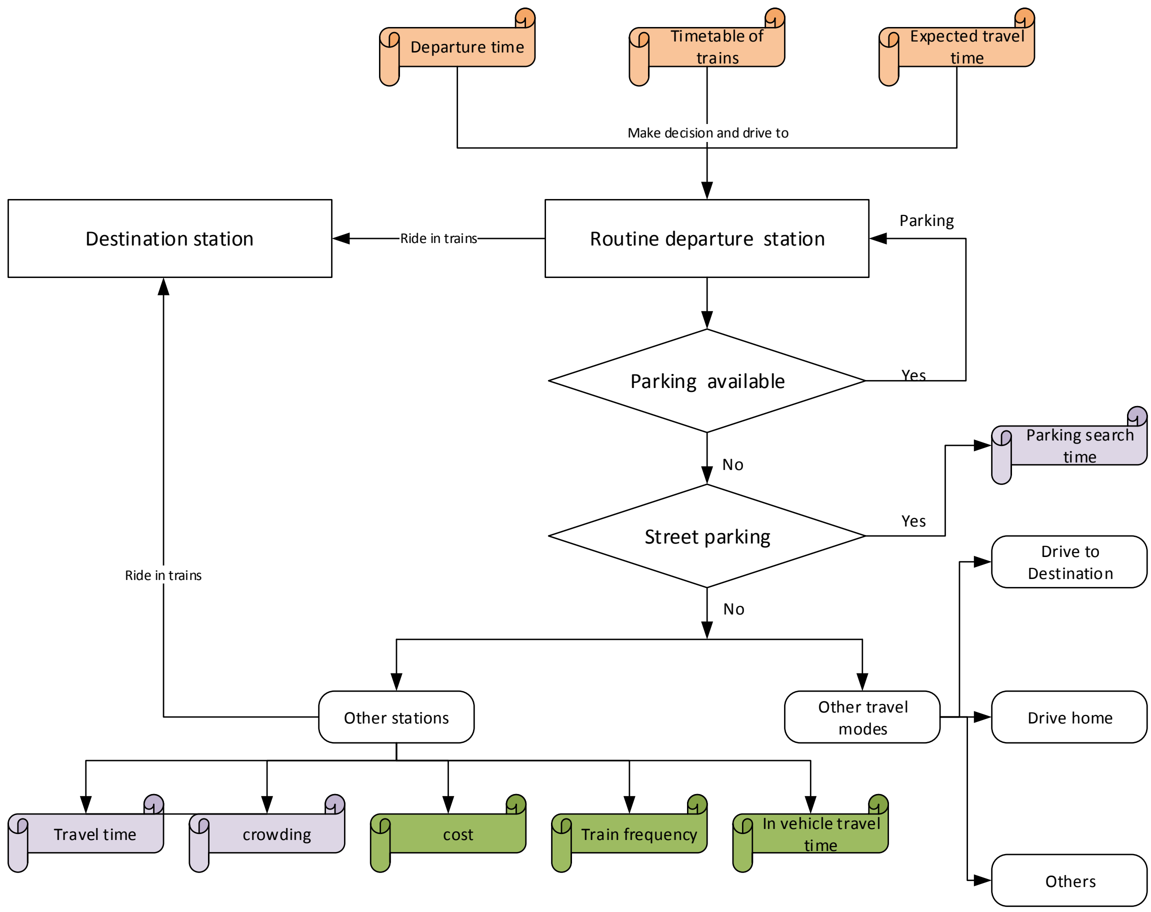


Figure 3.3 Decision-making process of station choice for P&R users

3. 4 Data

The data collected for and used in this research to develop the station choice model can be categorised into three types; train facilities data, observed travel data and train station choice data. The facilities data were mainly used to identify the attributes influencing station choice for P&R users. The observed travel data were used to determine the level of these attributes. The station choice data associated with those decisions made in the hypothetical situations that the researchers specified were mainly used to model the departure station choice of P&R users under uncertainty. Data related to the respondents' visual attention, i.e. the duration and frequency of the eyes' fixation on each attribute in the choice questionnaires, were collected through an eye tracking experiment. These data were used to improve the questionnaires and evaluate each attribute's significance on station choice.

3.4.1 Identification of train stations for collecting station choice data

The station choice model in the research was build using data from seven train stations in Perth, selected based on the following criteria.

(a) At least one station per train line should be selected

The current railway network in Perth has been gradually developed over nearly 140 years. The first line, Fremantle to CBD to Midland, opened in 1881, followed soon afterwards by the Perth to Armadale line (1889). After a hiatus of 90 years, the Perth to Joondalup line opened (1992), followed 15 years later by the Mandurian line (2007). There are significant differences in the locations, spacing, facilities/roles and land uses around the stations on the early, (pre-car) lines compared to those of the stations on the newer lines built in an age of high car ownership and in an attempt to attract people out of their cars. Therefore, to capture the influence and the specific characteristics of the different lines on P&R users' choice of departure train station, at least one station per train line should be investigated.

(b) Both stations in the middle and at the end of the line should be chosen, i.e. a middle station and the terminal station

Generally, train stations can be separated into two groups based on their locations, i.e. middle stations and terminal stations. Firstly, at terminal stations, all seats are available. At stations closer to the CBD, the availability of seats tends to reduce and, with it, the potential for having to stand, or possibly not being able to board, increases. In other words, a station further from the centre may be more attractive than a closer-in station. Secondly, the P&R parking capacities at the terminal stations are usually greater than at middle stations. Thirdly, the catchment areas that the terminal stations serve are usually larger than the areas served by the middle stations, indicating that the average travel times to the terminal stations are likely to be longer than the times to the middle stations. Fourthly, the station service quality at some terminal stations may be better than it at middle stations. These differences could lead to P&R users making a different choice of departure train station. Therefore, to explore the effects of crowding on trains, parking attributes, travel time to the station and service quality level on P&R users' choice of departure train station, both middle and terminal stations should be investigated.

(c) Two adjacent stations should be selected

As stations next to each other can compete for P&R users, a minimum of two adjacent stations on the same line, should be selected to determine the potential impact.

(d) Two stations on the same line but apart from each other should be selected

Stations on the same train line, even if well apart from each other, can still compete for patronage due to the differences in total travel time, their service quality, accessibility, parking availability, etc. Therefore, at least two stations on the same line but apart from each other (i.e. separated by a number of stations) should be investigated to capture these potential effects.

(e) Greater focus on the extended lines and newly opened stations

The Joondalup line serves the northern suburbs and was opened in 1992. Initially only three stations, Leederville, Edgewater, and Joondalup) were opened, with other stations opening in 1993. In the early years, the line had low passenger numbers but strong population growth in the northern corridor, congestion on the freeway and parking constraints in the city centre have resulted in increasing patronage and, eventually, overcrowding on some peak period trains and capacity issues at some P&R car parks. Measures have been taken to address these problems, including extending the line to Clarkson and later to Butler, adding more P&R spaces and opening a new P&R focussed station at Greenwood. In order to capture how the differences between a newly opened station and a station on the initial line affect P&R users' choice, at least two stations (i.e. one is on the initial line and one on the extended line) should be investigated.

(f) A range of car parking capacities should be selected

As the research aims to study P&R users' choice behaviour, only stations with a formal P&R car park were taken into consideration. Moreover, P&R parking capacity has a close relationship with parking search time for P&R users, which is one of uncertain factors we want to explore. Hence, a range of car parking capacities, (i.e. small, medium, and large), should be investigated.

(g) At least one station on the oldest line should be selected

As the Fremantle line was Perth's first railway line, some of its stations are the oldest on the Perth train network. While some improvements and upgrades have been made in the intervening years, their characteristics, locations, spacing, surrounding land uses, mode of operation and services are different to those of stations on the new lines. At least one of the oldest stations should be investigated to capture the impact that the above differences may have on how the P&R users using these oldest stations choose their departure station.

Based on the above criteria, the stations chosen to collect data were Claremont, Murdoch, Warnbro, Cannington, Greenwood, Warwick and Midland. Their locations and positions within the overall train network are shown on Figure 3.4, with Table 3.1 showing how each station satisfies the above criteria.

Table 3.1 The selection of stations for the study

Lines	Stations	Criteria						
		(a)	(b)	(c)	(d)	(e)	(f)	(g)
Fremantle	Claremont	✓	✓				✓	✓
Mandurah	Murdoch	✓	✓		✓		✓	
	Warnbro	✓	✓		✓		✓	
Armadale	Cannington	✓	✓				✓	
Joodalup	Greenwood	✓	✓	✓		✓	✓	
	Warwick	✓	✓	✓		✓	✓	
Midland	Midland	✓	✓				✓	

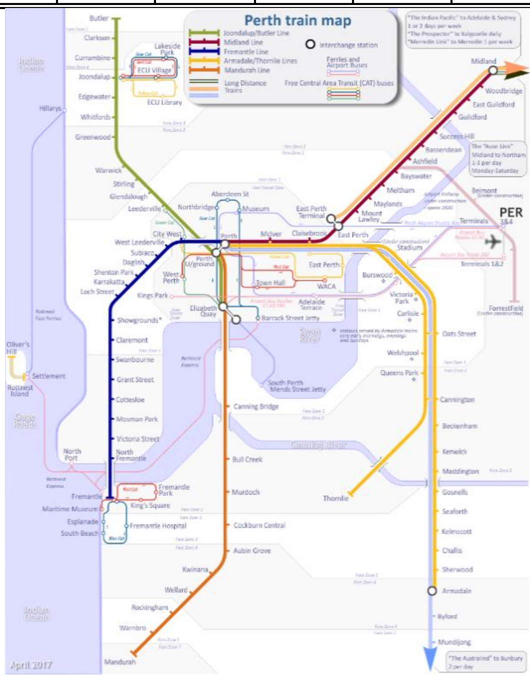


Figure 3.4 Locations of sampled train stations

Section 3.4.2 and Chapter 4 set out how the data were collected and collated.

3.4.2 Data Collection

The data used in the research were collected by three surveys: a train station's facilities survey, a train station survey and a train station choice survey. The first survey, conducted in April 2012, aimed to identify the facilities and information related to every train station on the Perth rail network, i.e. not just for the selected seven stations. The questionnaires used for the survey obtained the following: ① basic information on the train stations, including platform and station facilities, (such as toilets, seats, shelter, kiosk, etc.), bike and ride parking and the possibility of transferring to other travel modes; ② information on the P&R facilities, including the number of formal and informal, (e.g. adjacent streets), parking bays, safety facilities specific for P&R, etc.; and ③ the land-uses around the station, such as the type of land-use, the size of land parcels, etc. Example questionnaires can be seen in Appendices B-1 toB-3.

The second survey was called the train station survey. Its objective was to identify the factors influencing P&R users' choice and to determine the variations of all key factors. It was an intercept survey conducted at the seven stations (mentioned above) on 2 July 2012 jointly by Curtin University, the Western Australian University, the Department of Planning (DoP) and the Public Transport Authority (PTA). A sample of the questionnaire used in the survey is presented in Appendix C.

The third survey explored P&R users' choice of departure train station in an uncertain environment. It was conducted in November-December 2014 at the seven train stations and again was an intercept survey. The questionnaires were developed via a stated choice experiment which used the D-efficiency method, (see Chapter 4). Each respondent was presented with two hypothetical station choice situations and asked to choose a preferred station. The questionnaires can be seen in Appendix D. We conducted a number of pilot surveys to test and refine the SC questionnaires, (i.e. to test the questionnaires' readability and accuracy and their ability to correctly model P&R users' choice etc.), before carrying out the final station choice survey. These questionnaires can be seen in Appendices E-1 to E-5. The sample size and investigating time for each pilot survey can be seen in Table 3.2. More than 600 respondents were involved in the survey and around 2400 questionnaires were collected.

A summary of the surveys is shown in Table 3.2 with Figure 3.5 showing how they interrelate and their objectives.

Table 3.2 Summary of surveys

Name	Type	Stations	Sample number	Time period
Stations' facilities survey	Intercept	69		April 2012

Train station survey		Intercept	7	940	31/07/ 2012- 1/8/2012 (6:00AM - 4:00PM)
Station choice survey	Pilot surveys	Intercept	1-2	27-173	2013-2014
	Main survey	intercept	7	600	11/2014-12/2014

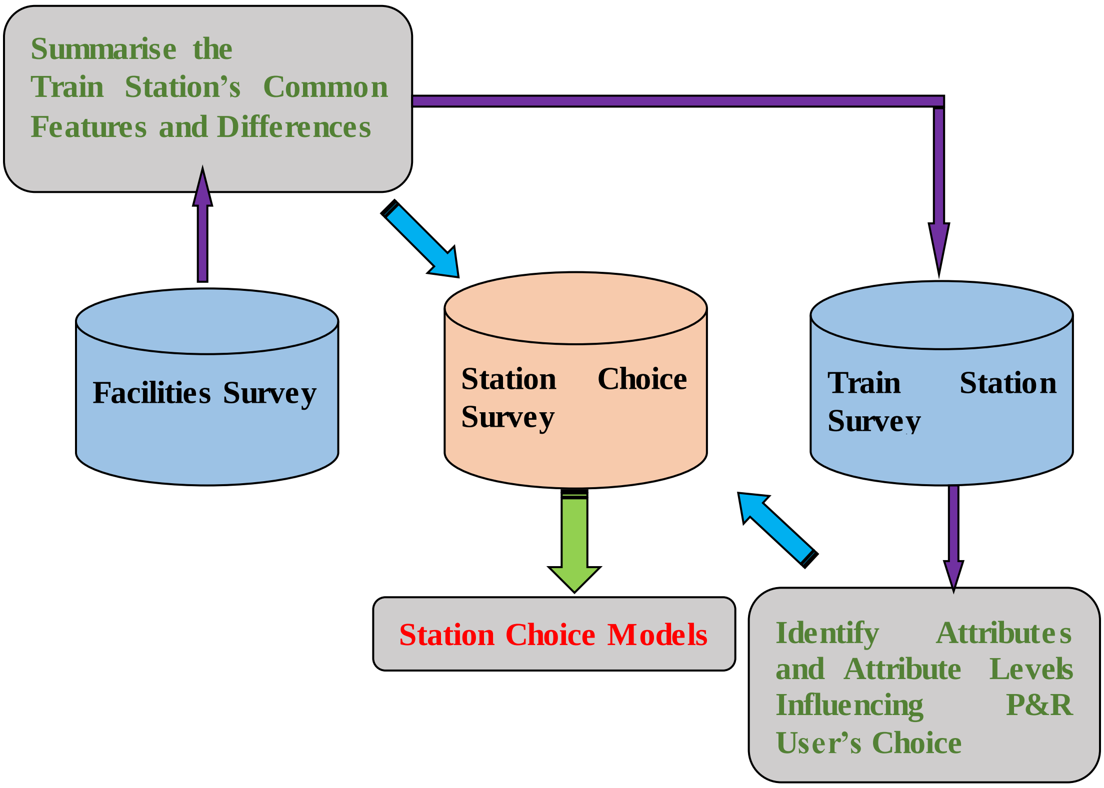


Figure 3.5 Relationship between the surveys and their objectives

3. 5 Data analysis and modelling

3.5.1 Modelling station choice under discrete choice theory

Station choice models were developed within a discrete choice framework in the research. There were two main reasons to choose this theory to quantitatively analyse and explain or predict P&R users' station choice behaviour. The first is that the choice set of train stations exhibits the three characteristics required of choice sets within a discrete choice framework, i.e. exclusiveness, exhaustive and finiteness. In practice, this means that P&R users can only choose one station once in a station choice set, station choice sets include all the stations P&R users can choose and the number of stations that will be chosen by P&R users can be determined (Fehr-Duda & Epper, 2011).

The second reason to choose a discrete choice framework is that we assumed that P&R users were rational and choose a station with an aim to maximise benefit or utility, which is consistent with the assumption in discrete choice models. Therefore, a discrete choice framework was used to develop the station choice models under uncertainty.

Discrete choice theory assumes that travellers' choice is aligned with their order of preference, which may be represented by a utility function (Train, 2003). A P&R user, labelled n, chooses a departure train station among J alternatives. The utility that the P&R user n obtains from alternative j is $U_{nj}, j=1,\cdots J$. He/she chooses the alternative that maximises his/her utility. Therefore, P&R user n chooses alternative i if and only if $U_{ni} > U_{nj} \quad \forall j \neq i$. In general, this utility is known to the P&R user but not to the researcher. The researcher only knows the component of the utility (V_{ni}) relating to the observed attributes at station i and of the P&R user n. As researchers cannot observe or measure every factor affecting the station choice made by the P&R user, V_{ni} is not equal to total utility U_{ni}. Thus, total utility can be decomposed as: $U_{ni} = V_{ni} + \varepsilon_{ni}$, where ε_{ni} captures factors not included in V_{ni}. As the researcher does not know the characteristics of ε_{ni}, it is treated as a random term. The joint density of these random terms can be denoted as $f(\varepsilon_n)$. With this density, the researcher can make probabilistic statements about the P&R users' choice. The probability that the P&R user n chooses departure train station i is (Train, 2003):

$$\begin{aligned}\pi_{ni} &= prob(U_{ni} > U_{nj}, \forall j \in j=1,\ldots,J, i \neq j) \\ &= prob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj,} \forall j \in j=1,\ldots,J, i \neq j), \\ &= prob(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}, \forall j \in j=1,\ldots,J, i \neq j)\end{aligned} \tag{3-1}$$

This probability is a cumulative distribution function and using the density $f(\varepsilon_n)$, can be rewritten as (Train, 2003) (Train, 2003):

$$\begin{aligned}p_{ni} &= \int_{\varepsilon}^{+\infty} I\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}, \forall j \in j=1,\ldots J; i \neq j\right) f\left(\varepsilon_n\right) d\varepsilon_n \\ &= \int_{\varepsilon}^{+\infty} I\left(\varepsilon_{nj} < \varepsilon_{ni} + V_{ni} - V_{nj}, \forall j \in j=1,\ldots J; i \neq j\right) f\left(\varepsilon_n\right) d\varepsilon_n\end{aligned} \tag{3-2}$$

where $I(\bullet)$ is the indicator function, equal to one when the term in parentheses is true and zero otherwise.

The discrete choice models mentioned above can be derived based on the multidimensional integrations of the density of the unobserved utility.

Different choice models can be derived based on the distribution of the unobserved factors. Logit and mix logit were also used in the research and are explained below.

(a) Logit model

Logit is the most widely used discrete choice model so far due to its convenience, i.e. the choice probabilities equation has a closed form. It is derived under the assumption that the unobserved factors (ε) are distributed IID, (i.e. independent and identical distribution), with extreme value I. In other words, the unobserved utility is uncorrelated over the alternatives and has the same variance for all alternatives. The density for each unobserved utility is (Train, 2003):

$$f\left(\varepsilon_{nj}\right) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}} \tag{3-3}$$

And the cumulative distribution is:

$$F\left(\varepsilon_{nj}\right) = e^{-e^{-\varepsilon_{nj}}} \tag{3-4}$$

where ε_{nj} is the unobserved utility obtained from alternative j by respondent n and can be written as the product of the individual cumulative distribution.

If ε_{ni} is given, then the individual cumulative distribution can be written as:

$$F\left(\varepsilon_{nj}\right) = e^{-e^{-\left(\varepsilon_{ni}+V_{ni}-V_{nj}\right)}} \tag{3-5}$$

Hence, the probability of P&R user n choosing alternative j is the product of the individual cumulative distribution. The specification of choice probability is shown as follows:

$$P_{ni} \left| \varepsilon_{ni} = \prod_{j \neq i} e^{-e^{-\left(\varepsilon_{ni}+V_{ni}-V_{nj}\right)}} \right. \tag{3-6}$$

In fact, ε_{ni} is not given, so the choice probability is the integral of the conditional probability overall all values of ε_{ni} weighted by its density. The choice probability formula can be seen in equation (3-7).

$$P_{ni} = \int \left(\prod_{j \neq i} e^{-e^{-(\varepsilon_{ni} + V_{ni} - V_{nj})}} \right) e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}} d\varepsilon_{ni} \tag{3-7}$$

After manipulating the integral, the choice probability within logit can be calculated based on a succinct and closed-form expression, as shown below:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \tag{3-8}$$

This approach is mainly used to develop travel time and crowding sub-models. The detailed applications can be seen in Chapters 5 and 6.

(b) Mixed logit (ML) model

Similar to any random utility functions in discrete choice models, the utility associated with alternative i in an ML model, as evaluated by respondent n in choice task t, can be expressed by equation (3-9) (Hensher & Greene, 2003; Mørkbak & Nordström, 2009).

$$U_{nit} = \beta_n X_{nit} + e_{nit} \tag{3-9}$$

where X_{nit} refers to the attributes of the i^{th} alternative faced by the respondent n in the choice situation t, which is a vector of independent variables that can be observed by the analyst. β_n and e_{nit} are unobserved and treated as random influences. In a logit context, the e_{nit} is limited to independent and irrelative distribution (IID) with extreme value tyre I, which means the error items of different alternatives are not correlated. However, that is not the case in reality. Therefore, researchers have taken it into account in other ways. One way is to split the random component into two parts. One assumes an independent and irrelative distribution for all alternatives and all individuals and the other is correlated over the alternatives and heteroskedastic, which means that the variance of the random variables is not the same for all alternatives. Based on this, equation (3-9) can be changed into (3-10):

$$U_{ni} = \beta^{'} X_{ni} + [\eta_{ni} + \varepsilon_{ni}] \tag{3-10}$$

where η_{ni} is random and has a mean equal to zero and distributed over individuals and alternatives based on parameters and observed data; ε_{ni} is also random and has zero mean, but IID over alternatives and is not supported by underlying parameters or data.

As the ε has IID extreme value I distribution, the conditional probability for choice is still logit for a given value of η, and its specification can be written as equation (3-11):

$$L_i(\eta) = \exp\left(\beta' x_i + \eta_i\right) \Big/ \sum_j \exp(\beta' x_j + \eta_j) \tag{3-11}$$

Actually, the η is not given, and therefore the choice probabilities should be integrated over all values of η, weighted by the density of η, as given in equation (3-12).

$$P_i = \int L_i(\eta) f\left(\eta | \Omega\right) d\eta \tag{3-12}$$

where $f\left(\eta|\Omega\right)$ is the density of η; and Ω are the fixed parameters of the distribution.

ML models allow the unobserved factors to correlate over all alternatives, follow any distribution and for an individual's taste to vary randomly. They are therefore considered to be the most flexible of discrete choice models.

Additionally, ML models can be used to analyse the preference heterogeneity over the sample population. Assuming β is not fixed but random, we can specify each β_n for each attribute of an alternative using a mean and a standard deviation. The standard deviation of a parameter β can indicate an individual's preference heterogeneity within the sample population. In order to obtain this heterogeneity through data segmentation, the individual's conditional distribution based on his/her choice should be derived first. Within Bayes Rule, the conditional distribution can be written as follows:

$$H_n\left(\beta|\Omega\right) = L_n\left(\beta\right) g\left(\beta|\Omega\right) \Big/ P_n\left(\Omega\right) \tag{3-13}$$

where $L_q\left(\beta\right)$ is the probability of an individual's choice if they had this specific β; $g\left(\beta|\Omega\right)$ is the distribution in the population of βs; and $p_q\left(\theta\right)$ is the choice probability function defined in open form as:

$$P_n\left(\theta\right) = \int L_n\left(\beta\right) g\left(\beta|\theta\right) d\beta \tag{3-14}$$

In the research, the ML approach was used to develop the parking search time sub-model and the overall model, and to analyse how an individual's preference heterogeneity influences his/her choice of departure train station. The detailed application can be seen Chapters 7 and 8.

3.5.2 Decision making theories under uncertainty (or risk)

"Risk" refers to situations where there is a range of possible outcomes and decision makers make their choice based on the probability of each outcome occurring and benefits or disbenefits arising from that outcome, while "uncertainty" relates to situations where the probabilities of each outcome occurring vary or are unable to be fully determined by the decision maker (Knight, 1921). Based on these concepts, P&R users' choice of departure train stations can be defined as decision making under uncertainty (or risk). Three key factors influencing station choice for P&R users, (parking search time, travel time to the departure train station and crowding on trains), can vary over time as travel environments around train stations and on trains change. However, as we cannot reliably and accurately predict these changes in advance, the variations of these three key factors being random. Correspondingly, the station choice for P&R users should be taken as a decision making under uncertainty (or risk) process.

Three decision making theories under uncertainty (or risk) were used to develop the utility function related to station choice and analyse respondents' attitudes towards the variation of uncertain factors. They are explained below.

(a) Expected utility theory

The basic economic theory for decision making under uncertainty (or risk) is expected utility theory. Initially proposed by von Neumann and Morgenstern (1947), it has dominated the analysis of decision making under uncertainty (or risk) for many years. It is generally accepted as the normative model of rational choice and is widely used as a descriptive model of economic behaviour. The theory assumes that all rational decision makers follow the axioms of the expected utility maximum. Therefore, the utility of any alternative can be written simply as the mathematical expectation of the utility of the outcomes. Its specification is shown in equation (3-15), when the expected utility is assumed as linear-additive:

$$E(U) = \sum_i p_i x_i \tag{3-15}$$

where $E(U)$ is the expected utility; x_i is the i^{th} outcome; and p_i is the objective probability that the i^{th} outcome occurs.

Furthermore, a power form utility functional has been suggested (see Eq. (3-16)), which can show not only decision making under uncertainty but also respondents' risk attitude.

$$E(U) = \sum_i p x_i^{\gamma} \tag{3-16}$$

where γ is an estimated parameter indicating decision makers' risk attitude towards the outcome i^{th}, when $\gamma < 1$ means risk aversion; $\gamma = 1$ is risk neutral; and $\gamma > 1$ is risk seeking.

The risk aversion is often obtained based on the EUT.

(b) Extended expected utility theory (EEUT)

Extended expected utility theory (EEUT), proposed by Li et al. (2009), is, as the name indicates, an extended version of the EUT. It still retains the axiom of expect utility maximum but differs from the EUT in how the probabilities used to weigh the utility are determined. Within the EUT, they are objective, i.e. reflect the actual frequencies at which the various outcomes occurred. However, within the EEUT, the probabilities are subjective and as perceived by respondents in terms of their experience and cognition. In practice, the probabilities transformed by respondents may overweight or underweight the objective probabilities in the EUT, especially for extreme situations, and a non-linear probability weighting function was introduced into the EUT to form the EEUT. Its specification is given in equation (3-17).

$$EE(U) = \sum_m \left(w(p_m) \times U \right) \tag{3-17}$$

where $w(\bullet)$ is a probability weighting function.

The theory is applied to the estimation of the utility of station choice based on the effects of the variation of crowding on trains (see Chapter 6).

(c) Prospect theory (PT) and cumulative prospect theory (CPT)

In contrast to the EUT, which aims to help individuals achieve better decisions, prospect theory (PT) and cumulative prospect theory (CPT) simply describe people's decision-making behaviour. Moreover, decision making within PT and CPT focus on the difference between the actual outcome and the expected outcome that is potential gains or losses, instead of the final outcome as evidenced in the EUT.

Prospect theory (PT) was proposed by Kahneman and Tversky (1979). It divided an individual's choice process into two stages: ① "editing" phase, where gains and losses relative to some neutral reference point (τ) were identified; and ② "evaluation" phase, where choice was made based on the outcome of alternatives by evaluating their value function $v(x)$ and weighting function $\pi\left(p\right)$. Based on this, the utility function of prospect n under PT is defined as follows (Timmermans, 2010; Tversky & Kahneman, 1992):

$$u_i^n = \sum_{j=1}^{J} \pi\left(p_l^n\right) v\left(x_j^n - \tau\right) \tag{3-18}$$

Prospect s^n is preferred to $s^{n'}$ *iff*

$$u_i^n > u_i^{n'} \qquad \forall s_i^n \neq s_i^{n'} \tag{3-19}$$

PT has two features: ① the value function $v(x)$ is concave for gains and convex for losses, with losses steeper than gains; and ② a nonlinear transformation of the individual's probability leads to overweight small probabilities and underweight high probabilities (Kahneman & Tversky, 1979).

Contrary to EUT, PT considers choices between high risk prospects with a small number of outcomes by transforming the individual probability using a non-linear function (Kahneman & Tversky, 1979). Later, Tversky and Kahneman (1992) extended PT by converting the individual probabilities into a cumulative probability distribution function and, thereby, formed cumulative prospect theory (CPT), which can explain not only uncertain behaviour but also the risk prospect for any number of outcomes.

Similar to PT, the uncertain prospect based on CPT is still the sum of the utility of gains and losses, each weighted by its own weighting function. The decision weights under CPT are the subjective weightings derived from the outcome probability given by equation (3-20) (Tversky & Kahneman, 1992):

$$\begin{aligned} w_1 &= p(p_1) \\ w_i &= p(p_1 + \cdots + p_i) - p\left(p_1 + \cdots + p_{i-1}\right) \quad \mathit{for}\ 1 < i < n \end{aligned} \tag{3-20}$$

where $p(p)$ is a monotonic risk weighting function restrained by $0 \le p(p) \le 1$.

Many different functional forms have been suggested for the risk weighting. The top four risk weighting function forms identified in a meta-analysis conducted by Stott

(2006), have been tested in the research. Their specifications are shown in equation (3-21) - (3-24).

TK $$\pi\left(p_i\right)=\frac{p_i^r}{\left(p_i^r+(1-p_i)^r\right)^{\left(\frac{1}{r}\right)}} \tag{3-21}$$

GE $$\pi\left(p_i\right)=\frac{sp_i^{\,r}}{sp_i^{\,r}+(1-p_i)^r} \tag{3-22}$$

Prl-I $$\pi\left(p_i\right)=e^{-(-\ln p_i)^r} \tag{3-23}$$

Prl-II $$\pi\left(p_i\right)=e^{-s(-\ln p_i)^r} \tag{3-24}$$

where p_i is the probability that the i^{th} outcome occurs; $\pi\left(p_i\right)$ is the subjective weighting function derived from the outcome cumulative probability; and r and s indicate the shape and location of the risk weighting functions.

Similar to the weighting functions, many different forms of value functions are available. Based on the examination of 256 combinations of value functions, weighting functions and stochastic choice models by Stott (2006), the best choice prediction model should use a combination of the value function with power form, the weighting function with TK form, and logit models. Therefore, the value function with power form was used to develop the station choice model. Its specification is given in (3-25).

$$v\left(x_j^n-\tau\right)=\begin{cases}\left(x_j^n-\tau\right)^{\alpha} & if\ \left(x_j^n-\tau\right)>0\\ -\lambda\left|x_j^n-\tau\right|^{\beta} & if\ \left(x_j^n-\tau\right)<0\end{cases} \tag{3-25}$$

where parameters α and β (less than or equal to one) measure the level of sensitivity to changes in both directions from the reference point, while parameter λ (≥ 1) captures the degree of aversion to loss. The value function under PT is usually S-shaped. It is generally concave for gains and commonly convex for losses; with losses steeper than gains if it describes loss aversion.

(d) Comparison of expected utility theory with extended expected utility theory and cumulative prospect theory

Even though expected utility theory (EUT), prospect theory (PT), cumulative prospect theory (CPT) and extended expected utility theory (EEUT) can all be used for understanding choice behaviour under uncertainty (or risk), PT/CPT and EEUT may be applied to develop better utility functions of the station choice under uncertainty than EUT.

EUT is a normative decision-making theory due to its simple and parsimonious format and consistency with evidence for most choice situations. The main criticism of EUT lies with its three key components:

- Expectation, which means the overall utility of prospect equals the sum of the expected utility of its outcomes;
- Assess integration, which refers to the utility function in EUT depending on the final states rather than gains or losses;
- Risk aversion, which means a person is risk averse if he prefers a certain prospect over a riskier prospect with a higher value, which implies that the utility function in EUT is concave.

These axioms have been violated by a serious of phenomena. PT/CPT is one of decision-making theories that can accommodate most of these violations and is good at explaining the choice between risky prospects with small probabilities of outcome. In contrast to EUT, the value function in PT is concave for gains, convex for losses and steeper for losses than for gains. Moreover, it uses a non-linear transformation of the probability scale, which overweighs small probabilities and underweights moderate and high probabilities. Cumulative prospect theory (CPT) is a new version of PT, which not only contains the functions of PT but also incorporates cumulative functionality into PT, so it can predict uncertain as well as risk prospects with any number of outcomes. Therefore, CPT should perform better than EUT in explaining station choice under uncertainty.

Another criticism of EUT is that the probability weighting function in EUT is objective, so it cannot recognise perceptual processing of respondents' decision making. However, EEUT and PT/CPT replace the objective probabilities in EUT with subjective probabilities with either cumulative or non-linear probability weighting, which entails elements of over and under-weighting, especially in an extreme or rare situation (de Finetti, 1937; Li et al., 2009; Tversky & Kahneman, 1992).

Based on the above, CPT and EEUT have been applied to explain the effect of different uncertain factors on station choice for P&R users under uncertainty.

3.5.3 Models measuring reliability (or variability) of uncertain factors

Reliability, or variability, has been widely used to assess the impact of uncertainty on travel choice (Carrion & Levinson, 2012). Actually, a choice related to travel is seldom affected by small variations in the uncertain factors (Hensher, Greene, & Li, 2011; Li et al., 2010; Nicholson & Du, 1997; Wong & Sussman, 1973). The standard deviation, as well as the difference between the 25th & 75th, and 90th & 50th percentiles of travel time, have been adopted as measurements of travel time variability by Jackson and Jucker (1982). Early research on travel time variability often utilised stated preference (SP) data (Department of Transport). Given that the SP data are collected by SC experiments, where respondents are making hypothetical choices which could be different to what they would actually do in a real situation, more recent research use a mix of SP and revealed preference (RP) data to explore the effect of variability of travel time on travel choice (Li et al., 2010).

Two approaches evaluating the influence of variations in travel time on travel choices were considered in the research, mean-variance deviation and scheduling delay.

(a) Mean-variance

The approach, proposed by Jackson and Jucker (1982), defines the utility of a choice option as a function of expected travel time (or mean travel time) (μ_T) and variability in travel time (standard-deviation) (δ_T). The commuter minimises the sum of these two terms when choosing a departure station. Its specification is shown in (3-26).

$$U = \gamma_1 \mu_T + \gamma_2 \delta_T \tag{3-26}$$

where γ_1, γ_2 are the estimated coefficients.

The model is usually estimated using discrete choice methods and has been successfully used in previous studies to predict a range of travel choices including route, mode and departure time choice.

(b) Scheduling delays

Similar to the mean-variance approach, the scheduling delay approach, proposed by Noland and small (1995), is based on a traditional utility maximisation framework. However, it focuses on the time constraints, (e.g. train departure time), a traveller may face and their associated costs due to an early or late arrival at the final destination. There is a desire to minimise the frequency of late arrivals and to maximise the time spent at home compared to waiting at the train station (Carrion, 2010; Carrion & Levinson, 2012; Gaver, 1968; Knight, 1974). The general formula is given by equation (3-27).

$$U\left(t_d;PAT\right)=\gamma_1 T+\gamma_2 SDE+\gamma_3 SDL+\gamma_4 DL \qquad (3\text{-}27)$$

where T refers to total travel time; SDE is the amount of time one arrives at a destination earlier than desired which is defined as $Max(0,PAT-\left[T+t_d\right])$; SDL is the amount of time one arrives later than desired which is defined as $Max(0,\left[T+t_d\right]-PAT)$; PAT is the preferred travel time; t_d is departure time; DL is a fixed penalty for any late arrival; and the γ s are estimated coefficients.

The model, historically, is mainly linked to departure time choice.

(c) Summary

Some researchers have found the two theoretical approaches to be broadly equivalent. In this research, we have chosen the mean-variance approach to measure the effect of variability of the uncertain factors on station choice due to the scheduling delay approach being more complex. Its application can be seen in Chapters 5, 6 and 7.

3.5.4 Station choice models under uncertainty for P&R users

Traditionally, station choice models were developed based on the effects of certain factors and linear utility functions were often used, with linear-additive properties assumed. However, this research aims to capture the effects of uncertainty and respondents' risk attitude towards the variation of the uncertain factors on station choice. Therefore, non-linear utility functions are used, which makes the station choice model more comprehensive. We divided the effects of factors on station choice into certain and uncertain effects, and tested the effects of each uncertain factor on station choice with non-liner functions. Then, these uncertain effects were taken as new indices and, together with the certain factors, were used to develop the linear utility function of the overall station choice model. In order to identify the effects of the uncertain factors on station choice, we established three sub-models using EEUT and CPT. The procedures for establishing each sub-model are explained below.

(a) Travel time sub-model

As travel time has been identified as an uncertain factor, the travel time sub-model was used to analyse station choice based on the effects of travel time under uncertainty.

The travel time sub-model was developed within discrete choice theory and the mean-variance method was applied to establish its utility function. In contrast to the classical measurement of variance, the variance used in the travel time sub-model was re-defined within cumulative prospect theory (CPT). The regular travel time, (defined as the travel time spent in most of days in one month), was taken as the reference point, with the difference between the regular travel time and travel time spent on good days treated as gains, and the difference between regular travel time and travel time spent on bad days as losses. Moreover, the frequencies that good days or bad days occurred in a week were taken as the objective probabilities that a gain or loss would occur. Therefore, the risk prospect, obtained based on gains and losses within CPT, was taken into account with the variation in travel time. The parameter coefficients in the sub-model were estimated using the Nlogit 5 Package.

The process to develop the sub-model is set out below.

- ***Step 1*:** Identifying unreliability of variation of travel time

Travel time in the research refers to the time spent in travelling from home to the departure station. It can be divided into two components based on a P&R user's routine travel, i.e. regular travel time and additional/reduced travel time. The former refers to the time spent in driving to the departure train station on most days. The latter refers to the change in travel time resulting from changes in the traffic conditions, which leads to unreliability of travel time.

In order to explore the effects of variations in travel time on station choice, we constructed a new variable within CPT to replace the variance in the mean-variance framework. It is made up of two components. The first is the variation caused by the difference between the good day travel time and regular travel time and its frequency and the second, the difference between the bad day travel time and regular travel time and its frequency. Assuming both are linear and additive, the new variable is their sum. Correspondingly, the coefficient of the variable can indicate P&R users' preference for variation of travel time.

- ***Step 2*:** Determining the value function form

Many value functions, combined with various weighting functions, can be used for CPT (Stott, 2006). According to the meta-analysis for combining value functions, weighting functions and stochastic choice models, the model made of the value function with power form, the weighting function with TK form and logit model is the best to make a prediction. Moreover, the shape of the value function with power form can indicate respondents' risk attitude, (as detailed in Chapter 5). Therefore, the value function with power form is used to measure the value of the outcomes in the sub-model.

- ***Step 3*:** Measuring respondent's risk attitude

Within CPT, the respondents' risk attitude is determined jointly by the value function and the weight function. Assuming a choice is between a prospect (x, p) with expected value (px), and the reference point is zero, a respondent's attitude is risk seeking when $\pi(p) > v(px)/v(x)$ and its shape is concave in the domain of gains. Based on this method, a respondent's risk attitude towards the variation of travel time can be determined.

- ***Step 4 :*** Analysing the impact of a respondent's real travel time experience on their risk attitude towards their station choice under uncertainty

We chose three train stations for this analysis, (Murdoch station, Greenwood station and Warnbro). Firstly, we geocoded the home end of ten P&R users using these stations and then used google maps to extract the shortest home-to-station travel time at 15 minute intervals between 6:00am and10:00am from Monday to Friday. From these data we obtained the daily variations in travel time for each station and ranked them. Next, we used the travel time sub-model to estimate respondents' risk attitude for each station and ranked them. We then compared both rankings.

Additionally, we calculated the difference between travel times estimated from Goggle Map and perceived travel times from the train station choice survey, and ranked them. From this, we drew conclusions on whether respondents who have higher differences between perceived and objective travel times tend to be more risk averse towards their station choice under travel time variability than those who have experienced less travel time variability.

(b) Parking search time sub-model

Similar to the travel time sub-model, the parking search time sub-model aims to check the effects of parking attributes, including variability of PST, on station choice for P&R users. It was developed using the mixed logit approach. The reactions of P&R users to the variability of PST were evaluated using the mean-variance approach and the variation of PST was measured within CPT. The main tasks for the sub-model are listed below:

- ***Task 1:*** Developing utility function of station choice based on the effects of variability of PST

The utility function is developed based on two situations, i.e. whether or not a parking space is available in the P&R car park. If parking is available, the variation of PST, based on the train service survey, is negligibly and can be ignored. Therefore, certain parking attributes, such as parking fee, parking capacity, parking availability in P&R parking lots, etc., were taken as the main factors contributing to utility of station choice. On the assumption that these factors are linear additive, a linear utility function was developed.

When parking is unavailable within the P&R car park, P&R users have to park their cars on the streets, verges, and other parking lots around the station. In this case the PST can vary significantly, as the alternative parking locations and parking availability vary with each station. Correspondingly, the day to day variation of PST should contribute to the disutility of station choice. Additionally, parking cost, determined based on the frequency of patrolling for illegal parking and the penalties, would lead to disutility of station choice as well. In summary, parking attributes only impact on the disutility of station choice when parking is unavailable in the P&R car park.

Based on this, the utility functions of station choice due to the effects of parking attributes should be established separately. For the former (i.e. parking is available in P&R car park), assuming all factors are linear additive, a linear utility function can be developed. For the latter, (that is parking is not available), the effect of variability of PST on station choice was measured with CPT first, then a linear utility function developed.

- ***Task 2:*** Determining parameters' distribution

Once the utility function of station choice was developed, the parking search time sub-model was established using a mixed logit (ML) approach.

As stated above, the ML model uses a range of βs to produce a mixed logit function $f(\beta)$, so it is necessary to identify parameters' distribution first. Normal distribution was tested. The recommended distribution is dependent upon the statistical results of the two models with different distributions.

- ***Task 3:*** Analysing P&R users' preference for each factor
- ***Task 4:*** Exploring P&R users' risk attitude towards variations in PST

Similar to the analysis of P&R users' risk attitude towards variations in travel time, their risk attribute towards variations in PST can be determined jointly based on the value function and weighting function of the PST as well as its shape.

(c) Crowding on trains sub-model

The sub-model is used to explore P&R users' station choice behaviour under the effects of variations in crowding on trains and to measure their risk attitude to crowding. A multinomial logit was used to develop the sub-model, with the utility function developed using the mean-variance method and the parameter coefficients estimated using the Nlogit 5 software package. Additionally, the relationship between the respondents' risk attitude and railway ridership was analysed based on the sub-model. The main steps to develop the sub-model are as follows:

- ***Step 1:*** Using stated choice (SC) survey to identify key factors measuring crowding

Both measurements of crowding, (i.e. seat availability and density of passengers standing in a train car), in previous questionnaires surveying respondents' reaction to crowding on trains were presented from above, (i.e. a bird's eye), so they can be used independently in the utility function. This does not reflect the actual view of a passenger standing on the platform waiting to board a train. They see the view from the front as the doors open and have to judge whether there are seats available and the level of crowding from that viewpoint. Therefore, to better represent reality, the crowding levels in this questionnaire were presented to respondents from the front view, i.e. the pictures in the questionnaires showed the situation you could see when the train doors open. Boarders cannot know clearly the density of those standing in the train when the train is crowded, a factor often exacerbated by people tending to cluster near the doors rather than moving down into the carriage. The crowding level boarders perceived may therefore not be the actual level. We needed to construct a new index to use in the MNL crowding sub-mode that reflects boarders perceived crowding level in the questionnaire, i.e. considers the interaction between seat availability and the density of people standing. Additionally, the interaction between in-vehicle travel time and the probability that seats have been taken, i.e. how long they have to stand, will affect P&R users' choice.

- ***Step 2***: Developing the crowding sub-model

The crowding sub-model was developed within discrete choice theory. Its utility specifications were established for three decision making theories under risk (or uncertainty) separately, from traditional and dominant EUT to CPT to Morden EEUT. Based on statistical indices of these models, the recommended model can be determined.

- ***Step 3***: Analysing respondents' risk attitude towards crowding on trains

Given that the value functions in these utility functions, developed under the three theories, were of the power form, both the power value and its shape can indicate respondents' risk attitude towards crowding.

- ***Step 4:*** Analysing the relationship between the respondents' risk attitude and train boardings

By comparing respondents' risk attitude for each train with their boarding number, we derived a relation between them. The risk attitude for each station was estimated using the recommended model based on the data from each station, with the boarding numbers provided by our cooperative partners.

- ***Step 5:*** Conducting sensitivity tests for identifying the effect of crowding on trains on station choice.

- ***Step 6:*** Analysing individual's preference heterogeneity

This section aims to identify how P&R users with different annual incomes choose their departure train station based on the effect of crowding on trains. A latent class model, with classes based on an individual's annual income, was used to calculate the probability that a station would be chosen by a particular income class. The differences in the conditional class probabilities indicate the effect of individuals' preference heterogeneity on station choice.

- ***Step 7:*** Elasticity analysis

- ***Step 8:*** Validating the model.

A Chi-squared (χ^2) test was used to validate the sub-model. We calculated the χ^2 based on the estimated probabilities from the recommended model and observed probabilities (frequencies) for each station (see Chapter 6). Then, the values were compared with the critical value to determine whether or not the model was acceptable, that is, validated.

3. 6 Software

3.6.1 The software to design stated preference experiment

(a) Evolver

Evolver, a sophisticated optimisation tool for spreadsheets, was used to identify the optimal scenarios in the SC experimental design. Its key role is to quickly solve problems modelled in Excel using innovative genetic algorithms (GA) and linear programming technology.

In the research, Evolver was used to determine the appropriate scenarios for conducting the train station choice survey.

(b) SketchUp

SketchUp is a simple, easy to use, 3D drafting software package used in a range of drafting and design fields. Moreover, it can modify and edit pictures imported from other software, such as CAD and BIM.

It was used to design the crowding levels in the questionnaires for the station choice survey. The indices reflecting crowding levels include the probability that seats have been taken and the density of passengers standing in a carriage. Therefore, we drew a 3D picture, showing the situation in a train car, based on the two indices and allowing respondents to clearly and easily identify crowding levels.

3.6.2 Software to analyse data

Statistical Package for the Social Science (SPSS) is a software package used for statistical analysis. Its roles included statistical analysis, data management, data documentation, etc.

The database (including RP and SP data) for the research was set up with SPSS and the descriptive statistics were also analysed using SPSS tools. The range and levels for each attribute used in the station choice survey were determined using classify tools within SPSS.

3.6.3 Software to develop model

NLOGIT is an extension of the software package LIMDEP. It can estimate a variety of discrete choice models (i.e. logit models) by full information maximum likelihood and provide programs for model simulation and analysis of discrete choice data. It is mainly used in discrete choice modelling based on panel data or real data from repeated observations of choice situations.

In this research, Nlogit 5 was mainly used to estimate the coefficients of the parameters in the station choice models, analyses individuals' preference heterogeneity and test sensitivity test.

3. 7 Chapter summary

This chapter has described the methodology used to develop station choice models under uncertainty for P&R users; to evaluate the effects of variation in the uncertain factors on station choice; to measure the variability of the uncertain factors and respondents' risk attitude to these uncertainties and to analyse individuals' preference heterogeneity.

In summary, all station choice models were developed within a discreet choice framework and the mean–variance approach was applied to establish the utility function, in which the effect of variation of crowding on station choice was estimated under CPT or EEUT. All the models were estimated using Nlogit 5. The respondents' risk attitudes towards the variation of uncertain factors were determined based on the combination of the value function with its weighting function.

As the process to model station choice under uncertainty is relatively complex, it was initially divided into a number of independent parts and then the parts were combined to produce a single model. In other words, we developed three sub-models of station choice based on the effect of three uncertain factors first, then, together with other certain factors, developed the overall model of station choice.

The study area and data collection were also introduced in the chapter together with the software used to design the stated choice (SC) experiment, and estimate and analyse the station choice model.

CHAPTER 4 EXPERIMENTAL DESIGN

The last chapter discussed a methodology framework for modelling station choice under uncertainty for P&R users. The first step is to collect sufficient relevant and reliable data. Therefore, this chapter identifies the data used in the research and then describes the design of the stated choice (SC) and eye tracking experiments used to collect the data required to develop and validate the model.

The chapter is structured as follows. Section 4.1 identifies the data used to develop and validate the station choice model. Section 4.2 presents a general SC experimental design workflow. Section 4.3 explains SC experimental design strategies. Section 4.4 designs an SC experiment to explore the station choice behaviour of P&R users. Section 4.5 designs an eye tracking experiment to validate the whole station choice model.

4. 1 Identification of data collection methods

Two types of data were used to explore station choice behaviour, revealed preference (RP) data and stated preference (SP) data. The former refers to the data related to what the respondent actually did, so they can provide analysts with the properties of reliability (Abdel-Aty, Kitamura, & Jovanis, 1995; Morikawa, Ben-Akiva, & Yamada, 1991). The latter are the data collected via responses to hypothetical situations to understand the preference of respondents (Ben-Akiva et al., 1994).

Most station choice studies use RP data, usually collected directly from passengers waiting at train stations or on-board the train (Lythgoe & Wardman, 2004; Mahmoud et al., 2014; Young & Blainey, 2018). Even though the use of RP data has certain methodological advantages over the use of SP data, they cannot be used for analysing choice behaviour for non-existing situations. Moreover, the collection of the required large sample size, compared with SP data, is costly and time-consuming. Certainly, SP data can be subjective, however, many studies have proven its authenticity (Abdel-Aty et al., 1995; Abdel-Aty, Kitamura, & Jovanis, 1997). A comparison of RP data with SP data can be seen in Table 4.1.

Table 4.1 Comparision between PR data and SP data

	RP data	SP data
Preference Information	• Reaction to actual situations • Align to the real behaviour • We can get "Choice" result	• Responding to hypothetical situations • Possibility of inconsistency with the behaviour in the real market

		• We can get "Ranking", "Rating", "Choice", etc.
Alternatives	• Real	• Real and hypothetical
Attributes	• Levels limited • Measurement error • May exhibit collinearity	• Levels extensive • No measurement error • Collinearity can be controlled
Number of Response(s) per respondent	• One	• One or more

Based on the above, SP data was used to model station choice in the research. They were collected through a stated choice (SC) experiment designed using the D-efficiency method. The experiment constructed twelve hypothetic situations. Each included two train stations with four groups of attributes, namely, travel time to the departure station, parking facilities, level of crowding on trains, and other relevant factors. Each attribute included two or three attribute levels determined based on the relevant data from the train survey. Pilot surveys were undertaken to test and refine the questionnaires. The final questionnaires covered 12 scenarios overall, with each scenario made up of simple and clear figures, numbers and words. The process to design the SC experiment is discussed in the following sections.

Additionally, an eye tracking experiment was conducted, based on these questionnaires, to evaluate the station choice models. The setup and implementation of the eye tracking experiment are described later in this chapter.

4. 2 SC experimental design workflow

The process to design an SC experiment generally starts with the refinement of the problem you need to solve. Then, the alternatives, attributes and attribute levels for each alternative used in the experiment should be identified and refined. Next, the choice set is determined based on the design approaches. The last step is to produce questionnaires based on the choice set. The process is illustrated in Figure 4.1.

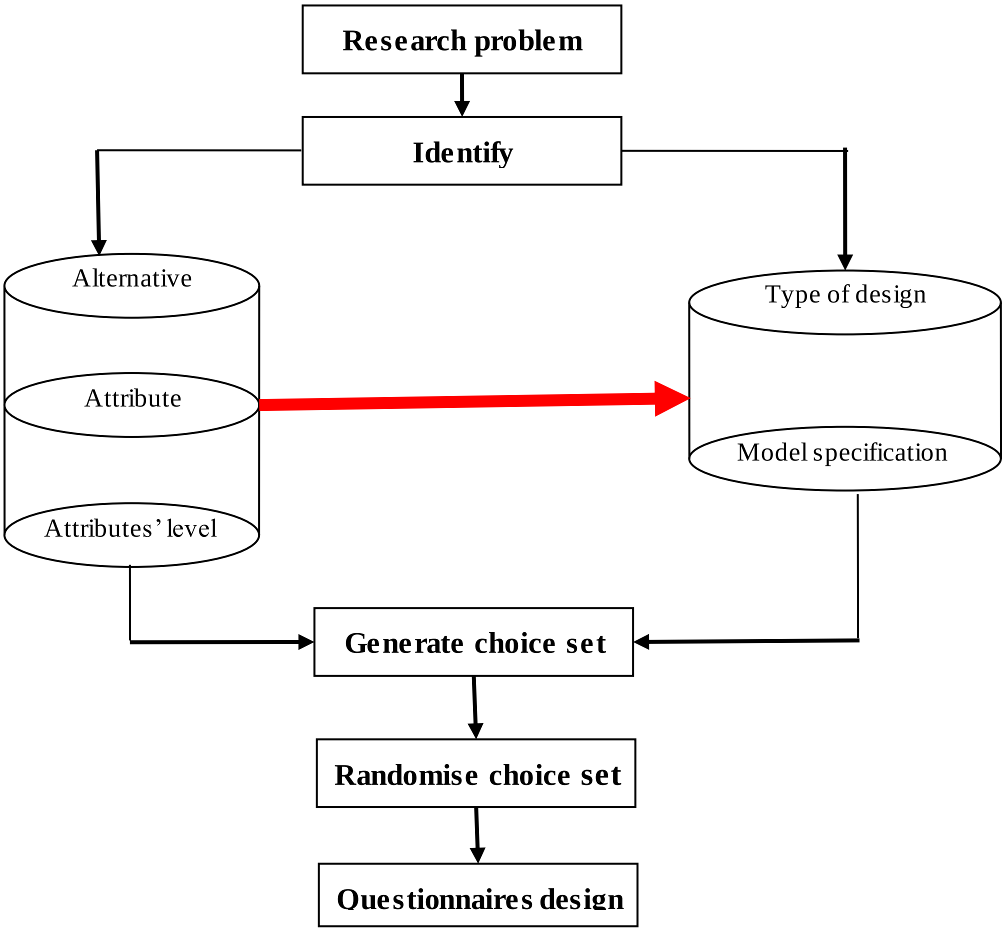


Figure 4.1 General process for an SC experimental design

4. 3 SC experimental design strategies

4.3.1 Considerations for the design of an SC experiment

In order to collect reliable SP data to develop station choice models specific to P&R users, this study took the following into consideration for the experimental design:

(a) Is the experiment labelled or unlabelled?

A labelled experiment refers to an experiment in which the alternatives' names have substantive meaning to respondents, such as private vehicle, public transport. An unlabelled experiment is the one in which the alternatives' names only convey their relative order of appearance, (e.g., route a, route b) or present their number and sequences, (e.g., station 1 and station 2). Generally, labelled experiments can be used to estimate not only generic parameters but also alternative specific parameters whereas unlabelled experiments only estimate generic parameters. This leads to differences in the minimum number of choice situations required and it is, therefore, very important to determine whether the experiment is labelled or unlabelled. Specific to the experiment in the research, the alternatives are train stations and are treated as unlabelled, i.e. station 1 or station 2

(b) Should the experiment keep attribute level balance?

Attribute level balance refers to ensuring that the different levels for each attribute are used the same number of times in the experiment, to minimise behavioural bias (Wittink, Krishnamurthi, & Nutter, 1982; Wittink, Krishnamurthi, & Reibstein, 1989). However, this property can lead to larger than necessary experimental designs and likely generate less efficient design (Hess & Rose, 2009). Hence, in the research, the SC experiment did not keep attribute level balanced.

(c) The number of attributes, attribute levels and the range for each attribute

The number of attributes and their levels will affect the minimum number of choice tasks in the experiment. The attributes in the research were identified based on previous literature and relevant surveys, (such as the station facilities' survey and the train station survey). Their range and levels were determined in two ways, the first from previous studies and the second using the Classify tools in SPSS. The process is detailed in Section 4.4.

(d) Design types.

Design types refer to the method used to design the SC experiment. The design type can determine the minimum number of choice tasks and which situation is chosen, based on the magnitude of the correlation among the attributes over the choice situations and the design efficiency. There is a number of available designs as described in Section 4.3.2 below.

4.3.2 Types of experimental design

The efficient choice (EC) method was chosen to design the experiment of station choice, even though many types of SC experimental design are presented in previous literature. The decision-making process to choose EC is described below.

Typically, SC design types can be divided into two classes, full factorial designs and fractional factorial designs.

(a) Full factorial designs

A full factorial design refers to an experiment that includes all possible combinations of each attribute level over all alternatives. Its advantages are that all effects, including main effects and interaction effects, can be estimated within the design and all attribute effects of interest can be guaranteed to be truly independent. However, its disadvantage is the size of the candidate set. Mathematically, the number of choice tasks in a full factorial design is determined by the number of attributes and the number of levels of each attribute. It can be calculated by the following specification (Rose & Bliemer, 2009):

$$CN = \prod_{k=1}^{K} L_k \tag{4-1}$$

where, CN is the number of choice tasks and L_k is the number of attribute levels for the k^{th} attribute.

As such, the design is only used when the number of attributes or attribute levels is relative small. Therefore, it is not suitable for a study with many attributes, and/or many attribute levels, to consider.

(b) Fractional factorial designs

Another SC experimental design is the fractional factorial design, which is made up of a subset of choice situations from the full factorial design. Obviously, this reduces the design size, but sacrifices the amount of information considered in the design. In order to address this problem, a number of ways to select the scenarios in a systemic manner have been proposed. Two popular approaches are described below.

(i) Orthogonal fractional factorial designs

The orthogonal design, the most famous fractional factorial design, is defined by Bliemer (2016) in his lecture as "a design in which every pair of levels occurs equally often across all pairs of attributes (or when the frequencies for level pairs are proportional instead of equal)" One of the competing orthogonal designs is Optimal Orthogonal Choice (OOC), which maximises the difference among the attribute levels across alternatives so that the maximum amount of information can be obtained from the respondents, by trading off all the attributes presented on the questionnaires (Street, Burgess, & Louviere, 2005). Generally, this design will be used when analysts cannot determine priors and the experiment has been treated as unlabelled. However, the OOC designs have some disadvantages. Firstly, the design is not a complete orthogonal design in that the orthogonal property cannot exist across alternatives, but there can be negative correlations across alternatives, and so it is usually applied to unlabelled choice experiments. Secondly, the number of attributes across the alternatives and the number of attribute levels for each attribute have to be the same. The OOC design was not used to design the SP experiment in the research even though all parameters in our experiment are likely to be treated as generic. The reasons were that different attribute levels were identified for different attributes and non-linear discrete choice models were applied to understand station choice behaviour, the desirable properties of which could be violated by the orthogonal fractional factorial design.

Based on this, the alternative fractional factorial design-Efficiency Choice (EC) design - is proposed, which is seen to easily outperform orthogonal designs (Rose & Bliemer, 2009).

(ii) Efficient choice (EC) design

In contrast to the OOC designs attempting to maximise the differences between attribute levels across alternatives, based on Rose and Bliemer (2005), efficient choice(EC) design refers to a design that aims to minimise the asymptotic standard deviations, (i.e. the square roots of the leading diagonal of the Asymptotic Variance-Covariance (AVC) matrix of a discrete choice model), so as to estimate more reliable parameters. The EC design is better than other methods as: ① prior values for the likely parameter estimates can be used as starting point values ② it can be used to estimate both generic and alternative specific parameters, and ③ it eases the requirements for attribute level balance across all attributes. Therefore, some researchers currently assert that the parameters estimates with the data from EC designs are more reliable than from orthogonal designs (given the availability of previous knowledge).

There are many types of EC designs, which one is best depends upon its efficiency. Based on previous literature, the most commonly used measure for efficiency is D-error. Usually, the D-error of an EC design will be low if the asymptotic (co)variances of the parameter estimates are low, and high if these (co)variances are high. Hence, the design with a minimum D-error is also called the D-optimal experimental design (Bunch, Louviere, & Anderson, 1996). However, the number of designs can increase exponentially with different combinations of attribute levels and it can be difficult to find the minimum D-error. Hence, an experiment with a low D-error is often, and more correctly, referred to as a D-efficient design.

Thus, we used the D-efficient approach to design the station choice experiment in the research.

4. 4 SC experimental design

According to Rose and Bliemer (2005), creating a stated preference experiment of station choice for P&R users includes three main steps: ① developing a utility function with all parameters to be estimated; ② determining an experimental design type based on the specification; and ③ constructing questionnaires (on paper, internet, CAPUI, etc.) based on the choice tasks determined in the experimental design. The detailed process is described below.

4.4.1 Specify the utility function

To develop a specific utility function of station choice, we should determine ① the attributes entering the utility function; ② which parameters should be generic and which parameters should alternative specific; ③ the attributes' format, e.g., dummy, effect codes, etc.; ④ which effects (i.e. main effects, interaction effects, or both) should be estimated; ⑤ the coefficients of the parameters in the model; and ⑥ the form of the econometric model obtained using the experimental design data. These were considered as follows:

(a) Determination of attributes

The criteria used to determine the attributes entering the utility function were ① previous literature; ② the objectives of the research; and ③ train station surveys.

(i) Literature review

A number of factors including travel time, parking capacity, cost, access time, accessibility to the station, etc. have been identified as attributes affecting station choice in past studies related to railway station choice. Hence, these attributes were directly considered in the research.

(ii) The objective of the research

The research aims to model station choice under uncertainty for P&R users. To achieve this objective, three uncertain environments influencing station choice were identified based on P&R users' decision-making process for their trips. These included the environment driving to the station, the environment on-board the train, and the parking environment. The uncertain measurement for each environment was determined based on previous literature (Li et al., 2009; Zheng Li & Hensher, 2013). Thus, three uncertain factors, (i.e. the travel time to the departure train station, parking search time and crowding on trains), influencing P&R users' choice for departure station were identified.

(iii) Surveys

The train station survey ranking the significance of various factors on station choice showed that a number of factors, (e.g. safety, ticket, train frequency, etc.), that had not been taken into account in previous studies can play a role in affecting P&R users' choice for departure train station. Therefore, these should also be included.

Based on the above, all the factors used in modelling station choice under uncertainty were obtained and are summarised in Table 4.2. Their relationship looks similar to the structure of the Milky Way Galaxy. Station choice is at the core of the structure, the first layer contains the main four environments. Each environment is made up of different factors, which comprise the second layer. The factors in the second layer are also made of different attributes, which form the third layer. The factors in each layer can directly affect the station choice and also indirectly influence it via the factors in their higher layer. Their relationship can be in Figure 4.2.

Table 4.2 Attributes entering the utility function

<table>
<tr><th colspan="2">Previous literature</th><th rowspan="2">Investigation</th><th rowspan="2">This thesis</th></tr>
<tr><th>Factors</th><th>Authors</th></tr>
<tr><td>• Location of station
• Access time
• Frequency of service
• Generalised cost</td><td>Kastrenakes (1988)</td><td rowspan="3">• Parking availability
• Parking search time
• Variation of parking search time
• Travel time to the station
• Variation of travel time
• Seat availability on trains
• Distance to station from home
• Facilities at the station</td><td rowspan="3">• Travel time to the station
❖ Variation of travel time to the station
❖ The frequency the travel time varies
• Parking availability
• Parking search time
❖ The variation of parking search time
❖ The frequency that parking search time varies
• Crowding on trains
❖ Seats availability on trains
❖ The density of standees in a carriage
❖ Transfer waiting time
❖ Access time</td></tr>
<tr><td>• Access mode to station
• Distance to station from home
• Additional facilities at the stations</td><td>(Fan et al., 1993)</td></tr>
<tr><td>• Travel time
• Distance
• Access time</td><td>Davidson and Yang (1999), Lythgoe et al. (2004), Lythgoe and Wardman (2004),Wardman (1997) and (Horner & Groves, 2007)</td></tr>
</table>

• Accessibility to the station • Rail service quality	Debrezion et al. (2009)		❖ Capacity of car parks • Safety • Ticket fare • Train frequency

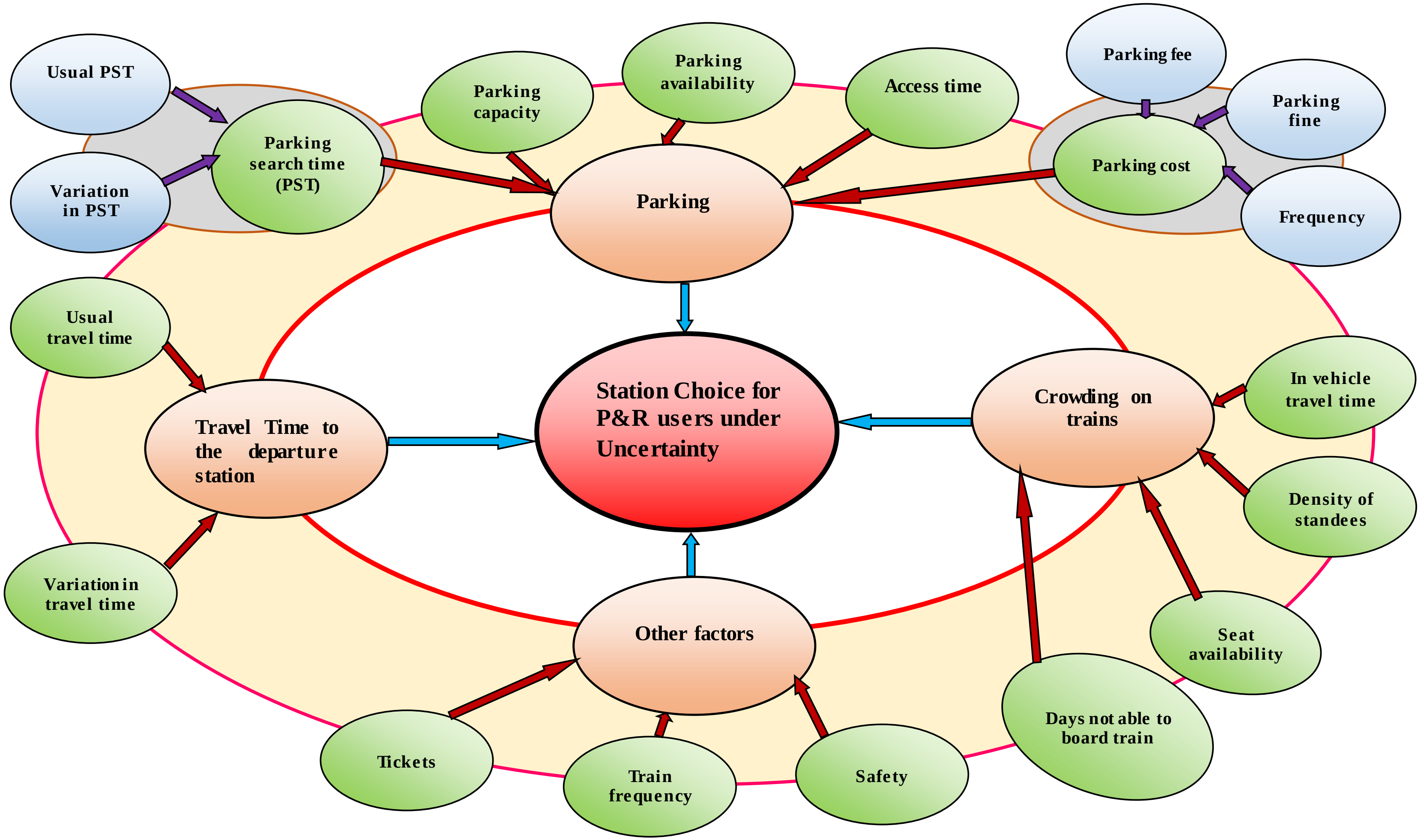


Figure 4.2 The relationship of factors influencing station choice for P&R users

(b) Attribute levels

Attribute levels in the experiment were determined in three ways. The first was from previous literature. For example, the levels for density of passengers standing in a train car and the probability that seats have been taken were determined based on the paper by Zheng Li and Hensher (2013). The second was from the field surveys. For example, the levels for safety, tickets, etc. were from the train station survey conducted on July 2, 2012. The last way, from a combination of the clustering approach and field survey, provided levels for factors such as travel time, parking search time, etc.

Clustering, conceptually, is to divide data into different groups, (clusters), based on their traits, the data in the same group being more alike than the data in other groups. Traditionally, clustering approaches can be divided into two types, relocation and hierarchical. The former refers to initially allocating data to different groups on a somewhat random basis then moving each record, on by one, to the group it is most alike. This process continues iteratively until no record can be moved to a better matching group. (Willett, 1984). The latter is a nested approach where the data are initially grouped into a small number of clusters (nests). Each nest is then divided into a number of sub-nests and so on, until the desired number of clusters is obtained. (Murtagh & Contreras, 2011). Both approaches can be efficient and accurate on small datasets but less so on large data sets. However, they can still be efficient if the datasets can be initially separated into smaller datasets, such as in a two-stage clustering method (Silva et al., 2017; Zhang, Ramakrishaman, & Livny, 1996). The first stage divides the records into initial clusters using a sequential clustering approach, in which the main component is the construction of a modified cluster feature (CF) tree. A CF tree consists of nodes with each node containing at least one branch with a leaf (record) (see Figure 4.3). The second stage is to divide the CF tree into the desired number of clusters.

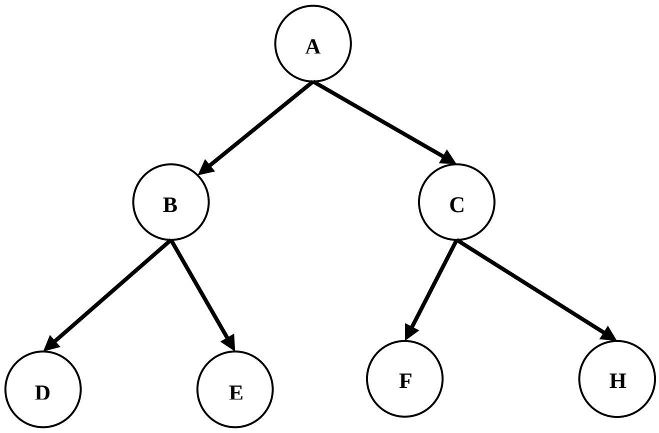


Figure 4.3 Example of a clustering feather tree

The SPSS software has developed a tool, Classify, which has a sub-tool, TwoSteps. We used this sub-tool to identify attribute levels for some factors, such as regular, best and worst travel times, and regular, slowest and fastest parking search times.

(i) Identifying attribute levels with TwoSteps tool in SPSS

Taking parking search time (PST) as an example, the first step is to construct a CF tree (see Figure 4.4) based on the data from the intercept survey carried out at Oats St, Cannington, Greenwood and Edgewater stations, (the questionnaire can be seen in Appendix F).

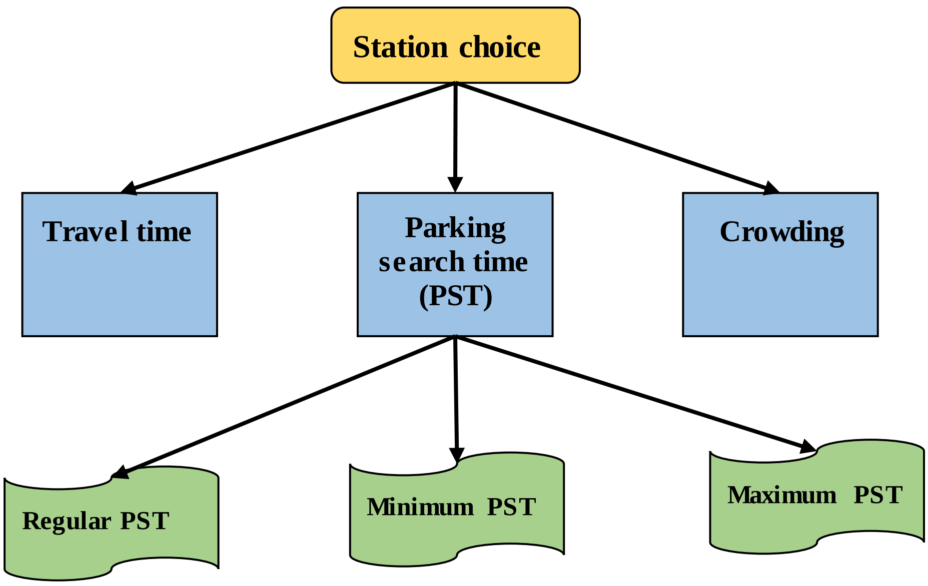


Figure 4.4 CF tree for identifying parking search time levels

Then the TwoStep tool was applied. Based on the clustering profiles, regular PST has three levels, namely, 1.5mins, 4.5mins and 10.0mins. The distribution of PST is also shown as follows.

Table 4.3 Clustering profiles

	PST min		PST max		PST regular	
	mean	SD	mean	SD	mean	SD
Cluster 1	2.5	.707	10.00	.000	4.50	2.121
Cluster 2	0.0	.000	5.00	.000	1.50	.707
Cluster 3	3.5	2.121	22.50	10.607	10.00	.000
combined	2.00	1.897	12.50	9.354	5.33	3.983

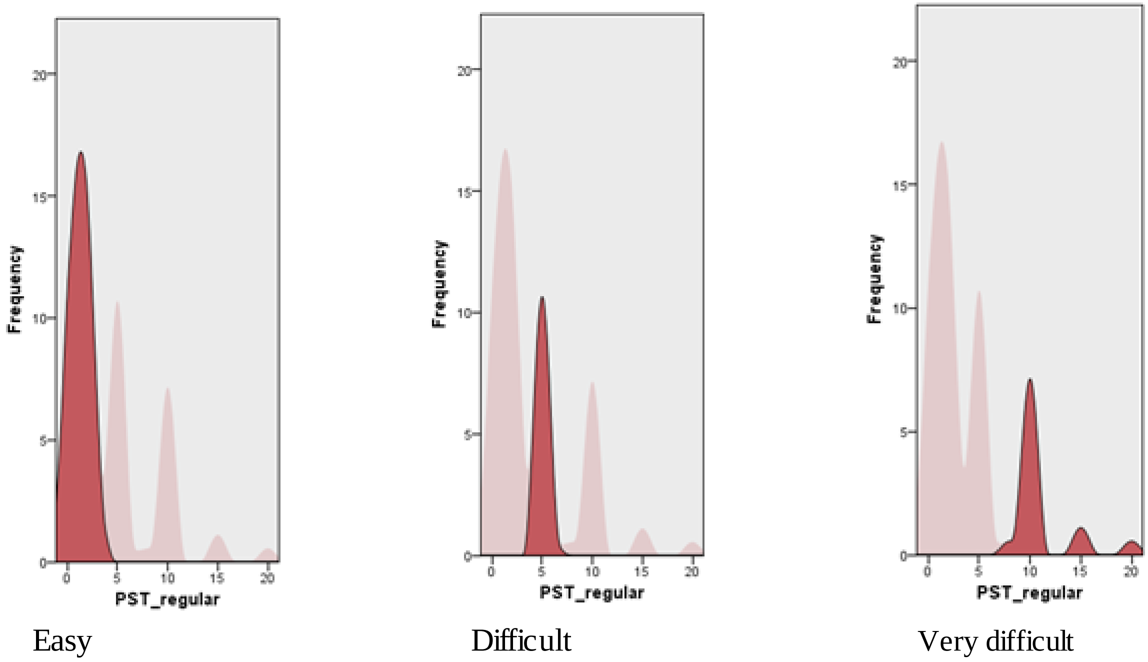


Figure 4.5 Distribution of regular parking search time

To test the results, the cumulative frequency of regular parking search time was calculated as well. The 25%, 50% and 75% thresholds correspond to PSTs of 0mins, 5mis and 10mins respectively. The histogram is shown in Figure 4.6

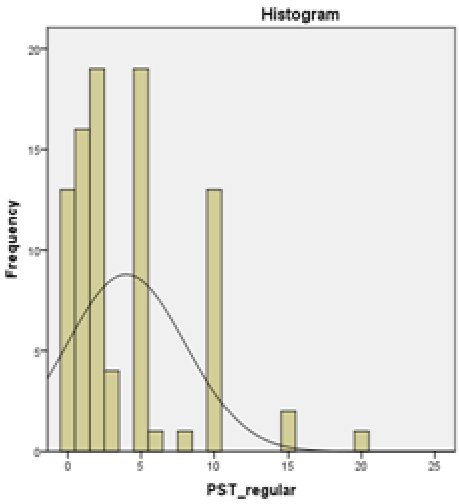


Figure 4.6 Histogram of regular parking search time frequency

Based on the above results and the requirement for even intervals, three levels were determined for the regular parking search time, namely, 1mins, 5mins and 10mins.

(ii) Identifying attribute levels from previous literature

The attributes whose levels were determined based on previous literature mainly included seat availability, density of standees in a passenger car and parking cost. The former two were determined based on the paper by Tirachini, Hensher, and Rose (2013) and Zheng Li and Hensher (2013). The last attribute was determined from Wang et al. (2004).

(iii) Identifying attribute levels by field surveys

The attributes whose levels were identified by field survey included safety, ticket fare, average journal time, train frequency, frequency of controlling illegal parking, parking capacity and parking availability. Their levels were determined based on their respective traits. For example, according to the train station survey, we found that safety at a train station is directly related to the frequency that security guards patrol the station. Therefore, we took the as a two-level dummy variable to measure perceived safety, i.e. safe or unsafe. The former refers to the situation where security guards patrol regularly and the latter to the situation where security guards are rarely, if ever, seen.

Another example is parking availability which was identified as having three levels. We used the number of parking spaces left when a P&R user arrived at a station as an index of parking availability in the research. We investigated P&R capacity left from 7:00am on 12 April, 2012 for all stations on the Armadale and Midland lines. The results indicated that, on average, about 10%-20% of parking spaces were still available between 7:00am to 8:00am. Moreover, almost all P&R facilities were full after 8:00am. Due to the duration we studied, between 7:00am to 8:30am, parking availability in this experiment was identified as having three levels, namely, 0% 10% and 20%.

All attribute levels are summarised in Table 4.4.

Table 4.4 Attributes and attribute levels

Number	Attribute	Levels	Description
1	Safety	2	• Regular security patrol • Irregular, or no, security patrol
2	Ticket fare	3	• \$2.5 • \$3.5 • \$4.5
3	Train frequency	2	• 5mins/train • 10mins/train
4	Regular travel time to the station from origin (home)	3	• 5mins • 10mins • 15mins
5	Increase in travel time to the station from origin compared to regular travel time	2	• 25% • 50%
6	Decreased in travel time to the station from origin compared to regular travel time	2	• 15% • 24%
7	Probability that the average slowest travel time occurs	2	• 5% • 20%
8	Probability that the average fastest travel time occurs	2	• 5% • 20%
9	Parking availability	2	• Unavailable • Available
10	Regular parking search time (PST)	3	• 1mins • 5mins • 10mins
11	Increase in PST compared to regular PST	3	• 100% • 400% • 800%
12	Probability that the longest PST occurs	2	• 5% • 20%
13	Capacity of park and ride car park	3	• 100 • 500 • 1000
14	Arrival time at the station	2	• 7:00am • 7:30am • 8:00am

15	The density of passengers standing inside (or in) the train	3	• 2passengers/m^2 • 4passengers/m^2 • 6passengers/m^2
16	The probability that seats have been taken	3	• 50% • 75% • 100%
17	In-vehicle travel time	3	• 10mins • 20mins • 30mins
18	The probability too crowed to board	3	• 0 • 20% • 40%

(c) Generic parameters

The parameters in the experiment can be considered to be generic if two criteria are met, i.e. if the experiment is unlabelled (Rose & Bliemer, 2009) and if all the attributes in the experiment are weighted equally in the utility functions over the different alternatives (Debrezion et al., 2009; Hensher, 2006). Specific to the station choice experimental design, only two unlabelled train stations were presented to respondents and we assumed all attributes could be weighted equally over different alternatives. Therefore, it is reasonable that the parameters for all attributes in the experiment, except the constant specific for station one, were considered to be generic.
Additionally, almost all attributes, except safety, in the utility specification were coded based on the value for each level; only safety was coded as a dummy variable, i.e. 1 for a regular patrol and 0 otherwise. Moreover, all attributes were assumed to be independent and only the main effects were considered.

(d) Choice tasks

The choice sets should be small enough for respondents to be able to choose without being overloaded. According to Hensher (2004) and Hensher (2006), the number of choice tasks should be determined jointly by degrees of freedom of the utility function, the number of alternatives for each choice task and the least common multiple (LCM) of the attribute levels. Their relationship is shown in equation (4-2).

$$\begin{aligned} &S \times (J-1) > K \\ &S = n \times LCM(AL_1, AL_2, \cdots AL_m) \qquad n \geq 1 \end{aligned} \tag{4-2}$$

where S is the number of independent choice tasks; J is the number of alternatives for each choice task; K is the number of parameters estimated; and AL_i is the number of attribute levels for i^{th} attribute.

(e) Model type

A multinomial logit model (MNL) was used in the thesis to understand the station choice behaviour under uncertainty because of its simplicity. More importantly, the data from the experiment designed by MML can be used to estimate the different utility specifications in the research (Louviere, Hensher, & Swait, 2000; Yang, Chen, Chen, Luo, & Ran, 2014).

(f) Utility specification estimated for the potential final model

Assuming an MNL model formulation, the utility specifications of station choice based on the considerations mentioned above can be given by equation (4-3).

$$\begin{aligned} U(st)_1 &= \sum_i \beta_i u_{j1} = \beta_0 + \beta_{TT} u_{TT1} + \beta_{PST} u_{PST1} + \beta_{Cr} u_{Cr1} + \beta_{sa} u_{sa1} + \beta_C u_{C1} + \beta_{tf} u_{tf1} \\ U(st)_2 &= \sum_i \beta_i u_{j2} = \beta_{TT} u_{TT2} + \beta_{PST} u_{PST2} + \beta_{Cr} u_{Cr2} + \beta_{sa} u_{sa2} + \beta_C u_{C2} + \beta_{tf} u_{tf2} \end{aligned} \tag{4-3}$$

where

$U(st)$ is the overall utility of the i^{th} train station;

u_{sa}, u_C, u_{tf} are the utility from safety, ticket cost, and train frequency at the i^{th} station; and

$\beta_0, \beta_{TT}, \beta_{PST}, \beta_{Cr}, \beta_{sa}, \beta_C, \beta_{tf}$ are estimated parameter coefficients.

The other parameters are complex and are explained below:

- u_{TT}

u_{TT} is the utility for travel time to the departure train station, which was calculated within CPT. According to Camerer and Ho (1994), the Power(Pwr) value function, together with a Tversky-Kahneman (TK) risk weighting function, can be the best fitted within CPT. Therefore, they were applied to develop the utility function of travel time. The utility function of travel time can be written as equation (4-4).

$$u_{TT,i} = \sum w(p_{TT,i}) v\left(\Delta_{TT,i}\right) \tag{4-4}$$

Where $\Delta_{TT,i}$ is the difference between regular travel time and bad (or good) travel time at station i; $v\left(\Delta_{TT,i}\right)$ is the value function of the i^{th} difference in travel time (its specification is given in equation (4-4)); $w(p_{TT,i})$ is the weighting function of the probability at which the i^{th} difference in travel time occurred in one month. It is given in equations (4-5) - (4-7).

$$v\left(\Delta_{TT,i}\right) = \begin{cases} \left(RTT_i - FTT_i\right)^{\alpha} \\ -\lambda\left(STT_i - RTT_i\right)^{\beta} \end{cases} \tag{4-5}$$

$$w(p_{fttf,i}) = \frac{p^{\gamma}{}_{fttf,i}}{\left(p^{\gamma}{}_{fttf,i} + \left(1-p_{fttf,i}\right)^{\gamma}\right)^{\frac{1}{\gamma}}} \quad (4\text{-}6)$$

$$w(p_{sttf,i}) = \frac{p^{\delta}{}_{sttf,i}}{\left(p^{\delta}{}_{sttf,i} + \left(1-p_{sttf,i}\right)^{\delta}\right)^{\frac{1}{\delta}}} \quad (4\text{-}7)$$

where

RTT_i is the regular travel time to station i from the origin for a P&R user;

FTT_i, STT_i are the fastest travel time on a good day or the slowest travel time on a bad day at station i;

$p_{fttf,i}, p_{sttf,i}$ are the probabilities of the fastest travel time (good day) or the slowest travel time, (bad day) occurring at station i; and

$\alpha, \beta, \lambda, \gamma, \delta$ are estimated parameters, with the former two indicating the respondent's risk attitude towards gains or losses respectively.

- u_{PST}

u_{PST} refers to the utility of the parking search time. It can be estimated under two situations, i.e. parking available and unavailable at the station (Baum & Epstein, 1978). For the first situation, only the capacity of P&R car parks and the access time are assumed to contribute to the utility. Moreover, the effects of their interaction on the utility can be assumed to be exponential, so that the higher the capacity and the earlier that a P&R user arrives the station, the greater the utility. Based on these, the utility function of parking search time under parking available can be written as (4-8):

$$u_{PST,i} = \exp\left(\max\left(PC_i - 100, 0\right)/10000 \times T_{a,i}\right) \quad (4\text{-}8)$$

where

PC_i is the capacity of the P&R car park at train station *i*; and

$T_{a,i}$ is the access time of P&R users arriving at station *i*.

When the demand for parking is greater than the capacity of the station car park, P&R users need to find an alternative parking space, e.g. on-street near the station or in a nearby shopping centre car park. The PST associated with this search could vary significantly depending upon the availability and locations of any alternative parking. Hence, the effects of PST on station choice when the P&R car park is full are a function of the variability of PST. Similar to travel time, CPT is also applied to evaluate the effects of variations of PST on station choice. The value function of parking search time was also developed based on the two-part power functions and TK weighting function. The gain for PST was the difference between the shortest PST and regular PST and the loss was defined as the difference between the longest PST and the regular PST. The gain part in the research was ignored because it was generally found to be too small to affect P&R users' station choice. The value function and weighting function of PST are given by equations (4-9) and (4-10) respectively.

$$\begin{aligned} v(\Delta_{PST,i}) &= -\lambda \times \left(-\left(PST_{R,i} - PST_{L,i}\right)\right)^{\beta} = -\lambda \times \left(-\left(PST_{R,i} - PST_{R,i} \times \left(1 + Var_{L,i}\right)\right)\right)^{\beta} \\ &= -\lambda \times \left(PST_{R,i} \times Var_{L,i}\right)^{\beta} \end{aligned} \tag{4-9}$$

$$\pi_i = p_{Lf,i}^{\theta} \Big/ \left[p_{Lf,i}^{\theta} + \left(1 - p_{Lf,i}\right)^{\theta} \right]^{\frac{1}{\theta}} \tag{4-10}$$

where

$PST_{R,i}$ is the regular PST at train station i;

$PST_{L,i}$ is the longest PST at train station i;

$Var_{L,i}$ is the percentage increase in PST (over the regular PST);

$p_{Lf,i}$ is the probability that the longest PST for station i occurs; and

λ, β, θ are the estimated parameters. Some researchers, such as Avineri (2004), Tversky and Kahneman (1992), etc., have proved that $\lambda = 2.25, \beta = 0.88, \delta = 0.69$ can best fit the data, so these values were taken as the priors of these parameters in the experimental design.

Integrating equation (4-8), (4-9) and (4-10), the utility function of PST is given by (4-11):

$$u_{PSTi} = \begin{cases} \exp\left(\max\left(C_{pi} - 100, 0\right) / 10000 \times T_{ai}\right) & if\ \ parking\ availablity = 1 \\ -\lambda (PST_{R,i} \times Var_{L,i})^{\theta} \times P_{Lf,i}{}^{\theta} \big/ [P_{Lf,i}{}^{\theta} + (1 - P_{Lf,i})^{\delta}]^{(1/\theta)} & if\ \ parking\ availablity = 0 \end{cases} \tag{4-11}$$

- u_{Cr}

u_{Cr} refers to the utility of crowding on trains.

We started with the following assumptions:

- When the percentage of seats occupied is less than or equal to 50%, there is no crowding on the train;
- When the percentage of seats occupied is greater than 50% and less than 100%, there may be some passengers who stand voluntarily; and
- When the percentage of seats occupied is equal to 100%, some passengers will be forced to stand.

Then, the utility function of crowding on trains is given by a three-part power function. The specification is shown in equation (4-12):

$$u_{Cr,i} = \begin{cases} 0 & if \quad P_{se} \le 0.5 \\ (p_{se,i} \times IVT_i)^{1-r} \big/ (1-r) & if \quad 0.5 < P_{se} < 1 \\ \left(D_{st,i} \times IVT_i + T_{wi}\right)^{(1-r)} \big/ (1-r) & if \quad P_{se} = 1 \end{cases} \tag{4-12}$$

where

$P_{se,i}$ is the probability of seats occupied at station i;

$D_{st,i}$ is the density of passengers standing in one carriage at station i;

IVT_i is the in-vehicle travel time from station i to the destination station

r indicates risk attitude of P&R users, $r = \begin{cases} 2 & IVT > 20 \min s \\ 0.5 & IVT \le 20 \min s \end{cases}$ and;

T_{wi} is the transfer waiting time at the station i , here $T_w = \begin{cases} 1 & D_{st} < 6 \\ T_w & D_{st} \ge 6 \end{cases}$.

By integrating equations (4-5), (4-6), (4-7), (4-11) and (4-12) into equation (4-3), the final utility specification was formed.

4.4.2 Generate the experimental design

After determining the form of the utility function, we commenced the generation of the experimental design. An experimental design generally aims to identify which hypothetical choice tasks should be presented to respondents in a stated choice experiment. Usually, an experimental design consists of a table, (matrix), with M rows and N columns, in which each row is a choice task. The process to develop the matrix with D-Efficiency choice design is as follows.

- **Step 1**, All the attributes with orthogonal coding are listed;

- **Step 2**, A matrix with orthogonal coding is randomly created. The same attribute in different alternatives must adopt different levels.

- **Step 3**, The coding matrix is converted using the actual attribute level values.

- **Step 4**, An initial matrix (X) is created and the probabilities that each alternative is chosen in the design are calculated. The parameter priors are determined based on previous literature.

- **Step 5**, An Asymptotic Variance Covariance (AVC) matrix is constructed. The value in the AVC matrix with generic parameters is calculated based on equation (4-13), as proposed by Rose and Bliemer (2005).

$$\frac{\partial^2 L\left(\beta^*,\beta\right)}{\partial\beta^*_{k1}\partial\beta^*_{K2}} = -\sum_{s=1}^{S}\sum_{j=1}^{J} x^*_{jk_1s}\, p_{js}\left(x^*_{jk_2s} - \sum_{i=1}^{J} P_{is}x^*_{ik_2s}\right), \quad \forall k_1,k_2 = 1,\cdots,K^* \qquad (4\text{-}13)$$

- **Step 6,** A Fisher information matrix is established. According to Rose and Bliemer (2005), the AVC matrix is the negative inverse of the expected Fisher information matrix. Thus, the Fisher information matrix is given by equation (4-14) (Rose & Bliemer, 2005):

$$\Omega\left(X,Y,\tilde{\beta}\right) = -\left[E\left(I_N\left(X,Y,\beta\right)\right)\right]^{-1} = -\left[\frac{\partial^2 L_N\left(X,Y,\tilde{\beta}\right)}{\partial\beta\partial\beta'}\right] \qquad (4\text{-}14)$$

where $I_N\left(X,Y,\beta\right)$ is the Fisher information matrix; N is the number of respondents; and $I_N\left(X,Y,\beta\right)$ is the log-likelihood function as shown in equation (4-15) (Rose & Bliemer, 2005).

$$I_N\left(X,Y,\beta\right) = \sum_{n=1}^{N}\sum_{s=1}^{S}\sum_{j=1}^{J} y_{jsn}\log p_{jsn}(X,\tilde{\beta}) \qquad (4\text{-}15)$$

- **Step 7**, The statistical efficiency of the design is evaluated using D-error. According to Rose and Bliemer (2005), it is equal to the determinant of the AVC matrix Ω. Based on the information on the priors, three kinds of D-error (i.e., D_z, D_p, D_b) can be produced. If there is no information on the priors, the D_z is preferred. If specific fixed and non-zero priors are used, D_p is preferred and if the designs are optimised with the priors with Bayesian distribution, D_b is preferred. Given that the priors in the station choice experimental design were determined based on previous literature, D_p is used to measure the efficiency. It can be calculated using equation (4-16) (Rose & Bliemer, 2005).

$$D_p - error = \det\left(\Omega\left(X, \tilde{\beta}\right)\right)^{\frac{1}{H}} \quad (4\text{-}16)$$

- **Step 8**, Optimising the design. Repeat step 2 to step 7 in order to get the minimum D_p-error. Evolver software, based on Genetic Algorithms (GA), is applied in this step. All the attributes in the first alternative, (except for the percentage increase in PTS compared to the regular PTS, the density of standees in a carriage and the transfer waiting time), were set to be adjustable. Moreover, the design should be as balanced as possible.

Following the above procedure, the efficiency of the experimental design for station choice was 0.000620159.

4.4.3 Construction of the questionnaires

(a) Construction

Based on the experimental design above, the individual questionnaires were constructed. In practice, each individual questionnaire was produced by selecting a subset of relevant values from the full range of values identified in the experimental design stage. In other words, we converted the table of numbers into words that respondents would be able to understand. Each row in the table was translated into a choice task and recorded, with each column indicating a different attribute level for each different choice task. For example, to obtain a respondent's choice between two hypothetical train stations based on crowding on trains, the experimental design could be per Table 4.5 and the questionnaire per Table 4.6.

Table 4.5 Example of the experimental design

	Station 1			Station 2		
	Probability that seats have been taken	Density of standees	Number of days per week on which trains are too crowded to board	Probability that all seats have been taken	Density of standees (passengers/m^2)	Number of days per week on which trains are too crowded to board
Case 1	100%	4	1	100%	4	2
Case 2	100%	4	1	75%	2	0
Case 3	100%	2	0	75%	2	0

Table 4.6 Sample questionnaire

Which of station would you choose?

Case 1

Factors	Scenario 1	Scenario 2
Crowding	 All seats taken and half of aisles filled	 All seats taken and half of aisles filled
Number of days per week on which the trains are too crowded to board	1	2
Tick the box for the scenario you would choose	☐	☐

Case 2

Factors	Scenario 1	Scenario 2
Crowding	 All seats taken and half of aisles filled	 Three quarters of seats taken
Number of days per week on which the trains are too crowded to board	1	0
Tick the box for the scenario you would choose	☐	☐

Case 3

Factors	Scenario 1	Scenario 2
Crowding	 All seats taken	 Three quarters of seats taken
Number of days per week on which the trains are too crowded to board	0	0
Tick the box for the scenario you would choose	☐	☐

The questionnaires used in the station choice experiment were tables made up of values (numbers), words, and pictures. In-vehicle travel time, transfer waiting time, ticket fare, train frequency and parking capacity were presented to the P&R users with values, and safety was explained only in words. The level of crowding on trains is more complicated. According to Li et al. (2012) and Hensher, Rose, and Collins (2011) visualization of the situation within the carriage, together with a description of the proportion of seats taken and the density of passengers standing in a train car, can provide a rich definition of the situation faced by passengers about to board. Therefore, the level of crowding situation in the experiment is described through pictures, numbers and words, (see Figure 4.7). The pictures were produced using Google SketchUp and showed the passengers density, (including the probability of all seats being occupied and the density of passengers standing in one carriage). In this experiment, each P&R user is presented with two pictorial representations, (see Figure 4.7), with different passenger densities.

All seats taken and half of aisles filled	Three quarters of seats taken

Figure 4.7 Example of pictorial representation of crowding

Initially, we also trialled using pictures to describe the parking availability at the station. Testing, however, indicated that a simple description in words was easier to understand. Three groups of attributes were used to show the difficulty of searching for and finding a parking bay. The following table shows an example of the questionnaire related to parking attributes.

Table 4.7 Example of questionnaire related to parking attributes

P&R Parking capacity	1000 bays	500 bays
Parking availability within P&R car park at 7:30 am	30%	30%
Parking search time (time spent searching for parking before giving up such as trying another)	Between 1 and 2 mins and only 5% chance to reach to 2 mins	Between 1 and 2 mins and only 20% chance to reach to 2 mins

Travel time to the station from the origin is described in words and numbers to allow more information to be included, such as the regular (usual), minimum and maximum travel times, the frequencies at which these occur, and the degree of variation from the regular travel time. An example of the travel time question format is provided in Table 4.8.

Table 4.8 Example of the travel time question format

Usual travel time	15mins	10mins
Travel time on a good day	About **7.5** mins on **4** days a month	About **7.5** mins on **4** days a month
Travel time on a bad day	About **18.6** mins **once** a month	About **12.4** mins on **4** days a month

The complete questionnaire is shown in Appendix D.

(b) Questionnaire testing

Testing of the questionnaires comprised two components. The first was to test their readability, understandability and reliability, and the second was to test their validity, i.e. to test whether the data collected by the questionnaires could correctly show respondents' preferences.

We conducted several pilot surveys at the Curtin University campus and at train stations to ensure that the questionnaires could be easily understood and completed by respondents within an acceptable, (short), time period.

We firstly aimed to test the questionnaires' readability and validity. We started with a sample questionnaire using words, numbers and figures, (see Appendix E-1), and tested it at Curtin University in July 2013. Given that the questionnaire was complex and technical, many respondents couldn't understand it correctly and it took up to 5 minutes to complete. To simply the questionnaire, we changed replaced some words with graphs or tables, (see Appendices E-2 to E-4), but the results were not much better. Therefore, we, then, designed a table with three columns and many rows, in which succinct words and clear pictures were used. The columns showed two stations and each row presented one attribute. The last row presented two choices for the respondent. An example of this questionnaire can be seen in Appendix E-5. Even though the questionnaire had been improved greatly, there were still some respondents who thought that the pictures were too technical. Therefore, we replaced some of the figures with words that were more easily understood, (see Appendix E-6) and conducted a pilot survey with it at Curtin University and at some trains stations. The results were satisfactory with respect to readability and time taken but the choice results did not align with our expectations. For example, respondents did not appear to prefer a station that was safer. Hence, we changed some of the pictures back to simple words again, with only the level of crowding being described through pictures. The revised questionnaire can be seen in Appendix E-7. We conducted a final pilot survey using these revised questionnaires and made further minor refinements to the wordings to make questions as clear as possible. This final questionnaire was used to conduct the main survey undertaken in November and December, 2014. The station choice model was developed with the data from the main survey.

4.4.4 Determination of sample size

There are several sampling methods to determine the minimum sample size requirements but almost none is appropriate for stated choice. This is because most current strategies to calculate minimum sample size requirements associated with SC experiments, are not concerned about how accurate and reliable the resulting parameter estimates are but instead concentrate on minimising the errors in the probabilities of the alternatives being chosen. However, the sampling theory proposed by Bliemer and Rose (2009) is completely different. They showed that the relationship between the AVC matrix and the square root of sample size was inversely correlated. Therefore, assuming the asymptotic standard deviations for the parameters estimates can be replaced by the square roots of the diagonal elements of the AVC matrix and the asymptotic t-ratios are equal to the estimations of the parameter coefficients divided by the asymptotic standard deviations, (see equation (4-17)), it is possible to derive the sample size based on the AVC matrix of the experimental design (Rose & Bliemer, 2009) .

$$t_{\beta_k} = \frac{\hat{\beta}_k}{\sqrt{\sigma^2_{\beta_k} / N_{\beta_k}}} \tag{4-17}$$

Rearranging equation (4-18),

$$N_{\beta_k} = \frac{t^2_{\beta_k} \sigma^2_{\beta_k}}{\hat{\beta}^2_k} \tag{4-18}$$

Assuming a two-sided confidence interval of 95% and given that the degree of freedom in this experiment is 5, the t-ratio is 1.96. The minimum sample size required for each parameter is listed in Table 4.9.

Table 4.9 Summary of sample size

	1	2	3	4	5	6
$\hat{\beta}_k$	0.3	0.08	-0.3	-0.2	0.2	0.3
σ_{β_k}	3.39	0.05	0.07	0.53	1.27	1.48
Sample size N_{β_k}	144	31	3	51	122	64
Minimum observations	1728	372	36	612	1464	768

Based on Table 4.9, the minimum sample size required for the station choice experiment should be 144. Another method to quickly calculate the sample size required, proposed by Orme (2005), was also used in the research. The sample size is determined based on experience, rules-of-thumb, and budget constraints. The equation is shown in (4-19) (Orme, 2005).

$$N = 500 \times \frac{l^*}{J \times S} \tag{4-19}$$

where

N is the sample size;

l^* is the largest number of attribute levels in the experimental design;

J is the alternative sets; and

S is the number of choice tasks determined in the design.

Substituting all of values in this experiment into equation 4-18, the sample size for the experiment is:

$$N = 500 \times \frac{l^*}{J \times S} = 500 \times \frac{4}{(2 \times 12)} = 84$$

Figure 4.8 shows the relationship between the asymptotic standard errors corresponding to the D-efficient design and the sample size. This means that, the larger the sample size, the smaller the asymptotic standard error and, correspondingly, the more efficient the experimental design. Therefore, 144 respondents is the minimum sample size for this experimental design with statistically efficiency, as each respondent can complete 12 choice situations. However, a respondent can only answer four choice situations in 3-5 minutes according to the results from the pilot surveys. Hence, the minimum sample size of respondents for the experiment was determined to be 432, (i.e. $144 \times 12 / 4$).

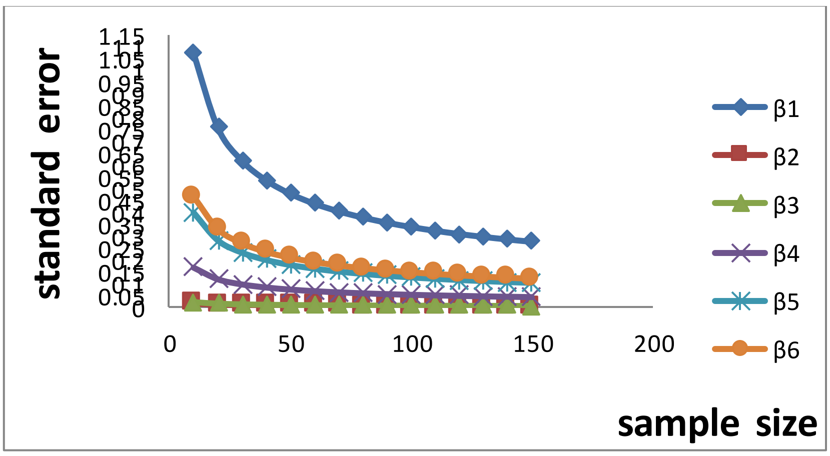


Figure 4.8 Standard errors when using the D-efficient design for different sample sizes

4.4.5 The bias in SP experiments

Stated preference (SP) experiments have been broadly used in many fields, such as marketing, transport, health, etc., for many years due to its ability to elicit behavioural responses and estimate individuals' preferences. However, they are still criticised due to the biases, especially hypothetical bias, existing in SP experiments.

(a) Sources of the bias in SP experiments

Previous literature has identified many biases in SP experiments. In this research, we summarised the biases occurring in the two phases of a SP experimental design: 1) designing choice scenarios and 2) testing these scenarios.

In the phase of designing choice scenarios, we can only include a limited number of scenarios in our questionnaires. The design of these scenarios may lead to biases, such as introducing implausible /unrealistic choice scenarios to the experimental design.

In the phase of testing these scenarios, we may encounter hypothetical bias, which has been studied in many papers (List & Gallet, 2001; Little & Berrens, 2004; Murphy, Allen, Stevens, & Weatherhead, 2005). It is defined as the inconsistent choice made by respondents in hypothetical versus real situations (Hensher, 2010). There is no evidence indicating that hypothetical bias exists in all SP experiments, due to the difficulty in, and cost of, testing for it (Fifer, Rose, & Greaves, 2014). Nevertheless, it is accepted that hypothetic bias is certainly an issue in many cases, e.g., hypothetical WTP typically exceeds the actual value by two or three times. Unfortunately, there is no widely accepted general theory to mitigate against this bias (Loomis, 2011).

(b) Solutions to mitigate the biases in a SP experiment

We spent half a year to design our SP experiment and adopted different methods to mitigate these biases.

Firstly, we adopted an iterative approach to mitigate the biases produced in the SP experimental design by applying a D-efficiency method (see section 4.4) and manually checking implausible /unrealistic choice scenarios. In this method, D-error is taken as the index indicating the magnitude of the bias in the experimental design. Smaller D-error indicates a better experimental design and fewer biases.

In order to address the hypothetic bias, we conducted a number of pilot studies using three different groups, including transport research group in the School of Earth and Planetary Sciences, Cafés and Labs at the Curtin University and respondents at train stations in the Perth Metropolitan area. To mitigate the biases in the testing phase, we asked respondents to explain the decision making process and propose suggestions for improving the questionnaires after completing the survey. Through these processes, we were able to identify any inconsistent choices made by respondents in hypothetical versus real situations. Based on these, we redesigned the questionnaires and repeated the process until we were confident with the questionnaires used in the main survey.

4. 5 Eye tracking experiment

4.5.1 Objective

Eye tracking is a process that follows and records the movement of the eyes, the point of gaze, i.e. where the eyes are looking, and the duration of that gaze. It is considered to be a good technique to analyse any kind of human behaviour and has been widely applied in many fields including psychology, medicine, marketing and engineering.
Eye tracking has been used in this research to test the validity of the questionnaires and the experimental design. It can explain the relationship between a human's visual attention to the factors in the questionnaires and their decisions. In other words, we used the level of consistency of the results from the station choice model with the results from eye tracking experiment to validate the model.

4.5.2 Hypothesis

The eye tracking experiment started with a number of hypotheses:

- The higher the frequency and the longer the time a respondent's eyes fixed on a factor, the more likely that factor is to influence the choice made;
- The greater the complexity of the questionnaire, the longer the duration of the fixation before a decision is made;
- Repeating the questionnaires will result in shorter average fixation times by the respondents.

4.5.3 Experimental setup

The experimental setup consisted of a 60Hz Remote Eye-Tracking Device (RED) and a laptop, as shown in Figure 4.9. The RED is a sophisticated and contactless electronic device with sensitive cameras mounted on the laptop's screen. It allows the head to move freely within the range of 40cm (side to side) x 20cm (up and down) at a distance of 70cm. It is widely used in many disciplines including for adolescents with Autism Spectrum Disorders (ASDs) (Horlin et al., 2013).

Figure 4.9 Eye tracking device setup

4.5.4 Questionnaires

The questionnaires used for the eye tracking experiment were designed based on those used in the station choice experiment. Firstly, we took part of the information from the original questionnaires, based on our assumptions, and modified these to produce 14 new questionnaires, (see Appendix F). Questionnaires 1-6 were used to assess the complexity of the questionnaire. Questionnaires 7-12 assessed how participants responded to the scenarios being repeated and questionnaires 13-14 were designed to test the impact of changing the order of the variables.

From on these new questionnaires, we identified and drew the areas of interest (AOI), i.e. sub-areas on the forms where the eyes were fixated, based on the eye movement measurements.

4. 6 Chapter summary

This chapter described two experiments, one for collecting the SP data used to develop the station choice models and the other the eye tracking experiment used to validate these models. The first experiment was designed using the D-efficiency method. The utility function was developed within CPT and is a multinomial logit model with eighteen attributes. Each attribute has two or three levels. Twelve choice situations were determined by the SOLVER tools in EXCEL based on the value of the D_p-error. From these we constructed questionnaires with simple words, numbers and pictures.

The second experiment was an eye tracking experiment. We designed questionnaires by drawing the areas of interest on the questionnaire designed in the last experiment and used a 60Hz Remote Eye-Tracking Device (RED) connected to a laptop to record the responses to these questionnaires.

CHAPTER 5 STATION CHOICE SUB-MODEL FOR MERGING TRAVEL TIME UNCERTAINTY

The previous chapter described how the data used for developing and validating the station choice models under uncertainty were collected. In the next three chapters, we will analyse the effects of three uncertain factors, (i.e. travel time, parking search time and crowding on trains), on a P&R user's choice of departure train station, respectively. This chapter focuses on modelling station choice under uncertainty of P&R users based on the effect of travel time and measuring P&R users' risk attitudes towards variations in travel time. From these, the influence of travel time variability on a P&R user's station choice will be quantified.

It is worth noting that most of the chapter is drawn from the paper written by Chen et al. (2017). This paper has been published and my contribution for the paper is about 80%.

5. 1 Research context

As traffic congestion increases and climate change receives more attention from the public, promoting sustainable mobility is becoming a key objective of transport policymaking (Rietveld, 2010). Park and Ride (P&R), as one of the potential solutions, is widely recognised as an efficient travel mode, combining the private car with public transport and leading to reduced energy use and air pollution and better social equity. P&R is now used extensively by commuters throughout the world (Cairns, 1998; Ginn, 2009). Currently, the capacity of the P&R facilities in Perth, Western Australia, is 23,000, which is well below the level required to cater for the demand (Martinovich, 2008). In addition, as predicted by the Commonwealth Government Bureau of Infrastructure, Transport and Regional Economics (BITRE) in 2007, the cost of congestion in Perth could reach $2.1 billion by 2020 (Bureau of Transport and Regional Economics [BTRE], 2007). This means not only increased car travel times generally but also substantial variability in travel conditions on the road network, adversely impacting on travel time reliability.

A survey conducted by University of Western Australia and Curtin University in July 2012, covering a subset of Perth's train stations, confirmed that travel time variability was not only a major motivation for P&R users to combine car with train travel, but also affected their choice of departure train station.

Many factors were identified in the literature affecting station choice, including location of station, access/arrival time, frequency of service, degree of overcrowding on the train, parking search time, generalised cost, travel time, access modes, accessibility and railway network services (Carrion & Levinson, 2012; Davidson & Yang, 1997; Debrezion et al., 2009; Fan et al., 1993; Hunt & Teply, 1993; Kastrenakes, 1988; Lythgoe & Wardman, 2004; Lythgoe et al., 2004; Shao et al., 2015; Wardman, 1997; Whelan & Johnson, 2004). Travel time variability is one of most common measures of the reliability of a railway service (Carrion & Levinson, 2012; Rietveld, 2010). It includes two components, predictable (i.e. congestion), and unpredictable (e.g., incidents and weather conditions), variability. The predictable part of travel time is expected, and can be anticipated by travellers, and, therefore, they can adjust their behaviour to avoid its consequences (Small & Verhoef, 2007). However, with irregular or non-recurring conditions this is not possible. The unpredictable variability of travel time was divided into three elements by Wong and Sussman (1973): ① unexpected seasonal or daily changes of travel time; ② variability caused by unpredictable events, such as weather or crashes; and ③ variations related to each traveller's perception. One of earliest studies, conducted by Gaver (1968), found that individuals usually adjusted their departure time to compensate for uncertainty about the time needed to complete a trip. Later, Guttman (1979) and Menaske and Guttman (1986) modelled the effect of travel time uncertainty on access mode and route choice, by incorporating travellers' risk attitudes. They found that travellers were risk averse, i.e. travellers would tend to choose the modes and routes with the more certain, namely, less variable, travel times. Travel time variability was found to introduce penalties in a variety of choice situations, including mode and station choice. In the context of P&R, P&R users limit the variability in car travel time on the road network by replacing a large part of their trip with a more reliable mode timewise, i.e. train, which has dedicated infrastructure and scheduled services. What is still affected by road network travel time variability is the travel by car from home to the station, i.e. access to the station.

With respect to P&R, total travel time consists of access travel time, (i.e. from origin (or home) to the station), parking search time, waiting time for the train, in-vehicle (train) travel time and egress time (i.e. from alighting station to final destination). Recently, Li et al. (2010) examined the willingness to pay for travel time reliability (variability) and found that travellers were risk averse towards travel time unreliability. In other words, they are willing to pay more to avoid unreliable travel. Parking search time and the degree of overcrowding on trains have been identified as uncertainty by Hunt and Teply (1993) and Whelan and Johnson (2004). The former developed a nested logit model of parking location choice based on the uncertainty of parking search time and the latter established the PRAISE (Privatisation of Rail Services) rail operations model based upon the uncertainty of crowding on trains. Rietveld (2010) analysed the impact of travel time unreliability on rail passengers' access mode and departure station choice. He found that travel time reliability played an important role in station choice and that high travel time unreliability was related to a low public transport share. Uncertainty enters the decision process for choosing a station by way of a combination of day to day travel variability in the transport network and the travellers' unfamiliarity with the network conditions (Circella et al., 2005). This chapter focuses on understanding travel time to stations and its uncertainty attributed to each traveller's perception.

There are numerous methods to estimate choice behaviour under uncertainty: discrete choice models (Bates, 1987; Ben-Akiva & Steven, 1985; Greene & Hensher, 2003; Small, 1987; Truong & Hensher, 1985), von Neumann-Morgenstern expected utility theory (Savage, 1972) and non-expected utility theory, such as weighted expected utility (Chew & MacCrimmon, 1979), rank dependent utility (Quiggin, 1993), and prospect theory (Kahneman & Tversky, 1979). Carrion and Levinson (2012) conducted a systematic review of the value of travel time reliability and recognised the role of travel time reliability in different choice situations, such as departure time, route and mode. However, there is limited research in relation to station choice with travel time uncertainty.

The aim of this chapter is to develop methods to estimate a train station choice model for P&R users, based on the effect of travel time variability, by which P&R users' risk attitude towards variability of travel time and its effect on the station choice can be explained.

5.2 Research method

The method is based on the theoretical framework of discrete choice models and cumulative prospect theory. The mean-variance approach is used to measure access travel time reliability, which is incorporated into logit models to estimate the probability of station choice of P&R users. The risk attitude towards travel time reliability is captured using cumulative prospect theory. Stated preference and revealed preference data are used to estimate the value of reliability. The developed method is implemented using a case study of three train stations in Perth, Western Australia.

5.2.1 Data used

Two types of data were used, revealed preference (RP) and stated preference (RP). The former were obtained from the train station survey conducted in July 2012 at seven train stations in Perth, WA. The latter were obtained from the station choice survey, as described in Chapter four. The specific RP data used were the travel times between the origins and the departure train stations, and the SP data used were the station choice data (see Figure 5.1). The SP data were used to develop the sub-model and the RP data mainly used to identify the relationship between an individual's experience and his/her risk attitude.

Example

Below are two train scenarios for a park-and-ride(P&R)commuter travelling from home to work at, say 7:00AM, on a **typical weekday**
All things considered, which scenario appeals to you most?

Factors	**Scenario 1**	**Scenario 2**
Travel time from home to the station		
Usual travel time	**10**mins	**5**mins
Travel time on a good day	About **5** mins on **4** days a month	About **3.75** mins on **4** days a month
Travel time on a bad day	About**11.5** mins **once** a month	About **6.2**mins **once** a month
Time on the train		
Average journey time	10 minutes	10 minutes
Finding parking at the station		
Parking capacity in P&R lot	100bays	100bays
Parking availability in a P&R lot at 7:00AM	20%	20%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **1** and **8** minutes and only **5%**chance to reach to **8** mins	Between **1** and **8** minutes and only **20%**chance to reach to **8** mins
Parking cost		
Parking fee in a P&R lot	$4/day	$6/day
Parking fine (if you park illegally, you would be fined)	$40	$80
Frequency of controls for illegal parking	Once a **day**	Once a **month**
Crowding on the train		
Crowding	 **All seats taken and half of aisles filled**	 **All seats taken and half of aisles filled**
Number of days per week on which the trains are too crowded to board	2	1
Other things		
Safety	Regular station and car park security patrols	Irregular station and car park security patrols
Ticket fare	$2.50	$2.50
Train frequency	All-stop trains at 10 minute internals	Express and all-stop trains at 5 minute intervals
Tick the box for the scenario you would choose	☐	☐

Figure 5.1 The data source

5.2.2 Method to develop station choice model

The sub-model of station choice was developed within discrete choice theory, which was explained in Chapter three. Here we focus on which model is the best for developing the sub-model.

Based on previous literature, four of the most commonly used discrete choice models are logit, GEV, Probits and mixed logit. Their characteristics are listed in Table 5.1.

Table 5.1 Comparison of different discrete choice models (Train, 2003)

	Logit model	Generalized extreme value models (GEV)	Probits	Mixed logit model
Distribution that unobserved factors followed	IID extreme value distribution	Generalisation of the extreme value distribution	Jointly normal distribution	Any distribution
Advantage	The most widely used, closed form	Avoiding independence assumption	Flexibility in handling correlations over alternatives and time	High flexible and solved three limitation of standard logit models
Limitation	Independence assumption	The extreme value distribution assumption	Normal distribution assumption	No closed form

Based on table 5.1, the logit model has been chosen to capture the effect of travel time variability on station choice due to its closed form and being readily interpretable. Its specification is shown in equation (5-1) (Train, 2003):

$$P_{ni} = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^{J} e^{V_{nj}(\beta)}} = \frac{e^{\beta_n' x_{ni}}}{\sum_{j=1}^{J} e^{\beta_n' x_{nj}}} \tag{5-1}$$

where P_{ni} is the probability that respondent n chooses i^{th} alternative; x_i is the attribute value for alternative i; and β_i is an estimated parameter indicating decision-makers' preferences.

5.2.3 Method to establish the utility function of station choice

This study adopted the mean-variance approach to estimate the utility of train stations resulting from travel time. Mean-variance models are mostly known in the context of risk-return models in finance. Decision makers can maximise the option's return while minimising its associated risk (Fosgerau & Fukuda, 2012; Louviere, Beavers, Norman, & Stetzer, 1973). Jackson and Jucker (1982) introduced the approach into the transportation field. They took the expected travel time and its variability (or unreliability) as the main sources of disutility and assumed that travellers would choose the alternative that minimised the sum of the two terms. Its general specification is shown in equation (5-2):

$$U = \alpha_1 u_T + \alpha_2 \delta_T \tag{5-2}$$

where μ_T is expected travel time of the trip, taken as a measurement of travel time distribution; σ_T is travel time variability and usually uses standard deviation of the travel time distribution; and α_1 and α_2 are the estimated coefficients.

The decision maker ranks each alternative, with a risk prospect obtained from Eq. (5-2), and chooses the alternative with the highest expected utility. The model is usually estimated via a discrete choice method, with the linear-additive specification given in (5-2), and used for choice of route, mode and departure time.

5.2.4 Method to measure travel time variability

Even though the mean-variance approach has been widely used due to its simplicity and better performance (Brownstone & Small, 2005; Small, Winston, & Yan, 2005), it is still criticised on various grounds. One of the main criticisms is that the standard deviation of travel time is not an outcome of a trip. Therefore, the variability of travel time is not the standard deviation but rather the sum of the travel time variation. The variation is divided into two parts, faster variation, i.e. shorter travel time, which is the difference between mean travel time and travel time on a good day, and slower variation, i.e. longer travel time, which is the difference between the mean travel time and travel time on a bad day. Based on this idea, cumulative prospect theory (CPT) was used to measure the travel time variability. As explained in Chapter three, the choice process under the CPT can be divided into two stages: ① "editing" phase, where gains and losses relative to some neutral reference point (t) are identified; and ② "evaluation" phase, where choice is made based on the outcome of alternatives by evaluating their value function $v(x)$ and weighting function (subjective probability function) $\pi(p)$. The utility function of prospect n under CPT is defined in equation (5-3) (Tversky & Kahneman, 1992):

$$u_i^n = \sum_{j=1}^{J} \pi\left(p_l^n\right) v\left(x_j^n - \tau\right) \tag{5-3}$$

Prospect s^n is preferred to $s^{n'}$ if and only if

$$u_i^n > u_i^{n'} \qquad \forall s_i^n \neq s_i^{n'} \tag{5-4}$$

Based on the CPT, and assuming that the regular travel time is taken as the reference point, the differences between the travel times on good days or bad days and regular travel time can be taken as gains or losses respectively. Correspondingly, the frequencies (or probabilities) that bad days or good days occurred in a month can be taken as weightings for gains or losses. Therefore, the utility of the station estimated under CPT can measure the travel time variability.

Many different functional forms have been suggested for both the risk weighting and value functions within CPT. Based on a meta-analysis of different forms (Stott, 2006), this chapter tested the power value function and four popular risk weighting function forms. Their specifications are shown in equation (5-5):

Value function: $$v\left(x_j^n - \tau\right) = \begin{cases} \left(x_j^n - \tau\right)^{\alpha} & if\ \left(x_j^n - \tau\right) > 0 \\ -\lambda \left|x_j^n - \tau\right|^{\beta} & if\ \left(x_j^n - \tau\right) < 0 \end{cases} \tag{5-5}$$

where parameters α and β (less than or equal to one) measure the level of sensitivity to changes in both directions from the reference point, while parameter $\lambda \geq 1$ captures the degree of loss aversion. The value function under prospect theory is usually S-shaped, generally concave for gains and commonly convex for losses, and steeper for losses than for gains, if it describes loss aversion.

The four weighting functions used are shown in equations (5-6) – (5-9):

TK $$\pi(p) = \frac{p^r}{\left(p^r + (1-p)^r\right)^{\left(\frac{1}{r}\right)}} \tag{5-6}$$

GE $$\pi(p) = \frac{sp^r}{sp^r + (1-p)^r} \tag{5-7}$$

Prl-I $$\pi(p) = e^{-(-\ln p)^r} \tag{5-8}$$

Prl-II $$\pi(p) = e^{-s(-\ln p)^r} \tag{5-9}$$

where p_i is the probability that the i^{th} outcome occurs; $\pi(p_i)$ is the subjective weighting function derived from the outcome cumulative probability; and γ and s indicate the shape and location of the risk weighting functions.

5. 3 Station choice model under the effect of uncertain travel time

5.3.1 Framework

The logit model was applied for modelling station choice under travel time uncertainty for P&R users, given that it has closed form and is easy to interpret (Train, 2003). The travel time variability was established within CPT with the regular travel time taken as the reference point. The difference between regular travel time and travel time spent on good days was used to calculate the value function in the gain situation, and the difference between regular travel time and travel time spent on bad days to calculate the value function in a loss situation. As discussed, the value function used a power form, while the weighting function adopted four common forms. The parameters in the choice models were estimated using the Nlogit 5 Package (Li et al., 2010). The framework is shown in Figure 5.2.

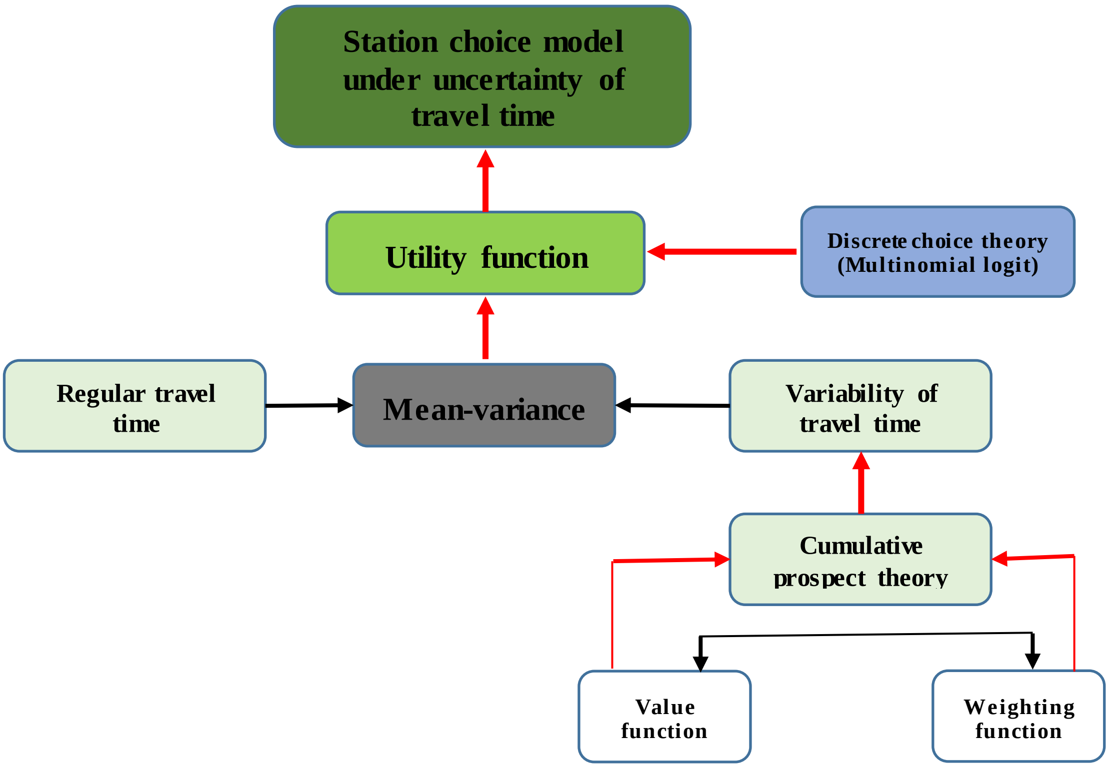


Figure 5.2 Framework for modelling station choice under travel time uncertainty

5.3.2 Station choice model

Based on the mean-variance approach, the observed utility of station choice i (V_i) is given by equation (5-10).

$$V_i = \alpha_1 RTT + \alpha_2 VTT \tag{5-10}$$

where RTT_i is the regular travel time to the station i; VTT_i is the travel time variability when accessing the station i; and α_1, α_2 are estimated parameters which represent a respondent's preference for the station i.

The travel time variability in Eq. (5-10) was redefined based on the CPT. We identified the regular travel time as the reference point, then defined the differences between regular travel time and travel times on good days ($FRTT$) as gains and the differences between regular travel time and travel times on bad days ($SRTT$) as losses. Therefore, the value function $v(\bullet)$ for estimating the effect of travel time variability on station choice under CPT can be written as equation (5-11).

$$v(TT_i - RTT_i) = \begin{cases} \left(-\left(FRTT_i - RTT_i\right)\right)^{\alpha} & \left(TT_i - RTT_i\right) \le 0 \\ -\alpha_3 \left|SRTT_i - RTT_i\right|^{\beta} & \left(TT_i - RTT_i\right) > 0 \end{cases} \tag{5-11}$$

Substituting Eq. (5-6) - Eq. (5-9) and Eq. (5-11) into (5-3), the utility for the travel time variability can be written as (5-12)-(5-15):

$$\begin{aligned} VTT_i = \sum_i v(\Delta x_i)\pi\left(p_i\right) = (RTT_i - FRTT_i)^{\alpha} \times \frac{frttf_i^{\delta}}{\left(frttf_i^{\delta} + (1 - frttf_i)^{\delta}\right)^{\left(\frac{1}{\delta}\right)}} \\ -\alpha_3 (SRTT_i - RTT_i)^{\beta} \times \frac{srttf_i^{\theta}}{\left(srttf_i^{\theta} + (1 - srttf_i)^{\theta}\right)^{\left(\frac{1}{\theta}\right)}} \end{aligned} \tag{5-12}$$

$$\begin{aligned} VTT_i = \sum_i v(\Delta x_i)\pi\left(p_i\right) = (RTT_i - FRTT_i)^{\alpha} \times \frac{frttf_i^{\delta}}{\left(frttf_i^{\delta} + (1 - frttf_i)^{\delta}\right)^{\left(\frac{1}{\delta}\right)}} \\ -\alpha_3 (SRTT_i - RTT_i)^{\beta} \times \frac{srttf_i^{\theta}}{\left(srttf_i^{\theta} + (1 - srttf_i)^{\theta}\right)^{\left(\frac{1}{\theta}\right)}} \end{aligned} \tag{5-13}$$

$$\begin{aligned} VTT_i = \sum_i v(\Delta x_i)\pi\left(p_i\right) = (RTT_i - FRTT_i)^{\alpha} \times {s_1 \times frtt_i^{\delta}} \Big/ {\left[s_1 \times frttf_i^{\delta} + \left(1 - frttf_i\right)^{\delta}\right]} \\ -\alpha_3 \left(-(RTT_i - SRTT_i)\right)^{\beta} \times {s_2 \times srttf_i^{\theta}} \Big/ {\left[s_2 \times srttf_i^{\theta} + (1 - srttf_i)^{\theta}\right]} \end{aligned} \tag{5-14}$$

$$\begin{aligned} VTT_i = \sum_i v(\Delta x_i)\pi\left(p_i\right) = (RTT_i - FRTT_i)^{\alpha} \times \exp(-s_1 \left(-\ln frttf_i\right)^{\delta}) \\ -\alpha_3 (SRTT_i - RTT_i)^{\beta} \times \exp(-s_2 \left(-\ln srttf_i\right)^{\theta}) \end{aligned} \tag{5-15}$$

Where $frttf_i, srttf_i$ are the probabilities that "good traffic" days or "bad traffic" days travel time to the station occur in one month; and $\alpha, \beta, \alpha_3, \delta, \theta, s_1, s_2$ are estimated parameters where α, β indicate a respondent's risk attitude to the good or bad variation in travel time respectively.

Substituting Eq. (5-12)-(5-15) into (5-10), the utility function based on the effect of travel time variability on station choice can be given by Eq. (5-16)-(5-19).

$$V_i = \alpha_1 RTT_i + \alpha_2 \times \left\{ \begin{aligned} &(RTT_i - FRTT_i)^{\alpha} \times \frac{frttf_i^{\delta}}{\left(frttf_i^{\delta} + (1 - frttf_i)^{\delta} \right)^{\left(\frac{1}{\delta}\right)}} \\ &- \alpha_3 (SRTT_i - RTT_i)^{\beta} \times \frac{srttf_i^{\theta}}{\left(srttf_i^{\theta} + (1 - srttf_i)^{\theta} \right)^{\left(\frac{1}{\theta}\right)}} \end{aligned} \right\} \tag{5-16}$$

$$V_i = \alpha_1 RTT_i + \alpha_2 \times \left\{ \begin{aligned} &(RTT_i - FRTT_i)^{\alpha} \times {s_1 \times frtt_i^{\delta}} \Big/ {\left[s_1 \times frttf_i^{\delta} + (1 - frttf_i)^{\delta} \right]} \\ &- \alpha_3 \left(-(RTT_i - SRTT_i) \right)^{\beta} \times {s_2 \times srttf_i^{\theta}} \Big/ {\left[s_2 \times srttf_i^{\theta} + (1 - srttf_i)^{\theta} \right]} \end{aligned} \right\}$$

(5-17)

$$V_i = \alpha_1 RTT_i + \alpha_2 \times \left\{ \begin{aligned} &(RTT_i - FRTT_i)^{\alpha} \times \exp(-(-\ln frttf_i)^{\delta}) \\ &- \alpha_3 (SRTT_i - RTT_i)^{\beta} \times \exp(-(-\ln srttf_i)^{\theta}) \end{aligned} \right\} \tag{5-18}$$

$$V_i = \alpha_1 RTT_i + \alpha_2 \times \left\{ \begin{aligned} &(RTT_i - FRTT_i)^{\alpha} \times \exp(-s_1 (-\ln frttf_i)^{\delta}) \\ &- \alpha_3 (SRTT_i - RTT_i)^{\beta} \times \exp(-s_2 (-\ln srttf_i)^{\theta}) \end{aligned} \right\} \tag{5-19}$$

5. 4 Results

The parameters in choice model Eq. (5-16)-(5-19) were estimated using a non-linear multinomial logit. The results are shown in Table 5.2.

Table 5.2 Parameters estimated in the four MNL station choice models

Variable	MNL model with TK weighting function	MNL model with GE weighting function	MNL model with Prl-I weighting function	MNL model with Prl- II weighting function
Regular travel time	-0.00278	-0.008096**	-0.08082**	-0.07929***
Variability of travel time	0.11929	3.03276	2.74289	5.17999
Bad variation of travel time	0.919943	13.6417	5.6	-44.7563
α	0.10195	0.29546	0.2963	0.03395
β	0.10248	-1.71542**	-1.72087**	0.0178
δ	0.02996	0.51817	0.26298	2.38435
θ	0.03375	3.40999	2.32824	-2.39625
s_1		2.43180		0.00340
s_2		10.8864		0.01647
Number of Obs.	2397	2397	2,397	2,397

AIC/N	1.392	1.37	1.37	1.375
Log likelihood	-1661.38706	-1635.533048	-1,635.329	-1,639.536
Chi-square	52.28218	52.28663	52.29	43.88

Note: ***, **, * ==> Significance at 1%, 5%, 10% confidence level.

Table 5.2 indicates that the model with the GE weighting function has the best fit, followed by the model with the Prl-I weighting function. Across the models, only the regular travel time and the parameter β in the last three models have a significant effect, (at the 95% confidence level), on station choice for P&R users. Moreover, according to the results from the first three models, a longer regular travel time and greater travel time variability on "bad traffic" days have negative effects on station choice. In gain situations, the more travel time travellers save, the greater the likelihood of choosing that station, which is consistent with the a priori expectations. The parameter α is not statistically significant, but it has some effect on the shape of the function in gain situations and can indicate a respondent's risk attitude towards time variability. Both values of the parameters α, β are less than 1, indicating that the shape of value of gains is concave and the shape of losses is convex (see Figure 5.3). As none of the parameters δ, θ, s_1, s_2 is significant, it is hard to draw conclusions on the shape of the weighting function. With the exception of the MNL model with TK weighting function, it is possible that the function has an inverse S shape (see Figure 5.4). The results also show that outcomes with low probabilities tend to be overweighed and the outcomes with high probabilities tend to be underweighted by respondents, which is aligned with the CPT assumption. Based on the data available here, the shape of the TK weighting function is convex and close to zero. The signs of the parameters $\alpha_2, \alpha_3, \theta$ in the last model in Table 5.2 are different from what we expected, so the MNL model with Prl-II weighting function has been excluded from further calculations. Based on all the results, the model with GE weighting function was chosen for modelling the travel time reliability.

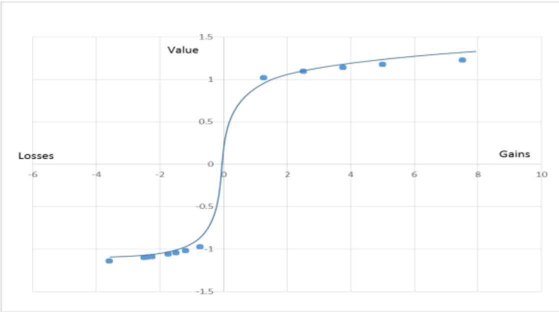


Figure 5.3 Shape of TK value function

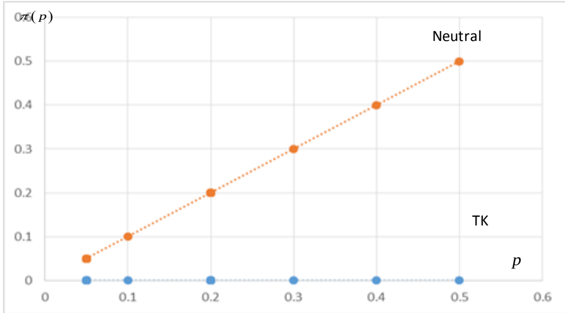


Figure 5.4.1. Shape of TK weighting function

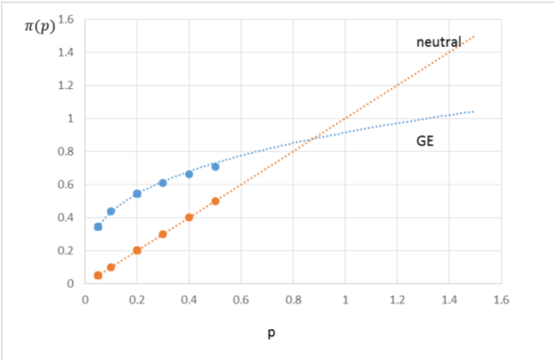


Figure 5.4. 2. Shape of GE weighting function

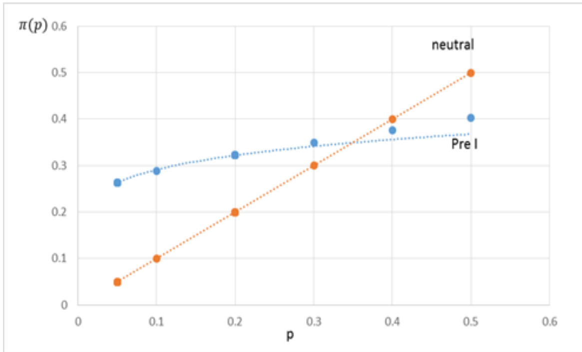


Figure 5.4. 3. Shape of Prl-I weighting function

Figure 5.4 Shape of weighting functions

After substituting the parameters presented in Table 5.2 into (5-12), we found that the relationship within cumulative prospect theory may indicate the risk aversion attitude of the respondents (see Eq.5-20).

$$\begin{aligned} &\pi(frttf_i) < v\big(frttf_i \times (RTT_i - FRTT_i)\big) / v(RTT_i - FRTT_i) \\ &\pi(srttf_i) > v\big(srttf_i \times (SRTT_i - FRTT_i)\big) / v(SRTT_i - RTT_i) \end{aligned} \tag{5-20}$$

5. 5 Impact of respondents' real travel time experience on their risk attitude towards station choice

In this section, two research hypotheses were set up to understand the impact of respondents' real travel time variation experiences on their risk attitude towards their station choice under travel time variability:

H_1: Respondents who have experienced higher travel time variations tend to be more risk averse towards their station choice than those who has experienced lower travel time variations.

H2: Respondents with higher differences between perceived and objective travel times tend to be more risk averse towards their station choice than those who have experienced less travel time variation.

In order to test these hypotheses, we set up three scenarios, which are based on three experiments conducted at three different stations. The access travel time for these three stations is different. The fundamental assumption in these experiments is that the objective travel time variation experienced by respondents and their perception of travel time could influence their perceptions of travel time variability. Table 4 summarises the parameter β value for the three train stations (Murdoch, Warnbro, and Greenwood). The respondents at Murdoch station were more risk averse compared to those at the other two stations. Figure 5.5 illustrates the relationships between the value of travel time variability and travel time variation at the three stations. The curve generated from the Murdoch station scenario is much steeper than the other two, which supports our conclusions derived from Table 5.3.

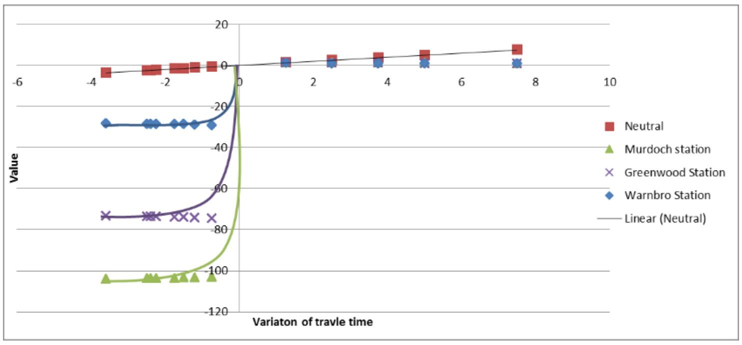


Figure 5.5 Comparison of respondents' risk attitude for the three train stations

5.5.1 Variation of travel time over a day

The origins of all trips for the P&R respondents at the three stations were geocoded and their access travel times checked using Google Maps for 5 weekdays (from Monday to Friday). We found that the variation in travel time over the 5 weekdays was highest at Murdoch station and lowest at Warnbro station (see Table 5.3), which is in line with the risk attitude results. Therefore, there is support for hypothesis 1 that *respondents who have experienced higher travel time variations tend to be more risk averse towards their station choice under travel time variability than those who has experienced lower travel time variations.*

Table 5.3 Summary of respondents' risk attitude towards bad variability and variation of travel time to the three train stations

	Murdoch station	Greenwood station	Warnbro station
β	0.00434	-0.1785	-0.01706
Variation in travel time	1.8	1.08	0.45

The difference between stated travel time by respondents and the estimated travel time by Google	Average (mins)	8.9	3.461	2.571
	S.D. (mins)	8.239	6.138	5.482
	Mean (mins)	7.5	3	1

5.5.2 Difference between perceived travel time by respondents and estimated travel time by Google

Based on the differences between the estimated travel time and the stated travel time (as perceived by the respondents), we found that P&R users using Murdoch station had the highest differences, whereas P&R users at Warnbro had the smallest differences (see Table 5.3), which is in line with the risk attitude. Therefore, hypothesis 2 that *Respondents who have higher differences between perceived and objective travel time tend to be more risk averse towards their station choice under travel time variability than those who have experienced lower travel time variations* appears to have some support.

5.5.3 Summary

Combining risk attitude, variation in travel time and the difference between perceived travel time by respondents and the estimated travel time by Google at three train stations, we found that P&R users at the stations with higher variations in access travel times and greater differences between perceived and estimated travel times, displayed more risk aversion than in other circumstances.

5. 6 Conclusions and limitations

In this section, P&R user station choice under travel time variability was analysed using multinomial logit models. According to our knowledge, this is the first attempt to understand P&R user station choice under travel time uncertainty based on a combination of cumulative prospect theory and discrete choice theory. The data used in the section were collected using a stated choice (SC) survey conducted at seven train stations in Perth, Western Australia. The questionnaires used in the survey were designed based on a D-efficiency approach. The SP data were modelled within discrete choice theory, and the utility function established by the mean-variance approach. The variability of travel time in the utility function was redefined within cumulative prospect theory. Four risk weighting functions were separately applied for the station choice model. The results showed that the effect of travel time variability is not statistically significant, but respondents may display risk aversion for travel time variability. The model with the GE risk weighting function was found to have the best fit model and obvious non-linearity exists in the risk weighting function.

A general conclusion from the results is that larger travel time variability in the losses situation, i.e. travel on "bad traffic" days, and longer regular travel time could lead to lower utility and hence a lower probability that a particular station is chosen. This result has two policy implications. Firstly, the station choice model indicated that P&R users showed risk aversion toward travel time variability. The highest risk aversion attitude was found among P&R users of Murdoch station, which has the highest variation in travel time. Therefore, we can conclude that risk aversion toward travel time variability has some influence on station choice. Secondly, we also investigated the impact of P&R users' experiences and perceptions of travel time variability, which could affect their attitude toward travel. P&R users who have experienced higher travel time variations tend to be more risk averse towards their station choice under travel time variability than those who have experienced lower travel time variations. This might indicate that a traveller's personal experience may play a vital role in risk attitude toward station choice under uncertainty. It could also indicate that a traveller's perceptions may play some role in risk attitude toward station choice under uncertainty. Although the SC survey was conducted based on hypothetical situations, the outcome of the choice model indicates that the choice model and cumulative prospect theory are powerful tools, able to capture choice under uncertainty behaviour. Further work is necessary to systematically review the factors affecting risk attitude in mode choice. In this study, the P&R users' risk attitude towards station choice considered only travel time variability. However, parking search time, parking availability, ticket fare, and crowding represent additional determinants of P&R station choice that should be tested.

Additional analysis is required to determine whether travel time variability is the main factor for station choice. One way is to use eye-tracking technologies to monitor the survey process. Visual attention could indicate the importance of factors, i.e. if participants pay more attention to a variable, this may suggest that it has more influence on the choice decision. Figure 5.6 shows a snapshot of a visual attention map (heatmap) of a participant. Each participant was shown two different questionnaires. For the questionnaire on the left, travel time was found to be the main centre of attention, i.e. with the longest visual attention. However, when we increased the number of variables, or factors, on the questionnaire, attention shifted to parking instead of travel time. Therefore, the best way to design the SP survey, and balance the number and order of variables and levels of attributes of variables, to estimate choice behaviour in a most efficient way is still open to debate and more evidence is needed.

Figure 5.6 Heatmap analysis for both questionnaires (different choice)

5. 7 Chapter summary

This chapter developed a sub-model of station choice based on the effect of travel time variability within a discrete choice theory framework. The mean-variance approach was used to estimate the utility of a station from the travel time attributes. CPT was applied to measure the effect of travel time variability. Based on the sub-model, we analysed P&R users' risk attitude towards the variation in travel time and impact of their real travel time experience on their risk attitude towards station choice.

The next chapter will analyse the effect of another uncertain factor on station choice for P&R users, i.e. crowding on trains. The concepts related to crowding on trains will be explained and the methods to develop the sub-model of station choice focusing on the effect of crowding discussed.

CHAPTER 6 STATION CHOICE SUB-MODEL MERGING THE EFFECT OF CROWDING ON TRAINS

The last chapter developed a station choice sub-model based on the effect of the variation of travel time on the choice of departure train station for P&R users. Using this the model, the influence of travel time variability on station choice was investigated, P&R users' risk attitudes towards variations in travel time were measured, and the impact of P&R users' experiences and perceptions of travel time variability on their risk attitude were also analysed. This chapter continues the modelling of station choice under uncertainty, mainly focusing on the effect of variations in the level of crowding on the train. Similar to the previous chapter, the influence of crowding variability on station choice is investigated and the P&R users' risk attitudes towards overcrowding are measured. The relationship between P&R users' risk attitudes and rail ridership, and P&R users' preference heterogeneity, led by individual differences, is also explored.

6. 1 Research context

Crowding on trains is technically known as, inter alia, "*passengers in excess of capacity*" (Government of UK, 2015), which is often taken as one of key service indicators for public transport (Hensher, Stopher, & Bullock, 2003). Crowding is the problem most frequently encountered by passengers in Australia and many other countries (Cox, Houndmont, & Griffiths, 2006; Thompson, L. Hirsh, S.Muller, & S.Rainbird, 2012). Generally, its effects can be analysed from two perspectives: the effects on passengers and the effects on operators. From a passenger's perspective, the first effect is on rail passengers' physiology and psychology. For example, Cox et al. (2006) and Katz and Rahman (2010) asserted that overcrowding on trains might make rail users more stressed. Lundberg (1976), Mahudin, Cox, and Griffiths (2011) also found that overcrowding on trains could increase anxiety and stress levels based on the rate of catecholamine excretion. Moreover, the feelings of discomfort might grow more intense as the density of passengers increased. Mahudin et al. (2011) found evidence that crowding led to more somatic symptoms such as sleeplessness, tension, headaches, etc.

The second effect is its influence on rail passenger journey times. For example, Fernández (2011) and Tirachini et al. (2013) discovered that crowding could increase riding time, boarding time, alighting time and waiting time. Moreover, the effect of crowding on alighting is greater than on boarding.

The third effect is its influence on travel choice. A study, conducted by Hensher and McLeod (1977), identified in-vehicle crowding as one of the main determinants of travel choice. After that, a number of researchers successively investigated its influence on specific travel choice behaviour. For example, Leurent and Liu (2009), Schmöcker, Fonzone, Shimamoto, Kurauchi, and Bell (2011) and Raveau, Muñoz, and Grange (2011) found that passengers adjusted their routes to minimise their travel cost, as in-vehicle congestion increased the route cost in the studies of transit assignment models. Sumalee, Tan, and Lam (2009) and Hamdouch, Ho, Sumalee, and Wang (2011) found evidence that transit commuters changed their departure time to reduce their in-vehicle congestion cost, using a seat allocation model developed within a schedule-based framework. In contrast to the above, Davidson, Vovsha, Abedini, Chu, and Garland (2011) identified the effect of crowding on waiting time and in-vehicle travel time, and hence on route choice, by developing a crowding module. Additionally, Kim, Lee, and Oh (2009) revealed that travellers did not always choose the first bus to arrive if it had a high occupancy rate. They observed that that the higher the level of crowding on a bus, the lower the probability that a public transit user would choose to board that bus.

From an operator's perceptive, overcrowding can affect operating speed and cost of providing the service, as well as public transport ridership. Therefore, level of crowding was often considered as an index for evaluating and improving public transport service operation. Intervention strategies, such as increasing service frequency and/or using larger vehicles, have been implemented to reduce crowding levels. Batarce, Muñoz, and Ortúzar (2016) found evidence of the effect of crowding on a public transport system's demand and user benefits by comparing outcomes from three transport policies that improved bus corridor operations. Tirachini et al. (2013) summarised the effect of crowding on public transport system reliability, optimal supply and pricing. Based on these, they suggested that public transport operators should determine the service frequency and capacity by considering, inter alia, the effects of crowding.

Whether for passengers or operators, it is very important to understand the mechanics of overcrowding, evaluate it and then develop strategies to reduce it. However, the literature related to crowding on trains is limited but can be divided into three broad classes: modelling the effect of crowding on travel choice (i.e. route choice, mode choice and departure time choice); measurements of crowding; and willingness to pay for reduced crowding.

(a) Travel choice models based on the effect of crowding

The studies can be traced back to the 1970s. Since Hensher and McLeod (1977) identified crowding as one of main determinants of travel choice, a number of authors has attempted to incorporate it into the utility related to travel choice. For example, Polydoropoulou and Ben-Akiva (2001) took seat availability as the indicator of crowding and introduced it into the utility related to public transport alternatives in their nested logit travel mode choice model. Hensher, Rose, et al. (2011) developed a stated choice experiment to explore commuters' choice between the proposed new Metro and existing available modal alternatives, in which crowding, indicated by seat availability and the density of passengers standing in a train car, was compared in the existing and the new public transport modes. Davidson et al. (2011) developed a crowding model that took crowding as a negative factor in the user perception of transit service quality, then, together with the effect of capacity and delayed vehicle arrival, incorporated it into the mode choice model. Debrezion et al. (2009) developed a nested logit model to analyse rail passenger choice of departure train station based on the effect of accessibility into which crowding measurements were integrated.

In summary, the research related to travel choice modelling that considered the effect of crowding is very limited and most were used to explore mode choice. Therefore, to efficiently reduce the effect of crowding on train trips, systematically exploring the effect of in-vehicle crowding on other travel choices, such as P&R access station choice, departure time choice, etc., should be one direction for future research.

(b) Measurements of in-vehicle crowding

Generally, two types of crowding measurements have been identified, objective and subjective (Cox et al., 2006; Day & Day, 1973; Evans & Wener, 2007; Li & Hensher, 2011; Mahudin, Cox, & Griffiths, 2012; Sundstrom, Busby, & Asmus, 1975; Turner, Corbett, O'Hara, & & White, 2004; Zheng Li & Hensher, 2013). The main difference is whether the measurement can reflect an individual's perception or not. Objective measurement refers to the metrics that can be used to objectively and quantitatively assess material circumstances. They include load factor, i.e. the ratio of the actual number of passengers inside a vehicle to the number of seats (Whelan and Crockett, 2009); passenger loading in terms of levels of service (Lam, Cheung, & Lam, 1999); the percentage of passengers standing (Blunden et al., 2011); standing passenger area (i.e. space (m^2) per standing passenger); number of standing passengers per square meter, and rolling hour average loads and length of standing time (Zheng Li & Hensher, 2013). Among these measures of crowding, the density, i.e. the ratio of passengers to space, is taken as the most common measurement of crowding on a public transport system (Li & Hensher, 2011; Mahudin et al., 2012), even though it still cannot fully capture an individual's perception of crowding in a given space (Cox et al., 2006; Day & Day, 1973; Turner et al., 2004). In reality, a rail passenger's perception is subjective and may be affected by many factors, such as physical antecedents, inter-personal, individual personality and previous experience (Evans & Wener, 2007; Sundstrom et al., 1975; Turner et al., 2004). Therefore, subjective crowding should be measured based on the objective crowding measurement plus an individual's previous experiences with crowding, tolerance and personal opinion. However, few quantitative metrics of subjective crowding have been identified (Zheng Li & Hensher, 2013). In this context, limited research used subjective measurements of crowding, even though it does directly influence choice behaviour.

(c) Willingness to pay

Other studies of crowding estimated the value of crowding (VOC) and explored public transport passengers' willingness to pay (WTP) for reduction of crowding. For example, Lu and Wardman (2008) estimated the VOC with two groups of stated preference data. One was from the experiment by adding task complexity and another was by adding cheap talk, (i.e. a term implying that nothing needs to be paid for the communication). The results showed that the VOC from the former experiment was larger than from the latter. Li and Hensher (2011) reviewed the published studies related to value of crowding and the results indicated that the average value of crowding was £7.23 per person hour in the UK and $9.92 per person hour in Australia. Overall, the research related to the VOC is relatively scarce and mainly focused on UK and Australia.

Studies related to crowding have been conducted for many years but there are still a number of gaps including how to measure crowding, how in-vehicle crowding affects train station choice and how much passengers are willing to pay for reduced crowding.

In this chapter, we identify a more efficient indicator of crowding, model train station choice under the effect of crowding for P&R users, analyse the respondents' risk attitudes towards crowding and measure the effect of an individual's personality on station choice.

6. 2 Research method

6.2.1 Introduction

This section starts with the identification of crowding measurements, then its characteristics, (certain or uncertain) are distinguished. Based on these, appropriate decision making theories to explore station choice under crowding uncertainty are determined and choice models are developed to meet the research objectives.

The objective measurements of crowding mentioned above, i.e. the probability that seats taken and the density of passengers standing in a carriage, are used in this chapter. However, in contrast to the load factor metric used in previous literature, they are independent, but interactive in the research. In the literature to date, the pictorial display of crowding shown to survey respondents was from above (i.e. a bird's eye view), so that both load factors can be taken as two independent indices and separately analysed (see Figure 6-1). However, the pictorial display of crowding used in this research is a horizontal view, (i.e. the view passengers waiting to board have when the carriage doors open), (see Figure 6-2). In this situation, it is difficult for respondents to clearly identify both load factors, especially when the carriage is over-crowded or people are clustered near the doorways rather than moved down inside the carriage. The level of crowding boarding passengers view, and therefore perceive, is an interaction between both load factors, as standing passengers can block the view of the carriage and any spare seats. The crowding measures in the research consider this interaction between the two load factors. The unit is passengers/m^2

Figure 6.1 Crowding displayed from bird's eye view

Figure 6.2 Crowding displayed from horizontal view

According to Knight (1921), risk is a situation where there is a range of possible outcomes faced by the decision maker. Given that the new crowding measures vary as the probability of seats taken and the density of passengers standing in a carriage change throughout the day and from day-to-day, it is reasonable to identify crowding as a risk factor for a boarding passenger. Correspondingly, station choice based on the effect of variations in the level of crowding can be defined as choice under risk. Two decision making theories under risk or uncertainty, namely expected utility theory (EUT) and extended expected utility theory (EEUT), have been applied to model station choice behaviour under the effect of crowding on trains. Specific to the EEUT model, the value function adopted the power form and the probability weighting functions used four popular forms, i.e. Tversky-Kahneman (TK), Goldstein-Einhorn (GE), PrelecI (Pr1-I) and PrelecII (Pr1-II). The detailed framework to model station choice based on the effect of variation of crowding on the train is presented on Figure 6.3.

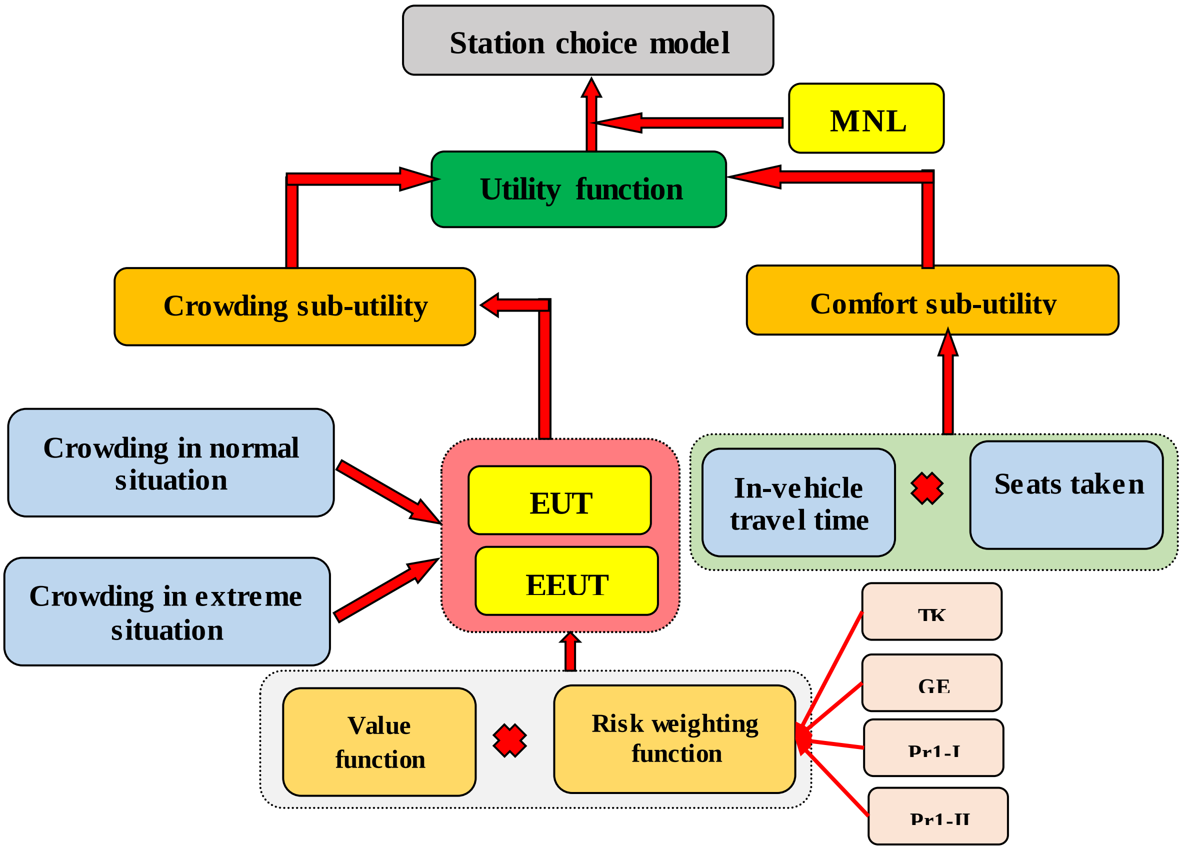


Figure 6.3 Framework to model station choice under crowding risk

6.2.2 Utility function related to station choice incorporating the effect of variations in crowding

Based on the findings of the station choice survey, we determined that crowding and discomfort on trains were the main factors contributing to the utility related to station choice under crowding risk, and that their effects were linear-additive. Thus, the utility specification for station choice under the effect of crowding can be written as equation (6-1).

$$U = \alpha_1 U(C) + \alpha_2 U(DCo) \tag{6-1}$$

where U is the utility related to station choice; C is the crowding on trains; DCo is the discomfort on trains; $U(C)$ and $U(DCo)$ are the utilities for crowding on train and discomfort on train respectively; and α_1, α_2 are estimated coefficients.

For the discomfort on train utility, we surmised that it could be measured by the interaction between the probability of seats taken and in-vehicle travel time, so the utility function for discomfort on train can be written as equation (6-2).

$$U(DCo) = IVT \times StP \tag{6-2}$$

where IVT is the in-vehicle travel time, i.e. the amount of time spent travelling on the train; and StP is the probability that all seats are taken in a train carriage.

We separately applied Expected Utility Theory (EUT) and Extended Expected Utility Theory (EEUT) to estimate the crowding on train utility, on the assumption that it introduced risk into the choice process. The EUT model and the EEUT model are discussed below.

(a) The utility function related to the effect of crowding on station choice within Expected Utility Theory (EUT)

Based on the EUT mentioned in Chapter 3, all reasonable decision makers choose the alternative with the maximum expected utility values. In detail, when they make a choice with risky or uncertain outcomes, they calculate the expected utility value for each alternative by multiplying the value of the outcome by the probability of it occurring. The simplest specification within EUT is shown in equation (6-3) (Einhorn & Hogarth, 1981; Neumann & Morgenstern, 1994).

$$E(U) = \sum_m (p_m x_m) \tag{6-3}$$

where $E(U)$ is the expected utility of choosing a train station; m is a set of outcomes; p_m is the objective probability of outcome m; and x_m is the utility of outcome m.

Based on the data we collected related to crowding, i.e. the crowding level on a typical week day(Ct), the extreme crowding level (Ce) and their relative probabilities, the linear utility function of crowding on trains within the EUT is the sum of the product of the crowding level and its probability. Its specification is shown in equation (6-4)

$$EU(C) = Ct \times (1 - p_t) + Ce \times p_e \tag{6-4}$$

where $EU(C)$ is the expected utility produced by crowding on trains; Ct is the crowding level on a typical week day, (equal to the production of the proportion of seats taken and the density of standees based on the definition of the new crowding measures in the section); Ce is the extreme crowding level, (here 8 passengers/m^2 and with all seats assumed to be taken); p_e is the probability that the extreme situation occurs, (calculated based on the number of days per week on which trains are too crowded to board (n), i.e. $p_e = 1 - n/5$); and p_t is the probability that the typical crowding level occurs (equal to $1 - p_e$).

EUT models also postulate non-linear functional forms. A power form utility function was used in the section (see Eq. (6-5)).

$$E(U) = \sum_m \left(p_m x_m^{\beta} \right) \tag{6-5}$$

where β is an estimated coefficient indicating a decision maker's risk attitude towards the outcomes.

Substituting the crowding levels and their relative probabilities from equation (6-4) into equation (6-5), i.e. combining the non-linear EUT model attributes with the linear EUT model and the risk attitude parameter, results in equation (6-6). This equation not only explains a P&R user's preference for the factors influencing their choice but also reveals their risk attitude towards these factors.

$$EU(C) = Ct^{\beta} \times \left(1 - p_t\right) + Ce \times p_e \tag{6-6}$$

(b) Utility function related to the effect of crowding on station choice within Extended Expected Utility Theory (EEUT)

Given that the probability weighted outcomes in EUT models use objective probabilities, they do not necessarily align with respondents' perceptions based on their experiences and knowledge. Therefore, non-linear probability weighting was introduced into the non-linear EUT models to produce an extended expected utility theory (EEUT) model, (as proposed by Li et al. (2009)), in which the probabilities can over-weight or under-weight the objective probabilities. Based on this, the utility function within the EEUT is given by equation (6-7).

$$EE(U) = \sum_m \left(w\left(p_m\right) \times x_m^{\beta} \right) \tag{6-7}$$

where $w(\bullet)$ is a probability weighting function.

Four popular probability weighting functions - TK, GE, Pr1-I, and Pr1-II - were used, as given by equations (6-8 to 6-11).

TK $$w(p) = \frac{p^{\gamma}}{\left(p^{\gamma} + (1 - p^{\gamma})\right)^{\left(\frac{1}{\gamma}\right)}} \tag{6-8}$$

GE $$w(p) = \frac{sp^{\gamma}}{sp^{\gamma} + (1 - p^{\gamma})} \tag{6-9}$$

Pr1-I $w(p)=e^{-(-\ln p)^{\gamma}}$ (6-10)

Pr1-II $w(p)=e^{-s(-\ln p)^{\gamma}}$ 6-11)

where γ, s are estimated coefficients.

Replacing the probabilities in equation (6-6) with those from equations (6-8) to (6-11), the utility functions of crowding within the EEUT can be written per equations (6-12) to (6-15) respectively.

$$EEU(C)_{TK}=Ct^{\beta}\times\frac{(1-p_t)^{\gamma}}{\left((1-p_t)^{\gamma}+p_t^{\gamma}\right)^{\frac{1}{\gamma}}}+Ce\times\frac{p_e^{\gamma}}{\left(p_e^{\gamma}+(1-P_e)^{\gamma}\right)^{\frac{1}{\gamma}}} \tag{6-12}$$

$$EEU(C)_{GE}=Ct^{\beta}\times\frac{s(1-p_t)^{\gamma}}{s(1-p_t)^{\gamma}+p_t^{\gamma}}+Ce\times\frac{sp_e^{\gamma}}{sp_e^{\gamma}+(1-P_e)^{\gamma}} \tag{6-13}$$

$$EEU(C)_{\mathrm{Pr1}-I}=Ct^{\beta}\times e^{-(-\ln(1-p_t))^{\gamma}}+Ce\times e^{-(-\ln p_e)^{\gamma}} \tag{6-14}$$

$$EEU(C)_{\mathrm{Pr1}-II}=Ct^{\beta}\times e^{-s(-\ln(1-p_t))^{\gamma}}+Ce\times e^{-s(-\ln p_e)^{\gamma}} \tag{6-15}$$

(c) Utility functions related to station choice incorporating crowding uncertainty

Substituting equations (6-2), (6-4), (6-6), and (6-12) to (6-15) into equation (6-1), the linear and non-linear utility functions of crowding within the EUT and the non-linear utility functions of crowding within the EEUT are presented in equations (6-16) to (6-21).

(i) Linear EUT model

$$U=\alpha_1\left[Ct\times(1-p_t)+Ce\times p_e\right]+\alpha_2 IVP\times Stp \tag{6-16}$$

(ii) Non-linear EUT model

$$U=\alpha_1\left[Ct^{\beta}\times(1-p_t)+Ce\times p_e\right]+\alpha_2 IVP\times Stp \tag{6-17}$$

(iii) Non-linear EEUT models

$$U_{TK}=\alpha_1\left[Ct^{\beta}\times\frac{(1-p_t)^{\gamma}}{\left((1-p_t)^{\gamma}+p_t^{\gamma}\right)^{\frac{1}{\gamma}}}+Ce\times\frac{p_e^{\gamma}}{\left(p_e^{\gamma}+(1-P_e)^{\gamma}\right)^{\frac{1}{\gamma}}}\right]+\alpha_2 IVP\times Stp \tag{6-18}$$

$$U_{GE}=\alpha_1\left[Ct^{\beta}\times\frac{s(1-p_t)^{\gamma}}{(1-p_t)^{\gamma}+p_t^{\gamma}}+Ce\times\frac{sp_e^{\gamma}}{p_e^{\gamma}+(1-P_e)^{\gamma}}\right]+\alpha_2 IVP\times Stp \tag{6-19}$$

$$U_{PreI}=\alpha_1\left[Ct^{\beta}\times e^{-\left(-\ln(1-p_t)\right)^{\gamma}}+Ce\times e^{-\left(-\ln p_e\right)^{\gamma}}\right]+\alpha_2 IVP\times Stp \tag{6-20}$$

$$U_{PreI}=\alpha_1\left[Ct^{\beta}\times e^{-s\left(-\ln(1-p_t)\right)^{\gamma}}+Ce\times e^{-s\left(-\ln p_e\right)^{\gamma}}\right]+\alpha_2 IVP\times Stp \tag{6-21}$$

6.2.3 Risk attitude and its measurement

Based on both EUT and EEUT, risk aversion implies that the probability of an alternative with a potentially poor outcome being chosen is lower than its objective probability. Conversely, risk seeking implies that the probability of an alternative with a potentially good outcome being chosen is greater than its objective probability. Based on these concepts, the value function is concave for risk aversion and convex for risk seeking. Figure 6.4 indicates risk aversion in the four different quadrants.

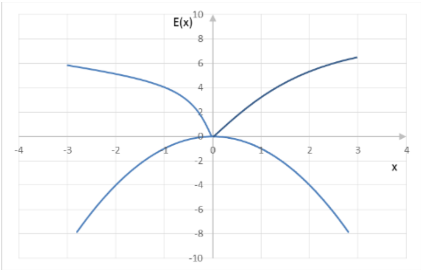


Figure 6.4 Risk aversion under EUT and EEUT

Usually, under EUT and EEUT, an individual's attitude is risk averse. Different methods have been developed to measure the scale of risk aversion. The most commonly and frequently used measures include the Arrow-Pratt measure of absolute risk-aversion (ARA) and relative risk-aversion (RRA) named by Pratt (1964) and Arrow (1971). Here, we use the coefficient of relative risk aversion (CRRA) to measure risk aversion, given that it is a unit-free measurement of sensitivity (Rubsinstein, 2006). Its specification is shown in equation (6-22)

$$CRRA(x)=-\frac{xu''(x)}{u'(x)} \tag{6-22}$$

where $u'(\bullet)$ and $u''(\bullet)$ are the first and the second derivatives of the utility function

Based on equation (6-22), the coefficient of relative risk aversion for crowding on trains in the section can be inferred as equation (6-23):

$$CRRA(Ct)=-\frac{Ctu''(Ct)}{u'(Ct)}=1-\beta \tag{6-23}$$

The larger its absolute value, i.e. the greater the curvature of the utility function, the more risk averse the respondent.

6.2.4 Analysis of individual heterogeneity

In the chapter, we use a latent class model to investigate the heterogeneity of individuals. The latent class model (LCM) is a semi-parametric approximation to the random parameter multinomial logit model that resembles the mixed logit model.
Within the LCM, each individual's behaviour is determined from the observable attributes and potential similarities within each class, (homogeneity within heterogeneity), for factors that cannot be unobserved, (known), by the assessor. Assuming that individuals are divided into Q classes and the mixing distribution $f(\beta)$ is discrete, then the choice probability can be calculated by equation (6-24).

$$P_{ni} = \sum_{m=1}^{Q} s_m \frac{e^{\beta_m' x_{ni}}}{\sum_{j=1}^{J} e^{\beta_m' x_{nj}}} \tag{6-24}$$

where p_{ni} is the probability that the n^{th} respondent chooses the i^{th} alternative; s_m is the share of the population in segment m; Q is the number of the class; and β_m is the estimated coefficient for the m^{th} class.

6. 3 Results

6.3.1 Estimation of coefficients

We started with the EUT models, then developed the non-linear EEUT models. All coefficients in the models were estimated within a multinomial logit model using the Nlogit5 package.

(a) The EUT models

The coefficients in the linear and non-linear EUT models are summarised in table 6.1.

Table 6.1 Estimation of coefficients for EUT models

	Linear EUT model	Non-linear EUT model
Crowding	-0.01341	0.000001
Discomfort	-0.01310***	-0.02070***
β		10.4431***
Constant specific for station one	0.11392***	0.08967***
Log Likelihood	-1648.25247	-1641.85502

Inf.Cr.AIC	3302.5	3291.7
Chi-square [d.f.]	15.65955[2]	37.85124[4]
Significance level	0.00040	0.00000

Note: ***, **,* == >Significant at 1%, 5% and 10% confidence levels

For the linear EUT model, all estimated parameters are statistically significant at the 99 percent confidence interval, with the exception of crowding on trains. The coefficient estimate for α_2 is negative, which means that the more seats taken or the longer the travel time, the lower the utility of the train station, which is aligned with expectations. The station one specific constant is positive, which indicates that the sampled respondents prefer the first alternative, all other factors being equal.

When compared to the linear EUT model, the non-linear EUT model delivered a similar behavioural response, even though the latter model introduced an extra parameter, i.e. a risk attitude parameter. The only difference between the results of the two models is the values of the parameter coefficients. The coefficients in the non-linear EUT model are a little larger than those in the linear EUT model. Focusing on the non-linear EUT model, coefficient β, indicating respondents' risk attitude, is significant at the 99 percent confidence interval, which means it does have an effect on station choice. Based on its value of 10.4431, the value function's shape is concave, (i.e. is in the lower right quadrant of Figure 6.4), and as shown on Figure 6-5. This indicates that P&R users' attitudes towards crowding on trains on a typical weekday should be risk averse. Even though both EUT models can fit the collected data based on their significance levels and their coefficient estimates are similar, we consider that the non-linear EUT model is better than the EUT model due to it having a smaller AIC index, (Akaike Information Criterion) - the smaller the AIC index the better the model. The index for the non-linear EUT model is 3291.7, which is smaller than the linear EUT model's 3302.5. Therefore, the non-linear EUT model was adopted as the preferred model to explore station choice behaviour for P&R users.

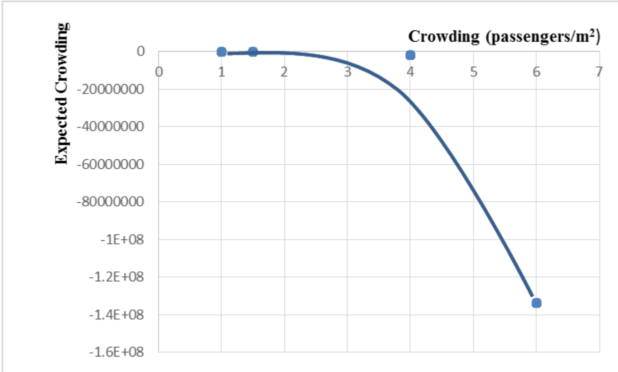


Figure 6.5 The value function in non-linear EUT utility function

(b) The EEUT model

The coefficients in the EEUT model with TK weighting risk function (see equation (6-18)) were estimated using Nlogit 5, with the results presented in Table 6.2.

Table 6.2 Estimation of coefficients in the EEUT model with TK weighting function

Choice	Coefficient	Standard Error	z	Prob. |z|>Z*	95% Confidence Intervals	
Crowding	-0.27669	0.57896	-0.48	0.6327	-1.41142	0.85805
Discomfort	-0.03335***	0.00680	-4.90	0.0000	-0.04667	-0.02002
β	2.65367***	0.81942	3.24	0.0012	1.04763	4.25970
γ	0.17229	0.16932	1.02	0.3089	-0.15957	0.50415
Constant specific for station one	0.05243	0.06224	0.79	0.4286	-0.0774	0.18226
Log Likelihood	-751.04433					
Inf. Cr. AIC	1512.1					
Chi-square	38.08438					
Significance level	0.000000					

Note: ***, **,* ==> Significant at 1%, 5% and 10% level

Similar to the non-linear EUT model, the sign of the specific constant for station one is positive and the parameters' coefficients, α and β, are significant at the 99 percent confidence level. Moreover, β is greater than 1, which means that respondents' attitudes towards crowding on a typical day in the EEUT model are also risk averse. Again, the curve of its value function is in the lower right quadrant (per Figure 6.4) and its shape is concave (see Figure 6.6). Additionally, the CRRA for the EEUT model with TK weighting function is -1.65367.

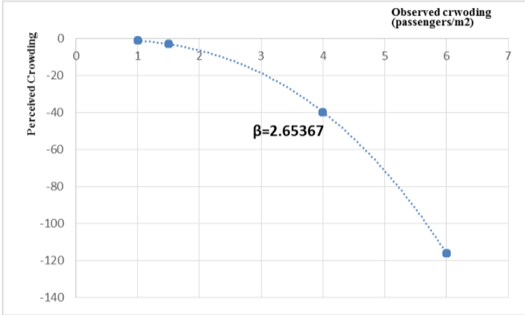


Figure 6.6 The value function in EEUT model

6.3.2 Comparison of the non-linear EUT model with the EEUT Model

Comparing the non-linear EUT model with the EEUT model with TK weighting function, we found their results were similar, including P&R users' responses to station choice and risk attitude towards crowding. Moreover, both models had a good fit with the collected data, (significance level is near zero). Nevertheless, the EEUT model was considered to be better than the non-linear EUT model for three main reasons. The first was that the coefficient of crowding on trains estimated under the non-linear EUT model was zero, which means that crowding has no influence on the station choice for P&R users. This is not aligned with what we surveyed and expected. Secondly, both the AIC and Log Likelihood statistical indices from the EEUT model are less than those from the non-linear EUT model (i.e. AIC: 1521.1 vs 3291.7 and Log Likelihood: │-751.0443│ vs │-1641.85502│), indicating that the EEUT model had a better fit with the collected data. Thirdly, the respondents' risk attitude measurements from the EEUT model were more reasonable.

Therefore, the following study will be carried out using the EEUT model with TK weighting function.

6.3.3 Comparison of the EEUT models

Three further popular weighting functions, GE, Pr1-I, Pr1-II, together with a value function in the power form, were also tested. The coefficients in these models were estimated within an MNL model, with the results presented in Table 6.3.

Table 6.3 Estimation of coefficients in the EEUT models

	EEUT model with TK weighting function	EEUT model with GE weighting function	EEUT model with Pr1-I weighting function	EEUT model with Pr1-II weighting function
Crowding	-0.27669	6.03511	-0.18290***	12.4712***
Discomfort	-0.03335***	-0.02117***	-0.02219***	-0.01590***

β	2.65367***	-0.19928	1.49265***	5.30915
γ	0.17229	0.24472	0.29140	-2.61314***
s		6.29036		0.30639D-04
Constant specific for station one	0.05243	0.07004	0.07101*	0.09837**
Log Likelihood	-751.04433	-1638.87008	-1638.81661	-1645.21150
Inf. Cr. AIC	1521.1	3289.7	3287.6	31.13828
Chi-square	38.08438	43.82113	43.92807	3302.4
Significance level	0.00000	0.0000	0.000000	0.00002

Note: ***, **,* ==> Significance at 1%, 5% and 10% level

The results in Table 6.3 show that the EEUT model with TK weighing function fits the collected data the best, based on all statistical indices, and specifically for AIC, which indicates how well the statistical model fits the observed data (Akaike, 1974). The AIC from the EEUT model with TK weighting function is 1521.1 which is much lower than its values for the other three EEUT models.

From Table 6.3, we also found that only the estimations from the models with TK and Pr1-I weighting functions were aligned with expectations. In detail, the effects of the attributes associated with crowding on station choice for P&R users in both models are negative, which implies that the more crowded a train is and the higher the level of discomfort, the lower the utility at the station. Moreover, the shape of value function is concave, (see Figure 6.7), and the respondents' attitude towards crowding on trains is risk averse, which are consistent with expected utility theory. Furthermore, the estimations of γ, (0.17229 and 0.29140 respectively), although not statistically significant, do indicate that they have an effect on station choice. Their common feature is that high probability outcomes are likely to underweight. However, the two models have some differences, one is the coefficients estimated by the model with TK weighting function are a little larger than those of the model with Pr1-I function. All indices from the model with TK weighting function are smaller than those from the model with Pre1-I weighting function, especially the AIC (TK 1521.1 vs Pre I 3287.6). This implies that the EEUT model with TK weighting function is better than the model with Pr1-I. Therefore, the EEUT model with TK weighting function was adopted as the preferred model to explain P&R users' choice of departure train station.

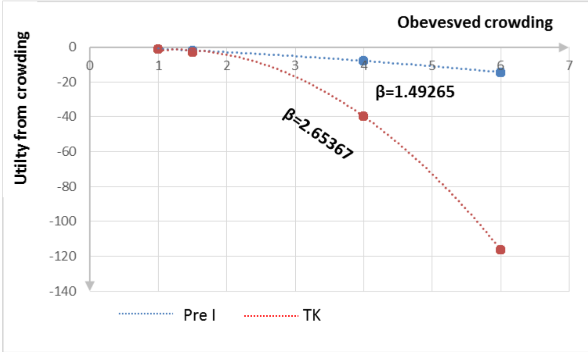


Figure 6.7 Value functions of both models

6.3.4 Analysis of respondents' risk attitudes

In this section, we apply the recommended model to measure respondents' risk attitudes towards crowding on trains on a typical week day by different train stations and explore the impact of risk attitude on station boardings. In detail, we re-estimated the coefficients in the recommended model using data for each train station separately and hence derived the risk attitude for each individual train station. From these, a relationship between the risk attitude parameter and station boardings was developed. The βs for each train station are summarised in Table 6.4. The respondents at all stations, except Warnbro and Warwick, showed risk aversion towards crowding on trains. Moreover, the respondents at Claremont station were most risk averse, while the respondents from Midland and Murdoch stations showed less risk aversion than others (see Figure 6.8). This might be a function of their locations. Midland station is located at the end of a train line and Warnbro station is the second to last station on the Mandurah line, (and with large spacing to adjacent stations), so that train users from these stations are more captive and therefore more tolerant of crowding.

Table 6.4 Coefficient estimation for seven train station with the recommended model

	Greenwood	Cannington	Claremont	Midland	Warwick	Warnbro	Murdoch
Crowding	-0.22580	-0.23691	-137.69	-0.83389	-0.05970	0.13158	-0.07884
Discomfort	- 0.04244***	- 0.03614***	-0.01658	-0.04853	- 0.02630***	-0.00658	-0.00954
β	2.64657**	2.59025*	3.95820	1.39738	0.13191	-0.62075	1.78158
γ	0.2163	0.20395	0.04882	0.68065	0.13327	0.52732	0.39311
Constant specific for station one	-0.26022**	0.39772***	0.06668	0.22340	0.21237**	-0.16785	0.13672
Log Likelihood	-219.86746	-223.80053	- 206.40577	- 74.36110	-249.37362	- 193.65504	- 343.99830
Inf. Cr. AIC	449.7	457.6	422.8	158.7	708.7	397.3	698.0
Chi-square	20.51481	34.82937	3.07678	6.54278	16.58065	11.94269	5.15059
Significance level	0.001	0.0000	0.68815	0.25692	0.00537	0.03558	0.39778

number of observation	332	348	300	112	516	288	500
CRRA	-1.64657	-1.59025	-2.95820	-0.39738			-0.78158

Note: ***, **,* ==> Significance at 1%, 5% and 10% level

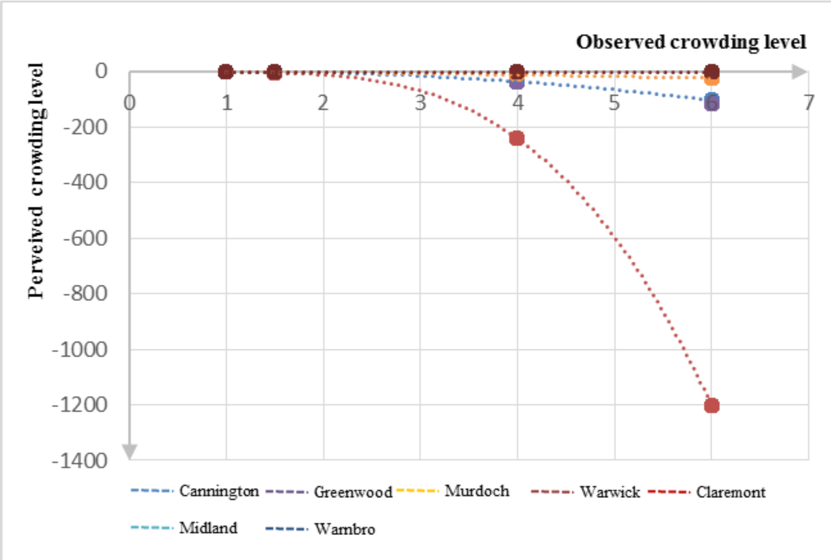


Figure 6.8 Risk attitude for respondents at different train stations

Additionally, we explored the relationship between crowding levels and the scale of respondents' risk attitude at the same train station. In order to test the relationship, we set up a hypothesis: the greater the level of crowding at the train station, the lower the aversion to risk exhibited by passengers at that station. To test the hypothesis, we compared train boardings by station on an average weekday with the β estimated separately using the data from four train stations (see Table 6.5). The results show that, the smaller the β' value for a station, the more train boardings at that station. Assuming that train boardings in the peak period as a proportion of daily boardings is broadly the same at the four stations, and the more train boardings at a station the more crowded trains stopping at the station are, the relationship shown in Table 6-5 implied that the lower risk-aversed passengers using a station are, the more crowded trains stopping at that station would be. Taking Claremont station as an example, it's β is the largest and train boardings is least among the four stations. This means the passengers at Claremont station maybe have the highest risk aversion to crowding, and that the trains at Claremont station are the least crowded, which is aligned with the hypothesis set up above.

Table 6.5 Relationship between risk attitude and rail ridership

	Claremont	Greenwood	Cannington	Murdoch
β	3.95820	2.64657	2.59025	1.78158
Average train boardings (on weekday)	1431	1921	2201	6369
Train frequency during peak hours	5mins	5mins	5mins	5mins

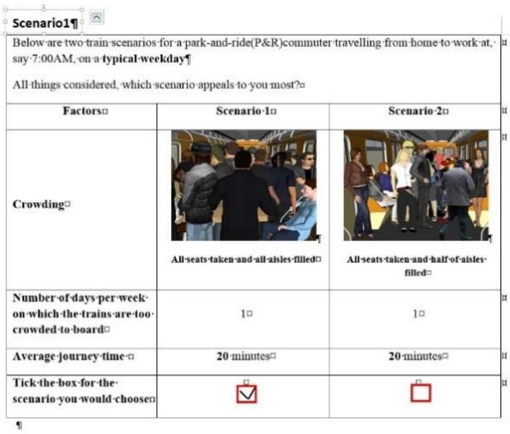
6.3.5 Sensitivity analysis for policy implication

A sensitivity test was conducted to identify the effect of crowding on trains on the probability that a train station was chosen, by holding other attributes at their mean. We set up two scenarios, each of which included two train stations, (see Figure 6.9). The only difference between the two scenarios was the crowding attributes level for station one. We set the density of standing passengers and the probability that seats had been taken for station one in scenario one as 6 passengers/m^2 and 100% respectively, with 2 passengers/m^2 and 0.75 for scenario two (see Table 6.6). This design aimed to test whether the probability that a station was chosen would change as the crowding level varied, i.e. to test our hypothesis that the greater the level of crowding on the train, the lower the probability that the station would be chosen by P&R users.

The crowding level for station one is 6 passengers/m^2 in scenario one and 1.5 passengers/m^2 in scenario two, which resulted in probabilities of choosing station one of 26% and 54% respectively. A comparison of the two groups of data revealed that the station choice probability was negatively correlated with crowding on trains and four times less crowding on trains led to a doubling of the probability of the station being chosen. Based on this, we can conclude that reducing crowding on trains would be an efficient way to increase a station’s attractiveness for P&R users.

Table 6.6 Relationship between crowding level and probability that a station is chosen

	Density of standing passengers in a carriage (passengers/m^2)	Probability that seats taken	Crowding level	Probability of choosing station 1
Scenario1	6	1	6 passengers/ m^2	26.06%
Scenario2	2	0.75	1.5 passengers/m^2	54.16%
Ratio of scenario 1 to scenario 2			4	2

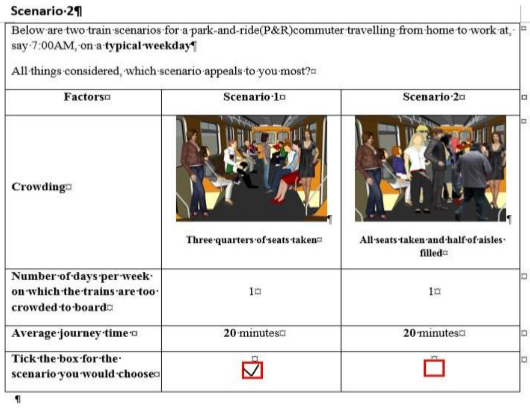


Scenario1

Below are two train scenarios for a park-and-ride(P&R)commuter travelling from home to work at, say 7:00AM, on a **typical weekday**

All things considered, which scenario appeals to you most?

Factors	Scenario 1	Scenario 2
Crowding	All seats taken and all aisles filled	All seats taken and half of aisles filled
Number of days per week on which the trains are too crowded to board	1	1
Average journey time	20 minutes	20 minutes
Tick the box for the scenario you would choose	☑	☐



Scenario 2

Below are two train scenarios for a park-and-ride(P&R)commuter travelling from home to work at, say 7:00AM, on a **typical weekday**

All things considered, which scenario appeals to you most?

Factors	Scenario 1	Scenario 2
Crowding	Three quarters of seats taken	All seats taken and half of aisles filled
Number of days per week on which the trains are too crowded to board	1	1
Average journey time	20 minutes	20 minutes
Tick the box for the scenario you would choose	☑	☐

Figure 6.9 The hypothetical scenarios

6.3.6 P&R users' preference heterogeneity led by individual personalities

In the section, we test the effect of personal heterogeneity, (in this case annual income), on an individual's choice of departure train station. In contrast to previous studies, annual income was not taken as an independent variable and simply introduced into the choice model. We took it as a latent variable and assumed that its effect on station choice would be indicated by differences in the responses to crowding on the train. Based on this, a latent class model was set up. After testing 2, 3, 4 and 5 classes, four latent classes were identified as giving the best fit. The resulting coefficient estimates are shown in Table 6.7.

Table 6.7 Estimated LCM: Utilities

	Multinomial logit model (MNL)	Class utility model			
		Class one	Class two	Class three	Class four
Crowding	-0.01341	8.19232	-13.1238	5.90785	-0.49485
Discomfort	-0.1310***	-2.33463	-0.45409	2.48048	-0.17198
Constant specific for station one	0.11392	-5.94053	-7.18746	19.9449	-0.98766
Class probability model					
Constant		-0.06271	-7.0121	-0.0265	0.0
Annual Income		-3.7879*	-0.40290	-0.03812	0.0
Log Likelihood	-1648.25247	-1622.73492			
Inf. Cr. AIC	3302.5	3281.5			

Note: ***, **,* ==> Significance at 1%, 5% and 10% level

The coefficient estimates for the four classes are completely different, which indicates that P&R users with different income levels have heterogonous preferences for each crowding attribute and, therefore, may make different choices. The results are aligned with our hypothesis. Additionally, we compared the probability of choosing a different station for the four classes by applying the LC model into the left scenario in Figure

6.9. We found that respondents in class one preferred station one with a higher level of crowding, while respondents in the other three classes tended to choose station 2 (see Table 6.8). The differences in choice probabilities amongst the different classes also indicate that an individual's heterogeneity would affect their choice, which further confirmed our hypothesis.

Table 6.8 Individuals' preference heterogeneity

		Class one	Class two	Class three	Class four
Chosen probability	Station 1	99.92%	0.00%	0.00%	14.44%
	Station2	0.08%	100.00%	100.00%	85.56%
Prior class probability		18.5%	9.5%	33.9%	38.2%
Posterior class probability		21.27%	9.63%	32.69%	36.41%
Annual income		1.186701	1.15487	2.018626	2.18566

We also analysed the characteristics of each latent class through the estimation of conditional class probability. The results showed that the estimated conditional class probabilities, (i.e. posterior class probability), are similar to the prior probabilities, which implies that the model is valid. Moreover, the results revealed that the individuals in latent class four were wealthier than those in the other three latent classes and had a greater preference for station 2, which had a lower level of crowding. The individuals in latent class one were poorer than other latent classes, (except class two), and preferred station 1.

Given the effect of individual preference heterogeneity, it is suggested that public transport providers consider the heterogeneity of passengers' personalities when developing their services, in order to attract more passengers and increase rail ridership. For example, pricing tickets differently based on crowding levels on trains could satisfy the different demands of passengers with different annual incomes.

6. 4 Model validation

According to Miller, Hui, and Tierney (1991), models developed for prediction need to be properly validated to reassure users of their output that they adequately perform the functions for which they are intended. The station choice model developed in this section is part of the demand model for prediction of rail ridership and therefore needs to be validated.

A Chi-square test was used to validate the station choice model developed in this chapter. Firstly, we predicated the probabilities with the model for the 12 scenarios used in the survey (see Table 6.9), then compared them with their observed outcomes, and got a Chi-square value of 0.98. Assuming a significance level of 0.05 and 11 degrees of freedom, the critical value is 19.675. Given that the critical value is greater than the Chi-square, the difference between prediction and observation is not statistically significant.

Table 6.9 Comparison of the observed and predicted probabilities that a station is chosen

Scenarios	Observation		Prediction	
	Station 1	Station 2	Station 1	Station 2
1	57%	43%	68%	32%
2	43.2%	56.9%	57.5%	42.5%
3	55.3%	44.7%	70.9%	29.1%
4	35.7%	64.3%	43%	57%
5	57.2%	42.8%	29.7%	70.3%
6	57.2%	42.8%	69.2%	30.8%
7	52.2%	47.8%	35.9%	64.1%
8	42.3%	57.7%	45.1%	54.9%
9	56.8%	43.2%	66.4%	33.6%
10	55.8%	44.2%	27.1%	72.9%
11	63.5%	36.5%	29.4%	70.6%
12	61.1%	49.9%	75.6%	24.4%

6. 5 Conclusions and limitations

Crowding is likely to threaten the health and safety of rail passengers. Therefore, it is very important to quantify its effect on passengers so that its adverse impacts can be mitigated. The sub-model of station choice incorporating the effect of variations in crowding levels is a multinomial logit model in which crowding and comfort on trains were taken as the main components of utility related to station choice. The effect of variations in crowding was evaluated within EEUT, in which the crowding measures were assumed as the interaction between the probability that all seats were taken and the density of standees. The value function and weighting function adopted, respectively, the power form and the TK form. Based on the sub-model, we can conclude that respondents' attitudes towards crowding were risk averse and the more crowded the train, and the longer the time that seats have been occupied by others, the lower the utility of that station.

Furthermore, the model was applied to the seven surveyed train stations, the results revealing that the greater the risk aversion displayed by the respondents, the lower the number of individuals boarding at the station.

Additionally, we found that the effect of crowding on respondents' train station preference was heterogeneous with respect to annual income. Individuals with higher incomes would be more likely to choose a station that had less crowding on trains.

This implies that commuters with higher incomes value a comfortable travel environment more than those with lower incomes, and would presumably be prepared to pay more to have it.

The sub-model still has a number of limitations. The first is that we adopted the interaction of both classic crowding measures as a new crowding measure in the research, given that the crowding is presented at eye level rather from above. However, we are not sure whether this is the best measurement to explore the effect of crowding on station choice, especially for the scenario in which the density of standees is low. Another limitation is that the sample size at some stations was insufficient and the scenarios for some stations did not meet the required number of choice tasks, (12), so that not all the parameter estimates displayed statistical significance. Therefore, further work would be to review the factors affecting station choice and test their effect on station choice, with sufficient data for each station.

The third limitation of the research is the data used in the paper were collected with other attributes in the same questionnaire. Therefore, we are not sure that the attributes mentioned above are the most important factors affecting P&R users' choice of access train station. To test this, we used eye-tracking equipment to monitor respondents' eyes, then identified the significance of each factor for station choice based on their visual attention. We assumed that the more attention participants paid to a variable, the greater the influence that variable had on the choice decision. Figure 6.10 shows a snapshot of a visual attention map (heatmap) of a participant. For the same participant, we displayed two different scenarios. For the questionnaire on the left the area with the crowding picture was found to be the centre of attention, i.e. had the longest visual attention.. However, when we increased the number of factors in the survey, attention shifted to other factors instead of crowding. Therefore, the best way to design the SC experiment and balance the number, order and levels of variables to explore choice behaviour in an efficient way is still open to debate and more evidence is needed.

Figure 6.10 Heatmap analysis for both questionnaires (different choice)

6.6 Chapter summary

This chapter reviewed the studies related to the effect of crowding on the public transport system, re-designed the pictorial display of crowding in the SP experiment and created a new crowding measure based on it. Given that the new measurement of crowding can vary randomly, the station choice based on the effect of crowding is defined as a choice behaviour under risk. Correspondingly, both EUT and EEUT were applied to explore the behaviour. Finally, the linear EUT model, the non-linear EUT model and the non-linear EEUT models, with four popular probability weighting functions, were developed and the non-linear EEUT model with power form value function and TK form weighting function was found to be the preferred model based on the results of the statistical analysis. The sub-model was estimated within MNL with Nlogit 5. Based on the model, respondents' risk was measured and the relationship between the respondents' risk attitude and boarding numbers at the same station was identified, which provides useful information for public transport operators to improve service quality.

The next chapter will analyse the effect of parking search time uncertainty on station choice for P&R users. The methods to develop the sub-model of station choice focusing on the effect of variation of parking search time will be discussed.

CHAPTER 7 STATION CHOICE SUB-MODEL WITH PARKING SEARCH TIME UNCERTAINTY

The previous chapter explored the effect of crowding on trains on P&R users' choice of departure train station, analysed their risk attitude towards variations in the level of crowding and proved individuals' preference heterogeneity for station choice. This chapter focuses on investigating the effect of parking attributes, in particular parking search time (PST), on P&R users' choices and estimating P&R users' risk attitude towards variations in PST. This provides a basis to improve station choice models under uncertainty and could assist public transport operators in improving the service quality of their P&R facilities.

It is worth noting that most of the chapter is from the paper written by Chen et al. (2015). This paper has been published and my contribution for the paper is about 80%.

7. 1 Research context

Parking search time (PST) is defined as the time spent searching for a parking bay after arriving at a train station and can be considered as a factor in choosing the departure train station when P&R demand is close to or above the capacity of the P&R facility. According to the Public Transport Authority (2012-2017), the total capacity of all the P&R facilities serving the Perth train system is about 21,000, (in 2016), which is well below the level required to cater for the latent demand estimated to be around 23,000. Therefore, the PST, from a theoretical viewpoint at least, could be a key factor in station choice for P&R users in Perth. Moreover, the train station choice survey conducted during July 2012 (see Chapter 3) also indicated that the PST was, in practice, one of the main factors influencing a P&R user's choice of departure train station. PST can vary significantly for users who arrive at the time when parking bays are full or close to full as they search for one of the few remaining spaces. If none is available they may then need to search for parking on the surrounding streets or at an alternative car park. To date, few studies have been undertaken to develop an understanding of the effect of variations in the PST on P&R users and their choice of departure train station.

Most of the literature related to the effect of PST on travel choice has focused on investigating its influence on route choice (e.g., Leurent and Boujnah (2012), Balijepalli, Shepherd, and Kant (2013), etc.) or parking type and location choice (e.g., Polak and Axhausen (1990), Hilvert, Toledo, and Bekhor (2012), etc.). Only very limited literature was found relating specifically to the effect of PST on station choice. In summary, the parking attributes included in previous station choice models were primarily parking charge (Miller & Cheah, 1991), parking capacity (Davidson & Yang, 1997; Fan et al., 1993; Vijayakumar et al., 2011), parking availability (Debrezion et al., 2007; Mahmoud et al., 2014; Vijayakumar et al., 2011) and parking cost (Mahmoud et al., 2014). Additionally, Kastrenakes (1988) initially took parking availability and parking fee into account, but then removed them from the model due to their counter-intuitive signs.

Therefore, it is considered to be time to study the effect of PST on station choice. This chapter explores the effect of parking attributes on departure train station choice for P&R users by testing the effect of variations in PST on station choice and measuring respondents' risk attitude towards these variations.

7. 2 Research method

Chapter 3 outlined the general methods used to develop the station choice models. This chapter focuses on the development of the utility function related to station choice under PST uncertainty. In other words, the effect of variations in the PST on station choice and the P&R users' risk attitude towards it are investigated.

7.2.1 Framework

On the assumption that parking availability is a constraint in station choice, a hybrid function has been used to measure the utility related to station choice from parking attributes. When parking spaces are available, (i.e. the P&R car park is not full), parking capacity, parking availability, access time to the station, and parking fee are considered to be the main components contributing to station choice utility. When parking is not available, (i.e. the P&R car park is full), the usual PST, the variability of the PST, the parking fine and the frequency of patrols for illegal parking were taken as the main components. The effect of the variability of the PST was calculated within the CPT and the coefficients in the utility function related to station choice under PST uncertainty were estimated with a mixed logit model, given that it can capture individuals' preference heterogeneity (Train, 2003). Based on this, the framework to model station choice under the PST uncertainty is given in Figure 7.1.

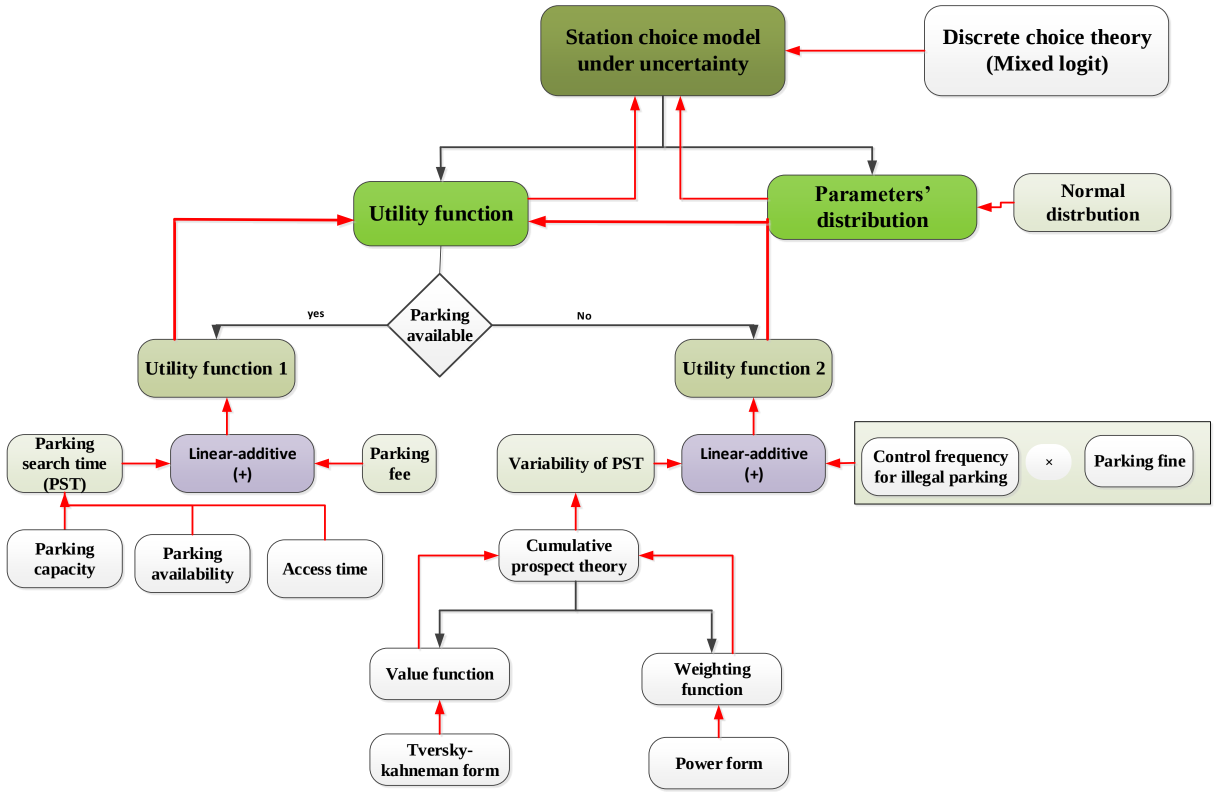


Figure 7.1 Framework to model station choice under PST uncertainty

7.2.2 Data used in the chapter

Per Chapters 5 and 6, the data used in this chapter are from the train station choice survey, which was described in Chapter four. The data used in the chapter are part of station choice data but focus on that related to parking attributes, (see Figure 7.2).

Example

Below are two train scenarios for a park-and-ride(P&R)commuter travelling from home to work at, say 7:00AM, on a **typical weekday** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Travel time from home to the station		
Usual travel time	10mins	5mins
Travel time on a good day	About 5 mins on 4 days a month	About 3.75 mins on 4 days a month
Travel time on a bad day	About11.5 mins once a month	About 6.2mins once a month
Time on the train		
Average journey time	10 minutes	10 minutes
Finding parking at the station		
Parking capacity in P&R lot	100bays	100bays
Parking availability in a P&R lot at 7:00AM	20%	20%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **1** and **8** minutes and only **5%**chance to reach to **8** mins	Between **1** and **8** minutes and only **20%**chance to reach to **8** mins
Parking cost		
Parking fee in a P&R lot	$4/day	$6/day
Parking fine (if you park illegally, you would be fined)	$40	$80
Frequency of controls for illegal parking	Once a **day**	Once a **month**
Crowding on the train		
Crowding	 All seats taken and half of aisles filled	 All seats taken and half of aisles filled
Number of days per week on which the trains are too crowded to board	2	1
Other things		
Safety	Regular station and car park security patrols	Irregular station and car park security patrols
Ticket fare	$2.50	$2.50
Train frequency	All-stop trains at 10 minute internals	Express and all-stop trains at 5 minute intervals
Tick the box for the scenario you would choose	☐	☐

Figure 7.2 The data source

7.2.3 Method to develop the utility function

According to the framework above, when parking is available in the P&R facility, the number of parking bays remaining at the time of access time and the parking fees represent the main components of utility; when parking is not available, (i.e. the car park is full), the usual PST and its variation, the parking fines for illegal parking and the parking violation patrol frequency are the determinants of utility considered in the station model. Thus, the observed part of the utility function in the mixed logit model is a hybrid function with two parts. Its specification is shown in equation (7-1).

$$V_i = \begin{cases} \beta_1 \times f\left(N_i^{pbay}\right) + \beta_2 \times fee_i^p & pa > 0 \\ \beta_3 \times V\left(VPST_i\right) + \beta_4 \times fine_i^p \times fre_i^{con} & Pa = 0 \end{cases} \tag{7-1}$$

where fee_i^p is the parking fee at station i; $fine_i^p$ is the fine for illegal parking around station i; fre_i^{con} is the frequency of patrols for illegal parking around station i; and N_i^{pbay} is the number of unoccupied parking bays at station i. The latter depends on the parking capacity (pc), parking availability (pa) and the time of access to the chosen station (at). Their relationship can be written as equation (7-2).

$$N_i^{pbay} = \left(pc \times pa\right) / g\left(at\right) \tag{7-2}$$

The demand for parking bays increases during the morning with most car parks being fully occupied within the AM peak. In other words, the later the time of arrival at the train station, the fewer available parking spaces and the greater the competition between P&R users for those remaining bays. Here, we took 7:00am as the time reference point and assumed the relationship between the variables to be per equation (7-3).

$$N_i^{pbay} = \left(pa_i \times pc_i\right)^{\left(at_{7:00}/at_k\right)} \tag{7-3}$$

where $at_{7:00}$ means the time of access to the chosen station is 7:00am; and at_k a time of access of k am.

In order to accurately identify the effect of remaining parking bays on station choice, we tested three forms of the utility function, namely, linear, power, and exponential. The linear form is preferred based on the performance of the model (see Table 7.1). Its formula is shown in equation (7-4).

$$V_i = \beta_1 N_i^{pbay} + \beta_2 fee_i^p \tag{7-4}$$

where β_1, β_2 are estimated coefficients.

In order to avoid overly large numbers during calculation, we divided N_i^{pbay} by 100 to decrease its magnitude.

Table 7.1 Statistical results for three specification forms of the utility function

	Linear	Power	Exponential
	$\beta_1 N_i^{pbay} + \beta_2 fee_i^p$	$\beta_1 \left(N_i^{pbay}\right)^{\alpha_3} + \beta_2 fee_i^p$	$\beta_1 e^{N_i^{pbay}} + \beta_2 fee_i^p$
Log likelihood	-817.89	-827	-885.76
Inf. Cr. AIC	1645.8	1663	1781.5

$VPST_i$ in equation (7-1) is the variability of PST at station i, which is the difference between the worst PST and the usual PST; $V(VPST_i)$ is a utility function of the variations in PST, developed under the CPT. The general function form of utility within CPT includes the value function $v\left(\Delta x\right)$ and the weighting function $\pi\left(p_i\right)$ as shown in equation (7-5).

$$V(x_i, p_i) = \sum_{i=1}^{n} \pi(p_i) v(\Delta x_i) \tag{7-5}$$

The value function in CPT is defined separately over gains and losses. Here, Δx is used to indicate the difference between the real value and a reference point with Δx greater than or equal to zero indicating a gain, and Δx less than zero a loss. p_i is the probability that the i^{th} outcome occurs, $\pi\left(p_i\right)$ is the subjective weighting function derived from the i^{th} outcome cumulative probability and $v\left(\Delta x_i\right)$ is a value function depending on gains or losses, with a power form. Their specifications are shown in equations (7-6) - (7-8).

$$v\left(\Delta x_i\right) = \begin{cases} x^{a} & if \ \ x \geq 0 \\ -\lambda\left(-x\right)^{\beta} & if \ \ x < 0 \end{cases} \tag{7-6}$$

$$\pi^{+}\left(p_i\right) = w^{+}\left(p_i + \cdots + p_n\right) - w^{+}\left(p_{i+1} + \cdots + p_n\right) \quad 0 \leq i \leq n \tag{7-7}$$

$$\pi^{-}\left(p_j\right) = w^{-}\left(p_{-m} + \cdots + p_{-j}\right) - w^{-}\left(p_m + \cdots + p_{j-1}\right) \quad m \leq -j < 0 \tag{7-8}$$

where α, β are the respondents' risk attitude towards gains and losses; $\pi^{+}\left(p_i\right)$ is cumulative probability that the i^{th} gain outcome occurs; $\pi^{-}\left(p_i\right)$ is cumulative probability that the j^{th} loss outcome occurs; and n, m are the number of gain outcomes and loss outcomes respectively.

Assuming that the usual PST is the reference point, gains refer to the difference between the best (i.e. shortest) and the usual PST, and losses are the difference between the worst (i.e. longest) and the usual PST. Because the data from the pilot and main surveys indicated very low gains, their influence has been disregarded as negligible in this research. According to Stott (2006), the weighting function with the Tversky-Kahneman (TK) form, together with the power form of the value function, can better model choice behaviour under the CPT. Avineri and Prashker (2004) also successfully applied them in their studies of route choice. Therefore, the weighting function here also adopted the TK form. Its specification is shown in equation (7-9).

$$w^{-}\left(p_i\right) = p_i^{\delta} \Big/ \left[p_i^{\delta} + (1 - p_i)^{\delta}\right]^{\frac{1}{\delta}} \tag{7-9}$$

where $w^{-}\left(p_i\right)$ is the weighting function for the probability that the i^{th} loss or gain happens; and δ is the estimated coefficient.

7. 3 Utility function related to station choice incorporating variations in the PST

7.3.1 Utility function incorporating variations in PST

Based on equations (7-6) and (7-9), the value function and weighting function of the variations in PST can be written as equations (7-10) and (7-11).

$$v\left(VPST_i\right) = \lambda\left(-VPST_i\right)^{\beta} = \lambda\left(wpst_i - upst_i\right)^{\beta} \tag{7-10}$$

$$w^{-}\left(wpstf_i\right) = p_i^{\delta} \Big/ \left[wpstf_i^{\delta} + \left(1 - wpstf_i\right)^{\delta}\right]^{\frac{1}{\delta}} \tag{7-11}$$

where $wpst$ is the worst PST; $upst$ is the usual PST; λ, β are estimated coefficients with β indicating P&R users' risk attitude towards variations in the PST; and $wpstf_i$ is the frequency at which the worst PST occurs in one month at station i.

Substituting equations (7-10) and (7-11) into equation (7-5), the utility function for the variations in PST is shown as equation (7-12).

$$V(VPST_i) = \pi\left(p_i\right)v\left(\triangle x_i\right) = \lambda\left(wpst_i - rpst_i\right)^{\beta} \times wpstf_i^{\delta} \Big/ \left[wpstf_i^{\delta} + (1 - wpstf_i)^{\delta}\right]^{\frac{1}{\delta}} \tag{7-12}$$

7.3.2 Utility function related to station choice

Substituting equations (7-4) and (7-12) into equation (7-1), the observed part of the utility function related to station choice can be written as equation (7-13).

$$V_i = \begin{cases} \beta_1 \times N_i^{pbay} + \beta_2 \times fee_i^{p} & pa > 0 \\ \beta_3(wpst_i - upst_i)^{\beta} \times P_i^{\delta} / [P_i^{\delta} + (1 - P_i)^{\delta}]^{(1/\delta)} + \beta_4 \times (fine_i^{p} \times fre_i^{p}) & pa = 0 \end{cases} \tag{7-13}$$

Assuming that respondents park their cars in the P&R area when spaces are available, the overall utility function can be written as equation (7-14).

$$U_i = V_{pa,i} + V_{npa,i} + \varepsilon_i = \beta_1 \times N^{pbay} + \beta_2 \times fee_p + \beta_3(wpst_i - upst_i)^{\beta} \times wpstf_i^{\delta} / [wpstf_i^{\delta} + (1 - wpstf_i)^{\delta}]^{(1/\delta)} + \beta_4 \times (fine_i^{p} \times fre_i^{p}) + \varepsilon_i \tag{7-14}$$

where $V_{pa,i}$ is the observed utility at station i when pa_i is greater than zero; $V_{npa,i}$ is the observed utility when pa_i equals zero; and ε_i is the unobserved utility at station i.

7. 4 Results

7.4.1 Estimation of coefficients

The coefficients of the station choice model were estimated using multinomial logit and mixed logit models with the Nlogit 5 software (Hensher, Rose, & Greene, 2005). In the mixed logit model, coefficients β_1, β_2 are assumed to be random and to follow the normal distribution. The results are summarised in Tables 7.2 and 7.3.

Table 7.2 Multinomial logit model with linear utility function

Choice	Coefficient	Standard error	z	Prob. $z > \|Z\|$	95% Confidence Interval	
Unoccupied parking bays	1486***	0.00534	********	0.0000	1486.85	1486.87
Parking fee	-1.04886***	0.05379	-19.50	0.0000	-1.15429	-0.94342
Parking search time	-0.020488***	0.00548	-3.74	0.0002	-0.03123	-0.00974
Parking fine	-0.00279	0.00548	-1.05	0.2927	-0.008	0.00241
Constant specific for station one	-0.16444**	0.07789	-2.11	0.0348	-0.3170	-0.01177
Inf. Cr. AIC			1873.6			
Log likelihood			-931.79661			
Chi-square			86.83265			

Note: ***, **, * = >Significance at the 1%, 5%, 10% confidence levels

Table 7.2 indicates that the number of parking bays in the P&R facilities, the parking fee and variability of the PST all have a significant effect on station choice for P&R users. A high availability of parking bays, low parking fee and lower variability of PST at a station increase the probability of that station being chosen.

Table 7.3 Mixed logit model with random parameters following normal distribution

Choice	Coefficient	Standard error	z	Prob. $z > \|Z\|$	95% Confidence Interval	
			Random parameters in utility functions			
Unoccupied parking bays	91.0325	691.5102	0.13	0.8953	-1264.3027	1446.3676
Parking search time	-0.1165D-04	0.000091	-0.01	0.9898	-0.17989D-02	0.177566D-02
			Non-random parameters in utility functions			
Parking fee	-0.15764***	0.05926	-2.66	0.0078	-0.27380	-0.04148
Parking fine	-0.00503	0.00371	-1.35	0.1757	-0.0123	0.00225
β	6.51412	6.33082	1.03	0.3035	-5.98406	18.92229
δ	0.30832	2.35560	0.13	0.8959	-4.30858	4.92521
			Distns. ofRPs. Std. Devs or limits of normal			
$Ns\beta_1$	10.2233	769.2380	0.01	0.9894	-1497.4554	1517.9020
$Ns\beta_3$	0.56768D-04	0.00444	0.01	0.9898	-0.86746D-02	0.87611D-02
Inf. Cr. AIC			1644.4			
Log likelihood			-814.4886			

Note: ***, **, * = >Significance at 1%, 5%, 10% confidence levels

The coefficients in Table 7.3 were estimated for a mixed logit model with β_1, β_2 following the normal distribution. Even though the values of the parameter coefficients are different to those in Table 7.2, their signs remains the same. The model has better goodness-of-fit results, so it is recommended in the research, even though they suggest that the parameters' coefficients may not be random or normal random.

7.4.2 Respondents' risk attitude

As for extended expected utility theory, the coefficient over the gains or losses value can indicate respondents' risk attitude for gains or losses. In the recommended model, parameter β can show the respondents' risk attitude for larger variations in PST. In the model, the estimate of β is 6.5. Although it is not statistically significant, it does have an effect on the shape of the value function can be seen in Figure 7.3, in which zero is taken as the reference point and the losses are calculated as differences between the worst PST and the usual PST. The shape of the value function is concave for losses but the figure can still show that: ① the higher the loss, the lower the value function; and ② the risk neutral value, ($\beta = 1$), is greater than the value for losses. Therefore, the respondents are risk averse based on the data used in the research.

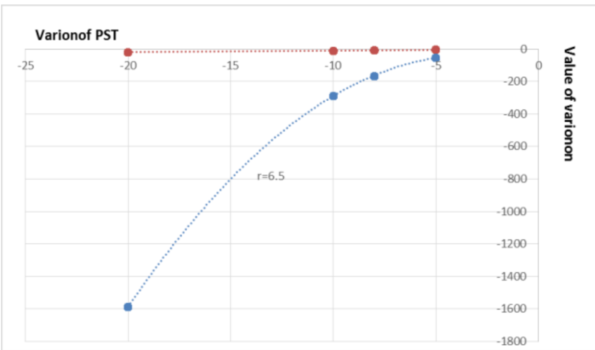


Figure 7.3 Non-linear value function

Furthermore, the estimate of δ is 0.30832. As for β, it is also not statistically significant, but it does have an impact on the shape of the risk weighting function (see Figure 7.4). The results suggest that the outcomes with low probabilities tend to be slightly overweighed and the outcomes with high probabilities tend to be underweighted by respondents.

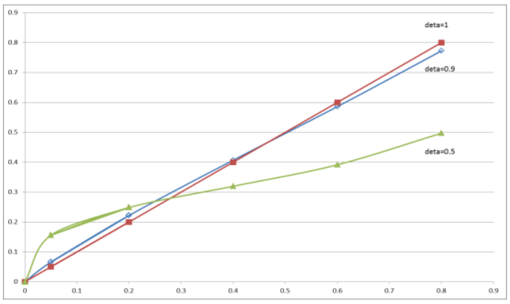


Figure 7.4 Non-linear risk weighting function

7. 5 Conclusions

In this chapter P&R users' station choice was analysed using multinomial and mixed logit models. According to our knowledge, this is the first attempt to understand P&R users' station choice under uncertain PST using a combination of cumulative prospect theory and discrete choice theory.

In the chapter, the utility function was established separately for two situations, (i.e. parking available and not available in the P&R areas). When parking is available, a linear function was used to capture the effect of the number of available parking bays remaining in the P&R facilities at a given access time on station choice for P&R users. The mixed logit model with parameters of normal distribution was found to be better fit model, although the results suggest non-random parameters for parking availability and parking fee. When parking is not available in the P&R facilities, variations in PST, parking fine and the frequency of patrolling for illegal parking were considered in the model. In order to capture the effect of variations in PST, part of utility function was developed within CPT and the coefficients were estimated. The results showed that the effect of variations in PST is not significant but that respondents may display risk aversion for variations in PST and very weak non-linearity in the risk weighting function. Larger variations in PST, higher parking fines and more frequent patrols for illegal parking, lead to lower utility functions and smaller probabilities that the station is chosen. The results of our study could provide useful insights for implementation of public transport policies (such as ticket pricing policy).

7. 6 Chapter summary

This chapter developed a sub-model of station choice focusing on the effect of variations in PST by combining discrete choice theory, cumulative prospect theory and the mean-variance approach. It measured P&R users' risk attitude towards variations in PST based on the sub-model.

The next chapter tests the other factors in the train station choice survey and, together with the three sub-models developed in the previous chapters, develops an overall station choice under uncertainty. Additionally, the analysis of individuals' preference heterogeneity on station choice is included.

CHAPTER 8 OVERALL MODEL OF STATON CHOICE UNDER UNCERTAINTY FOR PARK AND RIDE USERS

The last three chapters respectively explored the effects of variations in travel time to station, parking search time and crowding on trains on station choice for P&R users, and developed three different sub-models of station choice under uncertainty. This chapter develops and evaluates an overall model to understand how these three factors combined affect P&R users' choice of departure train station and the influence of individuals' preference heterogeneity on their choice. Using this information, transport planners and transit operators would be able to make better policy decisions, (such as location and size), to improve the overall efficiency, operation and effectiveness of the P&R facilities.

8. 1 Research context

Given that the three models mentioned in Chapters 5 to 7 addressed the effects of the factors related to one uncertain situation individually, it is worthwhile investigating the combined impact of these three factors on a P&R user's choice of departure train station. Therefore, we developed an overall station choice model by combining these three models into a single model. Our review of current literature indicated that research relating to station choice is very limited and also only considered the effects of a few factors. For example, the station choice model developed by Kastrenakes (1988) includes four factors: whether the station is located in residential areas, access time to the chosen station, train frequency at the station and generalised cost. Fan et al. (1993), Davidson and Yang (1997), Wardman and Whelan (1999), Davidson and Yang (1999), Lythgoe and Wardman (2004), Lythgoe et al. (2004) Fox et al. (2011) and Givoni and Rietveld (2014) modelled access mode and station choice, and included more attributes, such as access mode, parking capacity, accessibility, railway network, etc. In general, the attributes related to parking, crowding on trains, safety, train frequency, etc. were barely covered in the previous literature. Therefore, this study establishes an overall model that considers a broader range of factors. More importantly, we can identify the degree of significance of each factor on a P&R user's station choice, which could provide public transport operators with advice on how, (and where), best to improve the rail service, especially where budgets are limited.

Moreover, we also identified only limited research into station choice under uncertainty in the previous studies. However, a consensus has been reached in the literature that variations in the factors do affect people's choice (Cascetta, 1989; Chang & Mahmassani, 1988; Iida, Akiyama, & Uchida, 1992). In the previous three chapters, we also identified the impacts of variations in travel time to the departure train station, parking search time and crowding on trains on the choice of departure train station, and therefore the overall station choice model considers both certain and uncertain factors.

8. 2 Research methodology

8.2.1 The method to develop the overall models

The two key objectives in developing the overall model are to identify the degree of effect of each factor on a P&R user's choice for departure train station and to explore the effect of individuals' preference heterogeneity on station choice. In order to achieve both objectives, we needed to develop a mixed logit (ML) model of station choice. The theory related to ML models can be seen Chapter 3. The tasks undertaken to develop the ML model were:

- Identifying the factors to be used in the ML model;
- Selecting the random coefficients (or parameters);
- Specifying the distribution of the random coefficients; and
- Estimating the coefficients in the model.

The detailed process to develop the overall model of station choice under uncertainty is shown in Figure 8.1 and discussed below.

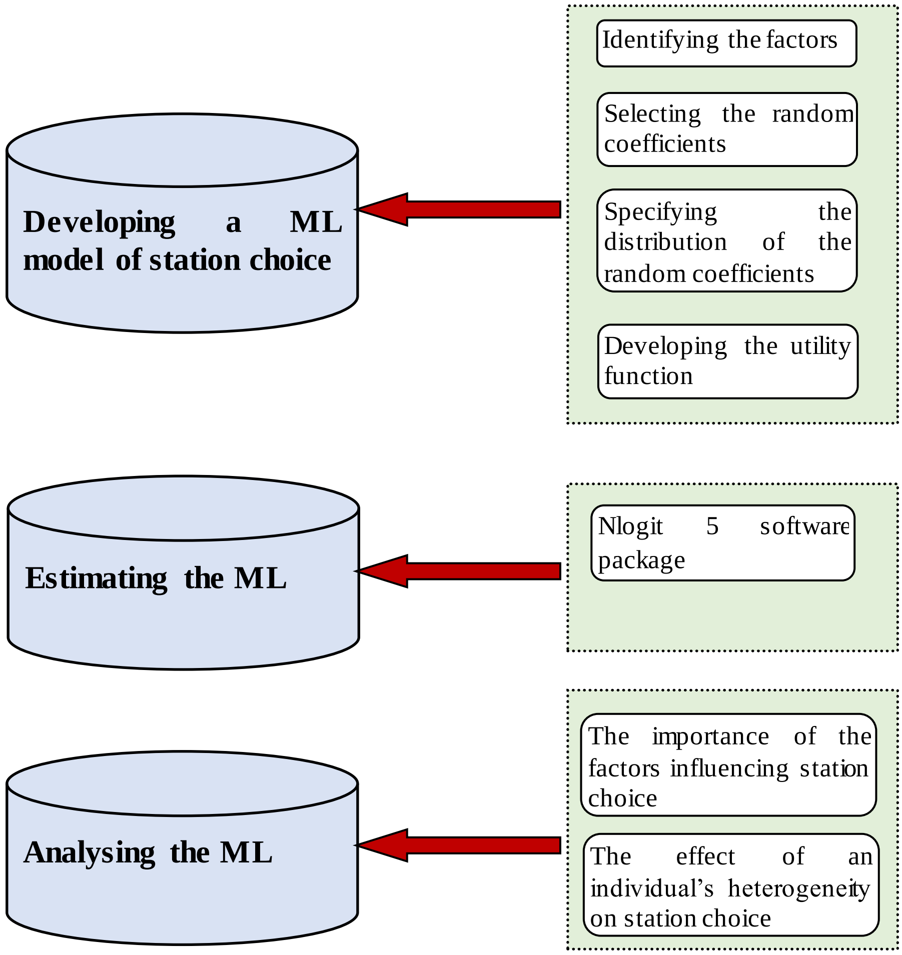


Figure 8.1 The flowchart to develop the main model

(a) Identifying the factors used in the ML model

An intercept survey was conducted to identify the factors to be used in the model of station choice. We conducted the survey on 2 July 2012 at the seven Perth train stations, (see Appendix C). We asked P&R users to rank a list of potential factors influencing their choice, as prepared by us, and to also write down any other factors affecting their choice, i.e. not included in our list. Based on the survey results, we identified 15 factors that could materially influence a P&R user's choice of departure train station, including travel time to the station, (divided into usual travel time, good day travel time and bad day travel time), parking searching time, parking capacity, parking availability, parking fee, parking fine, the frequency of patrols for illegal parking, crowding on trains, number of days per week on which the trains are too crowded to board, in-vehicle travel time, safety, ticket fares and train frequency.

We divided these factors into four types, namely, the factors related to travel time to the station, the factors related to parking, the factors related to crowding on trains, and all other factors. Each type included more than one factor. The effects of the first three types of factors on station choice have already been tested, separately, in the three independent models. Here we have integrated these three models into a single overall model. We took the disutility of station choice under travel time uncertainty, under parking search time uncertainty and under crowding on train uncertainty as three different independent variables, and added three other variables to estimate overall utility of station choice for P&R users. The three other variables are safety, ticket fare and train frequency. The three models were, therefore, considered as three sub-models to the overall model. The relationship between the factors influencing station choice and the station choice models can be seen in Figure 8.2.

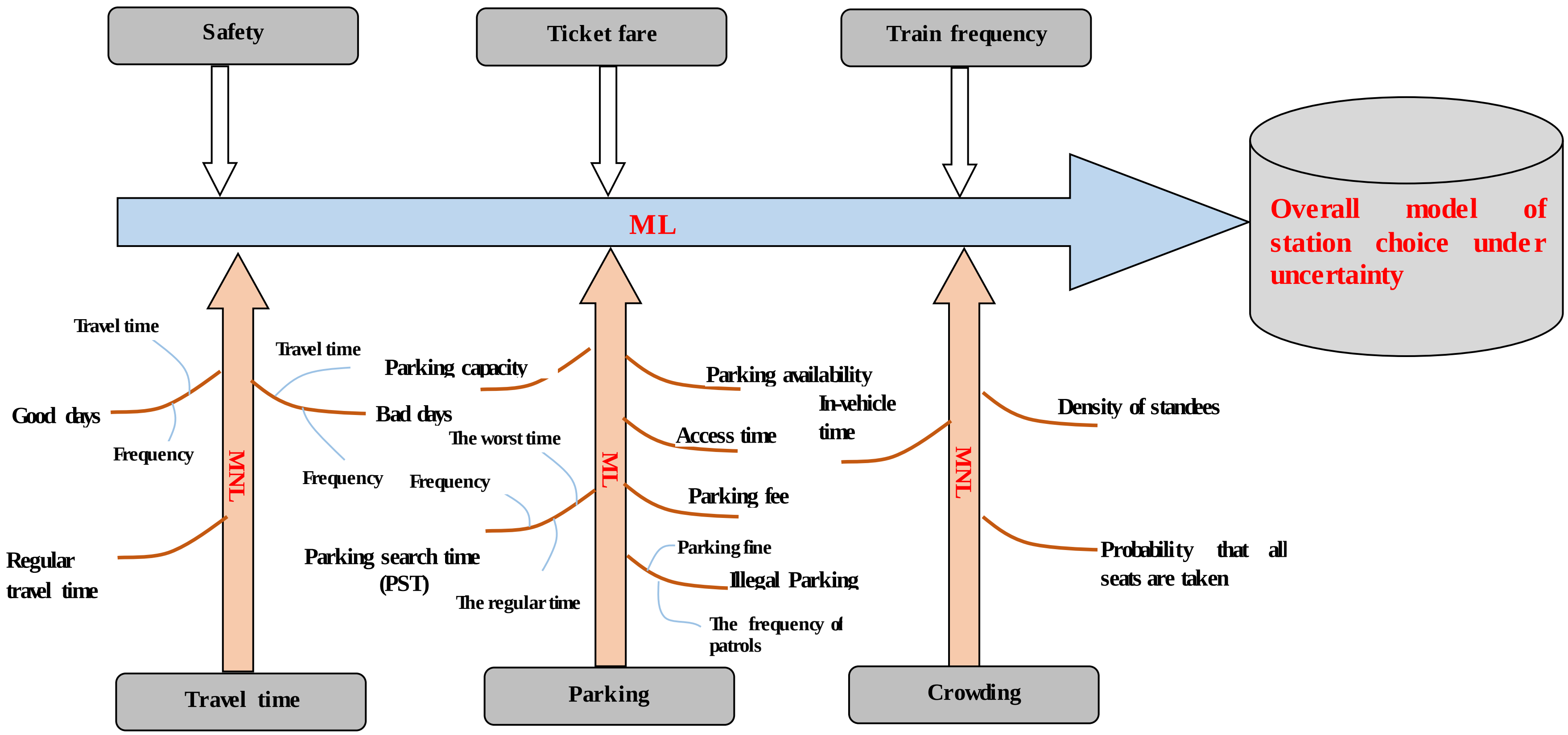


Figure 8.2 Attributes used to develop the overall model of station choice

(b) Selecting the random parameters

The random parameters in an ML model can reflect not only the degree of unobserved heterogeneity, (by the standard deviation of the parameters), but also an individual's preference heterogeneity, (based on an individual's personality and social-economic status, e.g. income, age, gender, education level etc.). Currently, the most popular way to select random coefficients is from the statistical results for the estimation of parameters in an ML model in which all attributes have random parameters (Hensher & Greene, 2003). In other words, the statistical index of parameter estimates can indicate whether the parameter is random or not.

Specific to this research, only two parameters were considered as random. There were two reasons for this: ① the focus of the research is to explore the effects of uncertain factors on station choice, so we only assumed the parameters of the three uncertain attributes to be random; and ② the estimation of parameters within the ML model (see Table 8.1) showed that the parameter of the effect of crowding is statistically insignificant. Thus, only two uncertain attributes, (i.e. the effect of travel time and parking search time), were taken as factors having random parameters.

Table 8.1 Estimation of random parameters within ML assuming normal distribution

	Coefficient	Standard Error	z	Prob. │z│>Z*	5% Confidence Interval	
	Random parameters in utility functions					
Effect of travel time(TT)	5.0552**	2.04581	2.47	0.0135	1.04581	9.06523
Effect of parking (P)	8.64750**	3.59659	2.4	0.0162	1.59832	15.69668
Effect of crowding on trains(CR)	-0.31488	2.88679	-0.11	0.9131	-5.97288	5.34312
	Non-random parameters in utility functions					
Safety	2.76134***	1.06889	2.58	0.0098	0.66636	4.85633
Ticket fare	-8.09251**	3.41569	-2.37	0.0178	- 14.78715	-1.39787
Headway (train frequency)	-0.86976	1.16262	-0.75	0.4544	-3.14845	1.40894
Station one reference	-0.84469**	0.41679	-2.03	0.427	-1.66158	-0.0278
	Distns. Of RPs. Std. Devs or limits of triangular					
NsNUTT	3.22349	5.60302	0.58	0.5651	-7.75823	14.20521
NsNUP	0.6147	1.61639	0.38	0.7037	-2.55337	3.78276
NsNUC	6.05980	4.91894	1.23	0.2180	-3.58113	15.70074
Log likelihood	-814.69719					
Chi-square [10 d.f.]	253.19337					
Significance level	0.0000					
Pseudo- R^2	0.1344922					
Inf.Cr.AIC	1649.4					

Note: ***, **,* == >Significance at 1%, 5% and 10% confidence levels respectively

(c) Specifying the distributions of the random parameters

Before developing the ML model, we needed to define the distribution of the selected random parameters. Currently, the four most popular distribution forms are normal, uniform, triangular and lognormal. However, we cannot readily determine which one is the best because all are in common use and each has its advantages and disadvantages. Empirically, the uniform distribution is often used for dummy variables and the lognormal distribution is often used when the parameter needs to be positive (Hensher & Greene, 2002). Focusing on the overall model, given that the two variables with random parameters are not dummy variables and the parameter estimates do not have to be positive, we only tested the normal and triangular forms for the random parameters. The general specifications in terms of normal and triangular distribution are shown as follows:

$$Normal: \quad \beta_n(i) = \beta_{mean} + \sigma_{SD} \times N \tag{8-1}$$

$$Triangular: \quad \beta_t(i) = \beta_{mean} + \sigma_{SP} \times t \tag{8-2}$$

where

$\beta_n(i), \beta_t(i)$ indicate the i^{th} random parameter with normal distribution or triangular distribution for its mean;

β_{mean} is the mean of parameter estimates;

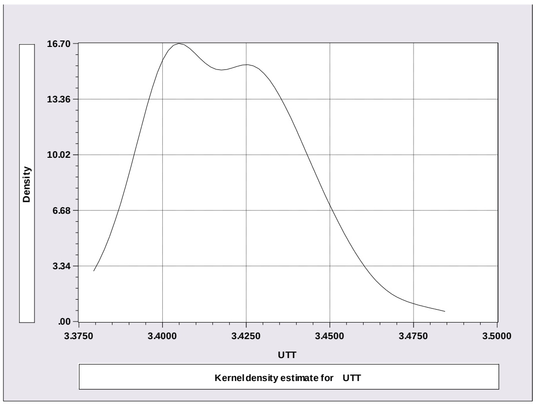
σ_{SD} is the standard deviation of the normal distribution; and

σ_{SP} is the spread of the triangular distribution.

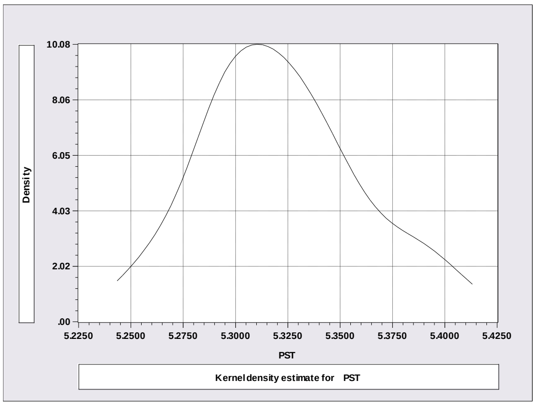
Our challenge was therefore to determine whether the random parameter distributions should be normal or triangular. Two approaches were considered. The first was an empirical method, namely, the random parameter's distribution was determined based on the shape of individual parameter estimates over the sample population. In practice it was difficult to directly estimate each individual's parameter due to the limitations of the data size. Moreover, most of the parameter estimates based on individual data were not aligned with what we would have expected. In order to resolve these issues, we applied the method suggested by Hensher and Greene (2003), namely, we estimated a model with data from all but one sampled respondent, only one respondent was removed each time and the model was re-estimated. The difference between the parameters estimates of the model based on the full sample size N and the model based on $N-1$ sample size can give us a clue about individuals' preference heterogeneity.

Another way to specify the distribution of random parameters is based on statistical indices for the models with different distributions. In order to use this approach, a hypothesis needs to be set up, namely, the better the ML model, the closer the distribution of random parameters in the model is to their real behavioural profile. Then, the statistical indices among the ML models are compared with the combination of different distributions of random parameters. Based on this comparison, the distribution used in the best model was considered as their real distribution.

Focusing on the research, we firstly used the empirical method proposed by Hensher and Greene (2003) to estimate individual parameters, then plotted the parameter estimates of the effects of travel time and parking using the kernel density estimator in Nlogit 5. The results are presented in Figures 8-3-1 and 8-3-2.



8-3-1. The effect of travel time



8-3-2 The effect of parking

Figure 8.3 Kernel density estimates for the random parameters

Based on the empirical shape of the random parameters, we can conclude that both random parameters broadly follow the normal distribution.

Next, we applied another method to test the results. We assumed four situations: ① that both random parameters followed the normal distribution; ② that both followed the triangular distribution; ③ that the parameter of the effect of travel time followed the normal distribution and that of parking the triangular distribution; and ④ the parameter of the effect of travel time followed the triangular distribution and parking the normal distribution. Then, we estimated the parameters within these ML models with different combinations of distribution of random parameters. The results are presented in Table 8.2 and show that the model in which the random parameters followed the normal distribution is the best, based on the Chi-square statistical index, even though the parameter estimates in the different models are similar, which aligns with the findings of Hensher and Greene (2003).

Table 8.2 Estimation of parameters within ML assuming different distribution forms

	(n,n)	(t,t)	(n,t)	(t,n)
	Random parameters in utility functions			
Effect of travel time (TT)	4.07279***	6.12289*	7.29343	5.12048***
Effect of parking(P)	6.22644***	8.81689**	9.96410	7.63670***
	Non-random parameters in utility functions			
Effect of crowding on trains(CR)	0.55589	0.15883	1.04859	0.53073
Safety	2.11804***	3.21333**	3.78058	2.71194***
Ticket fare	-5.62747***	-8.20432**	-9.68540	-7.18964***
headway	-1.13206	-1.31669	-1.04859	-1.26238
Station one reference	-0.57***	-0.78572*	-0.97143	-0.70972**
	Distns. Of RPs. Std. Devs or limits of triangular			
NsNUTT	3.36388	18.9962	9.62430	14.6579
NsNUP	0.10998	0.46078	1.74519	0.25584
	Statistical indices			
Log likelihood	-815.99316	-815.13434	-814.53984	-815.85821
Chi square	250.60143	252.31906	253.50806	250.87132

Note: ***, **,* ==> Significance at 1%, 5% and 10% confidence levels respectively

Based on these results, we specified the distribution of the effect of travel time and parking as the normal form.

(d) The analysis of heterogeneity around the mean of the random parameters

According to Hensher and Greene (2002), whether there is heterogeneity around the mean estimate of the random parameter depends on the interaction between the mean parameter estimate and a covariate. It is reasonable to believe the heterogeneity around the mean from the observed covariates exists if the interaction is statistically significant, otherwise we can only assume that the heterogeneity around the mean can exist but be from other unobserved covariates.

8.2.2 Data sources

Two sets of data were used to develop the overall model of station choice. The first contained the data directly related to the railway service, such as safety, train ticket fare, and train frequency, the effects of which on station choice were not tested in the three sub-models. They were collected from the train station choice survey mentioned in Chapter four. The second set was the output from the three utility models developed in the Chapters 5 to 7.

8. 3 Results

8.3.1 Estimation of parameters

The overall model of station choice is made up of six attributes. The effects of travel time and parking have random parameters and both follow normal distributions. All parameters in the overall model were estimated within a mixed logit model by Nlogit 5. The results are shown in Tables 8-3 and 8-4.

Table 8.3 Estimation of parameters in the overall model within initial MNL

	Coefficient	Standard Error	Z	Prob. │z│>Z*	95% Confidence Interval	
Effect of travel time(TT)	3.42090***	0.65297	5.24	0.0000	2.14109	4.70070
Effect of parking search time(PST)	5.33058***	0.97410	5.47	0.00000	3.42138	7.23979
Effect of crowding on trains(CR)	0.84618	1.72063	0.49	0.6229	-2.52620	4.21857
Safety	1.76494***	0.25315	6.97	0.0000	1.26878	2.26110
Headway(train frequency)	-1.14915*	0.69025	-1.66	0.0959	-2.500203	0.20372
Ticket fare	-4.81797***	0.74505	-6.47	0.000	-6.27824	-3.35771
Station one specific constant	-0.51107***	1.14779	-3.46	0.0005	0.80073	-0.22140
Log Likelihood	-816.833377					
Chi-square [6 d.f.]	237.21264 (6)					
Inf. Cr. AIC/N	1.213					

Note: ***, **,* ==>Significance at 1%, 5% and 10% confidence level respectively

Table 8.4 Estimation of parameters in the overall model within ML with two random parameters

	Coefficient	Standard Error	z	Prob. │z│>Z*	95% Confidence Interval	
	Random parameters in utility functions					
Effect of travel time(TT)	4.62924***	1.56510	2.96	0.0031	1.56170	7.69678
Effect of parking(P)	7.02911***	2.08879	3.37	0.0008	2.93515	11.12306
	Non-random parameters in utility functions					
Effect of crowding on trains(CR)	0.48521	2.32328	0.21	0.8346	-4.06834	5.03876
Safety	2.43142***	0.75117	3.24	0.0012	0.95915	3.90370
Ticket fare	-6.37722***	1.88068	-3.39	0.0007	-10.06329	-2.69115
Headway	-1.21434	0.94453	-1.29	0.1986	-3.06558	0.63689
Station one specific constant	-0.62974**	0.24909	-2.53	0.0115	-1.11795	-0.14154
	Distns. Of RPs. Std. Devs or limits of triangular					
NsNUTT	4.77377*	2.75816	1.73	0.0835	-0.63123	10.17966
NsNUP	0.10679	1.29036	0.08	0.9340	-2.42228	2.63585
Log likelihood	-815.09673					
Chi-square [9 d.f.]	252.39428					
Significance level	0.0000					
Pseudo- R^2	0.1283					
Inf. Cr. AIC/N	1.213					

Note: ***, **,* ==>Significance at 1%, 5% and 10% confidence levels respectively

The results started with the parameter estimates within an MNL model, (see Table 8.3), which are the initial values of the parameters in the ML model and required by Nlogit 5. The initial MNL model is statistically significant, (Chi-square value of 237.21264 with 6 degrees of freedom), and all the parameters are statistically significant, except for the effect of crowding on train with a *p*-value of 0.6229, and met our expectations / assumptions. In other words, the effects of travel time, parking and safety all have positive signs. This means that the higher the utility for travel time and parking and the safer a train station is perceived to be, the more likely that station is to be chosen. In the case of the effect of travel time above, the "utility" has a negative value and is usually expressed as a disutility, i.e. the larger its negative value, the greater the disincentive to travel, and the less the station's utility, which aligns with the findings of early studies related to travel time (Li et al., 2010; Mackie et al., 2003). Thus, the positive coefficient of the effect of travel time means that the larger the disutility of travel time to a particular station, the lower the probability that that station would be chosen, which is as we would expect.

The parameters for ticket fare, headway and station one specific constant have negative signs. The first two imply that the more expensive the ticket and the longer the time between trains at a particular station, the lower the probability of that train station being chosen.

Based on Table 8.3, only the parameter of the effect of crowding on trains is statistically insignificant. However, it does still have some impact on P&R users' choice. Similar to the effect of travel time, its values are negative and crowding should also be taken as one of the components of overall disutility of station choice for P&R users. Therefore, the positive parameter of the effect of crowding on trains means that the greater the disutility of crowding on trains, the lower the probability that the station would be chosen.

In general, the signs of these parameters estimated in the MNL model are reasonable and align with expectations.

Based on the initial parameter estimates within the MNL model, Nlogit 5 re-estimated all parameters in the overall model within the ML model. The results are shown in Table 8.4. The model is statistically significant, with a Chi-square value of 252.39428 with 9 degrees of freedom, a $p-$value of 0.0000 and a pseudo-R^2 of 0.13, which indicate that the data can fit the overall ML model. However, when compared to the MNL model used to estimate the starting values for the parameter estimates, we are unable to conclude that the fitted ML model is any better, based on the results of the χ^2 test. This is because the Chi-square value, calculated with the two log likelihood values in the two models, is 3.46 (i.e. ($-2\times\left(-816.83377-\left(-815.09672\right)\right)$)), which is less than the Chi-square critical value with two degrees of freedom of 5.991. Additionally, we compared the AIC/N values of the two models and found their values to be almost the same. Therefore, there were no obvious differences between the two models.

Based on Table 8.4, we also concluded that the mean of both random parameters over the sampled population is statistically different to zero, i.e. the $p-$value for the effect of travel time parameter is 0.0031 and for the effect of parking is 0.0008, which are less than alpha equal to 0.05, (i.e. the 95% confidence interval). Table 8.4 also shows that the parameter estimates for the derived standard deviations for the effect of travel time are statistically significant, which implies the parameter coefficients around the mean parameter estimate over the sample population is heterogeneous, (i.e. parameter estimates is individual-specific and may be different from the sample population mean parameter estimate). However, the dispersion of the effect of parking parameter is statistically insignificant, which indicates parameter estimates may be same over sample population and can be captured within the mean. Therefore, we had to re-estimate the overall ML model, retaining only the effect of travel time parameter as a random parameter. The results are given in Table 8.5.

Table 8.5 Estimation of parameters within ML based on only retaining the effect of travel time as a random parameter

	Coefficient	Standard Error	z	Prob. |z|>Z*	5% Confidence Interval	
	Random parameters in utility functions					
Effect of travel time(TT)	4.60317***	1.50224	3.06	0.0022	1.65883	7.54750
	Non-random parameters in utility functions					
Effect of parking (P)	7.00052***	2.03170	3.45	0.0006	3.01846	10.98258
Effect of crowding on trains(CR)	0.47763	2.30914	0.21	0.8361	-4.04819	5.00346
Safety	2.41836***	0.71987	3.36	0.0008	1.00745	3.82928
Ticket fare	-6.34539***	1.80603	-3.51	0.0004	-9.88513	-2.80564
Headway	-1.20726	0.932338	-1.29	0.1954	-3.03470	0.62018
Station one specific constant	-0.62653***	0.24238	-2.58	0.0097	-1.10158	-0.15148
	Distns. Of RPs. Std. Devs or limits of triangular					

NsNUTT	4.72447*	2.65998	1.78	0.0757	-0.48901	9.93794
Log likelihood	-815.10021					
Chi-square [8 d.f.]	252.38753					
Significance level	0.0000					
AIC/N	1.212					

Note: ***, **,*==> Significance at 1%, 5% and 10% confidence levels respectively

Comparing Table 8.5 with Table 8.4, we found that there were no significant differences between the two models, with just the parameter estimates in Table 8.5 being less than the ones in Table 8.4. Nevertheless, we considered that the model with one random parameter was still better than the model with two random parameters, based on the model fitness (AIC/N: 1.212 vs 1.213). Thus, the effect of the travel time parameter (ETT) can be written per equation (8-3):

$$ETT = 4.60317 + 4.72447 \times N \tag{8-3}$$

8.3.2 Revealing preference heterogeneity

The above section has proven the existence of heterogeneity in the mean of effect of travel time parameter estimates over the sampled population. In this section, we explore the possible sources of the heterogeneity. We compared the effect of travel time random parameter with each individual's personal attributes, (such as age, gender and annual income), to test whether the heterogeneity in the effect of travel time parameter is the result of differences in individuals' attributes. The results are shown in Table 8.6.

Table 8.6 Individuals' preference heterogeneities of the random parameter

	Coefficient	Standard Error	z	Prob. |z|>Z*	5% Confidence Interval	
	Random parameters in utility functions					
Effect of travel time(TT)	4.92254***	1.77899	2.77	0.0056	1.43641	8.40867
	Non-random parameters in utility functions					
Effect of Parking (P)	5.51980***	1.93371	2.85	0.0043	1.72980	9.30979
Effect of crowding on trains(CR)	2.43479	2.73706	0.89	0.3737	2.92976	7.79933
Safety	2.36643***	0.77899	3.04	0.0024	0.83963	3.89322
Ticket fare	-5.73641***	1.83821	-3.12	0.0018	-9.33924	-2.13357
Headway	-2.0618*	1.15097	-1.79	0.0732	-4.31767	0.19406
Station one specific constant	-0.69657**	0.28292	-2.46	0.0138	-1.25109	0.14205
	Heterogeneity in mean, Parameter: Variable					
NUTT:AGE	-0.18146	0.13907	-1.3	0.1920	-0.45403	0.09111
NUTT:GEN	0.68795	0.63418	1.08	0.2780	-0.55502	1.93092
NUTT:INC	0.01295	0.2810	0.46	0.6449	-0.04212	0.06802
	Distns. Of RPs. Std. Devs or limits of triangular					
NsNUTT	5.15902	3.19925	1.61	0.1068	-1.11140	11.42944
Log likelihood	-676.58094					
Chi-square [11 d.f.]	193.94263					

Significance level	0.0000
Adjust R-square	0.1167
Inf. Cr. AIC/N	1.232

Note: ***, **,* ==> Significance at 1%, 5% and 10% confidence levels respectively

Table 8.6 shows that the overall model is statistically significant, (Chi-square value of 193.94263 with 11 degrees of freedom). The overall model fit, obtained from the pseudo- R^2, is 0.12 which is statistically acceptable for this class of model.

Focusing on the output pertaining to the random parameters, the dispersion of the effect of travel time parameter is statistically insignificant, indicating that the differences in the marginal utilities held for the effect of travel time cannot be explained completely by the differences in these individuals' attributes (i.e. age, gender and annual income). However, the results do indicate a relationship between them as the values of NUTTs are not zero. The heterogeneity in the mean parameter estimates for age of -0.18 implies that sensitivity to the effect of travel time, over the sampled population, decreases as an individual's age increases, given that the effect of travel time is negative. In other words, younger individuals are more sensitive to travel time than elder individuals. The heterogeneities in the mean parameter estimates for income and gender are positive, which means the higher-income individuals tend to be more time-sensitive than those with smaller incomes and that females are more sensitive to the effect of travel time than males, due to gender being represented as a dummy variable with male = 1 and female = 0.

8.3.3 Elasticity analysis for key factors

Elasticity is a ratio of the variation in percentage of the dependent variable to the percentage change of some independent variables, which is generally used to analyse the relationship between the variation of percentage of the quality demanded and variation of percentage of some factors (Hensher et al., 2005).

In the research, we used it to investigate the relationship between the percentage change of station choice probabilities and percentage change of the six factors. The results are shown in the following table.

Table 8.7 Elasticity analysis

	Partial Effect	Standard Error	z	Prob. │z│>Z*	95% Confidence Interval	
Safety	0.23067***	0.02215	10.42	0.0000	0.18726	0.27407
Ticket fare	-0.57283***	0.10240	-5.59	0.000	-0.77354	-0.37212
Headway	-0.47128***	0.04429	-10.64	0.000	-0.55808	-0.38448

Effect of travel time	0.54741***	0.04917	11.13	0.000	0.45104	0.64377
Effect of crowding on trains(CR)	0.43404***	0.06336	6.85	0.000	0.30985	0.55822
Effect of Parking (P)	0.23401***	0.02683	8.72	0.000	0.18143	0.28659

Note: ***, **,* ==> Significance at 1%, 5% and 10% confidence levels respectively

Based on the table 8.7, all of the elasticities are statistically significant at the 99% interval confidence level. The elasticities for safety, the effect of travel time, the effect of crowding on trains and the effect of parking are positive, which imply that a 1 percentage increase in these attributes will increase the probability of station choice separately by 0.23, 0.55, 0.43, and 0.23. The elasticities for ticket fare and head way are negative, which means that a 1 percentage of the both attributes will decrease the choice probabilities of train stations by 0.57 and 0.47. Comparing all absolute elasticity values, we found the elasticity for ticket fare is biggest, which means the effect of its variation on station choice is biggest, next is the effect of travel time, the third and the fourth are headway and the effect of crowding on trains respectively, the safety and the effect of parking are relative inelastic. This comparison can provide transport planners and railway industries with evidence to improve rail patronage.

8. 4 Evaluation of experiment and the main model using eye tracking methods

8.4.1 Method

In order to evaluate the overall model, we designed an eye tracking experiment based on the questionnaires used in the station choice survey mentioned in Chapter 4. The data from the experiment are the measurements of visual attention given to each attribute in the questionnaire by each respondent. With these data, we can rank the level of significance of each attribute for P&R users' choice within the discrete choice model. A comparison of these rankings with the results of the overall model, developed in terms of SP data, allowed us to evaluate the model and determine its consistency with the observations.

The process for evaluating the overall model was as follows:

- **Step1**: Data collection

The data used to develop the new station choice model were from an eye tracking experiment and used on the basis that respondents in the eye tracking experiment and those in the station choice survey are affected in a similar way by the factors influencing station choice. We firstly designed the questionnaires based on our objectives, as mentioned in Chapter 4, and the questionnaires used in the station choice survey. They were then input into the eye tracking instrument, consisting of a 60Hz Remote Eye Tracking Device (RED) and a laptop. The experiment was implemented at Curtin University. Thirty-five respondents were invited to attend. Two types of visual attention data were obtained, the durations of the fixations on each particular area of interest (AOI) on the questionnaire and the frequencies of these fixations. A sample questionnaire showing the AOIs is presented in Figure 8.4.

Figure 8.4 Sample questionnaire with AIOs

- **Step 2:** Creating an index for visual attention

The new station choice model was developed based on the visual attention data. In order to simplify the model, avoid missing any information and minimise the bias that each datum could produce, we created a new index, (called average fixation), by combining both duration and frequency of fixation. For each AOI, we summed the duration of all the fixations on that AOI, then divided by the number of times (frequency) the eye fixed on that AOI, i.e. to produce the mean fixation time. The results were presented graphically using a heatmap which displays a respondent's visual attention in colour, ranging from green, (low attention), to red, (high attention). An example heatmap can be seen in Figure 8.5.

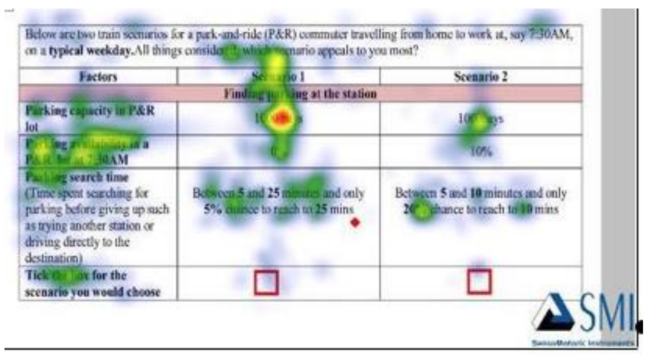


Figure 8.5 Example Heatmap

- **Step 3:** Estimating the station choice model and ranking the factors

Based on the visual attention data, we developed a multinomial logit model to explore station choice behaviour. All of the parameters' coefficients were estimated with Nlogit 5, which indicated the weights for each factor in the model. Then, all the factors were ranked based on their weights.

- **Step 4**: Ranking the factors in the overall model developed based on the SP data

- **Step 5**: Comparison of both rankings

8.4.2 Station choice model based on the eye tracking data

Using the data from the eye tracking experiment, we developed a multinomial logit model exploring station choice behaviour. In contrast to the SP data in the station choice survey, the factors in the eye tracking experiment were determined based on the AOIs in the questionnaires used in the eye tracking experiment. To compare the model with the eye tracking and the overall model with the SP data, we reorganised the eye tracking data based on the variables in the overall model developed based on the SP data. In other words, we divided the eye tracking data into six groups with the same variables as in the overall model based on the SP data, i.e. travel time, parking, crowding, safety, headway (train frequency) and ticket. Based on these eye tracking variables, the utility function related to station choice can be written as follows:

$$U_i = \beta_{TT} u_{TTi} + \beta_p u_{pi} + \beta_{Cr} u_{Cri} + \beta_{sa} u_{sai} + \beta_C u_{Ci} + \beta_{tf} tf_i \tag{8-4}$$

where $\beta_{TT}, \beta_p, \beta_{Cr}, \beta_{sa}, \beta_C, \beta_{tf}$ are the estimated coefficients, the values of which indicate the weight of the relevant parameter.

8.4.3 Results and discussion

This model was also developed using Nlogit 5, with the results shown in Table 8.8.

Table 8.8 Estimation of parameters in the model with the duration of fixation

	Coefficient	Standard Error	Z	Prob. $\|z\| > Z*$	95% Confidence Interval	
The effect of travel time	-0.00026	0.00040	-0.65	0.5171	-0.00105	0.00053
The effect of Parking	0.00034	0.00027	1.28	0.2010	-0.00018	0.00086
The effect of Crowding	0.00155***	0.00060	2.59	0.0096	0.00038	0.00273

Safety	0.00052	0.00077	0.67	0.4997	-0.00099	0.00204
Ticket fare	0.0045***	0.00121	3.71	0.002	0.00212	0.00688
Headway	0.00084**	0.00034	2.46	0.0138	0.00017	0.0015
Station one specific constant	-0.08560	0.15777	-0.54	0.5874	-0.39481	0.22362
Log Likelihood	-120.17314					
Chi-square [6 d.f.]	36.89259					
Pro chi-squared > value	0.000000					

Note: ***, **,*==>Significance at 1%, 5% and 10% confidence levels respectively

Table 8.7 shows that the model is statistically significant, (Chi-square of 36.89259 with 6 degree of freedom and a $p-$value of 0.000000). The overall fit of this model is adequate. However, only three attributes are statistically significant and only one uncertain factor is included. Among the six attributes, the most important factor is the train frequency, followed by safety, ticket fare (cost), then the other three uncertain factors in which parking is the most important, then travel time and lastly crowding. Similar to the ranking process above, we also ranked the factors based on the parameter estimates from the model (4-3) with the MNL approach. The results are also shown in Table 8.9. The ranking sequence based on the both experiments, except the crowding, is completely different, so that we cannot conclude that the longer respondents fix their eyes on an attribute, the more important that attribute is for respondents in choosing a train station. This can be the fact that the eye tracking experiment was conducted at the university and it cannot be assumed that all the respondents were P&R users. Additionally, we analysed the uncertain factors and the results were consistent, i.e. parking, travel time and crowding are ranked as first, second and third place respectively. This aligned with our expectations as P&R users are car users and the attributes affecting their parking and driving of cars should be critical to P&R travel. Therefore, it is reasonable that the effect of parking attribute is the most important and the travel time is ranked second.

Table 8.9 Comparison of the significance of attributes between the eye tracking data and SP data

	1	2	3	4	5	6
Model (4-3) SP Data	Parking	Ticket fare	Safety	Travel time	Train frequency	Crowding
Model (8-4) Eye Tracking Data	Train frequency	Safety	Ticket fare	Parking	Travel time	Crowding

8. 5 Chapter summary

In this chapter, an overall model of station choice under uncertainty, combining the three sub-models established in the previous chapters and other factors in the station choice survey, has been developed. The ML approach was applied to estimate the model, in which the effect of travel time was taken as a random parameter, and its distribution assumed to be normal, based on individuals' parameter estimates.

With this model, we found that the mean of the effect of travel time random parameter is statistically different to zero over the sampled population, which indicated that preference heterogeneity existed. Moreover, we found that age, gender and annual income do affect preference heterogeneity even though their estimates are statistically insignificant. This means that the young people, females and individuals with higher incomes are more sensitive to the effect of travel time.

Moreover, elasticity analysis was also conducted in the chapter. The results showed that the percentage increase in safety, the effect of travel time, the effect of crowding and the effect of parking will increase the probability of station choice, while the percentage increase in ticket fare and headway will decrease the probability of station choice, which can provide transport planners and railway industries with the evidence to improve rail patronage.

Additionally, we used an eye tracking experiment to test which factors are important to respondents in their choice of departure train station. The experiment assumed that the higher the frequency and the longer the duration of the eyes' fixation on an attribute, the more that attribute is likely to influence a respondent's choice behaviour. Based on this assumption, an MNL model was developed using the respondents' duration of fixation on the factors.

The results showed that the rankings for the factors, (except crowding), in the two models are different, which implies that the factors respondents pay more time looking at are not as important as other factors in the process for choosing departure train stations. Possible reasons include the questionnaire design being complex and the wording for some attributes may be ambiguous or difficult to understand, requiring respondents to spend more time on that particular attribute than would be expected, based on its actual importance.

CHAPTER 9 CONCLUSIONS AND RECOMMENDATIONS

9. 1 Introduction

This chapter summarises the major findings and key achievements of the research into modelling station choice under uncertainty for P&R users, and discusses the limitations in data collection, sample size and model development. Recommendations for improvements and directions for future research are also included. Additionally, the research objectives and relative sub-objectives are revisited to determine whether they have been achieved or not.

The research primarily aims to explore the station choice behaviour of P&R users under uncertainty, which was proposed and explained in Chapter 1. In order to achieve this objective, Chapter 2 first reviewed the previous research related to station choice, P&R mode and travel choice under uncertainty. Based on this, the research gaps in station choice modelling were identified and the need to undertake the research confirmed. Then, Chapter 3 outlined the methodology used to develop the station choice models for P&R users under uncertainty. In order to implement the station choice models developed by this study, an SC experiment and an eye tracking experiment were set up and discussed in Chapter 4. Using these data, Chapters 5, 6, and 7 separately outlined the effects of three different uncertain factors, (namely, travel time to departure train station, parking search time and crowding on trains), on station choice and measured the respondents' risk attitude towards each uncertainty. The effect of individuals' experiences on their risk attitude and the relationship between the respondents' risk attitude and boarding numbers at the same train station were also analysed. The output of these models may be very useful in assisting transport planners to more accurately and reliably predict P&R demand, local governments to make investment decisions and price P&R facilities, and public transport operators to improve the service quality of P&R facilities. Lastly, Chapter 8 developed an overall station choice model under uncertainty by combining the effects of these three uncertain factors with the three certain factors. By using the model, the probability that a train station is chosen can be calculated, individuals' preference heterogeneity for the effect of travel time to stations can be analysed and elasticity analysis can be investigated.

9. 2 Revisiting the thesis objectives

The research aimed to model the choice of departure train station for P&R users under the uncertainties of travel time to station, parking search time and crowding on trains. In order to achieve the objective, four sub-objectives were set up. They are checked below to see that they have been met.

(a) Develop a novel framework for estimating P&R users' station choice under uncertainty

While studies involving station choice modelling have been conducted for more than 40 years, there is only limited research available in the literature. Most previous station choice models were estimated using discrete choice techniques, the utility functions of these models being linear, and they didn't consider the effects of variations in the factors affecting station choice and respondents' risk attitudes to these variations. However, the previous literature did identify that uncertainty or variations in variables such as travel time, parking search time and crowding on trains can affect station choice and that an individual's personality, (such as attitude to risk), may also influence station choice. Therefore, it is necessary to develop a new framework to explore P&R users' choice of departure train station under uncertainty. The framework is shown on Figure 9.1 and described in detail in Chapter 3. In order to explore station choice under uncertainty for P&R users, we developed four station choice models including three sub-models and an overall model. Generally, these models were developed within discrete choice models, the specifications included the multinomial logit model, the latent class model and the mixed logit model. The utility functions within them were non-linear and developed with the mean-variance approach, in which the utility contributed by the variation of uncertain factors was evaluated based on decision making theories under uncertainty (or risk), i.e. cumulative prospect theory and extended expected utility theory. With these models, the probability that a departure train station is chosen and a respondent's risk attitude towards variations in the uncertain factors are measured. Their influence on station choice is also estimated and an individual's preference heterogeneity for station choice can be captured.

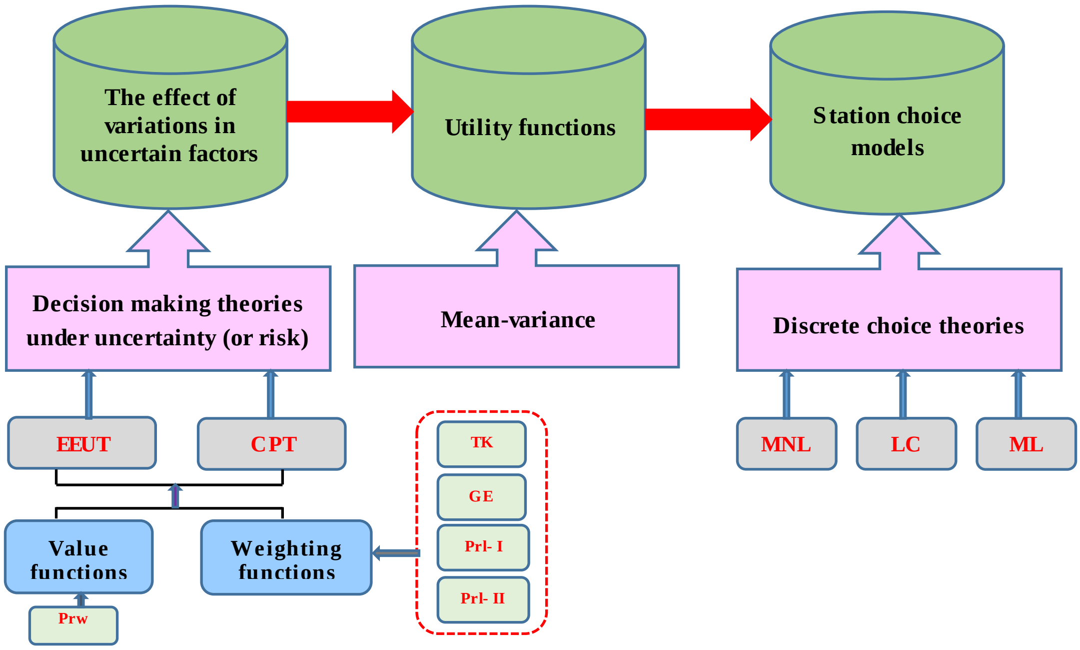


Figure 9.1 Framework for modelling station choice under uncertainty

(b) Investigate and identify the key factors, (both certain and uncertain), affecting train station choice for P&R users

In the research, two ways were used to identify key factors influencing train station choice for P&R users. They were introduced in Chapters 2 and 3. Firstly, Chapter 2 summarised the key factors used in previous station choice models. Then, we conducted a train survey, in which we listed all factors in the previous literature and asked respondents to identify any other factors they considered important but were not included in the previous literature, (see Appendix C). In Chapter 3, we built up a decision tree to summarise all the factors influencing P&R users' choice for departure train station, which further tested the completeness of the factors. By these actions, possible factors influencing P&R users' choice for departure train station were identified.

(c) Develop station choice models under uncertainty for P&R users

Four station choice models were developed in the research. They are the station choice sub-model under travel time uncertainty, the station choice sub-model based on the effect of variability of crowding on trains, the station choice sub-model for the variability of parking search time and the overall station choice model under uncertainty. The method to develop the models, their specifications and their results were presented separately in Chapters 5, 6, 7 and 8 respectively.

(d) Implement and evaluate the station choice models.

These were investigated in Chapters 5, 6, 7 and 8. Chapter 5 applied the sub-model of travel time uncertainty to three train stations in Perth and found that respondents' attitudes towards travel time uncertainty is risk averse. Moreover, it revealed that a P&R user's experience and perception could play vital roles in their risk attitude toward station choice under uncertainty, which provides useful information for transport planners to improve the local network around train stations. Chapter 6 validated the station choice sub-model of the effect of variability of crowding on trains with the chi-squared test. Moreover, a sensitivity test of the sub-model and an analysis of the effect of individuals' preference heterogeneity on station choice were conducted. The results revealed that the wealthier an individual is, the less likely he/she is to use crowded trains, which provides public transport authorities with evidence to propose a graded service. For example, differently priced ticket fares (e.g. business first and second class), could be offered according to the level of crowding on the train, and passengers could choose different tickets based on their income levels. Chapter 7 applied the station choice sub-model under parking search time uncertainty to measure respondents' risk attitude towards parking search time uncertainty, which provides a basis for government to invest in P&R facilities. Chapter 8 validated the overall station choice model under uncertainty with an eye tracking experiment.

In summary, the thesis has modelled station choice under uncertainty for P&R users and all four sub-objectives derived from the overall goal have been achieved.

9. 3 Conclusions of the research

This section highlights the major conclusions and findings of the research.

9.3.1 The introduction of new factors into station choice models

Station choice decision making is complex and can be influenced by many factors. Previous studies identified some, such as, in-vehicle travel time, train frequency, safety, travel time, cost, parking cost, parking capacity, etc. In the research, new explanatory variables were included in the station choice models, such as crowding attributes, (i.e. the probability that all seats have been taken and the density of passengers standing in a carriage), parking search time, parking availability, etc. Given that some of them have been identified as uncertain, the attributes reflecting their variations were also introduced. For example, travel time to train station, in the prior studies, was a fixed value for each level, but in the research, we divided travel time into three states, i.e. normal, bad days and good days. For each normal travel time, bad day and good day travel time and their frequencies varied across all alternatives. Based on these, four extra attributes reflecting travel time uncertainty were introduced. Similar to travel time to train station, regular parking search time, the worst parking search time in a month, and their relative frequencies, and regular crowding and the frequency that trains were too crowded to board in a month were introduced to respectively reveal the variations in parking search time and crowding on trains. With these new factors, we can not only better understand station choice behaviour, but more importantly, can better predict station choice under uncertainty.

Generally, the station choice models, involving both the new added factors and the factors identified in the previous literature, can better explain the choice behaviour under uncertainty and make a basis for more accurately predicting P&R demand.

9.3.2 Development of new station choice models

All previous station choice models were developed within a discrete choice theory framework, but they are either multinomial logit or nested logit, (or cross-nested logit), models with access mode choice. All these models have a closed form and can readily calculate choice probability. However, they are restricted in the extreme value distribution, especially for multinomial logit (MNL), and it assumed that the unobserved utility is not correlated over individuals, which is not consistent with reality. Therefore, the research tried another advanced logit model (i.e. mixed logit) to explore station choice behaviour, as it is more flexible and can approximate any random utility model. Moreover, it resolved the three limitations of the MNL by considering individual variations in the random variables, and allowing patterns to be substituted without restrictions and the unobserved factors to be correlated with respect to time. Furthermore, the choice probabilities can be simply computed through simulation. The application of the mixed logit model for modelling station choice not only realistically explained the choice behaviour but also clearly revealed individual's preference heterogeneity.

Additionally, a Latent class (LC) model was firstly used to analyse the individual's heterogeneity on station choice. With this model, we successfully proved the effect of the heterogeneity of P&R users with different income levels on station choice.

9.3.3 Creation of a new decision framework for station choice

In contrast to the prior decision frameworks for station choice, the decision framework in the research took the effects of uncertain factors on station choice into account. As an example, in previous studies travel time to train station was an independent factor contributing to the utility related to station choice. Nevertheless, in the new framework, travel time was divided into different types, i.e. regular travel time, the best and the worst travel times. Thus, by assessing the effect of each type separately, the effect of variations in travel time was introduced into the station choice models. Given that the effect of uncertainties was being considered, decision making theories under uncertainty (or risk) were introduced into the framework of station choice. It is worth noting that expected utility theory (EUT), which is a normative decision making theory under uncertainty (or risk) and widely used in studies of travel choice, was not used. Instead cumulative prospect theory (CPT) and extended expected utility theory (EEUT) were applied to evaluate the effect of uncertainties on station choice. In contrast to the EUT, in which the probabilities weighting each outcome are objective, the probabilities weighting each outcome in both the CPT and the EEUT are more subjective. Hence, the utilities related to the variation of uncertainties evaluated within the CPT and the EEUT can reflect individuals' preference heterogeneity. Additionally, non-linear value functions developed within the CPT and the EEUT, similar to the EUT, can indicate respondents' risk attitudes. Therefore, the models developed within the new framework may better explore station choice behaviour than the previous station choice model due to consideration of the effect of variations in uncertainty, individuals' risk attitudes and individuals' preference heterogeneity.

9.3.4 Creation of a new picture display of crowding closer to the real situation

In previous literature, crowding levels in the stated choice (SC) experiments were displayed pictorially from a bird's eye view, so that seat availability and the density of standees were often taken as two independent crowding measurements presented to respondents. However, in this research in-vehicle crowding levels in the SP experiment were displayed from the front, i.e. the view respondents would see looking into the carriage when the doors open.

Moreover, a new crowding measure was created based on the pictorial display. Respondents cannot distinguish seat availability from the density of standees once the door is opened, especially in overcrowding situations or when passengers are clustered close to the doors rather than moving down inside the carriage. The crowding level they perceive is therefore a function of, and interaction between, both available seats and density of standees. Hence, the interaction between seat availability and density of standees was taken as the crowding measurement in the research.

9.3.5 Main conclusions from the station choice models

(a) Travel time sub-model of station choice

The travel time sub-model of station choice is a multinomial logit model, in which the utility function is established using a combination of CPT and the mean-variance approach. Based on the goodness of fit test with the SP data in the research, the non-linear utility specification with power value function and GE risk weighting function is the preferred model. With this model, we concluded that: ① greater travel time variability in the loss situation, i.e. travel on "bad traffic" days, and longer regular travel times could lead to lower utility and lower probability that a station is chosen; ② P&R users' attitudes towards travel time variability were risk averse; ③ the risk attitude toward travel time variability has some influence on station choice; and ④ P&R users who have personally experienced higher travel time variations and greater differences between perceived and estimated travel times tend to be more risk averse towards their station choice under travel time variability than those who have experienced or perceived less travel time variations. This indicates that P&R users' experiences and perceptions could play vital roles in risk attitude toward station choice under uncertainty.

(b) Crowding sub-model of station choice

The sub-model of station choice based on the effect of crowding variability is also a multinomial logit model, in which the utility function is developed with the combination of EEUT and the mean-variance approach. The non-linear specification developed in EEUT is made up of a power value and TK risk weighting functions, in which the crowding measurement is created based on the interaction between seat availability and density of standees. The results from the sub-model showed that: ① the more crowded a train is and the longer time individuals spend boarding the train at a station, the less utility that station has; ② respondents' attitude towards crowding on trains is risk averse; and ③ the greater risk aversion respondents displayed, the fewer individuals boarding at the train station. A latent class model of station choice under crowding variability was also developed. With this model, we found that individuals with higher incomes were likely to choose the station where the trains were less crowded. This indicates that commuters with higher incomes place a higher value on, and potentially are prepared to pay more for, a more comfortable travel environment than those with lower incomes, i.e. that respondents' preference for the effect of crowding on station choice under annual income is heterogeneous.

(c) Parking search time sub-model

The sub-model of station choice under the variability of parking search time is an ML model, in which the utility of a train station is measured by a hybrid function defined by two sub-functions for two situations, i.e. parking available or not available in the legal P&R areas. When parking is available, parking availability at a given access time and parking fee are the main considerations for P&R users in choosing their departure train stations. When parking is not available in legal P&R areas, (i.e. all spaces are occupied), the utility of a train station is a combination of the variation in parking search time (PST), (estimated by CPT based on the normal PST, the worst PST, and their frequencies in one month), the parking fine and the frequency of patrolling for illegal parking. With this sub-model, we found that P&R users showed risk aversion to the variations in PST. Larger variations in PST, higher parking fines and more frequent patrols for illegal parking, lowered the utility at a train station and the reduced the probability that the station was chosen, which provides useful insights for the implementation of public transport policies (such as ticket price policy).

(d) Overall model

The overall model of station choice under uncertainty for P&R users was developed with an ML approach based on the utilities from the above three sub-models and other factors in the station choice survey, in which the utility of travel time was tested as a random parameter with normal distribution. From the overall model, we obtained the ranking of factors influencing station choice for P&R users. The results showed that the effect of parking attributes on station choice is the most important for P&R users, followed by ticket fare and safety, then the effect of travel time to train station, train frequency, and lastly the effect of crowding on trains. This provides public transport operators with useful information to focus improvements to station service quality to attract more P&R users, especially when budgets are constrained. Moreover, the model results revealed that the young, females and individuals with higher incomes are more sensitive to the effect of travel time, which indicates that there is individual preference heterogeneity for the effect of travel time. Furthermore, the elasticity analysis based on the model provide transport planners and railway industries with the evidence to improve rail patronage.

Additionally, by eye tracking techniques, we found that the factors respondents pay more time looking at are not necessarily more important than other factors in the process for choosing departure train stations.

9.4 Research limitations

To our knowledge, this is the first attempt to model station choice behaviour for P&R users under uncertainty. The research has some limitations due to time constraints and limits in experience of the researcher. This section discusses the limitations in the identification of uncertain factors, data collection, sample size and the methods used.

(a) Inadequacies of uncertain factors considered in the station choice models

In the research, three factors, i.e. travel time to departure train stations, parking search time and crowding on trains, were taken as uncertain. These three factors prove that station choice is a decision under uncertainty and that models considering their variations can better explain station choice behaviour than the previous station choice models. As well as these three, there are a number of other uncertain factors that could influence station choice such as crime rate, (which can vary from time to time and from station to station), departure time, (which can vary by time-of-day and day-to-day depending on the required arrival time and the traffic situation on the road network around the station), and in-vehicle travel time, (which is normally fixed and calculated based on the train timetable but can be adversely affected by incidents and weather conditions). The station choice model under uncertainty could therefore be further improved by adding the effects of more of these uncertain factors.

(b) Limitation of survey implementation

The data used to model station choice under uncertainty in the research were collected by train station choice surveys conducted at the seven train stations from 9:00 am to 3:00 pm for every surveying day. Although the seven train stations were proposed by our industry and government partners, met the research's selection criteria and were representative, Perth has 70 train stations and we cannot assume that the survey covered the opinions of all P&R users at the seven stations or at the other stations on the Perth train network.

Another limitation is the time we were permitted by the Public Transit Authority to collect the data is inconsistent with the time that most of P&R users usually travel (i.e. before 9am). Legal P&R facilities in Perth are usually full or almost full before 8:30am every weekday. Therefore, not all of the respondents in the surveys were P&R users. Even though it is reasonable that train passengers were taken as potential P&R users, their choice or considerations may be somewhat different from those of actual P&R users. In the survey, we asked respondents to make a station choice assuming they are P&R users.

(c) Limitation of sample size

The total sample size used to model station choice under uncertainty is 600, which is greater than the minimum requirement (84) derived based on the sampling method proposed by Rose and Bliemer (2009) and the experience method proposed by Orme (2005). However, the sample size for some of the individual stations didn't meet the minimum requirement. For example, the sample size for Midland is 29, Warnbro 83, Greenwood 72. Given that these models were also applied for specific analyses focusing on single stations, some parameters estimated with the data from single train stations may not be significant. Therefore, more data may to be collected for each train station to test the hypotheses for each individual station in the research in the future.

(d) Limitations of modelling

In the research, the data used to estimate the sub-models and the overall model were from the same questionnaire, due to the limitations of time and resources. Generally, the data collected from the station choice survey can be divided into four parts, i.e. related to parking, travelling to train station, crowding on trains and other factors. The first three sets of data, together with the station choice, were used to estimate the three sub-models respectively. A choice made by an individual should be based on trading-off all the factors presented to him/her. In the research, all factors were listed in the same questionnaire. Therefore, without any specific techniques in the SC survey, we cannot assume that the choice in sub-models is made only based on the factors related to the sub-model, rather than other factors. Hence, studies using special techniques to remove, or at least decrease, the interference among multipurpose data should be conducted in the future.

9. 5 Recommendations and Future research directions

This section discusses recommendations and directions for future research related to modelling station choice and the opportunities for strengthening and expanding the current approach.

9.5.1 Future tasks related to data

As the requirements for comfort on trains and P&R demand have increased, the factors related to comfort, (such as the probability that all seats are taken, the density of standees in a carriage), and the parking conditions, (such as parking search time), were introduced into station choice models in the research. The model considering these new factors did explain station choice behaviour in the current context. As individuals' demand and preferences change over time, more new factors affecting station choice could be identified in the future.

The second task is to update the selection criteria of sample train stations, enlarge the sample station sets and collect data from these stations as Perth's railway system is further expanded and travel characteristics change.

The third task is to enlarge the sample size for each station by more scientific sampling methods and re-estimate the station choice models.

The final task is to undertake a household survey to collect data to be used to explore station choice behaviour in the future due to its complexity and the difficult in predicting it, especially within the short time period of this research.

9.5.2 Future work related to the models

Within CPT, a number of different forms have been proposed for the value functions and their associated probability weightings (Tversky & Kahneman, 1992). Value functions include linear, logarithmic, quadratic exponential, Beli, and Hara., Risk weighting functions include linear, power, Wu-Gonzalez, etc. (Stott, 2006). In the thesis, we only tested the power value and four popular risk weighting functions within CPT. Therefore, different functional forms of cumulative prospect theory based on the different combinations of value functions and risk weighting functions should be further investigated in future research with the station choice data.

There are many descriptive decision-making theories that can be used to analyse choice behaviour under uncertainty (or risk). In the thesis, only CPT and EEUT were used to explore P&R users' choice of departure train station under uncertainty. Therefore, the application of other decision making theories under uncertainty, (or risk), such as Savage's theory, rank dependent utility theory or dual theory of expected utility, etc., for modelling station choice behaviour should be included in future research.

9. 6 Chapter summary

It is very important to accurately model P&R users' station choice under uncertainty, which can not only provide a basis for improving current P&R demand models, but also provide public transport operators with evidence for improving the service quality level. However, station choice is a complex behaviour, which is affected by many factors. To date, none of the research can fully explain, and therefore reliably predict, it. In the research, more factors were identified, the effects of variations in uncertain factors were evaluated by the combination of the decision making under uncertainty (or risk) and mean-variance approach, the respondents' attitudes towards uncertainty were measured, individuals' preference heterogeneities were analysed, and respondents' willingness to pay for the deduction of one unit of effect of uncertainty was estimated. The results of the research have significant benefits from both scientific and socio-economic perspectives.

REFERENCES

Abbas, K., & Sergany, A. A. (2008). *Park and Ride Strategy for the Australian Capital Territory (Final report)*. Retrieved from http://tpteh.com/wp-content/uploads/2015/02/Park-and-Ride-Strategy-for-the-Australian-Capital-Territory-2008.pdf

Abdel-Aty, M. A., Kitamura, R., & Jovanis, P. P. (1995). Investigating effect of travel time variability on route choice using repeated-measurement stated preference data. *Transportation Research Record*,(1493), 39-45. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-0029333456&partnerID=40&md5=a1faa5cf1e778c537848b7eb9195a966

Abdel-Aty, M. A., Kitamura, R., & Jovanis, P. P. (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transportation Research Part C: Emerging Technologies, 5*(1), 39-50. http://dx.doi.org/10.1016/S0968-090X(96)00023-X

Abkowitz, M. D. (1981). An analysis of the commuter departure time decision. *Transportation, 10*(3), 283-297. http://dx.doi.org/10.1007/bf00148464

Adcock, S. J. (1997). *A passenger station choice model for the British rail network*. Paper presented at the European Transport Conference Cambridge

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723. http://dx.doi.org/10.1109/TAC.1974.1100705

American Associaton of State Highway and Transportation Officials. (2004). *Guide for the deisgn of park-and-ride facilities*. United States: American Association of State Highway and Transportation Officials.

Arne, R. H. (2004). Forecasting the demand for an employee Park and Ride service using commuters' stated choices. *Transport Policy, 11*(4), 355-362. http://dx.doi.org/10.1016/j.tranpol.2004.04.003

Arrow, K. J. (1971). *Essays in the theory of risk-bearing*. Amsterdam, London: North Holland Pub.Co.

Assiociation of Train Operating Companies. (2002). Passenger demand forcasting handbook. Retrieved from http://www.atoc.org/about-atoc/commercial-activities/passenger-demand-forecasting-council/the-passenger-demand-forecasting-handbook.

Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES). (2017). Catchment Scale Land Use of Australia - Update September 2017.

Austroads. (2017). Guide to Traffic Management Part 11: Parking (2017 Edition). Retrieved from https://www.onlinepublications.austroads.com.au/items/AGTM11-17

Avineri, E. (2004). A cumulative prospect theory approach to passengers behavior modeling: Waiting time paradox revisited. *Intelligent Transportation Systems, 8*, 195-204. http://dx.doi.org/10.1080/15472450490523856

Avineri, E., & Bovy, P. H. L. (2014). Identification of Parameters for Prospect Theory Model for Travel Choice Analysis *Journal of the Transportation Rseach Board, 2082*, 141-147. http://dx.doi.org/10.3141/2082-17

Avineri, E., & Prashker, J. (2003). Sensitivity to Uncertainty: Need for a Paradigm Shift. *Journal of the Transportation Reserch Board 1854*(10), 90-98. http://dx.doi.org/10.3141/1854-10

Avineri, E., & Prashker, J. N. (2004). Violation of expected utility theory in route-choice stated preference *Transporation research record, 1894*(23), 90-98. http://dx.doi.org/10.3141/1894-23

Bajwa, S., Bekhor, S., Kuwahara, M., & Chung, E. (2008). Discrete choice modeling of combined mode and departure time. *Transportmetrica, 4*(2), 155-177. http://dx.doi.org/10.1080/18128600808685681

Balijepalli, N. C., Shepherd, S. P., & Kant, P. (2013). *Calibration and validation of parking search time function*. Paper presented at the the Institution of Civil Engineers - transport. http://dx.doi.org:10.1680/tran.11.00067

Batarce, M., Muñoz, J. C., & Ortúzar, J. d. D. (2016). Valuing crowding in public transport: Implications for cost-benefit analysis. *Transportation Research Part A: Policy and Practice, 91*, 358-378. http://dx.doi.org/http://dx.doi.org/10.1016/j.tra.2016.06.025

Bates, J. J. (1987). Measuring Travel Time Values with a Discrete Choice Model: A Note. *The Economic Journal, 97*(386), 493-498. http://dx.doi.org/10.2307/2232894

Baum, A., & Epstein, Y. (1978). *The gravity of crowding: application to social physics to the effects of others*. Erlbaum Hillsdale, NJ.

Bekhor, S., Ben-Akiva, M. E., & Ramming, M. S. (2001). *Estimating Route Choice Models for large Urban Networks*. Paper presented at the The 9th World Conference on Transport Research, Seoul, Korea

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., T, N., Oppewal, H., & V, R. (1994). Combining Revealed and Stated Preferences Data. *Marketing Letters, 5*(4), 335-349. http://dx.doi.org/10.1007/BF00999209

Ben-Akiva, M., & Francois, B. (1983). *Homogenous generalized extreme value model*. Cambridge, MA.:

Ben-Akiva, M., & Steven, R. L. (1985). *Discrete Choice Analysis:Theory and Application to Travel Demand*. Cambridge, Massachusetts, London, England: The MIT Press.

Ben-Elia, E., Erev, I., & Shiftan, Y. (2008). The combined effect of information and experience on drivers' route-choice behavior. *Transportation, 35*(2), 165-177. http://dx.doi.org/10.1007/s11116-007-9143-7

Ben-Elia, E., & Shiftan, Y. (2010). Which road do I take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice, 44*(4), 249-264. http://dx.doi.org/10.1016/j.tra.2010.01.007

Bhat, C. R. (1998). Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B: Methodological, 32*(6), 361-371. http://dx.doi.org/10.1016/S0191-2615(98)00004-6

Blainey, S., & Evens, S. (2011). Local station catchements: reconciling theory with reality *European Transport Conference held in United Kindom,*

Bliemer, M. C. J. (2016). *Orthogonal desings*.

Bliemer, M. C. J., & Rose, J. M. (2009). Efficiency and Sample Size Requirements for Stated Choice Studies. Retrieved from http://ws.econ.usyd.edu.au/itls/wp-archive/itls_wp_05-08.pdf

Blunden, S., Bray, D., Every, D., Hirsch, L., Holyoak, N., Rainbird, S., . . . Zito, R. (2011). A socioeconomic study of carriage and platform crowding in the Australian railway industry: literature review. Retrieved from file:///F:/school%20cumpter/ccm/2016/write%20paper/R2.104%20Literature%20Review.pdf

Bogers, E., Viti, F., & Hoogendoorn, S. (2005). Joint modelling of avacnced travel information service, havit, and learning impacts on route choice by laboratory simulator experiments. *Tansportaiton Research Record: Journal of the Tranpsortation Research Board, 1926*(1-8) http://dx.doi.org/10.3141/1926-22

Bolger, D., Colguhoun, D., & Morrall, J. (1992). Planning and design of park-and-ride facilites for the Calgary light rail transit system. *Transortatuib Research Record: the Journal of Tranpsortation Reserach Board, 1361*, 141-148. Retrieved from http://onlinepubs.trb.org/Onlinepubs/trr/1992/1361/1361-020.pdf

Bos, I. D. M., Heijden, R. E. C. M. V. d., Molin, E. J. E., & Timmermans, H. J. P. (2004). The Choice of Park and Ride Facilities: An Analysis Using a Context-Dependent Hierarchical Choice Experiment. *Environment and Planning A, 36*(9), 1673-1686. http://dx.doi.org/10.1068/a36138

Brownstone, D., & Small, K. A. (2005). Valuing time and reliabilty: assessing the evidence from road pricing demonstrations. *transportation reserch part A, 39*, 279-293.

Bunch, D. S., Louviere, J., & Anderson, D. (1996). *A comparison of experiemental design strategies for multinomial logit models: the case of generic attributes*. CA: Graduate School of Management,University of California.

Bureau of Infrastructure Transport and Regional Economics (BITRE). (2017). *Growth in Australian Road System*. Retrieved from https://bitre.gov.au/publications/2017/files/is_092.pdf

Bureau of Transport and Regional Economics [BTRE]. (2007). *Estimating urban traffic and congestion cost trends for Australian cities* Available from Bureau of Transport and Regional Economics Retrieved from https://bitre.gov.au/publications/2007/files/wp_071.pdf

Burrell, J. E. (1976). Multiple Route Assignment: A Comparison of Two Methods. *Traffic Equilibrium Methods, 118*, 229-239. http://dx.doi.org/10.1007/978-3-642-48123-9_12

Cairns, M. R. (1998). The development of park and ride in Scotland. *Journal of Transport Geography, 6*(4), 295-307. http://dx.doi.org/10.1016/S0966-6923(98)00016-7

Camerer, C. f., & Ho, T.-H. (1994). Violations of the Betweeness Axiom and Nonlinearaity in Probiaiblity. *Journal of Risk and Uncertainty, 8*, 167-196. http://dx.doi.org/10.1007/BF01065371

Cantarella, G., & de Luca, S. (2003). Modeling transportation mode choice through artificial neural networks*Fourth International on Uncertainty Modeling and Analysis, 2003, held in College Park, MD,* (pp. 84-90). http://dx.doi.org:10.1109/ISUMA.2003.1236145

Carrion, C. (2010). *Value of reliablity: actual commute experience revealed preference appraoch (Masters' thesis).* Retrieved from http://hdl.handle.net/11299/93565

Carrion, C., & Levinson, D. (2012). Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice, 46*(4), 720-741. http://dx.doi.org/10.1016/j.tra.2012.01.003

Cascetta, E. (1989). A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transportation Research Part B: Methodological, 23*(1), 1-17. http://dx.doi.org/10.1016/0191-2615(89)90019-2

Cascetta, E., Nuzzolo, A., Russo, F., & Vitetta, A. (1996). *A Modified Logit Route Choice Model Overcoming Path Overlapping Problems. Specification And Some Calibration Results For Interurban Networks.* Paper presented at the Transportation and traffic theory. Proceedings of the 13th international symposium on transportation and traffic theory, Lyon, FranceRetrieved from http://worldcat.org/isbn/0080425860

Chakour, V., & Eluru, N. (2014). Analyzing commuter train user behavior: a decision framework for access mode and station choice. *Transportation, 41*(1), 211-228. http://dx.doi.org/10.1007/s11116-013-9509-y

Chang, G. L., & Mahmassani, H. S. (1988). Travel time prediction and departure time adjustment behavior dynamics in a congested traffic system. *Tranportation Research Part B, 22*(3), 217-232. http://dx.doi.org/10.1016/0191-2615(88)90017-3

Chen, C., Jianhong(Cecilia), X., Smith, B., Olaru, D., Taplin, J., & Han, R. (2015). *Influence of parking on train station choice under uncertainty for park-and-ride users.* Paper presented at the 6th international conference on apllied human factors and ergonomics(AHFE 2015) and the affiliated conferences,AHFE 2015, Las vegas, USA. http://dx.doi.org:10.1016/j.trpro.2017.05.274

Chen, C., Xia, J., Smith, B., & Han, R. (2014). Development of a Conceptual Framework for Modeling train station choice under uncertainty for park and ride users*TRB 93rd Annual Meeting held in Washington DC, U.S.*: Transportation Research Board. Retrieved from http://hdl.handle.net/20.500.11937/13612

Chen, C., Xia, J., Smith, B., Olaru, D., Taplin, J., & Han, R. (2017, 2017/01/01/). *Influence of travel time variability on train station choice for park-and-rider users.* Paper presented at the World Conference on Transport Research - WCTR 2016 Shanghai, Shanghai, China. http://dx.doi.org:10.1016/j.trpro.2017.05.274

Cherrington, L. K., Brooks, J., Cardenas, J., Elagart, Z., Galicia, L. D., Hansen, T., . . . Coffel, K. (2017). *Decision-Making Toolbox to Plan and Manage Park-and-Ride Facilities for Public Transportation: Gudebook on Planning and Managing Park-and-Ride*. 500 Fifth, NW, Washington, DC 20001: Transporation Reserach Board.

Chew, S. H., & MacCrimmon, K. R. (1979). *Alpha-nu choice theory: A generalization of expected utility theory*: working paper 669, University of British Columbia.

Circella, G., Dell'Orco, M., & Sassanelli, D. (2005). *Uncertainty in Choice Behaviour Prediction: A Hybrid Approach to Combine Fuzziness and Randomness.* Paper presented at the European Transport Conference, Strasbourg, France

Cousins, E. (1977). A mathematical analysis of a park-and-ride scheme *Tranportation Research 12*, 247-252. http://dx.doi.org/10.1016/0041-1647(78)90066-7

Cox, T., Houndmont, J., & Griffiths, A. (2006). Rail passenger crowding, stress, health and safety in Britain. *Transportation Research Parrt A: Policy and Practice*, *40*(3), 244-258. http://dx.doi.org/10.1016/j.tra.2005.07.001

Daganzo, C. (1979). *Multinomial Probit: The Theory and Its Application to Demand Forecasting*. New York: Academic Press.

Davidson, B., Vovsha, P., Abedini, M., Chu, C., & Garland, R. (2011). *Impatct of capacity, crowding, and vehicle arrival adherence on public transport ridership: Los Angeles and Sydney experience and forecasting approach.* Paper presented at the Australasian Transport Research Forum 2011, Adelaide, AustraliaRetrieved from http://www.patrec.org/atrf.aspx

Davidson, B., & Yang, L. (1997). *Modeling of commuter rail station choice and access mode combinations.* Paper presented at the 12th International EMME/2 users' Conference, San Francisco, California, USA

Davidson, B., & Yang, L. (1999). *Modeling of commuter rail station choice and access mode combinations.* Paper presented at the Transportation research Board 1999 annual meeting, Washington D.C.

Day, A. T., & Day, L. H. (1973). Cross-national comparison of population density. *Science 181*(4104), 1016-1023. http://dx.doi.org/10.1126/science.181.4104.1016

de Finetti, B. (1937). La Prévision: Ses Lois Logiques, Ses Sources Subjectives. *Annales de l'Institut Henri Poincaré, 17*, 1-68.

de Palma, A., & Picard, N. (2005). Route choice decision under travel time uncertainty. *Transportation Research Part A: Policy and Practice, 39*(4), 295-324. http://dx.doi.org/10.1016/j.tra.2004.10.001

Debrezion, G., Pels, E., & Rietveld, P. (2007). Choice of departure station by railway users. *European Transport, 37*(78-92)Retrieved from https://www.openstarts.units.it/dspace/bitstream/10077/5959/1/Debrezion_Pels_Rietveld_ET37.pdf

Debrezion, G., Pels, E., & Rietveld, P. (2009). Modelling the joint access mode and railway station choice. *Transportation Research Part E: Logistics and Transportation Review, 45*(1), 270-283. http://dx.doi.org/10.1016/j.tre.2008.07.001

Department of infrastructure. (2005). *Understanding the market for public transport*. Merbourne:

Department of Transport. (2010). Car Parking at Stations.

Department of Transport. (2011). Public transport for perth in 2031. Retrieved from http://www.transport.wa.gov.au/ABOUT_P_PT_Plan2031.pdf

Department of Transport and Main Roads. (2015). *Public Transport Infrastructure Manual (PTIM) 2015*. Department of Transport and Main Roads. Queensland Government. Retrieved from file:///F:/school%20cumpter/ccm/2018/thesis/paper/ptimfull.pdf

Desfor, G. (1975). Binary station choice models for a rail rapid transit lines. *Transportation Research, 9*(1), 31-41. http://dx.doi.org/10.1016/0041-1647(75)90018-0

Dial, R. B. (1969). Algorithm 360: Shortest-path forest with topological ordering *Comn.ACM, 12*, 632-633. http://dx.doi.org/10.1145/363269.363610

Dial, R. B. (1971). A probabilistic multipath traffic assignment model which obviates path enumeration. *Transportation Research, 5*(2), 83-111. http://dx.doi.org/10.1016/0041-1647(71)90012-8

Dickins, I. (1991). Park-and-Ride facilities on light rail transit systems. *Transporatation, 18*, 23-36. http://dx.doi.org/10.1007/BF00150557

Dijk, M., & Montalvo, C. (2011). Policy frames of Park-and-Ride in Europe. *Journal of Transport Geography, 19*(6), 1106-1119. http://dx.doi.org/10.1016/j.jtrangeo.2011.05.007

Drrksrra, E. W. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematlk 1*, 269-271. Retrieved from https://link.springer.com/content/pdf/10.1007/BF01386390.pdf

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Reviews Psychol, 32*, 53-88. http://dx.doi.org/10.1146/annurev.ps.32.020181.000413

Eluru, N., Chakour, V., & El-Geneidy, A. M. (2012). Travel mode choice and transit route choice behavior in Montreal: insights from McGill University members commute patterns. *Public Transport, 4*(2), 129-149. http://dx.doi.org/10.1007/s12469-012-0056-2

Ettema, D., Tamminga, G., Timmermans, H., & Arentze, T. (2005). A micro-simulation model system of departure time using a perception updating model under travel time uncertainty. *Transportation Research Part A: Policy and Practice, 39*(4), 325-344. http://dx.doi.org/10.1016/j.tra.2004.12.002

Evans, G. W., & Wener, R. E. (2007). Crowding and personal space invsion on the train: please don't make me sit in the middle. *Journal of Environmental Psychology, 27*, 90-94. http://dx.doi.org/10.1016/j.jenvp.2006.10.002

Faghri, A., Lang, A., Hamad, K., & Henck, H. (2002). Integrated knowledge-based geographic information system for determing optimal location of park-and-ride facilities. *Journal of Urban Planning and Development 128*, 18-41. http://dx.doi.org/10.1061/(ASCE)0733-9488(2002)

Fan, K., Miller, E., & Badoe, D. (1993). Modeling Rail Access Mode and Station Choice. *Journal of the Transportation Research Board, 1413*, 49-59. Retrieved from http://onlinepubs.trb.org/Onlinepubs/trr/1993/1413/1413-006.pdf

Fan, W., Khan, M. B., Ma, J., & Jiang, X. (2013). Bilevel Programming Model for Locating Park-and-Ride Facilities. *Journal of Urban Planning and Development 140*(3) http://dx.doi.org/10.1061/(ASCE)UP.1943-5444.0000178

Farhan, B., & Murray, A. T. (2005). A GIS-Based Approach for Delineating Market Areas for Park and Ride Facilities. *Transactions in GIS, 9*(2), 91-108. http://dx.doi.org/10.1111/j.1467-9671.2005.00208.x

Fehr-Duda, H., & Epper, T. (2011). Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences. *Annual Review of Economics, 4*(1), 567-593. http://dx.doi.org/10.1146/annurev-economics-080511-110950

Fernandez, E., Cea, J. d., Florian, M., & Cabrera, E. (1994). Network Equilibrium Models with Combined Modes. *Transportation Science, 28*(3), 182-192. http://dx.doi.org/10.1287/trsc.28.3.182

Fernández, R. (2011). Experimental study of bus boarding and alighting times*European Transport Conference held in Glasgow,*

Fifer, S., Rose, J., & Greaves, S. (2014). Hypothetical bias in Stated Choice Experiments: Is it a problem? And if so, how do we deal with it? *Transportation Research Part A: Policy and Practice, 61*, 164-177. http://dx.doi.org/10.1016/j.tra.2013.12.010

Florian, M., & Los, M. (1979). Determining intermediate origin-destination matrices for the analysis of composite mode trips. *Transportation Research Part B: Methodological, 13*(2), 91-103. http://dx.doi.org/10.1016/0191-2615(79)90027-4

Fosgerau, M., & Fukuda, d. (2012). Valuing travel timeVariability: characteristics of the travel time distribution on an urban road *Transportation Research Part C: Emerging Technologies, 24*, 83-101. http://dx.doi.org/10.1016/j.trc.2012.02.008

Fox, J. (2005). Modelling park and ride in the PRISM model for the West Midlands region*AET European Transport Conference held in Strabourg, France* Retrieved from http://abstracts.aetransport.org/conference/index/id/11

Fox, J., Andrew Daly, Bhanu Patruni, & Milthorpe, F. (2011). *Extending the Sydney Strategic Model to represent toll road and park-and-ride choices* Paper presented at the Australasian Transport Research Forum 2011, Adelaide, Australia

Frederick R. Harris, I. (2012). *State Park-and-Ride Guide* [Florida Department of Transportation Office of Freight, Logistics and Passenger Operations
Transit Office]. 605 Suwannee Street, MS, Tallahassee: Retrieved from http://www.fdot.gov/transit/pages/finalparkandrideguide20120601.pdf

García, R., & Marín, A. (2005). Network equilibrium with combined modes: models and solution algorithms. *Transportation Research Part B: Methodological, 39*(3), 223-254. http://dx.doi.org/10.1016/j.trb.2003.05.002

Gaver, D. P. (1968). Headstart Strategies for Combating Congestion. *Transportation Science, 2*(2), 172-181.

Ginn, S. (2009). *The applicaiton of the park & ride and TOD concepts to develop a new framework that can maximise public transport praronage.* Master degree Queensland University of Technology Australia Retrieved from http://eprints.qut.edu.au/31827/1/Simon_Ginn_Thesis.pdf

Givoni, M., & Rietveld, P. (2014). Do cities deserve more railway stations? The choice of a departure railway station in a multiple-station region. *Journal of Transport Geography, 36*, 89-97. http://dx.doi.org/10.1016/j.jtrangeo.2014.03.004

Government of UK, R. E. (2015). Rail passenger numbers and crowding on weekdays in major citites in England and Wales:2014. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/466412/rail-passengers-crowding-2014.pdf

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological, 37*(8), 681-698. http://dx.doi.org/10.1016/S0191-2615(02)00046-2

Guttman, J. M. (1979). Uncertainty, the Value of Time, and Transport Policy. *Journal of Transport Economics and Policy, 13*(2), 225-229. Retrieved from http://www.jstor.org/stable/20052540

Hamdouch, Y., Ho, H. W., Sumalee, A., & Wang, G. (2011). Schedule-based transit assignment model with vehicle capacity and seat availability. *Tranportation Research Part B: Methodological, 45*(10), 16. http://dx.doi.org/10.1016/j.trb.2011.07.010

Harata, N., & Ohta, K. (1986). *Some findings on the application of disaggragate nested logit model to railway station and access mode choice.* Paper presented at the The 4th World Conference on Transport Resarch Vancouver, BC, Canada.

Harris, C., Cooper, B., & Whitfield, S. (1998). *The Travel Effects of Park and Ride* [Report for UnitedKingdom Government Department of Environment, Transport and the Regions]. Epsom, UK:

Heggie, I. G., & Papoulias, D. (1975). *Forecasting the number and characteristics of park-and-ride users* [working paper]. University of Oxford

Heggie, I. G., & Papoulias, D. (1976). *Operational performance of park-and-ride: objectives and achivements* [working paper]. University of Oxford

Hensher, D. A. (2004). Accounting for stated choice design dimensionality in willingness to pay for travel time savings. *Jounal of Transport Economics and Policy, 28*, 425-446.

Hensher, D. A. (2006). How do respondents process stated choice experiments? Attribute consideration under varying information load. *Journal of Applied Econometrics, 21*(6), 861-878. http://dx.doi.org/10.1002/jae.877

Hensher, D. A. (2010). Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B: Methodological, 44*(6), 735-752. http://dx.doi.org/10.1016/j.trb.2009.12.012

Hensher, D. A., & Greene, W. H. (2002). The mixed logit model: the state of practice. Retrieved from https://www.researchgate.net/profile/William_Greene2/publication/242076892_The_Mixed_Logit_Model_The_State_of_Practice_and_Warnings_for_the_Unwary/links/004635180450fdbb60000000/The-Mixed-Logit-Model-The-State-of-Practice-and-Warnings-for-the-Unwary.pdf

Hensher, D. A., & Greene, W. H. (2003). The Mixed Logit model: The state of practice. *Transportation, 30*(2), 133-176. http://dx.doi.org/10.1023/a:1022558715350

Hensher, D. A., Greene, W. H., & Li, Z. (2011). Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel time savings. *Transportation Research Part B: Methodological, 45*(7), 954-972. http://dx.doi.org/10.1016/j.trb.2011.05.023

Hensher, D. A., & McLeod, P. B. (1977). Towards an integrated approach to the identification and evaluation of the transport determinants of travel choices. *Transportation Research, 11*(2), 77-93. http://dx.doi.org/10.1016/0041-1647(77)90106-X

Hensher, D. A., Rose, J. M., & Collins, A. T. (2011). Identifying commuter preferences for existing modes and a proposed Metro in Sydney, Australia with special reference to crowding. *Public Transport, 3*(2), 109-147. http://dx.doi.org/10.1007/s12469-010-0035-4

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). *Applied Choice Analysis: a Primer*. Cambridge: the University Press in the United Kindom.

Hensher, D. A., Stopher, P., & Bullock, P. (2003). Service quality—Developing a service quality index in the provision of commercial bus contracts. *Tranportation Research Part A: Policy and Practice, 37*(6), 18. http://dx.doi.org/10.1016/S0965-8564(02)00075-7

Hess, S., & Rose, J. M. (2009). Some lessons in stated choice survey design. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.678.5504&rep=rep1&type=pdf

Hilvert, O., Toledo, o., & Bekhor, S. (2012). Framework and Model for Parking Decisions. *Journal of the Transportation Research Board, 2319*, 30-38. http://dx.doi.org/10.3141/2319-04

Hong Kong Transport Department (HKTD). (1995). *Parking Demand Study Final Report* Hong Kong SAR Government, Hong Kong.

Horlin, C., Falkmer, M., Fitzgerald, P., Leung, D., Ordqvist, A., & Falkmer, T. (2013). The influence of static versus naturalistic stimuli on face processing in children with and without Asperger syndrome or high-functioning autism. *Research in Autism Spectrum Disorders, 7*(12), 1617-1624. http://dx.doi.org/10.1016/j.rasd.2013.09.012

Horner, M. W., & Groves, S. (2007). Network flow-based strategies for identifying rail park-and-ride facility locations. *Socio-Economic Planning Sciences, 41*(3), 255-268. http://dx.doi.org/10.1016/j.seps.2006.04.001

Horner, M. W., & Grubesic, T. H. (2001). A GIS-based planning approach to locating urban rail terminals. *Tranportation 28*, 55-77. http://dx.doi.org/10.1023/A:100520401

Hunt, J. D., & Teply, S. (1993). A nested logit model of parking location choice. *Transportation Research Part B: Methodological, 27*, 253-265. http://dx.doi.org/10.1016/0191-2615(93)90035-9

Huntley, P. (1993). Park and ride: Strategic modal shift or cheap overflow parking. *Local Transport Today*, 11-12. Retrieved from https://scholar.google.com.au/scholar?hl=en&as_sdt=0%2C5&q=Park+and+ride%3A+Strategic+modal+shift+or+cheap+overflow+parking&btnG=

Iida, Y., Akiyama, T., & Uchida, T. (1992). Experimental analysis of dynamic route choice behavior. *Transportation Research Part B: Methodological, 26*(1), 17-32. http://dx.doi.org/10.1016/0191-2615(92)90017-Q

Islam, S. T., Liu, Z., Sarvi, M., & Zhu, T. (2015). Exploring the Mode Change Behavior of Park-and-Ride Users. *Mathematical Problems in Engineering 2015*, 1-9. http://dx.doi.org/10.1155/2015/282750

Ison, S., & Rye, T. (2008). *The implementation and Effectiveness of Transport Demand Management Measures*. 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN 711 Third Avenue, New York: Routledge.

Jackson, B. W., & Jucker, J. V. (1982). An empirical study of travel time variability and travel choice behavior. *Transportation Science, 16*(6), 460-475. http://dx.doi.org/10.1287/trsc.16.4.460

Jong, G. d., Daly, A., Pieters, M., Vellay, C., Bradley, M., & Hofman, F. (2003). A model for time of day and mode choice using error components logit. *Transportation Research Part E: Logistics and Transportation Review, 39*(3), 245-268. http://dx.doi.org/10.1016/S1366-5545(02)00037-6

Jou, R.-C., Kitamura, R., Weng, M.-C., & Chen, C.-C. (2008). Dynamic commuter departure time choice under uncertainty. *Transportation Research Part A: Policy and Practice, 42*(5), 774-783. http://dx.doi.org/10.1016/j.tra.2008.01.017

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263-292. http://dx.doi.org/10.1142/9789814417358_0006

Karamychev, V., & Reeven, P. v. (2011). Park-and-ride: Good for the city, good for the region? *Regional Science and Urban Economics, 41*(5), 455-464. http://dx.doi.org/10.1016/j.regsciurbeco.2011.03.002

Kastrenakes, C. R. (1988). Development of a rail station choice model for NJ transit *Tranportation Research Record, 1162*, 16-21. Retrieved from http://onlinepubs.trb.org/Onlinepubs/trr/1988/1162/1162-003.pdf

Katz, D., & Rahman, M. M. (2010). Levels of overcrowding in bus system of Dhaka, Bangladesh. *Journall of the Tranportation Research Board, 2143*, 6. http://dx.doi.org/0.3141/2143-11

Kim, J.-K., Lee, B., & Oh, S. (2009). Passenger choice models for analysis of impacts of real-time bus information on crowdedness. *Journal of the Tranportation Research Board, 2112*, 7. http://dx.doi.org/10.3141/2112-15

Knight, F. H. (1921). *Risk, Uncertainty and Profit*: Boston, MA:Hart, Schaffner & Marx; Houghton Mifflin Company.

Knight, T. (1974). An approach to the evaluation of changes in travel unreliaiblity: a safety margin hypothesis. *Transporation, 3*, 393-408.

Knight, T. E. (1974). An approach to the evaluation of changes in travel unreliability: A "Safety margin" hypothesis. *Transportation, 3*(4), 393-408. http://dx.doi.org/10.1007/bf00167968

Kono, Y., Uchida, K., & Andrade, K. (2014). Economical welfare maximisation analysis: assessing the use of existing Park-and-Ride services. *Transprotation 41*(4), 839-854. http://dx.doi.org/DOI

10.1007/s11116-013-9498-x

Krasić, D., & Lanović, Z. (2013). Park & Ride facility planning. Retrieved from file:///F:/school%20cumpter/ccm/2018/thesis/paper/JCE_65_2013_2_2_1037_EN.pdf

Lam, T. (2000). Route and scheduling choice under travel time uncertainty. *Tansportaiton Research Record: Journal of the Tranpsortation Research Board, 1725*(1), 71-78. http://dx.doi.org/10.3141/1725-10

Lam, W. H. K., Cheung, C.-Y., & Lam, C. F. (1999). A study of crowding effects at the Hong Kong light rail transit stations. *Transportation Research Part A: Policy and Practice, 33*(5), 401-415. http://dx.doi.org/10.1016/S0965-8564(98)00050-0

Lam, W. H. K., Holyoad, n. M., & Lo, H. P. (2001). How park-and-ride schemes can be successful in eastern asia. *Journal of Urban Planning and Develpment, 127*(2), 63-78. http://dx.doi.org/10.1061/(ASCE)0733-9488(2001)

Leurent, F., & Boujnah, H. (2012). Traffic Equilibrium in a Network Model of Parking and Route Choice, with Search Circuits and Cruising Flows. *Procedia - Social and Behavioral Sciences, 54*, 808-821. http://dx.doi.org/10.1016/j.sbspro.2012.09.797

Leurent, F., & Liu, K. (2009). On seat congestion, pasenger comfort and route choice in urban transit: a network equilibrium assignment model with application to paris*88th Annual Transportation Research Board Meeting held in Washingtong, D.C.*, January 2009 Retrieved from https://hal.archives-ouvertes.fr/hal-00348513/

Li, Z.-C., Lam, W., Wong, S., Zhu, D.-L., & Huang, H.-J. (2007). Modelling park-and-ride services in a multimodal transport network with elastic demand. *journal of the Transportation Resesarch Board, 1994*(1), 101-109. http://dx.doi.org/10.3141/1994-14

Li, Z., & Hensher, D. A. (2011). Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal. *Transport Policy, 18*(6), 880-887. http://dx.doi.org/10.1016/j.tranpol.2011.06.003

Li, z., Hensher, D. A., & Rose, J. M. (2009). Valuation of Travel Time Reliability in an Extended Expected Utility Theory Framework. Retrieved from https://iatbr2009.asu.edu/ocs/custom/resource/Keynote_LiHensherRose.pdf

Li, Z., Hensher, D. A., & Rose, J. M. (2010). Willingness to pay for travle time reliabilty in passenger transport: Areview and some new empirical evidence. *Tranportation Research Part E: Logistics and Transportation Review, 46*(3), 384-403. http://dx.doi.org/10.1016/j.tre.2009.12.005

Li, Z., Tirachini, A., & Hensher, D. A. (2012). Embedding Risk Attitudes in a Scheduling Model: Application to the Study of Commuting Departure Time. *Transporation Science 46*(2), 170-188. http://dx.doi.org/10.1287/trsc.1110.0389

Li, Z. C., Lam, W. H. K., Wong, S. C., Zhu, D., & Huang, H. (1994). Modeling Park-and-Ride Services in a Multimodal Transport Network with Elastic Demand. *Jounrnal of the Transporation Reserach Board, 14*, 101-109. http://dx.doi.org/10.3141/1994-14

Liou, P. S., & Talvitie, A. P. (1974). Disaggregate access mode and station choice models for rail trips. *Journal of the Transportation Research Board, 526*, 42-65. Retrieved from https://shareok.org/bitstream/handle/11244/3649/7406969.PDF?sequence=1

List, J. A., & Gallet, C. A. (2001). What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values? *Environmental and Resource Economics, 20*(3), 241-254. http://dx.doi.org/10.1023/a:1012791822804

Little, J., & Berrens, R. (2004). Explanining disparites between acutal and hypothetical stated values: further investigation suing meta-analysis. *Economics Bulletin, 3*(6), 1-13. Retrieved from http://www.economicsbulletin.com/2004/volume3/EB−03C90005A.pdf

Liu, T.-L., Huang, H.-J., Yang, H., & Zhang, X. (2009). Continuum modeling of park-and-ride services in a linear monocentric city with deterministic mode choice. *Transportation Research Part B: Methodological, 43*(6), 692-707. http://dx.doi.org/10.1016/j.trb.2009.01.001

Loomis, J. (2011). What's to know about hypothetical bias in stated preference valueation studies? *jounal of economic surveys, 25*(2), 363-370. http://dx.doi.org/10.1111/j.1467-6419.2010.00675.x

Louviere, J. J., Beavers, L. L., Norman, K. L., & Stetzer, F. C. (1973). *Theory, methodology, and findings in mode choice behavior*. The University of Lowa, Lowa City:

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated Choice Methods: Analysis and Application* (Vol. null). the Unite Kingdom: Cambridge Univerity Press.

Lu, H. F., A.S., & Wardman, M. (2008). Amending the incentive for strategic bias in stated preference studies: a case study in users' valuation of rolling stock. *Journal of the Transportation Research Board, 2049*, 128-135. http://dx.doi.org/10.3141/2049-15

Lundberg, U. (1976). Urban commuting:Crowdedness and catecholamine excretion *Journal of Human Stress, 2*(3), 7. http://dx.doi.org/10.1080/0097840X.1976.9936067

Lythgoe, W. F., & Wardman, M. (2004). Modelling passenger demand for parkway rail stations. *Transportation 31*(2), 125-151. Retrieved from https://link.springer.com/content/pdf/10.1023%2FB%3APORT.0000016468.14434.24.pdf

Lythgoe, W. F., Wardman, M., & Toner, J. P. (2004). Enhancing rail passenger demand models to examine station choice and access to the rail network*the AET European Transport Conference, held in PTRC, London*, Retrieved from http://www.its.leeds.ac.uk

Mackie, P. J., Wardman, M., A.S.Fowkes, G.Whelan, J.Nelthorp, & Bates, J. (2003). *Value of travel time savings in the UK-Summary Report*. Retrieved from http://www.its.leeds.ac.uk/downloads/VOTSummary.pdf

Mahmoud, M. S., Habib, K. N., & Shalaby, A. (2014). Park-and-Ride access station choice model for cross-regional commuting case study of greater Toronto and Hamilton Area, Canada. *Journal of the Transportation Research Board, 2419*, 92-100. http://dx.doi.org/10.3141/2419-09

Mahudin, N. D. M., Cox, T., & Griffiths, A. (2011). Modelling the spillover effects of rail passenger crowding on individual welll being and organistational behaviour. *WIT Transactions on the built environment, 116*, 227-238. Retrieved from https://www.witpress.com/Secure/elibrary/papers/UT11/UT11020FU1.pdf

Mahudin, N. D. M., Cox, T., & Griffiths, A. (2012). Measuring rail passenger crowding: Scale development and psychometric properties. *Transportation Research Part F: Traffic Psychology and Behaviour, 15*(1), 38-51. http://dx.doi.org/10.1016/j.trf.2011.11.006

Martinovich, P. (2008). The integration of rail transit and land use in western australia*Conference on Railway Engineering (CORE),, held in Perth, Australia* Railway technical society of australiasia: Engieners asutralia

Matthieu, L. D., & Ben-Akiva, M. (2014). Risk Aversion in Travel Mode Choice with Rank-Dependent Utility. *Mathematical Population Studies, 21*(4), 189-204. http://dx.doi.org/10.1080/08898480.2013.836415

McFadden, D. (1978). Modeling the choice of residential location. *Journal of Tansportaiton Research Record* (673), 72-77. Retrieved from http://onlinepubs.trb.org/Onlinepubs/trr/1978/673/673-012.pdf

McFadden, D., Talvitie, A., Cosslett, S., Hasan, I., Johnson, M., Reid, A., & K., T. (1977). *Demand Model Estimation and Validation*. Retrieved from https://eml.berkeley.edu/wp/utdfp/vol5/front.pdf

McFadden, D., & Train, K. (2000). Mixed MNLl models for discrete response. *Journal of Applied Econometrics, 15*, 447-470. http://dx.doi.org/10.1002/1099-1255(200009/10)

Meek, S., Ison, S., & Enoch, M. (2009). Stakeholder perspectives on the current and future roles of UK bus-based Park and Ride. *Journal of Transport Geography, 17*(6), 468-475. http://dx.doi.org/10.1016/j.jtrangeo.2008.10.007

Meek, S., Ison, S., & Enoch, M. (2011). Evaluating alternative concepts of bus-based park and ride. *Transport Policy, 18*(2), 456-467. http://dx.doi.org/10.1016/j.tranpol.2010.09.006

Menaske, E., & Guttman, J. M. (1986). Uncertainty Continuous Modal Split, and Value of Travel Time in Israrl. *Journal of Transport Economics and Policy, 20*(3), 369-375. Retrieved from https://www.jstor.org/stable/20052794

Miller, E. J., & Cheah, L. S. (1991). *Development of an Operational Peak-Period Mode Split Model for Metropolitan Toronto, Volume I: Final Project Report*. Toronto, Canada, :

Miller, M. E., Hui, S. L., & Tierney, W. M. (1991). Validation Techniques for Logistic Regression Models. *Statistics in Medicine* (10), 1213-1226. http://dx.doi.org/10.1002/sim.4780100805

Mingardo, G. (2013). Transport and environmental effects of rail-based Park and Ride: evidence from the Netherlands. *Journal of Transport Geography, 30*, 7-16. http://dx.doi.org/10.1016/j.jtrangeo.2013.02.004

Mogridge, M. J. H. (1990). *Travel in Towns: Jam Yesterday, Jam Today and Jam Tomorrow?* Basingstoke:Macmillan

Morikawa, T., Ben-Akiva, M., & Yamada, K. (1991). Forecasting Intercity Rail Ridership Using Revealed Preference and Stated Preference Data *Journal of Transportation Research Record 1328* 30-35. Retrieved from https://core.ac.uk/download/pdf/59190063.pdf

Mørkbak, M. R., & Nordström, J. (2009). The Impact of Information on Consumer Preferences for Different Animal Food Production Methods. *Journal of Consumer Policy, 32*(4), 313-331. http://dx.doi.org/10.1007/s10603-009-9106-9

Murphy, J. J., Allen, P. G., Stevens, T. H., & Weatherhead, D. (2005). A Meta-analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics, 30*(3), 313-325. http://dx.doi.org/10.1007/s10640-004-3332-z

Murtagh, F., & Contreras, P. (2011). *Methods of Hierarchical Clustering*.

Neumann, J. V., & Morgenstern, O. (1994). *Theory of Games and Economic Behaviour*: Princeton University Press.

Nicholson, A., & Du, Z.-P. (1997). Degradable transportation systems: An integrated equilibrium model. *Transportation Research Part B: Methodological, 31*(3), 209-223. http://dx.doi.org/10.1016/S0191-2615(96)00022-7

Noel, E. C. (1988). Park-and-Ride: Alive, Well, and Expending in the United States. *Journal of Urban Panning 114*(1), 2-13. http://dx.doi.org/10.1061/(ASCE)0733-9488(1988)

Noland, R. B., & small, K. A. (1995). Travel-Time Uncertainty, Depature Time Choice, and the Cost of the Morning Commute. *the journal of transport reserach record, 1493*, 150-158. Retrieved from http://128.200.36.2/its/publications/papers/ITS/UCI-ITS-WP-95-1.pdf

Olaru, D., Smith, B., Xia, J. C., Chen, C., Lin, T., Ying, H., & Han, R. (2013). Ticket to ride": factors affecting park-and-ride travel in Perth, WA*The 42nd European Transport Conference, held in Frankfurt, German*, Retrieved from http://hdl.handle.net/20.500.11937/47658

Olaru, D., Smith, B., Xia, J. C., & Lin, T. G. (2014). Travellers' Attitudes Towards Park-and-Ride (PnR) and Choice of PnR Station: Evidence from Perth, Western Australia*XVIII Congreso Panamericano de Ingeniería de Tránsito, Transporte y Logística (PANAM 2014), held in Espana*, 2014/12/19/ (pp. 101-110). http://dx.doi.org:10.1016/j.sbspro.2014.12.190

Orme, B. K. (2005). *Getting started with conjoint analysis: stategies for product design and pricing reserach* LLC: Research Pulishers.

Ortúzar, J. d. D., & Willumsen, L. G. (2011). *Modelling Transport, 4th Edition*. Chichester, UK: John Wiley &Sons.

Outram, V. E., & Thompson, E. (1977). *Driver Route Choice.* Paper presented at the PTRC Annual Meeting, London

Pan, X., & Zuo, Z. (2013). Analysis of travel choice behaviour under congestion pricing based on cumulative prospect theory*1st International Conference on Public-Private Partnerships held in Austin,*

Parkhurst, G. (1995). Park and ride: Could it lead to an increase in car traffic? *Transport Policy, 2*(1), 15-23. http://dx.doi.org/10.1016/0967-070X(95)93242-Q

Parkhurst, G. (1996). The Economic and Modal-split Impacts of Short-Range Park and Ride Schemes: Evidence from Nine UK Cities. Retrieved from https://www.researchgate.net/publication/228611988_The_Economic_and_Modal-split_Impacts_of_Short-range_Park_and_Ride_Schemes_Evidence_from_Nine_UK_Cities

Parkhurst, G. (2000). Influence of bus-based park and ride facilities on users' car traffic. *Transport Policy, 7*(2), 159-172. http://dx.doi.org/10.1016/S0967-070X(00)00006-8

Parkhurst, G., & Meek, S. (2014). The Effectiveness of Park-and-Ride as a Policy Measure for more Sustainable Mobility. In S. Ison & C. Mulley (Eds.), *Transport and Sustainability* (Vol. 5): Emerald Group Publishing Limited.

Pearson Education Ltd. (2014). Park and Ride ... good or bad Retrieved from http://www.contentextra.com/bacconline/bacContentFiles/essFiles/MonthlyUpdates/pdfs/August---2014.pdf

Phang, S.-Y., & Toh, R. S. (2004). Road Congestion Pricing in Singapore: 1975 to 2003. *Transportation Journal, 43*(2), 16-25. Retrieved from http://www.jstor.org/stable/20713563

Pickett, M. (2005). PARK & RIDE. *BPA PARKING PRACTICE NOTES 2,*

Polak, J., & Axhausen, K. (1990). *Parking Search Behaviour: A Review of Current Research and Future Prospects*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.622.6739&rep=rep1&type=pdf

Polydoropoulou, A., & Ben-Akiva, M. (2001). Combined Revealed and Stated Preference Nested Logit Access and Mode Choice Model for Multiple Mass Transit Technologies. *Journal of the Transportaion Research Board, 1771*, 38-45. http://dx.doi.org/10.3141/1771-05

Population 2018. (2018). Perth Population. Retrieved from http://www.population.net.au/perth-population/

Pratt, J. W. (1964). Risk Aversion in the Small and In the Large *Econometrica, 32*(1/2), 122-136. Retrieved from http://www.jstor.org/stable/1913738

Public Transport Authority. (2011-2017). Annual Reports. Retrieved from http://www.pta.wa.gov.au/about-us/priorities-and-performance/annual-reports

Public Transport Authority. (2012-2017). Annual reports. Retrieved from http://www.pta.wa.gov.au/about-us/priorities-and-performance/annual-reports

Qin, H., Guan, H., & Zhang, G. (2012). Analysis of the Travel Intent for Park and Ride Based on Perception. *Discrete Dynamics in Nature and Society, 2014*, 1-14. http://dx.doi.org/10.1155/2012/516197

Quiggin, J. (1993). *Generalized expected utility theory: the rank-dependent model.* Netherlands Kluwer Academic Publisher.

Rasmidatta, I. (2006). *Mode choice models for long distance travel in USA* the degree of doctor of philosophy The University of Texas at Arlington in Partial Fulfillment Retrieved from https://rc.library.uta.edu/uta-ir/bitstream/handle/10106/49/umi-uta-1429.pdf?sequence=1&isAllowed=y

Rasouli, S., & Timmermans, H. (2014). Applications of theories and models of choice and decision-making under conditions of uncertainty in travel behavior research. *Travel Behaviour and Society, 1*(3), 79-90. http://dx.doi.org/10.1016/j.tbs.2013.12.001

Raveau, S., Muñoz, J. C., & Grange, L. d. (2011). A topological route choice model for metro. *Tranportation Research Part A: Policy and Practice 45*(2), 10.

Rietveld, P. (2010). The impact of travel time unrelaiblity on the choice of rail, access mode and departure station *the 12th WCTR, held in Lisbon,Portugal,* Robert, J., & Spillar, P. E. (1997). *Park-and-Ride Planning and Design Guidelines*. One Penn Plaza, New York, New York 10119: Retrieved from http://www.sciencedirect.com/science/article/pii/S0038012106000115#bib1

Robinson, C. (1994). Bristol park and ride to assist in reducing traffic congestion. *Joual of institution of highways and transportation, 41*(3), 14-15.

Rose, J. M., & Bliemer, M. C. (2005). Constructin efficient choice experiements. Retrieved from http://ws.econ.usyd.edu.au/itls/wp-archive/itls_wp_05-07.pdf

Rose, J. M., & Bliemer, M. C. J. (2005). Stated preference experimental design strategies. In D. A. Hensher & K. J. Button (Eds.), *Handbook of Transport Strategy, Policy and Institutions* (First ed., Vol. 6). UK: Emerald Group Publishing Limited.

Rose, J. M., & Bliemer, M. C. J. (2009). Constructing Efficient Stated Choice Experimental Designs. *Transport Reviews, 29*(5), 587-617. http://dx.doi.org/10.1080/01441640902827623

Rubsinstein, A. (2006). *Lecture Notes in Microeconomic Theory* 41 William Street, Princeton, New Jersery: Princeton University Press

Savage, L. J. (1972). *The Foundations of Statistics*. United States: Dover Publications, Inc.

Schmöcker, J.-D., Fonzone, A., Shimamoto, H., Kurauchi, F., & Bell, M. G. H. (2011). Frequency-based transit assignment considering seat capacities. *Tranportation Research Part B: Methodological, 45*(2), 392-408. http://dx.doi.org/10.1016/j.trb.2010.07.002

Seik, F. T. (1997). Experiences from Singapore's park-and-ride scheme (1975–1996). *Habitat International, 21*(4), 427-443. http://dx.doi.org/10.1016/S0197-3975(97)00016-7

Sgurgaikar, M., & Deakin, E. (1927). Study of Park-and -Ride Facilities and Their Use in the San Francisco Bay Area of California. *Journal of the Transportation Reserach Board, 6*

Shao, C., Xia, J. C., Lin, T. G., Goulias, K. G., & Chen, C. (2015). Logistic regression models for the nearest train station choice: A comparison of captive and non-captive stations. *Case Studies on Transport Policy, 3*(4), 382-391. http://dx.doi.org/10.1016/j.cstp.2015.06.002

Shao, H., Lam, W. H. K., & Tam, M. L. (2006). A Reliability-Based Stochastic Traffic Assignment Model for Network with Multiple User Classes under Uncertainty in Demand. *Networks and Spatial Economics, 6*(3), 173-204. http://dx.doi.org/10.1007/s11067-006-9279-6

Shirgaokar, M., & Deakin, E. (2005). Study of park-and-ride facilites and their use in the San Francisco Bay Area of California. *Transportation Research Record 1927, 1*, 46-54. http://dx.doi.org/10.3141/1927-06

Shmueli, D., Salomon, I., & Shefer, D. (1996). Neural network analysis of travel behavior: Evaluating tools for prediction. *Transportation Research Part C: Emerging Technologies, 4*(3), 151-166. http://dx.doi.org/10.1016/S0968-090X(96)00007-1

Shukla, N., Ma, J., Wickramasuriya Denagamage, R., Huynh, N. N., & Perez, P. (2015). *Tour-based travel mode choice estimation based on data mining and fuzzy techniques*. Paper presented at the International Symposium for Next Generation Infrastructure (ISNGI 2014), United Kingdom: University College London (UCL)

Sikka, N. (2012). *Understanding travelers' route choice behavior under uncertainty*. The University of Iowa The University of Iowa The University of Iowa

Silva, E. J., Jr, A. S. B., Oliveira, L. S., Enembreck, F., Sabourin, R., & Koerich, A. L. (2017). A Two-Step Cascade Classification Method. Retrieved from https://www.etsmtl.ca/Unites-de-recherche/LIVIA/Recherche-et-innovation/Publications/Publications-2017/Silva_ijcnn_2017.pdf

Siu, B. W. Y., & Lo, H. K. (2014). Punctuality-based route and departure time choice. *Transportmetrica A: Transport Science, 10*(7), 585-621. http://dx.doi.org/10.1080/18128602.2012.751681

Small, K. A. (1978). The scheduling of consumer activities: work trips*the Econometric Society Annual Meeting, held in Chicago,*

Small, K. A. (1987). A Discrete Choice Model for Order Alternatives. *Econometrica, 55*(2), 409-424. Retrieved from http://www.socsci.uci.edu/~ksmall/Ordered%20GEV%20working%20paper.pdf

Small, K. A., & Verhoef, E. T. (2007). *The economics of urban transportation*: Routledge.

Small, K. A., Winston, C., & Yan, J. (2005). Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability. *Econometrica, 73*(4), 1367-1382. http://dx.doi.org/10.1111/j.1468-0262.2005.00619.x

Stott, H. P. (2006). Cumulative Prospect Theory's Fucntional Menagerie. *J risk Uncertainty 32*(2), 101-130. http://dx.doi.org/ 10.1007/s11166-006-8289-6

Street, D. J., Burgess, L., & Louviere, J. J. (2005). Quick and easy choice stes: constructin optimal and nearly optimal stated choice experiments *International Journal of Research in Marketing, 22*(null), 459-470. http://dx.doi.org/ 10.1016/j.ijresmar.2005.09.003

Sumalee, A., Tan, Z., & Lam, W. H. K. (2009). Dynamic stochastic transit assignment with explicit seat allocation model. *Tranportation Research Part B, 43*(8-9), 17. http://dx.doi.org/10.1016/j.trb.2009.02.009

Sundstrom, E. B., Busby, P., & Asmus, C. (1975). An experimental study of crowding:effects of room size, instrusion, and goal blocking on nonverbal behaviour, self-disclosure, and self-reported stress. *Journal of Personality and social Psychology, 32*, 645-654.

Thompson, K., L. Hirsh, S.Muller, & S.Rainbird. (2012). *A Socio-economic study of carriage and platform crowding in the Australian railway industry: Final report*. Brisban,Australia: Retrieved from https://www.researchgate.net/publication/264558266_A_socio-economic_study_of_carriage_and_platform_crowding_in_the_Australian_railway_industry_Qualitative_research_summary

Tian, H., & Gao, S. (2013). A Process Model for Route Choice in Risky Traffic Networks. *Procedia - Social and Behavioral Sciences, 80*, 764-778. http://dx.doi.org/10.1016/j.sbspro.2013.05.041

Timmermans, H. (2010). On the (Ir)relevance of Prospect Theory in Modelling Uncertainty in Travel Decisions. *EJTIR, 10*(4), 368-384.

Tirachini, A., Hensher, D. A., & Rose, J. M. (2013). Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. *Transportation Research Part A: Policy and Practice, 53*, 36-52. http://dx.doi.org/10.1016/j.tra.2013.06.005

Topp, H. H. (1995). A critical review of current illusions in traffic management and control. *Transport Policy, 2*(1), 33-42. http://dx.doi.org/10.1016/0967-070X(95)93244-S

Train, K. (2003). *Discrete Choice Methods with Simulation*. Shaftesbury Rd, Cambridge CB2 8RU, UK: Cambridge University Press.

Truong, T. P., & Hensher, D. A. (1985). Measurement of Travel Time Values and Opportunity Cost from a Discrete-Choice Model. *The Economic Journal, 95*(378), 438-451. Retrieved from https://www.jstor.org/stable/2233219

Turner, S., Corbett, E., O'Hara, R., & & White, J. (2004). Health and safety effects of rail crowding – Hazard identification. Retrieved from http://www.rssb.co.uk/SiteCollectionDocuments/pdf/reports/Research/T307_rpt_final.pdf

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323. Retrieved from http://psych.fullerton.edu/mbirnbaum/psych466/articles/tversky_kahneman_jru_92.pdf

Vijayakumar, N., El-Geneidy, A. M., & Patterson, Z. (2011). Driving to Suburban Rail Stations: Understanding Variables that Affect Driving Distance and Stantion Demand *Transportation Reserach Record: Journal of the Transportation Resarch Board, 2219*, 97-103. http://dx.doi.org/10.3141/2219-12

Vincent, M., & Hamilton, B. A. (2007). *Park and ride :Characteristics and demand forecasting* Retrieved from https://www.nzta.govt.nz/assets/resources/research/reports/328/docs/328.pdf

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. 41 William Street,Princeton, NJ: Princeton University Press.

Wang, D., Liao, F., Gao, Z., & Timmermans, H. (2018). *A scheduling utility budget–based user equilibrium model under travel time uncertainty*. Paper presented at the Transportation Research Board 97th Annual Meeting, Washington DC, United States

Wang, J. Y. T., Yang, H., & Lindsey, R. (2004). Locating and pricing park-and-ride facilities in a linear monocentric city with deterministic mode choice. *Transportation Research Part B: Methodological, 38*(8), 709-731. http://dx.doi.org/10.1016/j.trb.2003.10.002

Wardman, M. (1997). Inter-urban rail demand, elasticities and competition in Great Britain: Evidence from direct demand models. *Transportation Research Part E: Logistics and Transportation Review, 33*(1), 15-28. http://dx.doi.org/10.1016/S1366-5545(97)83195-X

Wardman, M., & Whelan, G. A. (1999). *Using Geographical information systems to improve rail demand models*. Retrieved from http://gtr.rcuk.ac.uk

Whelan, G., & Johnson, D. (2004). Modelling the impact of alternative fare structures on train overcrowding. *Transpoort Management, 2*(1), 51-58. http://dx.doi.org/10.1016/j.ijtm.2004.04.004

Willett, P. (1984). Evaluation of Relocation Clustering Algorithms for the Automatic Classification of Chemical. *J. Chem. Inf. Comput. Sci, 24*(1), 29-33. http://dx.doi.org/10.1021/ci00041a006

Williams, R. (1999). Oxford's park-and-ride system. *Municipal Engineer, 133*(3), 127-135. http://dx.doi.org/10.1680/imuen.1999.31935

Wiseman, N., Bonham, J., Mackintosh, M., Straschko, O., & Xu, H. (2012). Park and ride: An Adelaide case study *Road & Transport Research 21*(1), 39-52. Retrieved from http://hdl.handle.net/2440/76438

Wittink, D. R., Krishnamurthi, L., & Nutter, J. B. (1982). Comparing Derived Importance Weights Across Attributes. *Journal of Consumer Research, 8*, 471-474. Retrieved from https://www.jstor.org/stable/pdf/2489037.pdf

Wittink, D. R., Krishnamurthi, L., & Reibstein, D. J. (1989). The Effects of Differences in the Number of Attribute Levels on Conjoint Results. *Marketing Letters, 2*, 113-123. http://dx.doi.org/10.1007/BF00435295

Wong, H.-K., & Sussman, J. M. (1973). Dynamic travel time estimation on highway networks. *Transportation Research, 7*(4), 355-370. http://dx.doi.org/10.1016/0041-1647(73)90023-3

Xie, C., Lu, J., & Parkany, E. (2003). Work Travel Mode Choice Modeling Using Data Mining: Decision Trees and Neural Networks *Journal of the Transportaion Research Board, 1854*, 50-61. http://dx.doi.org/10.3141/1854-06

Xiong, P., & Yang, D. Y. (2008). Application of disaggregated model in park and ride choice behavior. *Journal of Shenzhen University Science and Engineering, 25*(2), 206-210.

Yang, F., Chen, L., Chen, Y., Luo, X., & Ran, B. (2014). An Empirical Study of Parameter Estimation for Stated Preference Experimental Design. *Mathematical Problems in Engineering, 2014*, 1-12. http://dx.doi.org/10.1155/2014/292608

Yang, H., & Huang, H.-J. (1997). Analysis of the time-varying pricing of a bottleneck with elastic demand using optimal control theory. *Transportation Research Part B: Methodological, 31*(6), 425-440. http://dx.doi.org/10.1016/S0191-2615(97)00005-2

Yang, J., & Jiang, G. (2014). Development of an enhanced route choice model based on cumulative prospect theory. *Transportation Research Part C: Emerging Technologies, 47*, 168-178. http://dx.doi.org/10.1016/j.trc.2014.06.009

Young, M., & Blainey, S. (2018). Railway station choice modelling: a review of methods and evidence. *Transport Reviews, 38*(2), 232-251. http://dx.doi.org/10.1080/01441647.2017.1326537

Zhang, T., Ramakrishaman, R., & Livny, M. (1996). BIRCH: An Efficient Data Clustering Method for Very Large Databases. *ACM SIGMod Record, 25*(2), 104-114. http://dx.doi.org/10.1145/233269.233324

Zheng Li, & Hensher, D. A. (2013). Crowding in Public Transport: A Review of Objective and Subjective Measures. *Journal of public transportation, 16*(2), 107-134. http://dx.doi.org/10.5038/2375-0901.16.2.6

APPENDIX

APPENDIX A PERMISSION FROM CO-AUTHORS

Chunmei Chen <chun.mei.chen2@gmail.com> 8 Aug 2018, 13:07 (10 days ago)
to Cecilia, Brett, Renlong, Doina, John

Dear Co-authors,

I am going to put three of our published conference papers into my PhD thesis, which is going to submit soon. Could I please get your permission here to put them into my thesis? Those are:

Chen, C.; Xia, J.; Smith, B.; Han, R. (2014). "Development of a conceptual framework for modelling train station choice under uncertainty for park and ride users". Paper presented at the TRB 93rd Annual Meeting, USA
Chen, C.; Xia, J.; Smith, B.; Olaru, D.; Taplin, J.; Han, R. (2015). "Influence of parking on train station choice under uncertainty for park and ride users". Paper presented at the 6th International Conference on Applied Human Factors and Ergonomics (AHFE2015) and the Affiliated Conferences, AHFE 2015, USA
Chen, C.; Xia, J.; Smith, B.; Olaru, D.; Taplin, J.; Han, R. (2016). "Influence of travel time variability on train station choice for park-and-rider users". Paper presented at World Conference on Transport Research - WCTR 2016 Shanghai, China.

Regards,

Chunmei Chen

Han, Renlong 8 Aug 2018, 13:33 (10 days ago)
to me, Cecilia, Brett, Doina, John

Hi Chunmei,

I do not have any problem for my side.

Regards
Dr Renlong Han
Team Leader Transport Modelling | Integrated Transport Planning | Department of Transport
140 William Street, Perth WA 6000
Tel: (08) (08) 6551 6610 | Fax: (08) 6551 6492
Email: Renlong.Han@transport.wa.gov.au | Web: www.transport.wa.gov.au

Department of Transport

Clear Direction • Fresh Thinking • Excellent Service • Great People

---

Doina Olaru 8 Aug 2018, 14:12 (10 days ago)
to Renlong, me, Cecilia, Brett, John

No problem Chunmei.
All the best,
Doina

Brett Smith 8 Aug 2018, 14:35 (10 days ago)
to Renlong, Cecilia, John, Doina, me

Okay with me.

Brett

John Taplin 8 Aug 2018, 14:58 (10 days ago)
to me

That's fine with me Chunmei. You certainly have the right to include these in your thesis.

Regards John Taplin
Emeritus Professor John H E Taplin
Senior Honorary Research Fellow
Transport and Logistics
Business School M261
The University of Western Australia

Cecilia Xia 8 Aug 2018, 20:06 (10 days ago)
to Renlong, Brett, Doina, me, John

Dear Chunmei

It is fine with me.

Cheers,
Cecilia

Sent from my iPhone

On 8 Aug 2018, at 1:33 pm, Han, Renlong <Renlong.Han@transport.wa.gov.au> wrote:

APPENDIX B STATIONS' FACILITIES SURVEY QUESTIONNIARES

APPENDIX B-1 The Basic Information of Train Stations on Armadale Line (Example only)

Number	Name of station	Platform																Park and Ride		Bike 'n'Ride			Bus	Rail
		Ticket machine	Ticket checking machine	emergency	lighting	Public telephone	information	chairs		Undercover waiting room	Toilet	office	Security personal	cleaner	Disable path/lift	Convenience store	Water foundation	unlocked	Locked	Have / not	room	number		
								Have/not	number															
1	Perth																							
2	McIver																							
3	Claserbrook																							
4	East perth																							
5	Mount																							
6	Lawley																							
7	Meltham																							
8	Bayswater																							
9	Ashfield																							
10	Bassendean																							
11	Success Hill																							
12	Guildford																							
13	East Guildford																							
14	Woodbridge																							
15	Midland																							

APPENDIX B-2 The Information about P&R Facilities on Armadale Line (Example only)

Number	Name of station	Investigating time	Legal bays										Illegal bays
			The number stopping	Number of Disable bays	The number of taxi	The number of PTA bays	The number of M/C	Locked	unlocked	Dv	lighting	charge	
1	Perth												
2	McIver												
3	Claserbrook												
4	East perth												
5	Mount												
6	Lawley												
7	Meltham												
8	Bayswater												
9	Ashfield												
10	Bassendean												
11	Success Hill												
12	Guildford												
13	East Guildford												
14	Woodbridge												
15	Midland												

APPENDIX B-3 Land Use around Train Stations on Armadale Line (Example only)

Number	Name of station	shops	Social facilities	Recreational facilities	Residential	factory	restaurant	wholesale	tafe	Lunch bar	school	childcare	Fast food	market	supermarket	Library	company	Food service	Gas station	center	park	Park lots	hostiptal
1	Perth																						
2	McIver																						
3	Claserbrook																						
4	East perth																						
5	Mount																						
6	Lawley																						
7	Meltham																						
8	Bayswater																						
9	Ashfield																						
10	Bassendean																						
11	Success Hill																						
12	Guildford																						
13	East Guildford																						
14	Woodbridge																						
15	Midland																						

APPENDIX C TRAIN STATION SURVEY (EXAMPLE ONLY)

THE UNIVERSITY OF WESTERN AUSTRALIA Curtin University Department of Transport Public Transport Authority Planning Western Australia

Train Station Survey *(Example completed)*

Project title: Modelling and Evaluating the Joint Access - Mode and Train Station Choice° **(LP110201150).** *This research has the UWA Human Research Ethics Office approval number RA/4/1/5370 from May 2012.*

Project description: *We are interested in the way people get to the train station.*

The survey will be completed on paper while waiting on the platform (ensure 3 min before train). Select (circle) the appropriate option.

Trips today from: Warwick, **Greenwood**, Bull Creek, Kwinana, Midland, Beckenham, or Claremont; Time: **8:10am**

Q1. How did you travel to the train station? **Car PnR**, Car kiss-and-ride, Bus, Walking, Cycling, Other (please specify) *taxi*

Q2. Where did you start that trip (street, suburb, landmark)? **Glencarry St., Padbury, 6015**

Is that: **Home**, Work, School, Shop, Gym, Pub, Other (please specify)?

Q3. What is the destination of this trip: **St George's Tce, Perth CBD**

Q4. What is the purpose of your travel *(open question)*? **dentist**

For those who used PnR:

Q5. Why did you choose the PnR today? *(open question?)* **lack of parking in the city**

(some pointers: no parking at destination, costly parking at destination, traffic congestion, convenience of riding the train, other activities on the return trip, environmentally responsible)

Q6. Did you easily find parking available when you arrived? Yes/**No**

6b. If No, please specify how long were you searching (min): **5**

Q7. Where did you park? Free parking bays, paid parking bays, residential streets, shopping centre parking, other (please specify)

Q8. How much would you pay to secure a parking bay? **$ 3.00**

Q9. If you could not find an available parking bay at this station, what would you do?

Arrive earlier, Use the bus, Get dropped-off, **Travel to other PnR stations**, Drive to your destination, Other (please specify)

9a. If considering changing the PnR station, please specify which ones: **Stirling**

Perception and attitudinal questions

Q10. What are the most important features/facilities for you at a train station? Please rate them on a scale 1 to 7, where: 1 = Not at all important at all to 7 = Utmost important

Facilities available at a train station:	1	2	3	4	5	6	7
Free park-and-ride bays					☑		

THE UNIVERSITY OF WESTERN AUSTRALIA | Curtin University | Department of Transport | Public Transport Authority | Planning Western Australia

Facilities available at a train station:	1	2	3	4		6	7
Secure/locked car park					☑		
Secure/locked bike storage				☑			
Information on Transperth's timetables and communication of changes							☑
Shops *(convenience stores)*					☑		
Vending machines						☑	
Seating available while waiting for the train					☑		
Seating on the train *(in the carriage)*				☑			
Emergency services							☑
Easy access to the platform for people with disabilities						☑	
Lighting					☑		
Other (please specify):							

Q11. Could you please rate the facilities at this train station? 1 = Very poor to 7 = Excellent

The quality of facilities available and services at this train station is:	1	2	3	4	5	6	7
Free park-and-ride bays						☑	
Secure/locked car park						☑	
Secure/locked bike storage						☑	
Information on Transperth's timetables and communication of changes							☑
Shops *(convenience stores)*			☑				
Vending machines				☑			
Seating available while waiting for the train					☑		
Seating on the train (in the carriage)				☑			
Emergency services						☑	
Easy access to the platform for people with disabilities						☑	
Lighting							☑
Train frequency at present						☑	
Service provided by train station staff					☑		
Arriving according to the timetable						☑	
Other (please specify)							

Thank you for your time.

Gender *(to be recorded by interviewer after the survey)*: **M/F**

Age group *(to be recorded by interviewer after the survey)*: Young adult, **Middle aged**, Senior

Interviewer code: **13** Signature interviewer: .... Questionnaire ID: 0001

APPENDIX D TRAIN STATION CHOICE MAIN SURVEY QUESTIONAIRES (EXAMPLE ONLY)

Example

Below are two train scenarios for a park-and-rid(P&R)commuter travelling from home to work at, say 7:00AM, on a **typical weekday**
All things considered, which scenario appeals to you most?

Factors	Scenario 1	Scenario 2
	Finding parking at the station	
Parking capacity in P&R lot	100 bays	100bays
Parking availability in a P&R lot at 7:00AM	10%	10%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **10** and **20** minutes and only **5%**chance to reach to **20** mins	Between **5** and **10** minutes and only 20**%**chance to reach to **10** mins
	Parking cost	
Parking fee in a P&R lot	$6	$8
Parking fine (if you park illegally, you would be fine)	$80	$60
Frequency of controls for illegal parking	1day /month	2days/month
	Travel time from home to the station	
Usual travel time	10mins	15mins
Travel time on a good day	About **5** mins on **4** days a month	About **11.25**mins on **4** days a month
Travel time on a bad day	About **12.4** mins on 1 days a month	About **18.6**mins on **4** days a month
	Time on the train	
Average journey time	10 minutes	10 minutes
	Crowding on the train	
Crowding	 **All seats taken and half of aisles filled**	 **All seats taken**
Number of days per week on which the trains are too crowded to board	1	0
	Other things	
Safety	Irregular station and car park security patrols	Regular station and car park security patrols
Ticket fare	$2.50	$2.50
Train frequency	Express and all-stop strains at 5minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose	☐	☐

Gender: (··) M (··) F

Age group: (··) 18-24 (··) 25-29 (··) 30-34 (··) 35-39 (··) 40-44 (··) 45-49
(··) 50-54 (··) 55-59 (··) 60-69 (·) 70-79 (··) >80

Annual personal income:

- Less than $10,000 (···) $50,000 - $59,999 (···) $90,000 - $99,999 (···)
- $10,000 - $19,999 (···) $60,000 - $69,999 (···) $100,000 - $150,000 (···)
- $20,000 - $29,999 (···) $70,000 - $79,999 (···) More than $150,000 (···)
- $30,000 - $39,999 (···) $80,000 - $89,999 (···) Prefer not to answer (···)
- $40,000 - $49,999 (···)

Thank you for your time and input

APPENDIX E STATION CHOICE PILOT SURVEY QUESTIONNIARES

APPENDIX E -1 Train Station Choice Pilot Survey Questionnaires 1 (Example only)

THE UNIVERSITY OF WESTERN AUSTRALIA Curtin University Department of Transport Public Transport Authority Planning Western Australia

Train Station Choice Survey of Park and Ride Users *(Example completed)*

Project title: Modelling and Evaluating the Joint Access - Mode and Train Station Choice (**LP110201150**). *This research has the UWA Human Research Ethics Office approval number RA/4/1/5370 from May 2012.*

Project description: *We are interested in the effect of variation of travel time to the station, crowding on the train and parking search time on the station choice for park and ride users.*

Q1. What was the location of your last activity (such as home, work, shopping) immediately before heading towards this train station? ________the nearest cross-street................Suburb.

Q2 How long did it take you to get to this station from the location of your last activity 5mins/10mins/15mins.

2a. If your regular travel time to a station is around 5 minutes, what was experienced variation of travel time to the station for you? Please fill in the worst travel time.........(mins) and the best travel time..........(mins) for you, then go to Q3.

2b. If your regular travel time to a station is around 10 minutes, what was experienced variation of travel time to the station for you? Please fill in the worst travel time..........(mins) and the best travel time ....(mins) for you, then go to Q3.

2c. If your regular travel time to a station is around 15 minutes, what was experienced variation of travel time to the station for you? Please fill in the worst travel time.........(mins) and the best travel time..........(mins) for you, then go to Q3.

Q3. Travel time to catch the train can vary depending on which station you choose. In order to guarantee a seat on trains, you might give up the nearest station from your last stop, and choose a station further away from your destination or a one on a different train line.

3a. If the regular travel time from your origin to the nearest station is about 10mins, what would be acceptable extra travel time for you to choose a station that can guarantee a seat for you?............................mins, then go to Q4.

3b. If the travel time from your origin to the nearest station is about 20mins, what would be acceptable extra travel time for you to choose a station that can guarantee a seat for you?............................mins, then go to Q4.

3c. If the travel time from your origin to the nearest station is about 30mins, what would be acceptable extra travel time for you to choose a station that can guarantee a seat for you?............................mins, then go to Q4.

1

THE UNIVERSITY OF WESTERN AUSTRALIA Curtin University Department of Transport Public Transport Authority Planning Western Australia

Q4. If you think the train is too crowded, how long would be acceptable for you to wait to board? ....____ Mins

Q5. How long would be acceptable for you to search for a parking spot? ____ mins

Q6. Next year, the parking fee in unlocked park and ride facilities in WA will be $2/day. Will you still use park and ride facilities? Yes/No

Q7. How much would you be willing to pay for a parking bay in the future if you are a regular park and ride user? $......./day.

Age group: 18-24 ☐ 25-29 ☐ 30-34 ☐ 35-39 ☐ 40-44 ☐ 45-49 ☐

50-54 ☐ 55-59 ☐ 60-69 ☐ 70-79 ☐ >80

Gender: M/F

Thank you for your time.

Survey date and time: ....(dd/mm/yyyy) and ....(hh:mm)

Survey locations: .... station

Interviewer code: ☐ ......

APPENDIX E - 2 Train Station Choice Pilot Survey Questionnaires 2 (Example only)

The University of Western Australia | Curtin University | Department of Transport | Public Transport Authority | Planning Western Australia

Train Station Choice Survey of Park and Ride Users *(Example completed)*

Project title: Modelling and Evaluating the Joint Access - Mode and Train Station Choice [°] **(LP110201150).** *This research has the UWA Human Research Ethics Office approval number RA/4/1/5370 from May 2012.*

Project description: *We are interested in the effect of variation of travel time to the station, crowding on the train and parking search time on the station choice for park and ride users.*

Q1. What was the time and location of your last activity immediately before heading towards this train station? ....(hh:mm) and ....street...............Suburb. ....

Q2. How long do you usually travelling to the station fromthis location? About 5mins/10mins/ 15mins.

Q3. Travel time to the station can vary daily. Assume your regular travel time to a station is around 5 mins, what would be very acceptable, acceptable, tolerable and intolerable variations of travel time for you? Please circle the minand highest travel time for each scenario.

...

Table 1. Regular travel time to the station is about 5 mins

Q4 → Now assume your regular travel time to a station is around 10 minutes, what would be very acceptable, acceptable, tolerable and intolerable variations of travel time for you? Please circle the lowest and highest travel time for each scenario.

Table 2. Regular travel time to the station is about 10mins

	Variation of travel time to the station
Very acceptable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 Mins
acceptable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 Mins
tolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 Mins
intolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 Mins

Q5 → Now assume your regular travel time to a station is around 15 minutes, what would be very acceptable, acceptable, tolerable and intolerable variations of travel

1

time for you? Please circle the lowest and highest travel time for each scenario. [20 mins is highlighted?]

Table 3. Regular travel time to the station is about 15mins

	Variation of travel time to the station
Very acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
tolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
intolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins

Q6. Travel time to catch the train can vary depending on which station you choose. Assume the travel time from your origin to the nearest station is about 10mins. In order to get a seat easily, you might have to choose a station further away from your destination or a different train line. what would be very acceptable, acceptable, tolerable, and intolerable extra travel time for you if you choose to do so. Please circle the extra travel time for each scenario.

Table 4. Travel time is 10mins from a origin to the nearest station

	Increased travel time to guarantee a seat
Very acceptable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
acceptable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
tolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
intolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins

Q7 Again, assume the travel time from your origin to the nearest station is about 20mins, in order to get a seat easily, you might have to choose a station further away from your destination or a different train line. What would be very acceptable, acceptable, tolerable, and intolerable extra travel time for you if you choose to do so. Please circle the extra travel time for each scenario

Table 5. Travel time is 20mins from a origin to the nearest station

	Extra travel time to guarantee a seat
Very acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins

tolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
intolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins

Q8 → Again, assume the travel time from your origin to the nearest station is about 30mins, in order to get a seat easily, you might have to choose a station further away from your destination or a different train line. What would be very acceptable, acceptable, tolerable, and intolerable extra travel time for you if you choose to do so. Please circle the extra travel time for each scenario.

Table 6. Travel time is 30mins from a origin to the nearest station

	Extra travel time to guarantee a seat
Very acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
tolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
intolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins

Q9. If you give up boarding on the train you think is too crowded, how long during PEAK hours would be very acceptable, acceptable, tolerable, and intolerable for you to wait. Please circle the waiting time for each scenario.

Table 7. Peak hours

	Waiting time
Very acceptable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
acceptable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
tolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
intolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins

Q10. If you give up boarding on the train you think is too crowded, how long during OFF PEAK hours would be very acceptable, acceptable, tolerable, and intolerable for you to wait. Please circle the waiting time for each scenario.

Table 8. Off-peak hours

	Waiting time
Very acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
acceptable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
tolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins
intolerable	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 Mins

Q11. Did you easily find parking available [where? P&R or elsewhere?when you arrived? Yes / No

→ 6a. If yes, how much did you pay for parking? $0/day, $2/day

Q12. Parking search time can vary daily. What would be acceptable, tolerable and intolerable parking search time for you? Please circle the parking search time for each scenario.

	Parking search time
acceptable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
tolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins
intolerable	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 Mins

Q13. Next year, the parking fee in unlocked park and ride facilities in WA will be increased from free to $2/day. Will you still use park and ride facilities? Yes/No

Q14.What is the maximum amount you would pay to guarantee a parking bay in the future? Please circle this value on the following axis.

3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 $/day

Age group: 18-24 ☐ 25-29 ☐ 30-34 ☐ 35-39 ☐ 40-44 ☐ 45-49 ☐

→ 50-54 ☐ 55-59 ☐ 60-69 ☐ 70-79 ☐ >80

Gender: M/F

4

THE UNIVERSITY OF WESTERN AUSTRALIA Curtin University Department of Transport Public Transport Authority Planning Western Australia

Thank you for your time.

Survey date and time: ....(dd/mm/yyyy) and ....(hh:mm)

Survey locations: ....station

Interviewer code: ....

APPENDIX E - 3 Train Station Choice Pilot Survey Questionnaires 3 (Example only)

THE UNIVERSITY OF WESTERN AUSTRALIA — Curtin University — Department of Transport — Public Transport Authority — Planning Western Australia

Station choice Survey of Park and Ride Users *(Example completed)*

Project title: Modelling and Evaluating the Joint Access - Mode and Train Station Choice (**LP110201150**). *This research has the UWA Human Research Ethics Office approval number RA/4/1/5370 from May 2012.*

Project description: *We want to find out how the time spent looking for a car-parking place affects the train station that park-and-ride commuters choose.*

Q1. Where have you just come from (e.g. home, work, visiting friends, shopping) ______________ and its address, ______________ Street and ______________ Suburb? ______________

Q2. How long did it take you to this station? Hours ________ Minutes ________.

Q3. What is the departure time of the train you will catch? Hours ________ (e.g. 7:00am)

Q4. How long will it take you to travel from this station to your destination? Hours ______ Minutes ______.

Q5. What are your average and the worst parking search times (**i.e. *the time spent looking for a parking space, parking the car, and walking to a train platform***)? Please fill in the following fields and circle the choice (i.e., week, month and year).

	Average	The worst
Parking search time	________ Mins	________ Mins
How often?	① 60% per month ② 80% per month ③ 100% per month	Once per month

Q6. If you can't find parking in a Park-and-Ride facility, what are your alternatives? Please tick the appropriate box or complete the answer.

[ ] Parking on the street
[ ] Driving to other stations
[ ] Parking in nearby parking lots such as a shopping centres parking lots
Others ______________.

Q7 If you have experienced unusually long parking search times, why has this occurred? Please tick the appropriate box or complete the answer.

[ ] Left home later than usual
[ ] Traffic congestion
[ ] Delayed by the other activities on the way
Others ______________.

Age group: 18-24 ☐1 25-34 ☐2 35-44 ☐3 45-54 ☐4 55-69 ☐5 70-79 ☐6 >80 ☐7

Gender: M/F

Thank you very much for your time.

The end of the survey

APPENDIX E - 4 Train Station Choice Pilot Survey Questionnaires 4 (Example only)

THE UNIVERSITY OF WESTERN AUSTRALIA | Curtin University | Department of Transport | Public Transport Authority | Planning Western Australia

Train Station Choice Survey of Park and Ride Users *(Example completed)*

Project title: Modelling and Evaluating the Joint Access - Mode and Train Station Choice **(LP110201150).** *This research has the UWA Human Research Ethics Office approval number RA/4/1/5370 from May 2012.*

Project description: *We are interested to find the effect of the variations in travel time to the station, crowding on the train, and the time used to search for parking on the station choice for park and ride users.*

Q1. What was your last location (such as home, work, shopping, and see Figure 1) before heading to this train station to use the park and ride service? Please tell us the address: the nearest intersection between .... (street) and .... (street) and .... (Suburb).

Figure 1: Your last location before heading to the train station may have been your house, or after you dropped the kids at school or after the gym. Please consider this.

Q2. What is your regular, worst and best travel time to get to **this station** from the location you mentioned in Q1?

	Best	Regular	Worst
Travel time	______ Mins	______ Mins	______ Mins
Frequency	Once per (week, month, year)	___ Times per (week, month, year)	Once per (week, month, year)

Q3. In order to guarantee a seat on a train, you might give up the station nearest to the location you mentioned in Q1, and choose a station further away to get to your destination or one that is on a different train line (Figure 2).

Figure 2: This shows a situation where one might choose a station that is not as close as the nearest station, but guarantees them a seat.

Q3a. What is your travel time from the location in Q1 to **your destination** (e.g. workplace, see Figure 2) by using **the nearest station**? ________ (mins)

Q3b. In order to avoid crowding, what would be the travel time from the location in Q1 to **your destination** by using **the further-away station that relates to you and as shown in Figure 2**? ………………………(mins)

Q4. If you think the train is too crowded, how long would be acceptable for you to wait before boarding a train with a seat guarantee? ……………………(mins)

Q5. How long is acceptable for you to search for a parking spot? ________ Minutes

Q6. The parking fee for the next year in unlocked park and ride facilities in WA will be $2/day. Will you still use park and ride facilities? Yes/No

Q7.What is the **highest** amount you would pay for a parking bay next year if you are a regular park and ride user? $……/day.

Age group: 18-24 25-29 30-34 35-39 40-44 45-49

50-54 55-59 60-69 70-79 >80

Gender: M/F

Thank you very much for your time.

The end of the survey

Survey date and time: ………………(dd/mm/yyyy) and…………(hh:mm)

Survey locations: ……………………station

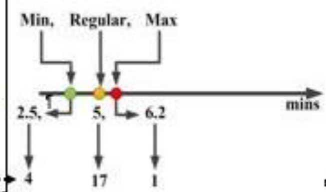
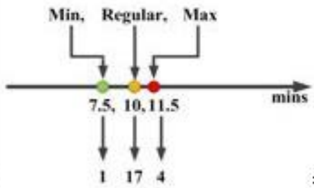
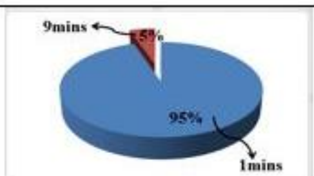
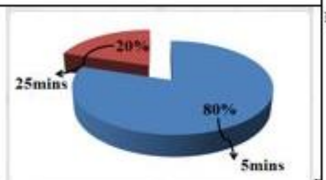
Interviewer code: ………

APPENDIX E - 5 Train Station Choice Pilot Survey Questionnaires 5 (Example only)

Example

Which station would you choose around7:00am?

Please make your choice based on the listed factors presented in this table.

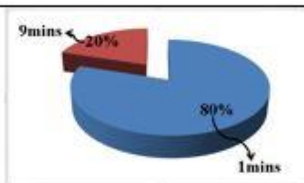
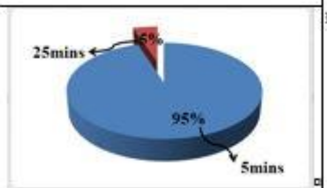
No.	Factors	Train station 1	Train station 2
1	Safety	Irregular security patrolls	Regular security patrolls
2	Ticket fare	$4.5	$2.5
3	Train frequency	Non-express trains stop at this station	Express and non-express trains stop at this station
4	In-vehicle travel time	30mins	10mins
5	Crowding on trains	 25% seats taken ≤2 passengers/m² standing	 100% seats taken 4-6 passengers/m² standing
6	Transfer waiting time	10mins	≤2mins
7	Parking capacity	100 Bays	500 Bays
8	Usual parking availability	 Parking available	 Parking unavailable
9 10 11	Regular travel time Range of travel time (the minimum time – the maximum time) Frequency (times per month)		
12	Parking search time in a month		
Which station would you choose?(tick)		☐	☐

Based on this scenario, please rank the top four factors(number) that contribute to your station choice, from the most important to the least improtant..

Next page

APPENDIX E - 6 Train Station Choice Pilot Survey Questionnaires 6 (Example only)

Example

Assuming you are a park and ride user, which station would you choose? *(around 7:00am)*			
Please make your choice based on the listed factors presented in this table.			
No.	**Factors**	**Train station 1**	**Train station 2**
Travel time to the staiton from your origins			
1	**Regular travel time /frequncy**	15mins/14 days per month	5mins/20 days per month
2	**The minimum travel time /frequncy**	7.5mins/4 days per month	3.75 mins /1 days per month
3	**The maximum travel time/frequncy**	17.25 mins/4 days per month	5.75 mins/1 days per month
In- vehicle travel time			
4	**In-vehicle travel time**	30mins	10mins
Crowding on trains			
5	**Crowding on trains**	 25% seats taken ≤ 2 passengers/m^2 standing	 100% seats taken 4-6 passengers/m^2 stand
6	**How many days in a week is too crowded to board trains ?**	2 days	3 days
Parking			
7	**Parking capacity**	500 Bays	1000 Bays
8	**Usual parking availability**	available	unavailable
9	**Parking search time in a month**	 	
Others			
10	**Safety**	Regular security patrolls	Irregular security patrolls
11	**Ticket fare**	$4.5	$2.5
12	**Train frequncy**	Non-express trains stop at this station[10mins/trains]	Express and non-express trains stop at this station[5mins/trains]
Which station would you choose?(tick one box)		☐	☐
Based on this scenario, please rank the top four factors from the most important to the least improtant……………………………………………………………			

APPENDIX E - 7 Train Station Choice Pilot Survey Questionnaires 7 (Example only)

Example

Below are two train scenarios for a park-and-ride commuter travelling from home to work at, say 7:00AM, on a **typical weekday**

All things considered, which scenario appeals to you most?

Factors	**Scenario 1**	**Scenario 2**
Travel time from home to the station		
Usual travel time	5mins	10mins
Travel time on a good day	About **3.75** mins on **4** days a month	About **5** mins **once** a month
Travel time on a bad day	About **5.75** mins on **4** days a month	About **12.4**mins **once** a month
Finding parking at the station		
Parking capacity in P&R lots	1000 bays	1000bays
Parking available in P&R lots at 7:00AM	Likely unavailable	Likely available
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **1** and **9** minutes and only 20% chance to reach to 9mins	Between **5** and **25** minutes and only 20% chance to reach to 25mins
Time on the train		
Average journey time	20 minutes	30 minutes
Crowding on the train		
Crowding	 **All seats taken and half of aisles filled**	 **Three quarters of seats taken**
Number of days per week on which the trains are too crowded to board	1	0
Other things		
Safety	Regular station and car park security patrols	Irregular station and car park security patrols
Ticket fare	$3.50	$4.50
Train frequency	All-stop trains at 10 minute internals	Express and all-stop trains at 5 minute intervals
Tick the box for the scenario you would choose	☐	☐

Thank you for your time and input.

APPENDIX F Parking Search Time Survey (2014.5.13)

The University of Western Australia | Curtin University | Department of Transport | Public Transport Authority | Planning Western Australia

Station choice Survey of Park and Ride Users *(Example completed)*

Project title: Modelling and Evaluating the Joint Access - Mode and Train Station Choice (**LP110201150).**

This research has the UWA Human Research Ethics Office approval number RA/4/1/5370 from May 2012.

Project description: *We want to find out how the time spent looking for a car-parking place affects the train station that park-and-ride commuters choose.*

Q1. Where have you just come from (e.g. home, work, visiting friends, shopping) ______ and its address, ______ Street and ______ Suburb? ______

Q2. How long did it take you to this station? Hours ______ Minutes ______.

Q3. What is the departure time of the train you will catch? Hours ______ (e.g. 7:00am)

Q4. How long will it take you to travel from this station to your destination? Hours ______ Minutes ______.

Q5. What are your average and the worst parking search times (**i.e.** ***the time spent looking for a parking space, parking the car, and walking to a train platform***)? Please fill in the following fields and circle the choice (i.e., week, month and year).

	Average	The worst
Parking search time	______ Mins	______ Mins
How often?	① 60% per month ② 80% per month ③ 100% per month	Once per month

Q6. If you can't find parking in a Park-and-Ride facility, what are your alternatives? Please tick the appropriate box or complete the answer.

[ ] Parking on the street

[ ] Driving to other stations

[ ] Parking in nearby parking lots such as a shopping centres parking lots

Others ______.

Q7 If you have experienced unusually long parking search times, why has this occurred? Please tick the appropriate box or complete the answer.

[ ] Left home later than usual

[ ] Traffic congestion

[ ] Delayed by the other activities on the way

Others ______.

Age group: 18-24 ☐1 25-34 ☐2 35-44 ☐3 45-54 ☐4 55-69 ☐5 70-79 ☐6 >80 ☐7

Gender: M/F

Thank you very much for your time.

The end of the survey

APPENDIX F TRAIN STATION CHOICE PILOT SRUVEY QUESTIONNIARES (EXMAPLE ONLY)

Questioner 1

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30AM, on a **typical weekday**. All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **5** and **25** minutes and only **5%** chance to reach to **25** mins	Between **5** and **10** minutes and only **20%** chance to reach to **10** mins
Tick the box for the scenario you would choose		

Questioner 2

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30AM, on a **typical weekday**. All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Finding parking at the station		
Parking capacity in P&R lot	1000 bays	1000bays
Parking availability in a P&R lot at 7:30AM	0%	10%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **5** and **25** minutes and only **5%** chance to reach to **25** mins	Between **5** and **10** minutes and only **20%** chance to reach to **10** mins
Tick the box for the scenario you would choose		

Questioner 3

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30 AM on a **typical weekday**. All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Finding parking at the station		
Parking capacity in P&R lot	1000 bays	1000bays
Parking availability in a P&R lot at 7:00AM	0%	10%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **5** and **25** minutes and only **5%** chance to reach to **25** mins	Between **5** and **10** minutes and only **20%** chance to reach to **10** mins
Parking cost		
Parking fee in a P&R lot	$2/day	$4/day
Parking fine (if you park illegally, you would be fined)	$60	$80
Frequency of controls for illegal parking	Once a **month**	Once a **week**
Tick the box for the scenario you would choose		

Questioner 4

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30AM, on a **typical weekday.** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Finding parking at the station		
Parking capacity in P&R lot	1000 bays	1000bays
Parking availability in a P&R lot at 7:00AM	0%	10%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **5** and **25** minutes and only **5%** chance to reach to **25** mins	Between **5** and **10** minutes and only **20%** chance to reach to **10** mins
Parking cost		
Parking fee in a P&R lot	$2/day	$4/day
Parking fine (if you park illegally, you would be fined)	$60	$80
Frequency of controls for illegal parking	Once a **month**	Once a **week**
Tick the box for the scenario you would choose	☐	☐

Questioner 5

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30AM, on a **typical weekday.** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Travel time from home to the station		
Usual travel time	**10**mins	**5**mins
Travel time on a good day	About **5** mins **once** a month	About **3.75** mins on **4** days a month
Travel time on a bad day	About **11.5** mins on **4** days a month	About **6.75** mins on **4** days a month
Time on the train		
Average journey time	**30** minutes	**20** minutes
Other things		
Safety	Irregular station and car park security patrols	Regular station and car park security patrols
Ticket fare	$4.50	$3.50
Train frequency	Express and all-stop strains at 5minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose	☐	☐

Questioner 6

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30AM, on a **typical weekday.** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Crowding on the train		
Crowding	 **Half of seats taken**	 **All seats taken and half of aisles filled**
Number of days per week on which the trains are too crowded to board	0	1
Other things		
Safety	Irregular station and car park security patrols	Regular station and car park security patrols
Ticket fare	$4.50	$3.50
Train frequency	Express and all-stop strains at 5minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose		

Questioner 7

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 8:00AM, on a **typical weekday.** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Time on the train		
Average journey time	**30** minutes	**10** minutes
Travel time from home to the station		
Usual travel time	**10**mins	**5** mins
Travel time on a good day	About **5** mins **once** a month	About **2.5** mins on **4** days a month
Travel time on a bad day	About **11.5** mins on **4** days a month	About **6.2** mins on **4** days a month
Finding parking at the station		
Parking capacity in P&R lot	500 bays	100bays
Parking availability in a P&R lot at 7:30AM	0%	0%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **10** and **20** minutes and only **20%** chance to reach to **20** mins	Between **10** and **20** minutes and only **5%** chance to reach to **20** mins
Parking cost		
Parking fee in a P&R lot	$4/day	$2day
Parking fine (if you park illegally, you would be fined)	$40	$40
Frequency of controls for illegal parking	Once a **day**	Once a **month**
Crowding on the train		
Crowding	 **All seats taken and all aisles filled**	 **All seats taken and all aisles filled**
Number of days per week on which the trains are too crowded to board	2	2
Other things		
Safety	Regular station and car park security patrols	Irregular station and car park security patrols
Ticket fare	$4.50	$2.50
Train frequency	Express and all-stop trains at 5 minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose		

Section Break (Next Page)

Questioner 8

Slide eleven

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 8:00AM, on a **typical weekday**. All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Time on the train		
Average journey time	**20** minutes	**20** minutes
Travel time from home to the station		
Usual travel time	**5**mins	**10** mins
Travel time on a good day	About **3.75** mins on **4** days a month	About **5** mins **once** a month
Travel time on a bad day	About **6.2** mins on **4** days a month	About **12.4** mins on **4** days a month
Finding parking at the station		
Parking capacity in P&R lot	500 bays	500bays
Parking availability in a P&R lot at 8:00AM	0%	0%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **10** and **20** minutes and only **20%** chance to reach to **20** mins	Between **10** and **20** minutes and only **5%** chance to reach to **20** mins
Parking cost		
Parking fee in a P&R lot	$4/day	$2/day
Parking fine (if you park illegally, you would be fined)	$60	$40
Frequency of controls for illegal parking	Once a **month**	Once a **day**
Crowding on the train		
Crowding	 **All seats taken and half of aisles filled**	 **Half of seats taken**
Number of days per week on which the trains are too crowded to board	1	0
Other things		
Safety	Irregular station and car park security patrols	Regular station and car park security patrols
Ticket fare	$3.50	$3.50
Train frequency	Express and all-stop trains at 5 minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose	☐	☐

Questioner 9

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:00AM, on a **typical weekday**. All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Finding parking at the station		
Parking capacity in P&R lot	100 bays	1000bays
Parking availability in a P&R lot at 7:00AM	20%	10%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between 1 and **8** minutes and only **5%** chance to reach to **8** mins	Between **5** and **10** minutes and only **20%** chance to reach to **10** mins
Parking cost		
Parking fee in a P&R lot	$6/day	$6/day
Parking fine (if you park illegally, you would be fined)	$80	$60
Frequency of controls for illegal parking	Once a **week**	Once a **day**
Travel time from home to the station		
Usual travel time	**15**mins	**15**mins
Travel time on a good day	About **7.5** mins on **4** days a month	About **11.25** mins on **4** days a month
Travel time on a bad day	About **18.6** mins **once** a month	About **18.6** mins on **4** days a month
Time on the train		
Average journey time	**10** minutes	**20** minutes
Crowding on the train		
Crowding	 **All seats taken and all aisles filled**	 **Three quarters of seats taken**
Number of days per week on which the trains are too crowded to board	2	0
Other things		
Safety	Irregular station and car park security patrols	Regular station and car park security patrols
Ticket fare	$2.50	$3.50
Train frequency	Express and all-stop strains at 5minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose	☐	☐

Questioner 10

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30AM, on a **typical weekday**. All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Crowding on the train		
Crowding	 **All seats taken and all aisles filled**	 **All seats taken and all aisles filled**
Number of days per week on which the trains are too crowded to board	1	2
Finding parking at the station		
Parking capacity in P&R lot	100 bays	500bays
Parking availability in a P&R lot at 7:30AM	10%	10%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **5** and **10** minutes and only **5%** chance to reach to **10** mins	Between **5** and **10** minutes and only **20%** chance to reach to **10** mins
Parking cost		
Parking fee in a P&R lot	$6/day	$6/day
Parking fine (if you park illegally, you would be fined)	$60	$40
Frequency of controls for illegal parking	Once a **month**	Once a **day**
Travel time from home to the station		
Usual travel time	**15**mins	**5** mins
Travel time on a good day	About **11.25** mins on **4** days a month	About **2.5** mins **once** a month
Travel time on a bad day	About **17.25** mins on **4** days a month	About **5.75**mins on **4** days a month
Time on the train		
Average journey time	**10** minutes	**30** minutes
Other things		
Safety	Irregular station and car park security patrols	Regular station and car park security patrols
Ticket fare	$2.50	$4.50
Train frequency	Express and all-stop trains at 5 minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose	☐	☐

Questioner 11

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:00AM, on a **typical weekday**. All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Travel time from home to the station		
Usual travel time	**10**mins	**5**mins
Travel time on a good day	About **5** mins on **4** days a month	About **3.75** mins on **4** days a month
Travel time on a bad day	About**11.5** mins **once** a month	About **6.2** mins **once** a month
Time on the train		
Average journey time	**10** minutes	**20** minutes
Finding parking at the station		
Parking capacity in P&R lot	100bays	100bays
Parking availability in a P&R lot at 7:00AM	20%	20%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **1** and **8** minutes and only **5%** chance to reach to **8** mins	Between **1** and **8** minutes and only **20%** chance to reach to **8** mins
Parking cost		
Parking fee in a P&R lot	$4/day	$6/day
Parking fine (if you park illegally, you would be fined)	$40	$80
Frequency of controls for illegal parking	Once a **day**	Once a **month**
Crowding on the train		
Crowding	 **All seats taken and all aisles filled**	 **All seats taken and all aisles filled**
Number of days per week on which the trains are too crowded to board	2	1
Other things		
Safety	Regular station and car park security patrols	Irregular station and car park security patrols
Ticket fare	$2.50	$2.50
Train frequency	All-stop trains at 10 minute internals	Express and all-stop trains at 5 minute intervals
Tick the box for the scenario you would choose	☐	☐

Questioner 12

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:30AM, on a **typical weekday.** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Finding parking at the station		
Parking capacity in P&R lot	1000 bays	1000bays
Parking availability in a P&R lot at 7:00AM	0%	10%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **5** and **25** minutes and only **5%** chance to reach to **25** mins	Between **5** and **10** minutes and only **20%** chance to reach to **10** mins
Parking cost		
Parking fee in a P&R lot	$2/day	$4/day
Parking fine (if you park illegally, you would be fined)	$60	$80
Frequency of controls for illegal parking	Once a **month**	Once a **week**
Travel time from home to the station		
Usual travel time	**10**mins	**5**mins
Travel time on a good day	About **5** mins **once** a month	About **3.75** mins on **4** days a month
Travel time on a bad day	About **11.5** mins on **4** days a month	About **6.75** mins on **4** days a month
Time on the train		
Average journey time	**30** minutes	**20** minutes
Crowding on the train		
Crowding	 **Half of seats taken**	 **All seats taken and half of aisles filled**
Number of days per week on which the trains are too crowded to board	0	1
Other things		
Safety	Irregular station and car park security patrols	Regular station and car park security patrols
Ticket fare	$4.50	$3.50
Train frequency	Express and all-stop strains at 5minute intervals	All-stop trains at 10 minute internals
Tick the box for the scenario you would choose	☐	☐

Questioner 13

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:00AM, on a **typical weekday.** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Parking capacity in P&R lot	500bays	500bays
Parking availability in a P&R lot at 7:00AM	20%	20%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **1** and **9** minutes and only **20%** chance to reach to **9** mins	Between **1**and **9** minutes and only **5%** chance to reach to **9** mins
Parking fee in a P&R lot	$4/day	$6/day
Parking fine (if you park illegally, you would be fined)	$40	$40
Frequency of controls for illegal parking	Once a **month**	Once a **week**
Tick the box for the scenario you would choose	☐	☐

Questioner 14

Below are two train scenarios for a park-and-ride (P&R) commuter travelling from home to work at, say 7:00AM, on a **typical weekday.** All things considered, which scenario appeals to you most?		
Factors	**Scenario 1**	**Scenario 2**
Parking fee in a P&R lot	$4/day	$6/day
Parking fine (if you park illegally, you would be fined)	$40	$40
Frequency of controls for illegal parking	Once a **month**	Once a **week**
Parking capacity in P&R lot	500bays	500bays
Parking availability in a P&R lot at 7:00AM	20%	20%
Parking search time (Time spent searching for parking before giving up such as trying another station or driving directly to the destination)	Between **1** and **9** minutes and only **20%** chance to reach to **9** mins	Between **1**and **9** minutes and only **5%** chance to reach to **9** mins
Tick the box for the scenario you would choose	☐	☐