# Modelling travellers' risky choice behaviour in revealed preference contexts: A comparison of EUT and non-EUT approaches

Guotao Hu

September 2013

A thesis submitted for the degree of

Doctor of Philosophy of the Imperial College London

Centre for Transport Studies

Department of Civil and Environmental Engineering

Imperial College London

# Abstract

Recent work on risky choice modelling has sought to address the theoretical shortcomings of expected utility theory (EUT) by using non-expected utility theoretic (non-EUT) approaches. To date, however, there is little evidence to show whether the complexity of non-EUT actually leads to better model performance. Moreover, almost all the relevant research has adopted stated choice data which, although flexible and cheap, has limited validity. This thesis empirically investigates the feasibility and validity of non-EUT approaches in revealed preference (RP) contexts, in which travel time distribution is extracted from historical travel time data to subsequently present systematic comparisons between EUT and non-EUT approaches. Additionally, this thesis also discusses implementations based on these empirical results and, in particular, highlights the influence of non-EUT on the valuation of travel time savings.

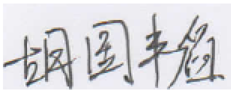
A risky choice framework is proposed so as to incorporate non-EUT into a Random Utility Maximization structure. The non-EUT approaches modelled in the thesis consist of Subjective Expected Value Theory, Subjective Expected Utility Theory, Weighted Utility theory, Rank Dependent Expected Value, Rank Dependent Expected Utility, Prospect Theory, and Cumulative Prospect Theory. The first dataset is collected from the SR91 corridor in California and involves a choice between a free flowing and reliable tolled facility and a congested and unreliable un-tolled facility. The second case study is based on the London Underground (LU) system and involves the choice between alternative competitive underground services linking pairs of stations.

This thesis provides insights into how EUT and non-EUT models perform in the real world. The RP methodology and risky choice framework offers an avenue for future research to identify a wider range of alternative choice theories using realistic data. The empirical results suggest that there are merits in applying non-EUT to the modelling of travellers' risky choice behaviours.

# Copyright Declaration

# Declaration of Originality

My research has been established on the collaboration with my supervisors, Professor John Polak and Dr. Aruna Sivakumar. Our collaborative research papers have been presented in several conferences and submitted for journal publication, and all these work have been properly referenced in this thesis. Therefore, I declare that all the work presented in this thesis have been carried out by myself.



-----------------------------------

Guotao Hu

September 2013

# Acknowledgements

I am grateful to many people for their support and advice during my stay at Imperial College London.

Among all the people I have remembered to appreciate here is Prof. John Polak, my supervisor, who selected me from tons of application forms, educated me, guided me, and offered me a constant source of inspiration. I learned travel demand modelling, econometrics, and even English from him. And most importantly, he let me know that I should not attempt to trade off quality with other factors, which will definitely guide my following career path. I feel really lucky to known John, and to work with him over the past four years. Without his assistance and guidance, it is impossible for me to achieve this dissertation. Hence, I would like to express my deepest appreciation to him.

I am also extremely grateful that my second supervisor, Dr. Aruna Sivakumar, offered valuable advice and tremendous help in my research. I cannot remember how many times she inspired me to overcome technical problems with her great patience. And I cannot imagine how my research would be without the help from Aruna, a professional, helpful, and elegant lady.

This study would not be possible without the financial support from 'China-UK Scholarships for Excellence' jointly operated by China Scholarship Council, the Department for Business, Innovation and Skills, and Imperial College London. I would like to thank all the people helping me apply for the scholarship and PhD, in particular Fionnuala Donovan who solved my applications problems by the deadline.

Warm thanks are due to Prof. Kenneth Small and his colleagues for providing access to datasets used in my SR91 case-study. This study has also benefited from the cooperation with London Underground. Here I feel very grateful to Elizabeth Gilliard, Sandra Weddell and Geoffrey Maunder for facilitating my work at London Underground, and similar appreciation belongs to Nigel Kelt and Nick Rusz who provided the access to the train performance data.

I have also benefited from discussions with many of my colleagues at Imperial. I am particularly grateful to Dr. Scott Le Vine and Dr. Jiang Han for their helpful advice on discrete choice modelling and the EM algorithm. Here, I would like to thank all the colleagues in CTS, especially the members in DeMoNs, for interesting discussions with them. Additionally, I would like to thank Jackie Sime, the key executive lady, for all her administrative assistance.

Finally, the biggest thanks of all my achievements are definitely given to my mother who taught me to be an honest man. I own you!

Guotao Hu
August 2013

# Publications

**Journal**

Hu, G., Sivakumar, A. and Polak, J. W. (2012) Modelling travellers' risky choice in a revealed preference context: a comparison of EUT and non-EUT approaches, Transportation, 1-17.

**Conference**

Hu, G., Sivakumar, A. and Polak, J. W. (2012) Modelling travellers' risky choice in a revealed preference context: a comparison of EUT and Non-EUT approaches, In The 87th Annual Meeting of Transportation Research Board, Washington D.C., USA.

Hu, G. (2012) Non-Expected Utility Theory Approaches to Measurement of Risk Attitudes and Valuation, paper presented to the 45th UTSG Conference, University of Oxford, 2-4 January

# Contents

# List of Tables

# List of Figures

# List of Acronyms

AIC: Akaike Information Criterion

ASC: Alternative Specific Constant

APCP: Average Probability of Correct Prediction

BIC: Bayesian Information Criterion

CARA: Constant Absolute Risk Aversion

CPT: Cumulative Prospect Theory

CRRA: Constant Relative Risk Aversion

DA: Theory of Disappointment Aversion

EVT: Expected Value Theory

EUT: Expected Utility Theory

GEV: Generalized Extreme Value

IIA: Independence from Irrelevant Alternatives

IID: Independently, Identically Distributed

JTM: Journey Time Metric

LL: Log-likelihood

LR: Likelihood ratio

LU: London Underground

LUL: London Underground Limited

MAPE: Mean Absolute Percentage Error

MMNL: Mixed Multinomial Logit

MNL: Multinomial Logit

NetMIS: Network Management Information System

NL: Nested Logit

Non-EUT: Non-Expected Utility Theory

OD: Origin-Destination

PAT: Preferred Arrival Time

PEDS: Pedroute Strategic Model

PR: Prospective Reference Theory

PT: Prospect Theory

RDEV: Rank-Dependent Expected Value Theory

RDEU: Rank-Dependent Expected Utility Theory

RODS: Rolling Origin Destination Survey

RMSE: Root Mean Square Error

RP: Revealed Preference

RR: Reliability Ratio

RT: Regret Theory

RUM: Random Utility Maximization

SEV: Subjective Expected Value Theory

SEU: Subjective Expected Utility Theory

SDE: Schedule Delay Early

SDL: Schedule Delay Late

SP: Stated Preference

SR91: The State Route 91 Toll Road

TD: Theory of Disappointment

TfL: Transport for London

TT: Travel Time

VOR: Value of Reliability

VTTS: Value of Travel Time Saving

WTA: Willingness to Accept an Increase in Travel Time in Return for Toll Savings

WTP: Willingness to Pay Extra Money in Return for Travel Time Savings

WUT: Weighted Utility Theory

# Chapter 1 INTRODUCTION

## 1.1 Background

Risky choice, or decision making under risk, plays a crucial role in people's daily life in contexts as varied as the monetary risk in gambling game and the risk to life in a new surgery, etc. These choice situation s entail a variety of possible outcomes and associated probabilities, where the decision maker is not certain about which outcome will eventually occur. In travel choices, travellers also often confront various risky choices, such as the risk associated with underground seat availability, and the risk of arrival time for flights. Among all the risky issues in transportation it is the unpredictability of travel time that most broadly influences travel behaviours. Travel time risk is derived from the randomness of actual travel time in a congested transportation system, and is often referred to as travel time variability or reliability (Bates et al., 2001, Small, 1982). The reasons for studying risky travel choice are twofold. Firstly, the uncertainty of travel time is a pivotal issue of concern to travellers. A number of empirical findings have revealed that, in making travel decisions, travellers not only consider the travel time savings but also the reduction of travel time variability (De Jong et al., 2004, Lam and Small, 2001). Secondly, it is not only a theoretically important issue, but also a significant focus for policy concerns. For instance, one of Transport for London's (TfL) main objectives for road network management is to reduce travel time variability (TfL, 2013 b).

A variety of competing theories exist for modelling individuals' risky choice behaviours. Expected Utility Theory (EUT) (Von Neumann and Morgenstern, 1947) has long been used as the standard theory of individual decision making in economics. Over the last four decades, however, there has been extensive research involving substantive experiments on decision making. The common aim of these theoretical efforts is to generate alternative theories to EUT. As a consequence of these studies and the associated development of experimental and behavioural economics, there has recently been a great deal of work on non-expected utility theories (non-EUT), and a growing interest in their application in transport. By taking psychological factors into account, non-EUT appears more theoretically sophisticated but, to the author's knowledge, very little evidence exists as to whether these theories really produce empirically better models. Given the prevalence of risky choice theories, it is worth testing which theory actually better characterizes and predicts travellers' choices under risk.

Another gap arises in regard to the validity of data, given that almost all the existing non-EUT studies simply rely on laboratory experimental data or stated preference (SP) data. These hypothetical data collection strategies have the merit of flexibility, but the problem arises in the external validity and generalizability of their results. Revealed preference (RP) data fully addresses the shortcomings of SP data, but, as has been discussed in a large body of literature, it is difficult to observe sufficient detail in RP exercises to model risky choice behaviour. To improve the validity of results, therefore, it is also worth investigating the method for the analysis in an RP context.

To bridge these gaps in current knowledge, this thesis investigates the performances of EUT and non-EUT approaches using two distinct RP datasets in the real world. It also follows the current trend in scholarship by using Random Utility Maximization (RUM) to estimate these risky choice models (Batley and Daly, 2004, Liu and Polak, 2007, Polak et al., 2008). In this way, it becomes possible to test and compare candidate models in terms of calibrations and predictions, and also, based on these results, to discuss methods of implementation.

## 1.2 Aims and Objectives

Building on the background presented in section 1.1, the aims and objectives of the thesis can be summarised as follows:

- To develop a modelling structure allowing the operationalisation EUT and non-EUT theories of risky choice in conjunction with RUM.
- To develop a method for RP data collection that is suitable for risky choice modelling.
- To identify how EUT and non-EUT models perform in the real transport choice contexts.
- To compare EUT and non-EUT models using RP data.
- To implement models with respect to the travel time savings and prediction.

## 1.3 Outline of the Thesis

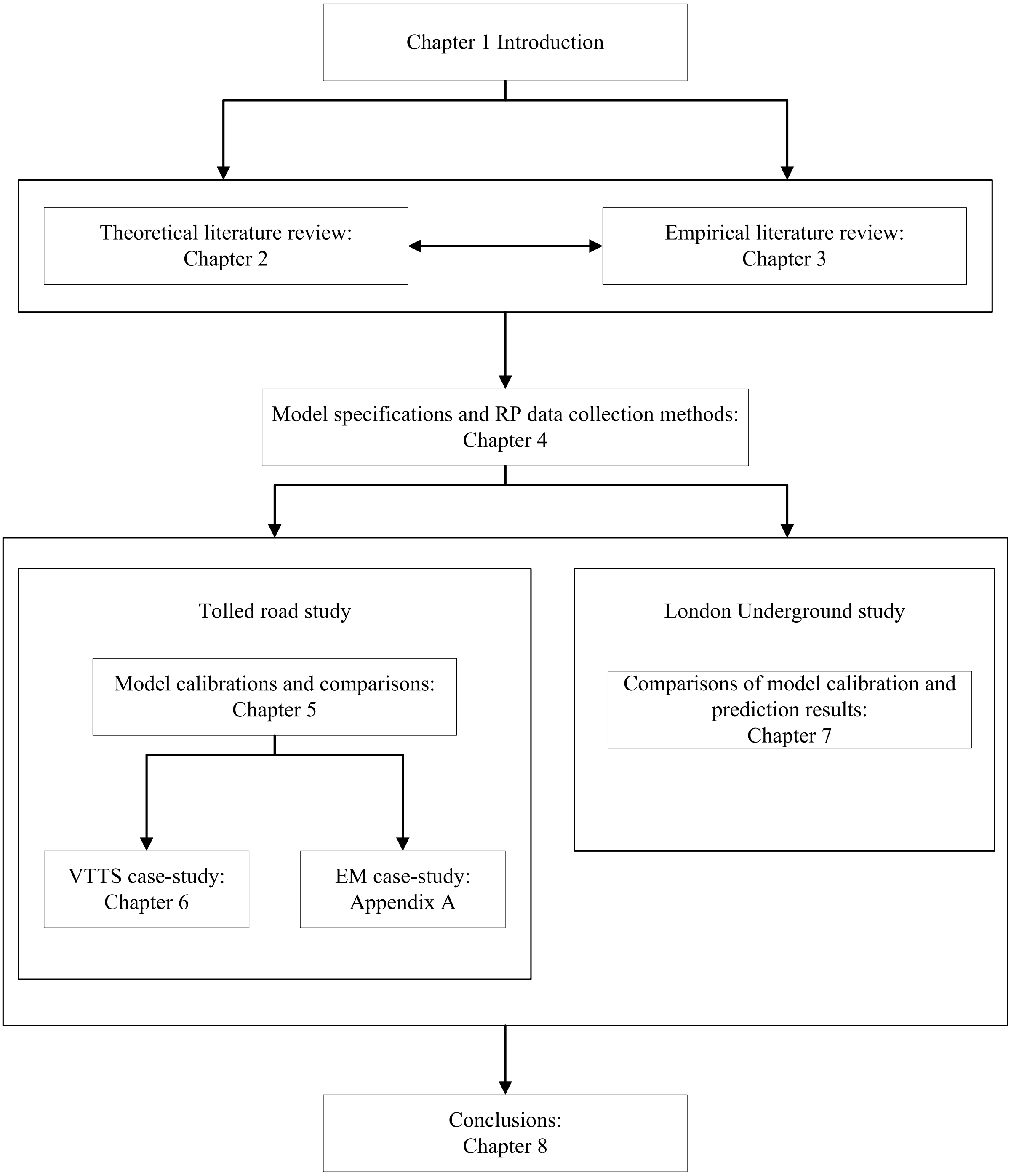



**Figure 1.1: Illustration of the work flow**

To provide an overview of the remainder of the thesis, we now briefly present the main content of each chapter.

- **Chapter 2** reviews the state-of-the-art of choice theory, and presents the evolution of risky choice theory. It highlights the motivations for using non-EUT approaches, and describes the detailed methods of several promising non-EUT methods which could be applied in this thesis.

- **Chapter 3** looks at the state-of-the-practice of risky choice modelling in transport. It first reviews both theoretical and empirical studies on risky travel choice, in particular on the choice model with travel time variability. After a thorough literature review, the chapter discusses the opportunities and challenges in relation to applying non-EUT approaches to modelling travellers' risky choice behaviours.
- **Chapter 4** develops the risky choice framework and RP data collection strategy to be used in this thesis. Moreover, the chapter pays particular attention to the detailed method of incorporating various non-EUT approaches, such as subjective expected utility theory (SEU), rank-dependent expected utility theory (RDEU) and Prospect Theory (PT), into the proposed risky choice framework. After a discussion of the feasibility of data collection, the State Route 91 (SR91) dataset and the London Underground (LU) dataset are selected for use in the applied part of the thesis.
- **Chapter 5** serves as the first chapter of the applied part in the thesis. It presents the findings of a binary route choice case-study between a toll road and a free road on the SR91 corridor in the US, making use of expected value theory (EVT), EUT, subjective expected value theory (SEV), SEU, rank-dependent expected value (RDEV), RDEU and PT models. The chapter explores candidate model specifications in an RP context, and systematically compares the calibration performances between EUT and non-EUT models. It highlights the behavioural benefits of applying non-EUT methods to addressing travellers' attitudes towards risk.
- **Chapter 6** looks at the application of models based on the estimation results of Chapter 5, especially focusing on the value of travel time savings (VTTS). The chapter identifies the influence of each non-EUT approach on the estimated VTTS.
- **Chapter 7** presents the findings of a case-study of route choice in the context of the London Underground, aiming to compare the model performances from estimation and prediction perspectives. By separately evaluating each candidate model, it identifies which non-EUT method is best able to improve model fit and prediction.
- **Chapter 8** provides a summary of the research presented in this thesis, and sets out several recommendations for future research.
- **Appendix A** presents a novel method to address multiple reference points by incorporating an expectation-maximisation (EM) algorithm into the PT model. The

reason for presenting this part as an appendix is that its key methodology (EM) is different from the mainstream methodology used in the thesis (non-EUT).

# Chapter 2 THEORETICAL BACKGROUND

## 2.1 Introduction

In recent research, the field of individual choice under risk has witnessed significant development from the normative method, i.e. expected utility theory (EUT), to newer alternatives, i.e. non-expected utility theories (non-EUT). These theoretical concepts have been developed in various contexts, generating substantive approaches to assist the understanding of individuals' choice behaviour. These behavioural approaches appear to differ greatly in the way in which they assume different choice processes, resulting in a wide range of model specifications. Before reviewing the potential applications from the perspective of the transport literature, therefore, it is essential to present a multi-disciplinary literature review covering the fundamental concepts, definitions, assumptions and specifications of choice behaviour theories.

Although most researchers in transport are well aware of the intuitive appeal of non-EUT, these behavioural approaches are not yet widely adopted (refer to Chapter 3 for details). This is a potential weakness in current research and more studies should be conducted to analyse and compare the actual performances of EUT and non-EUT approaches in a travel choice context. To address this gap, this thesis seeks to arrive at an unbiased judgement of the state-of-art in regard to these behavioural theories and, as a consequence, this chapter conducts a comprehensive analysis to understand how individuals make choices and, more importantly, how a set of non-EUT approaches can be synthesized into tractable model specifications using a Random Utility Maximization (RUM) Theory framework. The issues associated with empirical evidence in the field of transport studies will be presented more fully in Chapter 3 and, partly, in Chapter 4, for reasons of clarity. The outline of this chapter is shown in Figure 2.1.

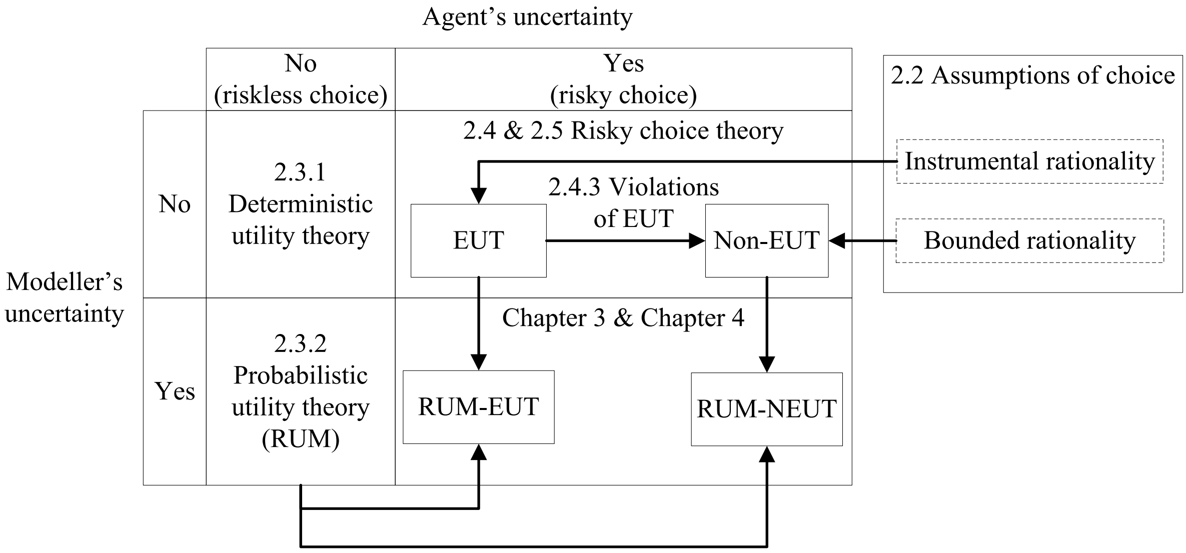



**Figure 2.1: Workflow of Chapter 2 (modified version based on Batley et al. (2008))**

The first uncertainty is from the decision maker's perspective. It should be noted that decision making under risk is different from decision making under uncertainty. According to Batley (2007), decision making under risk assumes that the individual is aware of the likelihood of the emergence of contingent events. By contrast, in the situation of decision making under uncertainty, the distribution of possible outcomes is neither knowable nor definable. Hence, in the case of risky choice, there exist two features of importance to researchers: firstly, that researchers should identify all possible outcomes; secondly, that it must be assumed that decision makers are aware of the probability distribution of each prospect. In some special cases, decision makers may not have the basic knowledge of the probability distribution, for instance, tourists may find it difficult to know the travel time distribution of all corridors since they do not have sufficient commuting experiences on these routes. As discussed above, this situation is more correctly referred to as decision making under uncertainty. This thesis, however, merely focuses on modelling choice under risk rather than uncertainty, although uncertainty will occasionally be used with the same meaning as risk but in reference to uncertainty from the decision-maker's perspective.

In reality, decision makers do not necessarily make consistent decisions in repeated choice problems. To explain this 'inconsistency', Thurstone (1927) suggested that modellers merely have limited observational capability and, as a result, the choice utility should be considered  and a random error term which reflects the modeller's uncertainty. A great deal of literature has attempted to offer constructive methods to address modeller's uncertainty, such as Block and Marschak (1960). This has led to the development of Random Utility Maximization (RUM) theory.

The discussion in this chapter is structured as follows. Section 2.2 reviews the economic assumptions and behavioural theories of the decision making process. Section 2.3 reviews the dominant theories on modelling riskless choice behaviour, including RUM. This is followed, in Section 2.4, by a discussion of risky choice theories. After a brief discussion of the limitations of EUT, a wide range of alternative theories are thoroughly reviewed in Section 2.5. Further discussion is extended to the position, relationship and features of different behavioural theories. The chapter ends with a summary in Section 2.6.

## 2.2 Theoretical foundation of Choice Theory

### 2.2.1 Terminology and notation

Before an extensive review of findings from economics and behavioural science, it is necessary to clarify briefly the elementary concepts of choice theory used in this thesis.

**Alternatives:** A set of choices that can be selected by decision makers. They are considered to be mutually exclusive in this thesis.

**Prospect:** In the risky choice framework, alternatives involved with risk are termed as prospects. We define prospect as the counterpart of alternative in order to discriminate risky choice and riskless choice.

**Outcomes:** Possible events as a result of choosing a prospect. They cannot be controlled by the decision maker.

**Attributes:** A set of inherent features of an alternative Lancaster (1966).

**Value:** The face value of an object which is not affected by people's attitudes or tastes.

**Utility:** A representation of preference in economics. In this thesis, utility is a subjective counterpart of value, incorporating the decision maker's attitude to risk and taste for attributes.

**Cardinal utility:** The magnitude of cardinal utility has some significance.

**Ordinal utility:** The counterpart of cardinal utility. It only captures the ranking rather than the magnitude difference.

**RUM utility:** This converts multiple attributes and taste parameters into a single metric, mostly, but not exclusively, applying a linear utility function.

**EUT and non-EUT utility**: These embody attitude towards risk by nonlinearly distorting an object's value.

Utility was originally conceived by economists as the indicator of a person's happiness. This scalar, therefore, potentially enables researchers conveniently to analyse individuals' preferences. The first definition of utility can be traced back to Bernoulli (1738): '*The determination of the value of an item must not be based on its price, but rather on the utility it yields…*'. However, it is not convincing that we can apply a single scalar to measure people's happiness and utility has been questioned for a long time due to this conceptual problem. Consequently, neoclassical economists routinely apply ordinal utility as a scalar that ranks different consumption bundles, whilst the change of utility magnitude does not matter as long as the ranks of preferences are constant. Utility function, from the point of view of ordinal utility, merely serves as a way of assigning a numerical utility to each alternative. The counterpart of ordinal utility is cardinal utility in which the numerical difference between cardinal utilities is supposed to have some significance.

It should be noted that although the terminology of 'utility' is adopted by both RUM and risky choice theories (i.e. EUT and non-EUT) the property of utility varies between them. For risky choice theory, the initial research has often adhered to simple monetary gambling experiments, in which the risky choice is determined by the weighted face value which can be converted to a cardinal utility by adding attitude towards risk. As a result, historically, the utility in EUT and non-EUT is associated with cardinal utility. The theoretical origin of RUM, meanwhile, considers ordinal utility as the main vehicle (Block and Marschak, 1960, Marschak, 1960), however, the later implementation of RUM also carries the property of cardinal utility (Cavagnaro et al., 2013, Fennema and Wakker, 1997). Credit for analysing the distinction between cardinal utility and ordinal utility in RUM should be given to Batley (2008), who revisited the development of RUM from ordinal utility and cardinal utility perspectives respectively. Though RUM shows the property of cardinal utility, especially in some special cases such as when considering marginal utility and consumer surplus change, Batley suggested that we can still treat cardinal utility as a special operation or representation of ordinal utility in RUM. Hence, he concluded that RUM is still in line with the original presentation of using ordinal utility.

For the purpose of better understanding risky choice theories, the notation that is applied throughout this chapter is set out here. It should be noted that the basic structure follows Liu and Polak (2007). $C$ represents the choice set with *N* alternatives faced by individual $i$[1], i.e., $C = \{s^n; 1 \leq n \leq N\}$ where all components in choice set *C* are regarded as exhaustive and mutually exclusive prospects. Each prospect, meanwhile, includes a set of possible outcomes or state of the world $s^n = \{s_k^n; 1 \leq k \leq K\}$, if there exist $K$ risky outcomes. Correspondingly, the observed vector of attribute values is expressed as $V^n = \{v_j^n; 1 \leq j \leq M\}$ if there are $M$ attributes. The risky choice implies that at least one attribute value in $V^n$ (such as uncertain travel time), say $v_m^n$, varies across different risky outcomes. In this case, each possible outcome is associated with a pair of utility and probability, i.e., $u_m^n$ and $p_m^n$.

### 2.2.2 Overview of choice theory

Choice behaviour in this thesis is defined as the mental process of thinking that leads to decision making between two or more possibilities. It is usually associated not only with reasoning and logic-analytic thinking, but also intuition and emotion. In social psychology, choice behaviour is often characterized by using a variety of system theories, such as Lieberman et al. (2002)'s reflexive system and reflective system, and Zajonc (1980)'s dual-process system. While it is hard to define a generally applicable theory to describe all choice behaviour, a broad picture, which will be helpful for understanding the structure of risky choice theories, can be attained from the dual system theory proposed by Hickey (2011), (Figure 2.2).

[1] The label of individual *i* is not embodied by our current notations for the sake of clarity, while it should also be noted that heterogeneity across respondents is essential in choice models.

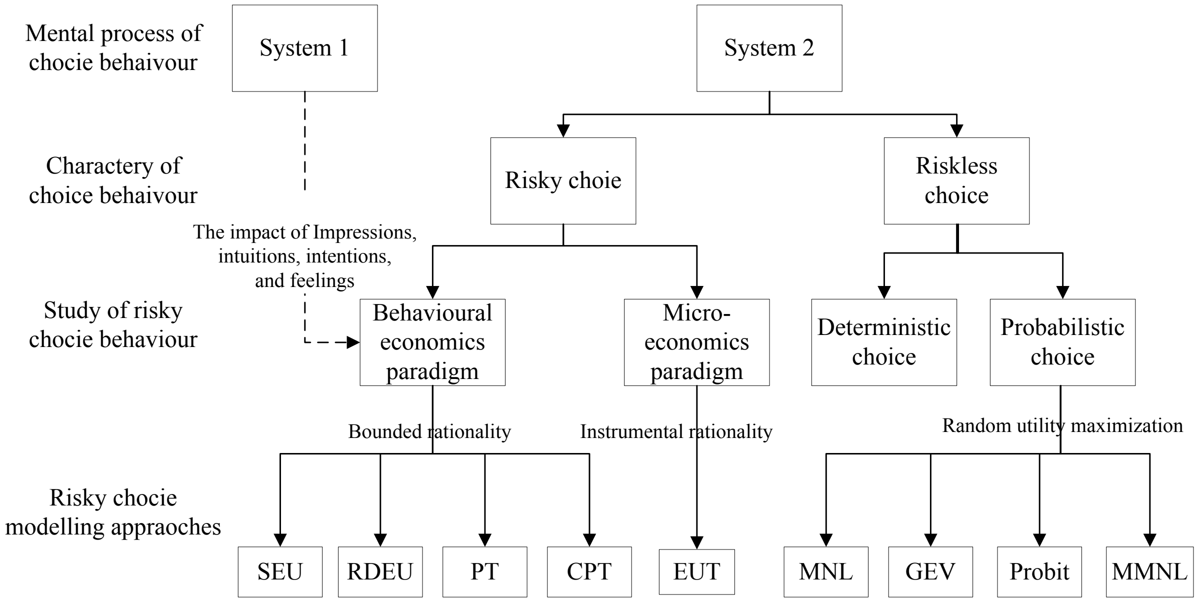



**Figure 2.2: The structure of choice theory used in this thesis**

According to Hickey (2011), the mental process of choice behaviour consists of two types of thinking systems: System 1 is unconscious and automatic, while System 2 is conscious, deliberate, and decision rules based. Kahneman and Frederick (2002) adopt this dual system theory, observing that System 1 is a fast response system with little effort required for computation, statistical analysis and voluntary control, e.g., comparing the colour of two apples; System 2 is a slow system with effort required for reasoning and even computation, e.g., comparing the technical features of two mobile phones. Conscious thinkers seem to pay more attentions to decision making than unconscious thinkers, and obviously the 'budget of attention' is limited. As a result, System 1 is often preferred when we face some decision problems that can be automatically solved by using simple choice strategies. Indeed, System 1 is capable of dealing with most everyday issues,[2] whilst we have to resort to System 2 when we confront important and complex problems that cannot be solved by System 1. Lastly but not least, System 1 delivers intuitions, emotions, feelings, attitudes, impressions to System 2, which treats these psychological factors as beliefs if it accepts them (Abrantes and Wardman, 2011).

Travel choice, in most cases, is not an unconscious process, especially in a congested and uncertain network. We assume that travel choice behaviour is based on rule-based

[2] Kahneman (2011) believes that System 2 is 'lazy' in that it requires effort for computation and reasoning. As a result, we tend to rely most on System 1 to deal with usual problems, and System 2 only acts as a monitor of the ideas constructed by System 1.

evaluation (utility function) accounting for various attributes, and limited by personal experience, memory and information. These features closely adhere to System 2. The following discussion of choice behaviour and related theories, therefore, is based on the structure of the System 2 process, although impacts from the System 1 process are taken account of as appropriate.

Various types of choice behaviour exist. For instance, based on the impact of choice making on the corresponding results, choice behaviour is defined as 'strategic decision making', 'tactical decision making', and 'operational decision making'. In this thesis, however, only the situation with risk is concerned and, therefore, choice need only be characterised as riskless choice and risky choice. This research specifically focuses on risky choice and the modelling strategies that follow the micro-economics and the behavioural paradigms.

The micro-economics paradigm is based the assumption of instrumental rationality (Hargreaves Heap, 1992). An instrumentally rational person is considered to be an omniscient decision maker who is clearly aware of his/her choice environment, including all of his/her objectives, choice sets, possible outcomes and associated probabilities. Moreover, he/she has the unlimited computational ability and stable order of preferences over all outcomes to choose their preference according to the maximum utility criteria. Following these assumptions, *utility* is selected to rank the desires by comparing the pleasures of satisfying individuals. Expected Utility Theory (EUT) is the classical representation of instrumental rationality. This modelling method is based on utility maximization and instrumental rationality and is usually referred to as the normative approach (Myerson, 2013). In EUT, the single scale of utility greatly simplifies rational choice modelling and thus we will frequently apply the concept of utility in this thesis. In the micro-economics paradigm, economists seem to prefer research on 'how people should make the best decision', and such instrumentally rational behaviour is often defended via an as-if argument (Hodgson, 2012). In other words, it is simply assumed that individuals tend to make choices "as if" making a trade-off between costs and benefits with the purpose of utility maximization, although they do not consciously maximize utility during decision making. On the basis of this as-if argument, this normative approach serves as the guideline for modelling risky choice behaviour.

The behavioural paradigm, meanwhile, highlights the impact of subjective issues on decision making. Empirical observations and experiments reveal that the computational capability of decision makers is far too imperfect to arrive at an optimal decision (see critiques by Gigerenzer and Todd (1999) and Stigler (1961)). Additionally, individuals are not instrumentally rational in that they are affected by their emotions, such as pessimism and optimism. As a result, the behavioural paradigm resorts to alternative assumptions about rationality, e.g. *limited rationality* (March and Simon, 1958), *contextual rationality* (Long, 1958), *game rationality* (Farquharson, 1969), *process rationality* (Edelman, 1985), *adaptive rationality* (Cyert and March, 1963) and *selected rationality* (Winter, 1965). Perhaps the most influential among these alternatives to instrumental rationalities is the concept of *bounded rationality* initially proposed by Simon (1955). Bounded rationality is often defined as the optimization procedure under some constraints, such as incomplete information (Conlisk, 1996, Sargent, 2011). It is also called *procedural rationality* since procedures or rules of thumb are employed to understand an individual's choice. That is, an individual would like to use rules of thumb to avoid the effort of searching for information which may lead to the better calculation of optimal options. This theory, therefore, abandons the concept of *the rational man*, leading to more complex models of behaviour (Salant, 2011, Simon, 1979). With the development of behavioural economics, there is a growing research focus on alternative choice theories which are no longer concerned with what rational people should do, but rather take into account how and why people think and act the way they do. These alternative methods are often referred to as non-Expected Utility Theories (non-EUT). Along with EUT, various non-EUT methods have resulted in the recent rich variety of choice theories.

In summary, a variety of choice theories exist that are derived from different assumptions on rationality. The non-EUT approaches, based on bounded rationality, benefit from their theoretical realism, although this plausible sophistication also renders them less mathematically tractable than EUT approaches. It is difficult properly to compare EUT and non-EUT approaches without empirical tests. Surprisingly, however, there is little research which empirically assesses the validity of non-EUT approaches.

## 2.3 Riskless Choice Theory

Traditional choice theory defines certainty as the specified outcome associated with a single state of the world. The domain of decision making under certainty, also called riskless choice, generally consists of deterministic choice (without uncertainty on the part of the researcher) and probabilistic choice (with uncertainty on the part of the researcher). The former is the theoretical starting point, while the latter serves as the operational method for model estimation in this thesis.

### 2.3.1 Deterministic choice theory

With the development of neoclassical economics, the deterministic choice model has been designed to characterize the trade-off between what an individual wants to buy and what that individual can afford. It is derived from consumer choice theory, and researchers have also frequently applied this theory, especially the concept of utility, in travel choice modelling. In consumer choice theory, it is assumed that individuals are capable of ranking different alternatives subject to their budget constraints. Consequently, each alternative $s^n$ is characterized by a set of attributes$(a_1^n, a_2^n \dots a_M^n)$ associated with this alternative, and ranking alternatives reveal the individual's preference. According to Hargreaves Heap (1992), we define that:

- Symbol > means one bundle is strictly preferred to another.
- Symbol ~ means the individual is indifferent between two bundles, or she has the same satisfaction from the two bundles.
- Symbol $\geq$ means the individual prefers or is indifferent between the two bundles, or she weakly prefers one bundle to another.

Provided any two alternatives, $s^m$ and $s^n$, the assumptions of deterministic choice theory is as follows:

- Completeness: any two alternatives can be compared, i.e., $s^m \geq s^n$, or $s^m \leq s^n$, or both.
- Reflexivity: any alternative is at least as good as itself, i.e., $s^m \geq s^n \geq s^m$.
- Transitivity: if $s^m \geq s^q$ and $s^q \geq s^n$, thus it should be $s^m \geq s^n$.

Furthermore, if we admit that more is better (the assumption of non-satiation), another assumption, called monotonicity of preferences, has to be followed. That is, if the attributes of the $n^{th}$ alternative $\mathrm{s}^{\mathrm{n}}$ are at least as much as the attributes of the $m^{th}$ alternative $s^m$, then preferences preserve the order $s^n \geq s^m$.

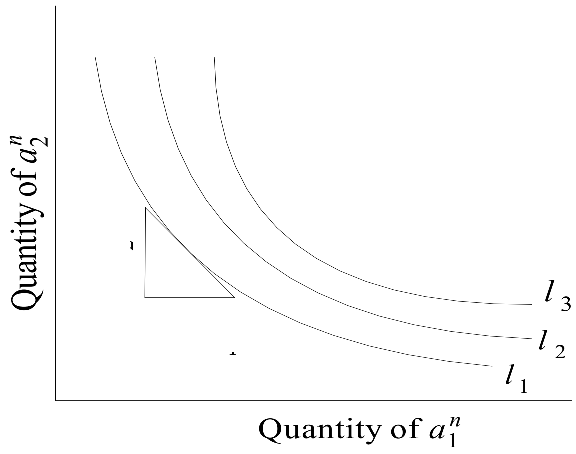



**Figure 2.2: Indifference curves and MRS**

One of the central concepts of preference is the marginal rate of substitution (MRS), which measures the rate at which an individual is willing to substitute attribute $a_2^n$ for another attribute $a_1^n$. In travel choice modelling, this is usually interpreted as marginal willingness to pay.[3] This essential concept can be visually illustrated by indifference curves showing bundles of attributes between which decision maker is indifferent. As shown in Figure 2.2, all the points on the same curve have the same utility, and decision maker has no preference between them. Accordingly, MRS is measured as the slope of an indifference curve: $\partial a_2^n / \partial a_1^n$. If curves are strictly convex, MRS turns out to decrease as $a_1^n$ increases. Thus, another important assumption for indifference curves is of diminishing MRS. This assumption has some intuitive appeal, since it is consistent with some of the other assumptions of expected utility theory and prospect theory.[4]

## 2.3.2 Probabilistic choice theory

The development of probabilistic choice theory is highly related to psychology (Luce and Suppes, 1965). This probabilistic mechanism explains the observed inconsistency in people's decision making. The first inconsistency arises when decision makers are found to select different alternatives in repetitions with the same choice context. Furthermore, decision makers have been found not to select the same alternative even if they have identical characteristics, choice sets, and attributes of alternatives. It can be argued that either people's

---

[3] Willingness to pay (WTP) is commonly applied to measure travellers' intrinsic valuations of travel time savings and plays an important role in this thesis. It is described in detail in Chapter 6.

[4] Diminishing marginal rate of substitution is similar to the concept of risk aversion in EUT and diminishing sensitivity in PT, which is introduced in sections 2.4 and 2.5 respectively.

decision making procedure is inherently probabilistic, or researchers do not have sufficient knowledge about the attributes actually perceived by decision makers. Although it is still impossible to reveal the accurate reasons for this inconsistency, various probabilistic approaches have been proposed to address the observed variation of choice behaviour. We adopt random utility maximization (RUM) approach as it is more consistent with consumer choice theory described in section 2.3.1 (Ben-Akiva and Bierlaire, 1999, Chorus et al., 2010, Manski, 1977).

In RUM, attribute values of alternatives and individuals' characteristics are observed and applied to the deterministic portion of utility $V^n$. The deterministic utility consists of three components: (1) that which is exclusively related to the alternative; (2) that which is exclusively related to the individual's characteristics; (3) the interactions between (1) and (2). Specifically, the first component addresses the level of service of the alternatives. The second component aims to measure the bias across individuals by introducing personal and household variables. These variables are straightforwardly related to social and demographic factors, such as the gender of the traveller, income, age etc. The other excluded variables are roughly measured by the alternative specific constant (ASC). Finally, the demographic variable, in some cases, interacts with variables relating to alternatives, such as the product of the dummy variable of male times mean travel time.

The deterministic utility is not the whole story of probabilistic choice theory. RUM shows that there is an unobserved component $\varepsilon^n$ which captures the factors that affect utility dramatically but is excluded in $V^n$. Therefore, the utility of the $n^{th}$ alternative becomes:

$$U^n = V^n + \varepsilon^n \tag{2.1}$$

Due to this random component $\varepsilon^n$, a joint density $f(\varepsilon)$ is adopted to make probabilistic statements regarding the decision maker's choice. The probability of selecting the $n^{th}$ alternative is:

$$P^n = P(V^n + \varepsilon^n > V^p + \varepsilon^p \quad \forall n \neq p) \tag{2.2}$$

And

$$P^n = P(\varepsilon^n - \varepsilon^p > V^p - V^n \quad \forall n \neq p) \tag{2.3}$$

The above equation reveals that the probability of selecting the $n^{th}$ alternative is the cumulative distribution of the random term $\varepsilon^p - \varepsilon^n$. Thus, the equation becomes:

$$P^n = \int_{\varepsilon} I(\varepsilon^n - \varepsilon^p > V^p - V^n) f(\varepsilon) d\varepsilon \quad \forall n \neq p) \tag{2.4}$$

where *I()* is the indicator function which equals 1 if $\varepsilon^p - \varepsilon^n < V^n - V^p$ and 0 otherwise. A wide range of distributions are applicable to represent $f(\varepsilon)$ over individuals and alternatives. As a result, different distributions lead to different discrete choice models. Different types of RUMs, which are critical approaches to evaluate individuals' preferences, are discussed in the following subsections.

#### *2.3.2.1 Multinomial Logit (MNL) Model*

The Logit formula was originally derived by Luce (1959), while Marschak (1960) subsequently showed that the model is in line with utility maximization. McFadden (1972) subsequently parameterized Luce's 'strict utility' and developed Multinomial Logit (MNL). McFadden explained that the form of the Logit formula implies independently, identically distributed (IID), extreme value, i.e. Gumbel or type I extreme value for the unobserved component of utility $\varepsilon^n$ (relevant proof can also be found in Johnson and Kotz (1969) and Domenich and McFadden (1975). MNL is an extension of the binary Logit model, and is also the widely used Logit model due to its simplicity. The MNL choice probability for alternative *n* is:

$$P^n = \frac{e^{V^n}}{\sum_{j=1}^{N} e^{V^j}} \tag{2.5}$$

This equation only depends upon parameters which are either observed or can be estimated, and the choice probabilities no longer contain the error terms $\varepsilon^n$. Moreover, such probabilities exhibit several desirable properties.

The most important property of MNL is the '*independence from irrelevant alternatives* (*IIA*)' assumption. According to Luce (1959), where the probability ratio for alternative *m* and *n* does not depend on any other alternatives other than *m* and *n*. The ratio of the MNL choice probabilities for selecting the $m^{th}$ alternative and the $n^{th}$ alternative is constant no matter what other alternatives or attributes are available.

The advantages of IIA are apparent. Firstly, the model can address the condition of different populations facing different sets of alternatives. Moreover, this model is capable of addressing the prediction of probability for new alternatives. On the other hand, it has been argued that the IIA property may not appropriately reflect a realistic situation since some alternatives are not simply irrelevant and independent from the other alternatives, which consequently leads to inaccurate predictive results. The famous 'red bus/blue bus paradox' is an extreme example (Train, 2003). To address this limitation, some substitutional models have been generated.

#### *2.3.2.2 Generalized Extreme Value (GEV) Model*

McFadden (1978) introduced the GEV family of models in which the error terms follow a joint generalized extreme value distribution rather than the IID extreme value distribution in MNL. It divides the choice set into various nests of alternatives, with the result that the error terms associated with alternatives are correlated with each other in every nest. The MNL model can be seen as a simple type of the GEV family with only a single nest and zero correlation between alternatives.

GEV highlights the generating function $G = G(Y^1, Y^2 \dots Y^N)$, where $Y^n = e^{V^n}$ for $n \in [1, N]$, $V^n$ is the observed part of utility for alternative $n$ . If *G* satisfies all the several conditions (refer to McFadden (1978) for details), the GEV choice probability for alternative *n* is given by:

$$P^n = \frac{Y^n G^n\left(Y^1, Y^2 \dots Y^N\right)}{G\left(Y^1, Y^2 \dots Y^N\right)} \tag{2.6}$$

where Let $G^n$ be the derivative of *G* with respect to $Y^n$, that is $G^n = \partial G / \partial Y^n$.

#### *2.3.2.3 Probit model*

The Probit model has its roots in psychology. Thurstone (1927) proposed the first derivation of the Probit model by using so called 'stimuli', which are interpreted as utility by Marschak (1960). Generally, it is assumed that the observed error term vector $\varepsilon = (\varepsilon^1, \varepsilon^2 \dots \varepsilon^N)$ has a normal distribution. Therefore, the density of $\varepsilon$ is:

$$f(\varepsilon) = \frac{1}{(2\pi)^{n/2} |\Omega|^{1/2}} e^{-\frac{1}{2}\varepsilon' \Omega^{-1} \varepsilon} \tag{2.7}$$

Where W is the covariance matrix. Because of the inclusion of the covariance matrix Probit models do not exhibit the three main restrictions of the Logit model. Also, we can define different substitution patterns for the Probit model by changing the structure of the covariance matrix. Therefore, all types of substitution patterns are available, including the one violating the IIA property. Moreover, random taste variation in the population is feasible in Probit models in that there is no assumption of independent error terms over decision makers. Finally, panel data can be accommodated in Probit models by considering the correlation in the unobserved part of utility.

On the other hand, there is no closed-form for the choice probabilities and, consequently, additional simulation is required to approximate for the multi-dimensional integral. These drawbacks mean that the majority of current researchers eschew the flexible but complicated Probit model in favour of sophisticated Logit models such as the Mixed Multinomial Logit models.

#### ***2.3.2.4 Mixed Multinomial Logit (MMNL) Models***

With recent developments both in methodology and information processing technology, MMNL models have become widely used in transport studies (Hess et al., 2005, Polak et al., 2008). Researchers are attracted by its flexible form and the capability of obviating the three limitations of MNL. Moreover, unlike Probit models, MMNL is not restricted to normal distributions, which means that the model is more appealing in terms of evaluation. According to Train (2003), the first applications of MMNL was Boyd and Mellman (1980) for automobile demand models. The utility in the MMNL is given by:

$$U^n = V^n + \eta^n + \varepsilon^n \qquad (2.8)$$

Similar to the MNL model, the error term $\varepsilon$ in MMNL is assumed to be a distributed IID extreme value over alternatives and decision makers, and $V$ is regarded as the observed utility. However, the additional unobserved utility $\eta$ leads to an integral without a closed-form solution. The mean of $\eta$ is zero, and no *a priori* constraints exist on the distribution of $\eta$, so this model is free of any restrictive assumptions, such as the IIA property. Another important consequence caused by the presence of $\eta$ is that a simulation process is required in the estimation of the model. Hence, the MMNL probability is the integral of standard Logit probability, and is explicitly expressed as:

$$P^n = \int L^n(\beta)\phi(\beta|b, W)d\beta \quad (2.9)$$

where $L$ is the standard Logit probability evaluated at parameter$\beta$.

$$L^n = \frac{e^{V^n(\beta)}}{\sum_m^N e^{V^m(\beta)}} \quad (2.10)$$

And $\phi(\beta|b, W)$ is the probability density function with the mean $b$ and covariance $W$. A great deal of literature has specified a range of suitable distributions of $\phi(\beta|b, W)$ for MMNL (refer to Revelt and Train, 1998; McFadden and Train, 2000 for details).

## 2.4 Expected Utility Theory (EUT)

The initial studies regarding decision making under risk were mostly carried out in a restricted context where participants were given the monetary value of lottery choices. As a result, $u(s_k^n)$ is simply the face value of monetary outcome $s_k^n$ without any nonlinear transformation of utility. In 1738, Bernoulli originally proposed Expected Utility Theory (EUT) which assumes that individuals use subjective utility instead of monetary value to measure gamble outcomes. He concluded that subjects would not choose alternatives based on the expectation of monetary wealth, rather on the expectation of utility. Von Neumann and Morgenstern (1947) developed EUT by extending it into Game Theory. Thereafter, EUT was generally developed by Marschak (1950), Herstein and Milnor (1953) and Fishburn (1970). As a result, EUT has operated as the normative approach to address choice under risk for more than 60 years. It should be noted that there has been an increasing interest in experimental economics which challenges the validity of EUT (De Palma et al., 2008, Kahneman, 2011).

EUT follows several assumptions of neoclassical economics, such as completeness, transitivity and reflexivity (these are also assumptions of preference ordering over prospects). According to Hargreaves Heap (1992), there exist four extra axioms that EUT can be derived from:

- Preference increasing with probability: if $s_1^n \leq s_2^n$ and $s^n = (s_1^n, s_2^n; p, 1-p)$ and $s^m = (s_1^m, s_2^m; p', 1-p')$, then $s^n > s^m$ if and only if $p < p'$.
- Continuity: for all prospects $s_1^n, s_2^n, s_3^n$ where $s_1^n \leq s_2^n \leq s_3^n$, there must exist some probability p such that $(s_1^n, s_3^n; p, 1-p) \sim s_2^n$. Combining the assumptions of ordering

and continuity implies that preference can be represented by the utility function which assigns a specific number to each prospect.

- Strong independence: given $(s_1^n, s_2^n; p, 1-p) \sim (s_1^n, s_3^n; p, 1-p)$ , if $s_3^n \sim s_4^n$ , then$(s_1^n, s_2^n; p, 1-p) \sim (s_1^n, s_4^n; p, 1-p)$, if this assumption holds, the utility function is determined to be additive across different states of the world.
- Usual rule for combining probabilities: for prospect $s^n = (s_1^n, s_2^n; p, 1-p)$ and $s^m = (s_1^m, s_2^m; p', 1-p')$, $s^n > s^m$ if and only if $u(s^n) \geq u(s^m)$.

Based on these assumptions, the EUT utility function is expressed as:

$$u(s^n) = \sum_{k=1}^{K} p_k^n v(s_k^n) \tag{2.11}$$

where $p_k^n$ is the associated probability of the $k^{th}$ outcome. Therefore, EUT provides a basic structure for decision making under risk by simply converting consequences and associated probabilities into the single scalar of utility. This normative method has been widely reported due to its intuitive appeal and mathematical capacity.

### 2.4.1 Attitude towards risk

According to Petty et al. (1983), the term attitude is defined as a "general, enduring, positive or negative feeling about some person, object or issue". Rosenberg and Hovland (1960), meanwhile, stated that attitude is a "*predisposition to respond to some class of stimuli with certain classes of response*". These so called *stimuli* correspond to risk in the domain of risky choice.

Psychologists and behavioural scientists have argued that "attitude towards risk" is an essential complementary idea in economics in so far as attitude towards risk enables a *rational man* in economics to behave more like a realistic man with different risk attitudes. Moreover, it is attitude towards risk that has evoked substantive research regarding nonlinear transformations of utility and probability.

Given a gamble between two scenarios with the same expected payoff, in which one provides certain payoff while the other one is uncertain, attitudes towards risk can be categorized as risk aversion, risk neutrality and risk proneness according to the different decisions made by individuals. Specifically, if an individual prefers the scenario with the certain payoff, he is categorised as risk averse; if he prefers the scenario with the uncertain payoff, he is categorised as risk prone; whereas if he shows indifference between the scenarios, he is risk neutral.

The shape of the utility function can also be interpreted behaviourally with respect to attitude towards risk. Specifically, the concavity, convexity and linearity of the utility function implies risk aversion, risk proneness and risk neutrality respectively.This, therefore, means that is straightforward to illustrate the connection between risk attitudes and the curvature of the Bernoulli utility function, as shown in Figure 2.3.

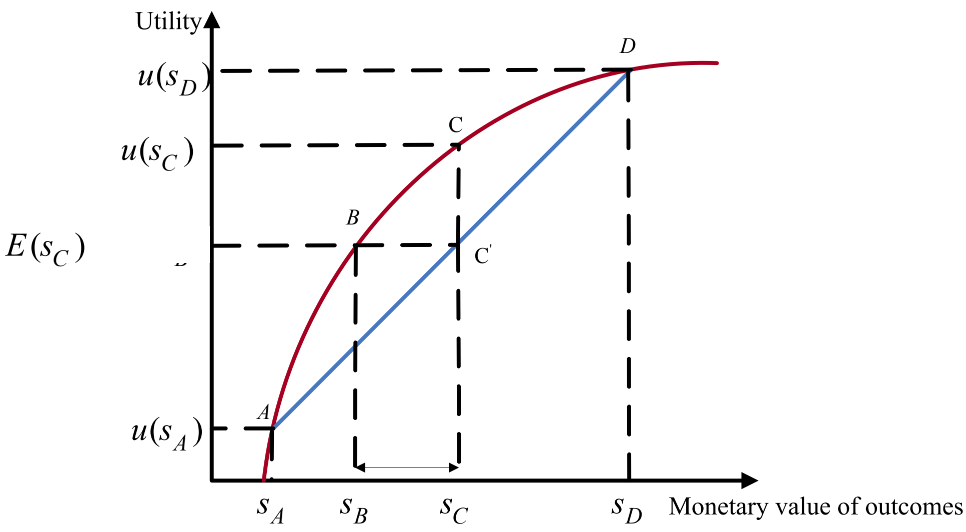



**Figure 2.3: Risk aversion and the effect of 'certainty-equivalent'**

Given two outcomes, $s_A$ and $s_D$, the curve represents the elementary utility of each certain outcome. Let $s_c$ be the random variable within the interval $[s_A, s_D]$, along with the associated probabilities $p_A$ and$1 - p_A$for outcome $s_A$ and $s_D$. Consequently, if we treat outcome $s_c$ as the probabilistic outcome of $s_A$ and $s_D$, the utility of $s_c$ is the expected value of outcome $s_c$ is $E(s_c) = p_A u(s_A) + (1 - p_A)u(s_D)$, as shown on the linear line *AD*. If we treat outcome $s_c$ as certain monetary outcome, the utility is $u(s_C)$. Notice that the point *C* is higher than *C'*, i.e., $u[p_A s_A + (1 - p_A)s_D] > p_A u(s_A) + (1 - p_A)u(s_D)$, which means that the decision maker prefers the alternative with certainty to the alternative with risk, i.e., the decision maker is risk averse. In this case, the risk attitudes can be expressed as follows:

- Risk aversion if $u(p_1 s_1 + p_2 s_2) > p_1 u(s_1) + p_2 u(s_2)$
- Risk proneness if $u(p_1 s_1 + p_2 s_2) < p_1 u(s_1) + p_2 u(s_2)$
- Risk neutrality if $u(p_1 s_1 + p_2 s_2) = p_1 u(s_1) + p_2 u(s_2)$

Notice that risk attitudes can be explained by the effect of 'certainty-equivalent' as well. Given an outcome $s_B$ with certainty, the utility of $\mathrm{s_B}$ is $u(s_B)$. Notice that $u(s_B) = E(s_C)$, while the monetary value of a certain outcome $s_B$ is less than the value of a risky outcome $s_C$. Therefore, $s_B$ with certainty is referred to as the '*certainty-equivalent lottery*', i.e., the uncertain lottery delivers the same utility as the certain lottery; $pre(s) = s_C - s_B$ is recognized as the '*risk premium*', i.e., the maximum quantity of income that a decision maker prefers to pay for an allocation without risk. Consequently, risk attitudes can also be interpreted as follows:

- Risk-aversion: if the utility function is concave or $pre(s) > 0$
- Risk-neutrality: if the utility function is linear or $pre(s) = 0$
- Risk-proneness: if the utility function is convex or $pre(s) < 0$.

The second derivative $u''$ is generally applied to represent the shape of the utility function, in particular, a linear function has $u'' = 0$; $u'' > 0$ for a convex function, whereas $u'' < 0$ for a concave function. Based on these features of utility function, Arrow (1965) and Pratt (1964) proposed a widely-used measure of risk-aversion, termed the '*Arrow-Pratt index of risk-aversion*'.

- The coefficient of absolute risk aversion is defined by $A(x) = -\frac{u''(x)}{u'(x)}$
- And the coefficient of relative risk aversion is defined as $R(x) = -\frac{xu''(x)}{u'(x)}$

where *x* represents the total wealth level or monetary value. For both $A(x)$ and $R(x)$, if the index is greater than zero, risk aversion is implied. Arrow and Pratt also proposed the concept of constant risk aversion. That is, if both $A(x)$ and $R(x)$ is constant in *x*, the decision maker has constant absolute or relative risk aversion.

## 2.4.2 Limitations of EUT

Allais (1953) was the first study to provide convincing counterexamples to challenge the validity of EUT. Here the variations of the *Allais paradox* are described, as set out by Kahneman and Tversky (1979).

| Problem 1 | | Problem 2 | |
|---|---|---|---|
| Prospect *A* (18%) | Prospect *B* (82%) | Prospect *C* (83%) | Prospect *D* (17%) |
| 33% chance to win 2500;<br>66% chance to win 2400;<br>1% chance to win 0. | 100% chance to win 2400. | 33% chance to win 2500;<br><br>67% chance to win 0. | 34% chance to win 2400;<br><br>66% chance to win 0. |

**Table 2.1: Allais paradox**

Notice that problem 2 is obtained from problem 1 by removing the common consequence of winning 2400 with a probability 66%. From the assumption of independence of EUT, individual's should display the preferences in problem 1 and problem 2. Nonetheless, the final data reveals that the utility for certain gain (prospect *B*) reduces more markedly.

Considerable violations of EUT are also apparent from other perspectives, such as inflating small probabilities, preference reversal, failure of description invariance et al. (Cox et al., 2011, Douglas, 2013, McFadden, 1999, Tversky and Kahneman, 1986). Most of these criticisms concentrate on the validity of the transitivity and independence axioms (Allais, 1979, Camerer et al., 2011, Manktelow, 2012). It seems that individuals' biases, misconceptions and errors affect their actual decision making. However, it should be noted that almost all the attempts were simply based on laboratory experiments without empirical evidences. This is unfortunate, whilst these experimental efforts have been already directed at developing alternatives to EUT, i.e., non-EUT.

# 2.5 Non-Expected Utility Theory (Non-EUT)

## 2.5.1 Subjected expected utility theory (SEU)

Diecidue and Wakker (2001) argued that the real probability distribution, in most cases, is not the same as the one perceived by decision makers. This suggests that decision makers' risk attitudes do not only nonlinearly transform utility, but also affect the perceived probability. Consequently, if individuals do have misperceptions of probability distribution, objective probability may not act as a proper measure of individuals' perceived likelihood of consequences, and thus subjective probability should be taken into account instead. This idea leads to the SEU used in this thesis.

#### *2.5.1.1 Objective probability vs. Subjective probability*

Objective probability is universal and replicable, and reflects the empirical frequencies of repeated events. Thus, different individuals would consider the objective probability of outcome $s_k^n$ as $p_k^n$ uniformly, insofar as they are all rational enough. Furthermore, provided one and only one outcome in alternative $s_k^n$ can occur, the following properties of objective probability should be satisfied:

- $p(s_k^n) \geq 0$
- $p(s^n) = 1$
- $s_k^n \cap s_q^n = \emptyset \Rightarrow p\left(s_k^n \cup s_q^n\right) = p(s_k^n) + p(s_q^n)$

Any function $p(.)$ satisfying the above properties is termed as a measure of objective probability.

Unlike objective probability, subjective probability is often referred to as the coefficient of plausibility, associated with unique events which rarely repeat themselves. In reality, different individuals would estimate subjective probabilities diversely due to their different tastes and perceptions. Moreover, even the same individual appears to change his/her assessment of the subjective probability of the same event in different choice contexts. Given the complex properties of subjective probability, it requires much more advanced techniques of assessment than objective probability. Pidgeon et al. (1992) designed a laboratory experiment to detect people's judgements on the likelihood of death. They found out that individuals tend to distort objective probabilities, subjecting them to specific contexts. For instance, individuals turned out to overestimate deaths from infrequent causes, while underestimating the deaths due to frequent causes. In practice, the survey is the traditional method to identify decision makers' subjective probabilities. Specifically, researchers ask respondents to assign a number from 0 to 1 to each event according to their perceived likelihood, with this specific number representing his/her subjective probability.

Another method to address the subjective attitude towards probability, from an endogenous perspective, is to transform objective probability into subjective probability via weighting functions. This is described in detail in the following subsection.

### *2.5.1.2 Subjected expected utility theory (SEU) evolution*

The earliest exploration of the endogenous method for subjective probability was conducted by Edwards (1995, 1962), in which it was suggested that individuals seem to have misperceptions of objective probability and tend to distort objective probability subjectively. Handa (1977) subsequently suggested that the probability associated with each outcome should be expressed by a specific decision weighting function. He proposed the following functional form:

$$u(s^n) = \sum_{k=1}^{K} \pi(p_k^n) v(s_k^n) \tag{2.12}$$

where $\pi$ represents the subjective probability function or decision weight function. A number of alternative weighting function forms $\pi$ have been evaluated by existing studies, but these will be explained in more detail in chapter 3. It should be noted, however, that in this earliest version of SEU, $v(s_k^n)$ is only the objective value of outcome $s_k^n$. Thus, the above formula is more like a subjective expected value model. It is easy to generalize SEU by incorporating subjective utility and decision weight:

$$u(s^n) = \sum_{k=1}^{K} \pi(p_k^n) u(s_k^n) \tag{2.13}$$

Again, the weighting function $\pi$ directly transforms a decision maker's objective probability $p_k^n$ into weight. Subjective probability has been criticised by economists, however, since the weighted utility function cannot satisfy monotonicity (Starmer, 2000). To see why, suppose we have a convex subjective probability, i.e., there exists $\pi(p) + \pi(1 - p) < 1$ and some $\varepsilon > 0$ which result in prospect $(s_1^n, 1)$ being preferred to prospect $(s_1^n, p; s_1^n + \varepsilon, 1 - p)$, even though all the outcomes in the latter are at least as good as the outcome $s_1^n$ in the former. This phenomenon was first pointed out by Fishburn (1978), who claimed that the violation of monotonicity, in theory, has been regarded as the fatal objection to using this kind of subjective probability.

Several studies have proposed possible resolutions to SEU from both conventional and non-conventional perspectives and there is evidence to show that the violation of monotonicity may not happen if the dominating and dominated prospects are simply transparent to decision makers (Tversky and Kahneman, 1986). Researchers have to conduct some pre-processing on all possible prospects before evaluation. This, therefore, is a typical non-conventional approach which requires special conditions on choice context. A famous example using this descriptive method is the Prospect Theory proposed by Kahneman and Tversky (1979) in which the editing phase is capable of removing the dominated prospects from the choice set before the evaluation phase. This editing phase highlights the individual's framing effect (e.g. reference dependence) and the procedure for simplifying the choice set. In the evaluation phase, nonlinear distortion of probability is applied to address an individuals' subjective attitude towards the likelihood of outcomes.

Despite the intuitive appeal of SEU, however, it is not strictly in line with the standards of economics since $\sum_{k=1}^{K} \pi(p_k^n)$ needs not be equal to unity. The probability transformation is consistent with conventional subjective probability (i.e. decision weights sum to unity), while still permitting violations of monotonicity.

## 2.5.2 Rank-Dependent Expected Utility Theory (RDEU)

One of the most natural and useful generalizations to SEU is rank-dependent expected utility theory (RDEU). The decision weight in RDEU is inherently different from simple subjective probability since the decision weight takes the rank of outcomes into account. Specifically, RDEU provides a weighting function that not only depends on the face value of probability but also on the ranking relative to other outcomes.

The theory was first proposed by Quiggin (1982) and subsequently developed by a number of other studies (Abdellaoui, 2000, Segal, 1990, Wakker, 1994, Yaari, 1987). Quiggin argued that monotonicity is so convincing that any violations of it should be regarded as mistakes. For this reason, while RDEU preserves the concept of subjective probability, the ranking of consequences also matters.

Let prospect $s^n = (s_1^n, s_2^n \dots s_K^n; p_1^n, p_2^n \dots p_K^n)$ where$s_1^n \leq s_2^n \leq s_K^n$. That is the ranking of possible outcomes are from the worst to the best. The following functional form can express the preference under RDEU.

$$u(s^n) = \sum_{k=1}^{K} w(s_k^n) u(s_k^n) \tag{2.14}$$

where the decision weight for all consequences $1 \leq \mathrm{k} \leq \mathrm{K}$ is defined by:

$$w(s_k^n) = \pi(p_k^n, p_{k+1}^n \dots p_K^n) - \pi(p_{k+1}^n, p_{k+2}^n \dots p_K^n) \tag{2.15}$$

and

$$w(s_k^n) = \pi(p_K^n) \text{ if } k = K \qquad (2.\ 16)$$

Here, $\pi(.)$ is an increasing weighting function with $\pi(0) = 0$ and $\pi(1) = 1$ ; $\pi(p_k^n, p_{k+1}^n \dots p_K^n)$ corresponds to the weight of obtaining the consequence $k$ or better than $k$; $\pi(p_{k+1}^n, p_{k+2}^n \dots p_K^n)$ means the decision weight of obtaining a consequence better than $k$. In RDEU, decision weight $w(s_k^n)$ is different from the simple transformation of probability that SEU performs. Instead, the decision weight in RDEU is regarded as the difference between the distortions of cumulative probabilities. It should be noted that it is this additive technique that ensures monotonicity. Moreover, it allows the decision weight attached to each consequence to be influenced not only by the objective probability, but also by the given ranking of consequences. Decision makers, therefore, are assumed to conduct pre-processing regarding 'how good' or 'how bad' each outcome is. This procedure, similar to the editing phase of Prospect Theory to some extent, is critical to the following evaluation phase. The sensitivity towards outcome ranking, however, is a debateable feature of RDEU in that the change of outcome utility might have extreme effects on preference. For instance, a very small change of outcome utility could dramatically influence the value of the induced decision weight if the rank of this outcome changed due to the change of outcome utility; on the other hand, a significant change of outcome utility may have no influence on the associated decision weight if the rank of this outcome is still constant.

The decision weighting function includes a new crucial component of risk attitude which is omitted in EUT. There is sufficient evidence to show that decision makers' subjective attitudes in probability do matter. Diecidue and Wakker (2001), for example, argued that in an RDEU context, the curvature of the weighting function $\pi(.)$ is a representation of pessimism if it is convex, and of optimism if it is concave. They concluded that optimism and pessimism is related to risk proneness and risk aversion.

In some circumstances, the weighting function is not simply concave only or convex only but can present a mixed shape with several interesting features. S shaped and inverse S shaped weighting functions are the most commonly reported mixed specifications. The inverse S shaped weighting function, as shown in Figure 2.5, over-weights small probabilities and under-weights large probabilities. No matter whether it is S shaped or inverse S shaped, there must be a crossover point where $\pi(p_k^n) = p_k^n$. The pattern of the weighting function is determined by this crossover point $p_k^n$. In the original version of RDEU, Quiggin (1982) proposed $p_k^n = 0.5$. However, the value of $p_k^n$ varies in different weighting functional forms and, therefore, the selection of each weighting function should be carefully evaluated in the specific choice context.

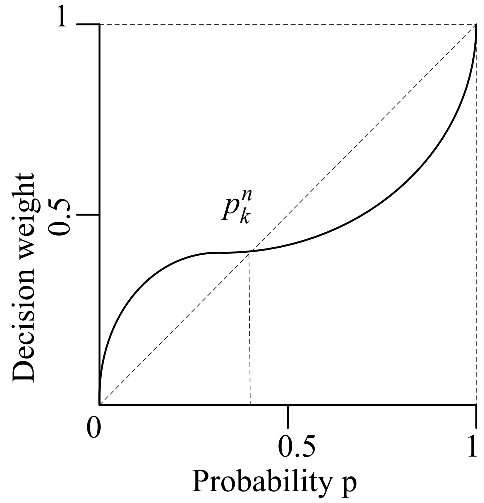



**Figure 2.5: Inverse S shaped weighting function**

## 2.5.3 Prospect Theory (PT)

The prospect theory proposed by Kahneman and Tversky (1979) distinguishes editing stage and evaluation stage in the decision making. In the editing stage, a variety of decision rules are adopted; In the evaluation stage, decision making is articulated by maximizing the specific utility function.

### *2.5.3.1 Editing phase*

There are a wide range of special decision heuristics to reorganize the choice context in the editing phase. In this research, these pre-operations are generally summarized as choice context interpretation and simplifications (Abeler et al., 2011).

- Context interpretation

An individual's preference not only depends on alternatives *per se*, but also the way that alternatives are present and described. A range of literature has attempted to explain this problem from different perspectives, but all these explanations highlight the influence of context interpretations, or the *framing effect* (i.e. that the way in which possible outcomes are interpreted is important).

According to Kahneman and Tversky (1979), the first to "propose that utility be defined on gains and losses rather than on final asset positions" was Markowitz (1952). In PT, outcomes are interpreted as gains or losses depending on their relative position to a specific reference point. Like the development of other behavioural theories, PT was initially studied by laboratory experiments in which the utility of each consequence was measured by monetary value. As a result, Kahneman and Tversky naturally defined an individual's current wealth position as the reference point. The original definition of reference point is intuitively appealing since individuals' seem to prefer both maintaining the status quo and comparing with the status quo (the endowment and anchoring effect).

- Context simplification

The original version of PT applied a variety of rules to simplify the choice context and convert it to a situation which is easier to deal with. One of these operations is *combination*, which simply combines the probabilities of identical outcomes. For instance, individuals can simplify the prospect $(s_1^n, s_2^n, s_1^n; p_1^n, p_2^n, p_3^n)$ as the updated prospect $(s_1^n, s_2^n; p_1^n + p_3^n, p_2^n)$. Another operation is *cancellation*, which cancels the same outcome and the associated probability shared by each prospect during comparison. The third method is a fuzzy way to perceive the value of outcome and probability. For instance, the prospect $(s_1^n, s_2^n; p_1^n, p_2^n)$ can be considered as $s_1^n$ for sure if $s_1^n$ and $s_2^n$ are quite similar, such as $s_1^n = 1$ and $s_2^n = 1.0001$. Arguably, the most important rule used in the editing phase is the elimination of dominance. This approach cancels the stochastically dominated prospects from the choice set before evaluation

### *2.5.3.2 Evaluation phase*

Individuals' preferences are ultimately determined by a utility function across edited prospects. Similar to EUT and most behavioural theories, the prospect with the highest utility is selected. The most distinguishing feature in evaluation is the way to assess outcome utility. The perceived utility of each outcome is no longer dependent on the evaluation of final quantity but on the changes or differences of outcome value relative to reference value. For simplicity, the prospect $s^n = (s_1^n, s_2^n \dots s_K^n)$ is here divided into losses $s^{n-} = (s_1^{n-}, s_2^{n-} \dots s_{i-1}^{n-})$ and gains $s^{n+} = (s_i^{n+}, s_{i+1}^{n+} \dots s_K^{n+})$ according to the relative location of each outcome to the reference outcome $s_{ref}^n$; and, the outcome utility is the increment of value. If the relative utilities of all the outcomes labelled as a gain are positive, and the value of $s^{n-}$ is negative. Then the outcome utility function is given as:

$$\begin{cases} u(s_k^{n+}) = \left(v(s_k^{n+}) - v\left(s_{ref}^n\right)\right)^{\alpha} \\ u(s_k^{n-}) = -\lambda \left(v\left(s_{ref}^n\right) - v(s_k^{n-})\right)^{\beta} \end{cases} \tag{2.18}$$

where $\lambda$ describes the degree of loss aversion, and $\lambda > 1$ if loss aversion holds, and $\alpha, \beta < 1$ captures different degrees of diminishing sensitivity respectively. As shown in Figure 2.6, the shape of the utility function is kinked at the reference point, while the slope of loss is steeper and looms larger than the corresponding gain ($h_l > h_g$), which implies loss aversion.

One of the most important contributions of PT is the finding of *loss aversion*. This pivotal concept has been demonstrated by a range of laboratory experiments regarding respondents' different valuations during the process of buying and selling products (Benartzi and Thaler, 1993, Brown, 2005, Kahneman et al., 1991, Knetsch, 1989, Shogren et al., 1994).

The level of loss aversion varies between populations with different characteristics; for instance, Schmidt and Traub (2002) and Brooks and Zank (2005) consistently identified that men are less loss averse than women. Similar experimental results were obtained by Booij and Van de Kuilen (2009) who measured the degree of loss aversion via the ratio of utilities between gains and losses. They reported that gender and educational level have a significant effect on the degree of loss aversion, i.e., women and people with lower levels of education are more loss averse than the remainder of the population. Some studies, however, have reported different findings on the heterogeneity of loss aversion. Gachter et al. (2007) found that the gap between WTA and WTP increased with income and age, which implies the increasing degree of loss aversion. Harrison and Rutström (2008), meanwhile, incorporated the socio-demographic attributes of gender, age and race into the assessment of loss aversion. Their cumulative prospect theory model using lottery data turned out to have no significant connection between the degree of loss aversion and any of the characteristic variables. Their later work (Harrison and Rutström, 2009) slightly modified the experiment but obtained similar results.

If loss aversion does exist, it is natural to consider what leads to this asymmetric preference and how to reduce this effect. Johnson et al. (2006) observed that the degree of loss aversion is much smaller for people who are well aware of the product attributes. List (2005) also reported that the degree of loss aversion seems to be diminished with increasing knowledge of the market. This finding seems plausible, since sufficient knowledge on the choice context can diminish errors of misconception. In the context of trading, such knowledge about choice context can be obtained by either learning from the market behaviour of others, or from the market discipline (Loomes et al., 2009, Loomes et al., 2003).

. Furthermore, we can observe an additional crucial feature from Figure 2.6: that is, gain is concave and loss is convex. This is interpreted as diminishing sensitivity, i.e. the marginal utility decreases with increasing distances from the reference point. This effect is consistent with the diminishing marginal rate of substitution in economics.

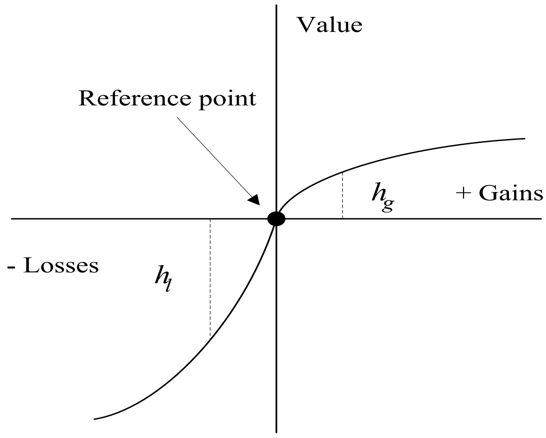



**Figure 2.6: Value function based on Kahneman and Tversky (1979)**

As discussed in the section of attitude towards risk, the concave utility function for gain implies risk aversion. Likewise, the convex utility function in the loss domain means risk proneness. This is explained by the reflection effect of PT. That is, that the convexity of loss is mirrored by the concavity of gain, and that the risk attitudes are also mirrored.

PT also adopts decision weight as the degree of belief, which is similar to the subjective probability in SEU. An axiom like monotonicity is held in editing phase via a series of decision heuristics, such as the elimination of dominated prospects. Nevertheless, using various decision heuristics makes the modelling of decision making less mathematically appealing. Furthermore, although the editing phase can eliminate those clearly dominated options which are already detected, it still permits the survival of unidentified dominated options. This was the main drawback of PT until the emergence of cumulative prospect theory at the end of the 1980s.

### 2.5.4 Cumulative Prospect Theory (CPT)

Tversky and Kahneman (1992) elaborated their PT model by employing the approach of rank-dependence from Quiggin (1982)'s *rank dependent expected utility theory*. This modification enables PT to become more conventional and operational and, more importantly, the property of ranked outcomes with cumulative probability avoids the violation of monotonicity. The functional soundness of RDEU and descriptive realism of PT are simultaneously captured by this updated version of PT, i.e., cumulative prospect theory (CPT) (Wakker and Zank, 2002).

CPT allows different weighting functions for gain and loss respectively. It should be noted that the outcomes in CPT are ranked similar to the way in which RDEU operates, such that the prospect $s^n$ consists of a range of outcomes with $s_1^n < s_2^n \dots s_K^n$. CPT also indicates a reference point which separates prospect $s^n$ into prospect of gain $s^{n+}$ and the prospect of loss $s^{n-}$. The $s^{n+}$ is the same as $s^n$ except that all the outcomes belonging to loss domain are replaced by zero. And the $s^{n-}$ is equivalent to $s^n$ except that all outcomes that are a gain are replaced by zero as well. Thus the utility of prospect $s^n$ is as follows:

$$u(s^n) = \sum_{l=1}^{i-1} w^-(p_l^{n-})u(s_l^{n-}) + \sum_{g=i}^{K} w^+(p_g^{n+})u(s_g^{n+}) \qquad (2.18)$$

The decision weight in the above utility function is not simply the transformation of probability. According to RDEU, it is the difference of cumulative probabilities.

$$w^-(p_l^{n-}) = \pi^-(p_1^{n-} + p_2^{n-} \dots p_l^{n-}) - \pi^-(p_1^{n-} + p_2^{n-} \dots p_{l-1}^{n-}) \qquad 2 \le l \le i \qquad (2.19)$$

$$w^+(p_g^{n+}) = \pi^+\left(p_g^{n+} + p_{g+1}^{n+} \dots p_K^{n+}\right) - \pi^+\left(p_{g+1}^{n+} + p_{g+2}^{n+} \dots p_K^{n+}\right) \quad i+1 \le l \le K-1 \quad (2.20)$$

For the special situation, $w^-(p_1^{n-}) = \pi^-(p_1^{n-})$ and $w^+(p_K^{n+}) = \pi^+(p_K^{n+})$. Tversky and Kahneman (1992) applied the weighting function with only a single weighting parameter, which turns out to mean that small probabilities are over-weighted and big probabilities are under-weightedA similarity shared by CPT and RDEU is that both theories weight the extreme outcome first. As a result it is possible to analyse the connection between attitude towards risk and decision weight, as was done with RDEU. If the weighting function is concave, i.e., the extreme outcome is over-weighted, risk aversion is found in gains and risk proneness exists in losses. In contrast, individuals would display risk aversion in relation to gains and risk proneness losses if the weighting function is convex (Schmidt and Zank, 2008). Regarding the relationship between PT and CPT, Fennema and Wakker (1997) analysed a couple of experiments conducted by Lopes (1993), and reported that CPT not only has the appeal of satisfying stochastic dominance but is also more suitable for modelling the effect of diminishing sensitivity.

### 2.5.5 Other non-EUT

Here we also briefly present other important non-EUT that potentially will be adopted in future research.

### *2.5.5.1 Theory of Disappointment (TD)*

Allais (1979) and Hagen (1979) are the earliest attempts to address common consequence and common ratio effects. They proposed the moments of utility in which the utility of a prospect is not only dependent on expected utility (the first moment), but also on the variance of utility regarding the mean (the second moment), and skewness (the third moment). The third moment of utility is expressed as $\sum_k^K p_k^n\,[u(s_k^n) - \bar{u}(s^n)]$, where $u(s_k^n)$ is the utility of the consequence $s_k^n$, and $\bar{u}(s^n)$ is the expected utility of the prospect $s^n$. Bell (1985) and Loomes and Sugden (1986) developed the theory of disappointment which is closely related to moments of utility. Under this theory, the preference over prospects can be represented as:

$$\sum_k^K p_k^n\,[u(s_k^n) + D(u(s_k^n) - \bar{u}(s^n))] \qquad (2.21)$$

where $D(.)$ is a non-decreasing function with $D(0) = 0$, and $\bar{u}$ is a measure of the 'prior expectation' of the utility from prospect $s^n$. If the outcome utility is worse than the expected utility (i.e., $u(s_k^n) < \bar{u}(s^n)$), a sense of disappointment is generated; if the outcome of the prospect is better than expected ($u(s_k^n) > \bar{u}(s^n)$), the consequence would produce elation. In a triangle diagram, TD also implies a fanning-out effect since individuals are assumed to be 'disappointment averse' ($D(.) < 0$) and 'elation prone' ($D(.) > 0$). It should be noted that this representation is less axiomatic compared to EUT, while it provides psychological insights in its intuitive interpretation (Loomes, 2010).

### *2.5.5.2 Theory of Disappointment Aversion (DA)*

The psychological concept of disappointment has also been applied in the theory of disappointment aversion. This behavioural theory, introduced by Gul (1991), is not only analytically tractable but also parsimonious, with only one more parameter in addition to those required by EUT. This extra parameter carries the intuitive meaning of individuals' disappointment aversion. From this intuitive point of view, DA is based on the individuals' *ex ante* evaluation incorporating *ex post* disappointment or elation. The feeling of disappointment is distinguished from elation depending on whether the actual consequence is worse or better than the individual's anticipation.

Let the certainty equivalent of prospect $s^n = (s_1^n, s_2^n \dots s_K^n)$ is $CE(s^n)$. If we assume that $s_i^n > s_{i-1}^n$ for all $i = 2,3 \dots K$, there exist some $1 < k < K$ such that $u(s_i^n) \leq CE(s^n)$ for all $i \leq k$ (disappointment), and $u(s_i^n) > CE(s^n)$ for all $i > k$ (elation). Hence, we can decompose the prospect $s^n$ into two prospects $r^n = (r_1^n, r_2^n \dots r_K^n; p_1^n, p_2^n \dots p_K^n)$ and $q^n = \left(q_{(k+1)}^n, q_{(k+2)}^n \dots q_K^n; p_{(k+1)}^n, p_{(k+2)}^n \dots p_K^n\right)$, i.e., $s^n = \alpha q^n + (1-\alpha) r^n$, where $\alpha$ represents a probability with the form $\alpha_i = \sum_{l=i+1}^{K} p(x_l)$. In this case, $s^n$ can be expressed by an elation/disappointment decomposition (EDD) with the form $(\alpha_k, q_k^n, r_k^n)$. Gul proposed the following functional form:

$$u(s^n) = u(\alpha_k, q_k^n, r_k^n) = (1-\pi(\alpha)) \underbrace{\sum_{i=1}^{k} p_i^n u(s_i^n)}_{\text{Disappointment}} + \pi(\alpha) \underbrace{\sum_{i=k+1}^{K} p_i^n u(s_i^n)}_{\text{Elation}} \tag{2.22}$$

where $\pi(\alpha)$ is the increasing transformation of probability $a$ with $\pi(0) = 0$ and $\pi(1) = 1$ (decision weights). Gul (1991) proved that his theory is validated only when $\pi(\alpha)$ has the following form:

$$\pi(\alpha) = \frac{\alpha}{1+(1-\alpha)\beta} \tag{2.23}$$

where the disappointment aversion parameter $\beta$ should be estimated according to the individual's actual choice. Gul observed that $\pi(\alpha) < \alpha$ if $\beta > 0$ , which implies disappointment aversion; it is elation loving if $-1 < \beta < 0$. Notice that this representation reduces to EUT if $\pi(\alpha) = \alpha$. Given that $u(s_i^n) < u(\alpha_k, q_k^n, r_k^n) < u(s_{i+1}^n)$ is satisfied if $i = k$, at most $k - 1$ steps are required to calculate EDD and the utility of prospect $s^n$.

#### *2.5.5.3 Prospective Reference Theory (PR)*

Viscusi (1989) proposed prospective reference theory in which preference is assumed to be expected-utility-maximizing with subjective posterior probabilities rather than only objective probabilities. This subjective posterior probability consists of two components, namely objective probability and prior probability. From an intuitive point of view, the prior probability can be interpreted as the *ex ante* judgement of the *ex post* likelihood of the state of the world, while objective probability, in the Bayesian sense, provides the information to update their priors. Viscusi assumed that all prior probabilities (he calls these reference risk levels) are $1/K$ for $K$ possible consequences. Then the subjective posterior probability can be written as:

$$\pi(p_i^n, K) = \alpha p_i^n + (1-\alpha)\frac{1}{K} \tag{2.24}$$

where $0 < \alpha < 1$ represents the relative weight which an individual gives to the objective probability. Notice that PR reduces to EUT if $\alpha = 1$. The objective probability $p_i$ is revised to be larger if $p_i < 1/K$, while it is revised to be smaller if $p_i > 1/K$. This feature is consistent with the fact that individuals tend to over-weight extremely low probabilities and under-weight high ones (Aliev et al., 2012).

#### *2.5.5.4 Weighted Utility Theory (WUT)*

Chew and MacCrimmon (1979) introduced the weighted utility theory which has been further developed by Chew (1989) and Chew (1983). The formulation of WUT can be expressed by the following:

$$u(s^n) = \sum_{k=1}^{K} p_k^n\, v(s_k^n) \left[\frac{w(s_k^n)}{\sum_{j=1}^{K} w(s_j^n) p_j^n}\right] \tag{2.25}$$

where $v(s_k^n)$ is the utility of consequence $s_k^n$ for prospect $s^n$; the real-value function $w(.) > 0$ assigns a positive weight to each consequence. The component in square brackets can be regarded as the weight associated with the consequence $s_k^n$. It is for this reason that this theory is called weighted utility theory. Fishburn (1982) extended WUT to a more general form called skew symmetric bilinear utility theory, while Dekel (1986) proposed the implicit weighted utility theory as another generalization of WUT. The common feature of the WUT family is the use of weights on consequences. The weight not only depends on the consequence *per se*, but also on the whole prospect. As a result, fanning-out and fanning-in effects are captured by WUT, and the extreme outcomes with small probabilities can be measured differently compared to the EUT method. For instance, provided $w(s_k^n) = v(s_k^n) > 0$, the weight for the extremely good outcome $s_1^n$ is the utility of $s_1^n$ divided by the expected utility of $s^n$. Thus, the extremely good outcome is over-weighted comparing with the outcomes with relatively small utility.

Chew (1989) subsequently derived this theory from three axioms, namely ordering, continuity, and the weakened form of independence. The latter can be interpreted as: if $y_1 \sim y_2$, then for each $p_j$, there exist some $p_k$ such that $p_j y_1 + (1-p_j) y_2 \sim p_k y_1 + (1-$

$p_k)y_2$. With respect to WUT, the indifference curves are linear and fanning out and, therefore, there must be a point at which all the indifference curves cross. Note that if $w(s_k^n)$ is constant for all consequences WUT reduces to the EUT.

#### *2.5.5.5 Regret Theory (RT)*

In the regret theory (RT), the term *regret*, serving as the counterpart of utility, is generally referred to as the induced emotion when the chosen alternative turns out to be worse than the other alternatives (Bell, 1982, Fishburn, 1982, Loomes and Sugden, 1982). The intuition behind this theory is based on two assumptions: first, decision makers are aware of the fact that the chosen prospect may turn out to be less attractive than the other prospects, and such 'mistakes' can result in the negative emotion of regret; second, decision makers have a basic knowledge regarding the distribution of possible regrets and aim to minimize the expected regret (Chorus et al., 2006b).

Several unique features of RT distinguish it from the family of utility-based theories. Firstly, unlike most behavioural theories, utility is replaced by the scalar of regret. Moreover, individuals make judgements based on the comparison between the attributes of alternatives, rather than within the considered alternative, as in EUT. That is, RT is capable of accounting for non-compensatory choice behaviour, i.e., the decrease of one attribute does not necessarily offset the increase of another attribute of the same prospect. Finally, the decision rule for the choice behaviour is no longer utility maximization, instead it is regret minimization. These behavioural theories have been widely used in different areas, such as psychology (Crawford et al., 2002), healthcare (Smith, 1996) and finance (Stoltz and Lugosi, 2005). In transport, RT has been applied recently to travellers' responses to information within the context of Advanced Traveller Information Services (Chorus, 2011, Chorus et al., 2006a, Chorus et al., 2007).

Notice that several experimental tests have observed violations of monotonicity and transitivity under the RT framework (Loomes et al., 1991, Tsalatsanis et al., 2010). Economists seem to reject the non-transitive situation in that it violates the whole theory of preference, although subsequent experimental results have revealed that these violations are largely due to experimental control, if one takes so called *event-splitting effects* into account (Starmer and Sugden, 1993).

## 2.6 Summary

The review of existing literature presented in this chapter provides theoretical insights into a wide range of choice behaviour theories. It has also shown the inconsistency between the rapid pace of theoretical development and the lack of empirical tests. The chapter began with a discussion about the basic assumptions of rational decision making theory in economics. This normative approach was rarely been questioned due to its consistency with the principles of individuals' preferences. With the development of experimental economics within the last 30 years, a number of non-EUT approaches have been proposed to explain the observed violations of EUT. In addition to the non-EUT approaches presented in this chapter, there are also other alternatives, e.g. Implicit Weighted Utility (Chew, 1989), Quadratic Utility Theory (Chew et al., 1991), Lottery-Dependent Expected Utility (Becker and Sarin, 1987) and so on. The theoretical progress of non-EUT is set to continue, however, we cannot take for granted that non-EUT naturally outperforms EUT.

In fact, existing literature have revealed a couple of crucial factors that researchers should pay special attention to. Firstly, each non-EUT approach is capable of explaining some perspectives that EUT fails to address, but none of them can deal with all. In this regard the progress of choice theory is no longer dependent on rationality *per se*, but on identifying the limits and scope of rationality (Hargreaves Heap, 1992). Secondly, almost all the non-EUT approaches were established on the basis of experimental observations, whilst few empirical evidences indicate whether non-EUT provides superior model performance. The same issue is extended to the problem in determining which model to be adopted in a specific context, and comparing the robustness of EUT and various non-EUT models. Furthermore, it also leads to the problem of whether the violations of EUT observed in laboratory experiments are also found in the real world.

Given the above issues, this thesis is worthwhile as we apply these candidate theories into modelling travel behaviour using revealed preference data. Both EUT and non-EUT are incorporated into a random utility maximization (RUM) structure, which allows model evaluation and validation. However, it is evidently a strong assumption if we simply transfer the monetary risk of the original EUT and non-EUT to the risk of travel time presented in this research. More research should be conducted to test the validity of travel time risk using EUT and non-EUT, and identify whether travellers actually perceive the risk in a way that we modelled the risk in this thesis. It also leads to the final problem relating to the gap between the state-of-art and the state-of-practice, in particular for travel choice behaviour. To investigate the operational plausibility of these theories, it is more credible to test them in different transport contexts in such a way that this explanatory choice theory is combined with some operational approaches, such as a discrete choice model. Chapter 3, therefore, concentrates on identifying empirical evidence from transport studies and aims to set out the existing techniques that have been applied to risky travel choice.

# Chapter 3 EXISTING MODEL STRUCTURES

## 3.1 Introduction

The previous chapter presents the state of knowledge relating to risky choice theories, and especially highlights the importance of non-EUT theories. This theoretical literature review explained what risky choice theories should be researched, and why, and correspondingly, this current chapter aims to show how these behavioural approaches have been applied to existing empirical transport studies.

The chapter begins with a presentation of existing research on travel time variability, from both theoretical and empirical perspectives. This can be considered to be the starting point of this empirical literature review, since the risk in the real world, as in the subject of this current research, is usually derived from travel time variability which is contains an inherent connotation of repeatability and uncertainty. The remainder of this chapter, therefore, describes how empirical studies apply risky choice theories to account for such unpredictable travel time, paying special attention to the state-of-practice and corresponding gaps.

In the previous chapter it was shown that EUT is still the dominant approach for modelling risky travel choice even though transport researchers have been well aware of the limitations of EUT. Whilst non-EUT approaches have been widely discussed in other fields, they have attracted only a handful of empirical studies in transport. In fact, it is this gap between state-of-art and state-of-practice that motivates this chapter.

The workflow of this chapter and its connection with the previous chapter is illustrated in Figure 3.1. Section 3.2 reviews both theoretical and empirical studies on risky travel choice and travel time variability, *inter alia*. Section 3.3 describes the method for modelling traveller's uncertainty in relation to travel time under a risky choice framework, in particular EUT. The transport literature relating to non-EUT is briefly reviewed in section 3.4, aiming at identifying the state-of-practice and challenges in this field. While the review presented in this chapter concentrates on existing empirical work regarding modelling risky travel choice behaviour, chapter 4 focuses on providing a more detailed discussion of non-EUT approaches in so far as they relate to the topic of this thesis, as well as providing original contributions on model specifications and data collection. The same strategy also applies to Chapter 6, including reviews on value of travel time savings (VTTS).

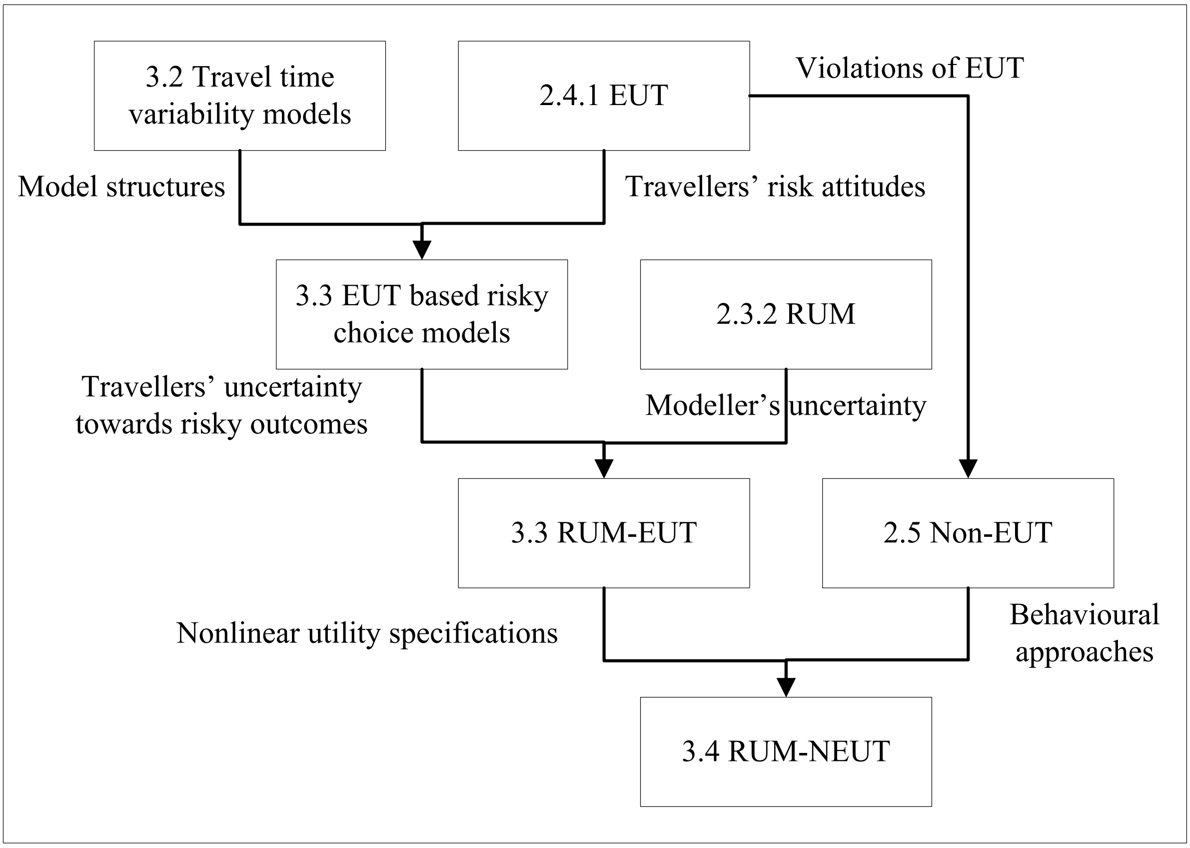



**Figure 3.1: The workflow of this chapter and its connection with previous chapter**

## 3.2 Theoretical and empirical issues in travel time variability

In a realistic transport context, there are a wide range of choice issues associated with uncertain situations, such as availability of seats, variation in tolls, unpredictable train delays, vehicle breakdown, flight cancellation, incidental traffic congestion, etc. Among these risky issues the one that has attracted the most research over the last twenty years is the relationship between uncertain travel time and travellers' responses. In this case, travel time variability leads to uncertain travel time in repeated journeys and therefore serves as the main risk when travellers make choices.

The existing literature has mainly concentrated on answering how travel time distribution is best measured (data collection and measurement), to what extent travel time and travel time variability contributes to travellers' decision making (modelling), and how travellers value travel time and its variability (valuation). Since appropriate research addressing these subjects can provide important insights for transportation planning and project appraisals, it is worth paying special attention to issues relating to such unpredictable variations in travel time. This section reviews the state-of-practice and identifies appropriate approaches for modelling travel time variability. The issues regarding value of travel time savings (VTTS) and value of reliability (VOR) are discussed in Chapter 6.

## 3.2.1 Basic concepts

Travel time is generally regarded as the time elapsed when a traveller moves from one location to another. It is one of the most critical level-of-service data in almost all the travel choice modelling, and provides an essential contribution to travellers' decision making procedures. Travel time, serving as one cost of choice, have been widely applied into existing studies in riskless situations, where travel time is constant and known to travellers. It is rare to observe a situation with constant travel time in the real-world and travellers generally have to confront either predictable or unpredictable travel time variations during their actual travelling experiences. In this risky case, travel time is not sufficient to characterize all the performance factors, due to its variation.

### *3.2.1.1 Congestion vs. Travel time variability*

Predictable variation of travel time is commonly found in regular traffic congestion. An important feature of this variation is that travel time can be properly expected by decision makers, and they are thus able to adjust their travel plans accordingly. For instance, an experienced commuter leaves home half an hour earlier to avoid arriving late at the office due to the morning peak congestion. This recurrent variation of travel time is of interest to traffic flow theory (Daganzo, 1997). Despite the existence of travel time variation in congestion traffic, it is still by no means a risky choice situation. That is travellers do not necessarily 'gamble' in the regular congestion scenario, since they are fully aware of this consistently congested network no matter how slow the traffic is.

Travel time variability adds another cost to trips and it is often interpreted in a statistical way (Asensio and Matas, 2008, Noland and Polak, 2002). Its positive counterpart is transport reliability which can also be found in transport literature (Noland et al., 1998, Tilahun and Levinson, 2010). Alternative definitions of travel time variability are routinely adopted in different transport services, such as unreliability and punctuality (Van Lint et al., 2008). Different interpretations of this travel time variation are confusing, although they are not mutually exclusive. Here, travel time variability and unreliability are closely related (negative interpretation), and reliability and punctuality are also correlated (positive interpretation). In this thesis travel time variability is considered as the fact that travel time is unpredictable and inconsistent, while reliability is defined as the ability of the level of transport service to be consistent with travellers' expectations.

It should be noted that travel time variability does not necessarily mean congestion and vice versa (Bates et al., 2001). In circumstances where there is no regular congestion on a road, however, if travellers experience unexpected delays (e.g. due to debris flow and engineering work) this is regarded as an unreliable road with a relatively big travel time variability. The above instances suggest that congestion and travel time variability are distinguished by predictability. The transport service is reliable as long as travel time can be fairly predicted.

Consequently, only situations with travel time variability associated with risk are of interest to this thesis. In these circumstances, travellers make travel decisions in an uncertain choice environment where they are not capable of predicting the exact travel time during their trip planning: they are, in effect, gambling between travel choices in which uncertain travel time serves as an important attribute for decision making. In the context of risky choice, travellers are able to perceive the travel time distribution, although they cannot predict which travel time outcome certainly happens. They can even be informed of the likelihood of all the possible consequences but, despite this awareness of distribution, risk has to be taken into account due to the random occurrence of travel time outcomes.

#### *3.2.1.2 Does travel time variability really matter?*

It is worth identifying why travel time variability matters before extending the discussion to how it influences the modelling of risky choice behaviour. To explain it in general way, both reasons and consequences of travel time variability are presented.

Figure 3.2 illustrates possible reasons for travel time variability. Notice only non-recurrent accidents are a related risky choice problem, some recurrent factors are not taken into account, such as peak-hour congestion, physical bottlenecks, etc. The occurrence of any of the factors in the 'black box' leads to the observed variation of travel time distribution.

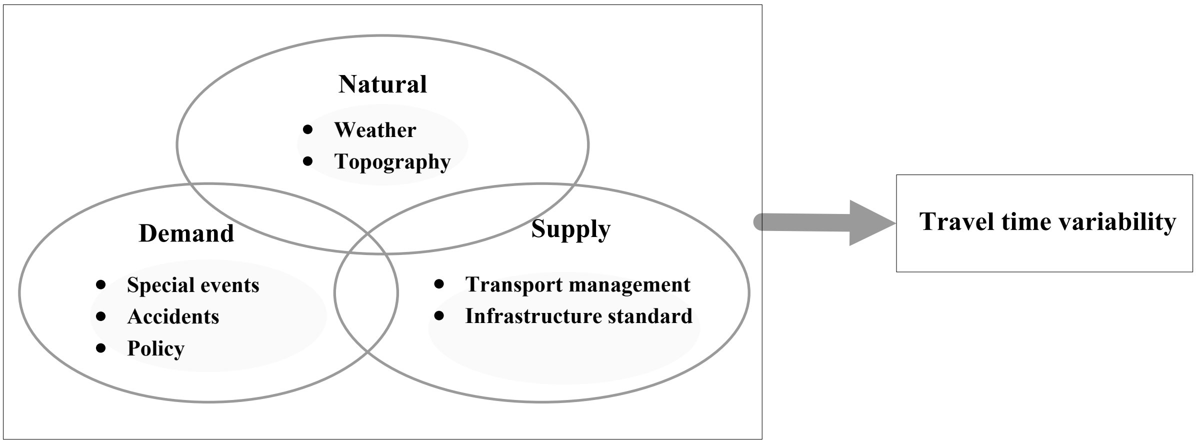



**Figure 3.2: The factors leading to travel time variability**

On the supply side, the standard of infrastructure and the level of maintenance are closely related to reliability. On the demand side, special events and accidents also have a direct connection with reliability. For instance, a road has a large travel time variability if it is close to a major arena where shows often cause serious traffic congestion. Some temporal policies also have a significant influence on travel time variability. In addition to the supply and demand sides, traffic reliability is also affected by natural incidents (e.g. floods, typhoon and debris flow) and external factors (e.g. terrorist attacks).

According to Bates et al. (2001), travel time variability leads to at least two consequences for travel choice. Firstly, travellers are extremely sensitive to the recurrent variability of travel time and the induced consequences. They have to allow extra travel time in an uncertain environment as a *safety margin* or *buffer* (Knight, 1974, Pells, 1987). This extra travel time to avoid unexpected delay reduces the efficiency of travelling and, therefore, provides an additional negative effect for travellers (Batley, 2007, Hollander, 2006).Secondly, uncertain travel time causes cognitive burdens, and travellers would like to value the decrease of variability monetarily. Few studies, however, have applied travel time variability to a transport demand model. This is unfortunate, since the model without accounting for travel time variability could potentially lead to inconsistent estimates related to transport service reliability (Bhat and Sardesai, 2006). Consequently, both the reasons and consequences of travel time variability suggest that the behavioural importance and explanatory effects of variability should be taken into account when modelling traveller's risky choice behaviour.

The following subsections focus exclusively on various model specifications of travel time variability. Before discussing risky choice models, it is essential to review these specifications, for three reasons. Firstly, the strand of research provides theoretical insights into how travellers make travel decisions in an unreliable network. Secondly, it explores important attributes associated with risky travel choice, in particular, travel time variability, in order to predict travellers' future behaviour. Thirdly, it identifies the mutual influences between model specifications and transport policy or infrastructure changes.

## 3.2.2 Mean-variance model

This well-developed model has its roots in finance for portfolio analysis and the valuation of risk assets (Lintner, 1965). This two-parameter approach is naturally transferred to transport in order to address travel choice behaviour under risk. The basic idea of the mean-variance model in transport is to extract the information not only on the centrality (expected travel time) but also the dispersion (variability of travel time) of the distribution. It should be noted that mean-variance model is not an EUT model unless the travel time is nonlinearly transformed into utility.

### *3.2.2.1 Linear model*

The first theoretical breakthrough was from Jackson and Jucker (1982) who proposed their formulation with expected travel time as well as variance.

$$disu^n = \alpha E(t^n) + \beta Var(t^n) \tag{3.1}$$

Unlike the utility maximization rule, the decision rule is to minimize the weighted additive disutility of the $n^{th}$ alternative $disu^n$. $E(t^n)$ reflects the expected travel time for each trip, and $Var(t^n)$ is the variance of observed travel time.

In addition to variance of travel time, standard deviation is also commonly regarded as a representation of travel time variability (Black et al., 1993). The function is expressed as:

$$u^n = \alpha E(t^n) + \beta SD(t^n) + \gamma C \tag{3.2}$$

where *SD* represents the standard deviation of travel time, and *C* is the travel cost. Abdel-Aty et al. (1995) who applied this model to fit their stated preference (SP) data collected from Los Angeles, and they found significant negative parameters for travel time and standard deviation.

### *3.2.2.2 Nonlinear model*

Notice that the expectation operator in the linear mean variance approach is merely the expected travel time. It is not based on utility space, therefore, but value space. Small et al. (1999) pointed out that linear model seems inappropriate to address all of the explanatory effect of travel time variability. Nonlinear utility is therefore required to understand real risky choice behaviour better (Sinn, 1983, Varian, 1992). The first significant exploration of a nonlinear mean variance model was conducted by (Polak, 1987). He proposed an alternative functional form of Jackson and Jucker's two-parameter model by using the following function:

$$u^n = \alpha E(t^n) + \beta E((t^n)^2) \tag{3.3}$$

Given $E((t^n)^2) = Var(t^n) + (E(t^n))^2$, Polak's utility function can be expressed as follows:

$$u^n = \alpha E(t^n) + \beta(Var(t^n) + (E(t^n))^2) = \alpha E(t^n) + \beta(E(t^n))^2 + \beta Var(t^n) \tag{3.4}$$

where $\beta$ indicates the decision maker's attitude towards risk, i.e., risk aversion (preferring the alternative with low variance), risk proneness (preferring the alternative with high variance), and risk neutrality (only considering expected travel time). Senna (1994), meanwhile, pointed out that the omission of the term $\left(E(t^n)\right)^2$could lead to inaccurate estimates of $\beta$.

In addition to the quadratic utility function, Polak (1987) also proposed another nonlinear formulation as:

$$u^n = -e^{\alpha t^n} \tag{3.5}$$

Senna (1994) proposed the general form of the mean variance model in an EUT framework:

$$u^n = \alpha(E(t^n))^\beta + \gamma C \tag{3.6}$$

Again, if the expectation operator is applied to the above function, the nonlinear utility form is expressed as:

$$u^n = \alpha(E((t^n)^{\frac{\beta}{2}}))^2 + \alpha Var((t^n)^{\frac{\beta}{2}}) + \gamma C \tag{3.7}$$

where travel time and variance apply the same scale parameter $\alpha$ and risk attitude parameter $\beta$. In this case, however, risk attitude is no longer measured by the taste parameter of variance. Instead, EUT allows the curvature of the nonlinear utility function to reflect attitude towards risk (the magnitude of $\beta$). Many subsequent studies have theoretically established various utility functional forms and empirically tested these alternative functions in transport (De Palma and Picard, 2005, Recker et al., 2005). Detailed EUT applications are presented in section 3.3.

### 3.2.3 Scheduling model

The scheduling approach, originally applied to modelling departure time choice (although it has also been applied to other choice problems), has long been regarded as the most important alternative approach for modelling travel time variability. Schedule delay and the omission of mean variance model, serves as the main carrier of travel time variability. The earliest explorations of this approach were carried out by Gaver (1968) and Vickrey (1969). They introduced a new theory for understanding travel time variability in the scheduling framework. Based on their theoretical work, Small (1982) formulated a departure time choice model of scheduling choice. His model highlights timing issues, which is the main difference from the mean variance model and the others. It is consistent with intuition since some time points, in particular preferred arrival time (PAT), are naturally regarded as key attributes for departure choice. Small (1982) defined schedule delay (SD) as the difference between PAT and actual arrival time:

$$SD = PAT - [T(t) + t] \tag{3.8}$$

$T(t)$ is the actual travel time determined by specific departure time *t*. Schedule delay late (*SDL*) occurs if $SD < 0$; schedule delay early (*SDE*) occurs if $SD > 0$. That is *SDE* and *SDL* is the amount of time by which the traveller arrives early or late comparing to *PAT*. In this case, we can evaluate the different contributions of *SDL* and *SDE* to modelling. It is assumed that disutility only occurs insofar as the traveller fails to arrive at the destination at their preferred arrival time, rather than occurring by itself as the mean-variance approach does. Small (1982) first specified the scheduling model for departure time choice using the following equation:

$$u(t^n) = \alpha t^n + \beta SDE^n + \gamma SDL^n + \theta D_L^n \tag{3.9}$$

where $D_L^n$ is the fixed penalty for late arrival. The parameters $\alpha$, $\beta$ , $\gamma$ and $\theta$ are the marginal utility for $t^n$, $SDE^n$, $SDL^n$ and $D_L^n$ respectively. It is expected that four parameters will be negative. The marginal disutility from $SDL^n$ may differ from the marginal disutility from $SDE^n$, suggesting travellers' different tastes towards arriving late and arriving early Small et al. (1999). Moreover, empirical findings also indicate that $SDE^n$ is preferred to $D_L^n$, and $D_L^n$ is preferred to $SDL^n$, i.e., $\beta > \theta > \gamma$.

Noland and Small (1995) were not merely concerned about the parameters of schedule delay but also with the real source of risk: i.e. travel time distribution. They believed that researchers should pay special attention not only to the difference between PAT and actual arrival time (consequence) but also the likelihood of being late or early (probability) in a statistical way. Similar to Small (1982)'s utility function, Noland and Small's scheduling expression is:

$$u(t^n) = \alpha E(t^n) + \beta E[SDE^n] + \gamma E[SDL^n] + \theta P_L^n \tag{3.10}$$

where $P_L^n$ is the probability of late arrival and it is evident that the value of $P_L^n$ is dependent on the assumed distribution.

## 3.2.4 Other approaches to modelling travel time variability

### *3.2.4.1 The safety margin approach*

At the early stage of travel time variability research, the cost of risk was simply derived from the extra travel time, i.e., the safety margin (Gaver, 1968, Knight, 1974, Thomson, 1968). It was assumed that travellers are capable of maximizing utility by selecting an earlier departure time with an acceptable 'slack time'. Reducing the disutility of travel time variability, therefore, corresponds to the reduction of 'slack time' allocated to the planned journey. Polak (1987) assumed that travellers consider two types of travel time: planned travel time, which is allowed by a traveller, and expected travel time which is based on the traveller's previous experience.

Similar indexes are the *buffer index* and *planning time index*, which are used in the US Federal Highway Administration's Urban Congestion Reports. The former is referred to as the extra proportion of journey time relative to the average journey time. For instance, a buffer index of 70% means that the traveller would like to spend an additional 70 minutes on a journey with an expected travel time of 100 minutes. The planning time index, meanwhile, is relative to the free flow travel time. The basic idea of both travel time variability indexes is that each traveller has an *a priori* idea of the average journey time.

### *3.2.4.2 Centrality-dispersion model*

A general method to account for travel time variability is to incorporate both the central tendency and dispersion of travel time distribution into models. This family of models is called the centrality-dispersion model. Note that the mean-variance model is a special case of centrality-dispersion models where there are a wide range of alternative measures for the central tendency and dispersion of travel time.

In order to consider the possible influence of extreme delays, it is useful to apply the quartile difference of travel time distribution. Lam and Small (2001) used the difference between the 90th percentile and the median travel time, denoted as dmp90. Their estimation results show that dmp90 provides a better explanatory power for risky route choice than standard deviation. Small et al. (2005b) conducted travel time variability research on the same corridor as Lam and Small (2001) (i.e. California State Route 91) but using different revealed preference data. They assumed that travellers are concerned about extreme delays, especially the upper tail of travel time distribution. The difference between the 80th and 50th percentile (dmp80) was employed in their research, and led to a better model fit than the other candidate models. Similar findings to the median-dmp80 approach can be found in Recker et al. (2005).

Other measures for travel time variability have also been evaluated in the transport literature. For instance, Senbil and Kitamura (2008) suggested that travel time variability can be measured by the difference between the maximum and minimum travel time. Other measures are also potentially useful, e.g. the ratio between standard deviation and mean travel time, and the percentage of observations that exceed the mean/median travel time, etc.

### *3.2.4.3 The mean lateness model*

This approach has been often referred to as the standard method for modelling the reliability of rail in the UK (ATOC, 2005). The original mean lateness model consists of two elements, namely scheduled journey time ($SchedT$) and mean lateness at destination ($L^+$). Thus, the utility function is expressed as:

$$u = \alpha SchedT + \beta L^+ \tag{3.11}$$

$\alpha$ and $\beta$ are the coefficients to be estimated, with both coefficients being expected to be negative, suggesting passengers' aversion attitudes to travel time and delays. Batley and Ibáñez (2009) and Batley and Ibáñez (2012) extended the mean lateness model by including three extra variables, i.e., mean lateness at boarding ($B^+$), standard deviation of in-vehicle journey time ($\sigma$), and train fare ($C$).

$$u = \alpha SchedT + \beta L^+ + \gamma B^+ + \pi\sigma + \theta C \tag{3.12}$$

This specification is regarded as a combination of mean-variance model and mean lateness model. In order to better understand the concepts underpinning the above measures of travel time variability, various journey time components are illustrated in Figure 3.3. And Table 3.1 illustrates several empirical studies with different models and data collection methods.

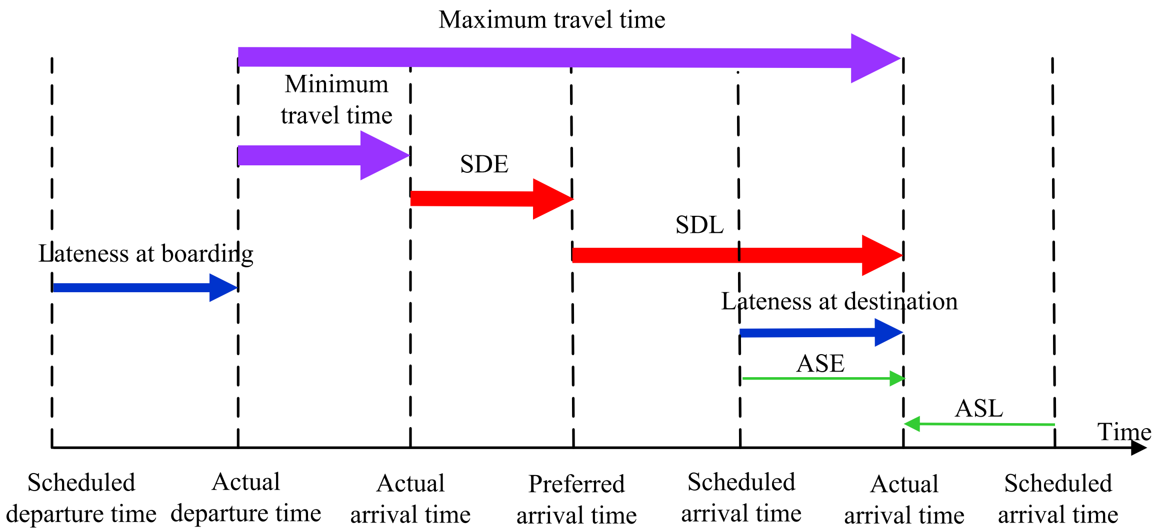



**Figure 3.3: The representations of travel time variability (modified version based on Batley and Ibáñez (2009))**

| Literature | Type of choice | Represent ation of risk | Sampling methods | Data on risky outcomes | Data on travellers' choices | Data on travellers' characteristics | Data on non-EUT |
|---|---|---|---|---|---|---|---|
| Ghosh (2001) | Mode | Dmp90 [a] | Database of the billing agency; random digit dialling | Estimated from loop detector data | Telephone survey | Telephone survey | NA |
| Lam & Small (2001) | Route & time-of-day | SDL & SDE | Database of the Department of Motor Vehicles; random digit dialling | Estimated from loop detector data | Telephone survey | Telephone survey | NA |
| Small, Winston & Yan (2005) | Route | Dmp80 [b] | Database of a market research firm. | Students drove on the free lanes repeatedly and clocked the travel time | Mail survey | Mail survey | NA |
| Bhat & Sardesai (2006) | Mode | Additional travel time | NR | Web-based survey (maximum travel time, usual travel time) | Web-based survey | Web-based survey | NA |
| Senbil & Kitamura (2008) | Mode & departure time | Time difference [c] | Randomly select drivers who passed a toll gate on R13 during morning peak periods. | Travel diary | Travel diary | Travel diary | NA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Senbil & Kitamura (2003) | departure time | Time difference [d] | Questionnaires were mailed to randomly selected resident drivers | Mail survey | Mail survey | Mail survey | Preferred arrival time [e] |
| Ettema & Timmermans (2006) | Departure time | SDL & SDE | NR | Trajectory-methodology | NR | NR | NA |
| Carrion-Madera & Levinson (2010) | Route | Standard deviation | Flyers and emails | Transponder and GPS logger. | Transponder and GPS logger. | Questionnaires | NA |

NA: Not applicable.

NR: Not reported.

a: 90$^{th}$ percentile travel minus median travel time.

b: 80$^{th}$ percentile travel minus median travel time.

c: The time difference between the fastest trip and the slowest trip.

d: The difference between actual arrival time and corresponding reference points.

e: Preferred arrival time is collected to locate the reference points

**Table 3.1: The RP studies modelling travel time variability**

## 3.3 Risky choice models in transport from an EUT perspective

The stream of studies discussed here, however, omits a key component of risky choice model, i.e., travellers' attitude towards risk,[5] thus, strictly speaking, it is not a risky choice model structure! Recent research has resorted to utility theory from the vast literature on microeconomics, as these risk attitudes can be explicitly embodied in the utility function. And the dominant theoretical paradigm for modelling traveller's risky choice behaviour is to combine expected utility theory (EUT) with random utility maximization (RUM) (Batley, 2007, De Palma and Picard, 2005, Liu and Polak, 2007). Consequently, modellers' uncertainty in terms of unobserved heterogeneity (from RUM's perspective) and decision makers' uncertainty towards travel time (from EUT's perspective) is jointly addressed by this mixed model.

### 3.3.1 Nonlinear utility functions

Pratt (1964) initially applied attitude towards risk into a utility function which is concave when the decision maker is risk averse, convex when the decision maker is risk prone, and linear for risk neutrality. This flexible specification enables attitude towards risk to be embedded into the model by nonlinearly transforming the utility function. Existing studies have provided considerable insights regarding various utility forms with respect to the von Neumann–Morgenstern (vNM) utility. The following only summarizes the four most 'popular' nonlinear utility formulas:

- Quadratic Utility Function: $u(x) = ax - bx^2$ , $u'(x) = a - 2bx$ , $u''(x) = -2b$ . Recall Arrow-Pratt index of risk-aversion introduced in chapter 2, the coefficient of absolute risk aversion is $A^a(x) = -\frac{u''(x)}{u'(x)} = \frac{2b}{a-2bx}$, and the coefficient of relative risk aversion is $A^r(x) = -\frac{xu''(x)}{u'(x)} = \frac{2bx}{a-2bx}$.
- Power Utility Function: $u(x) = \frac{ax^{(1-b)}}{1-b}$, $u'(x) = x^{-b}$, $u''(x) = -bx^{-(b+1)}$. In this case, $A^a(x) = -bx^{-1}$ and $A^r(x) = -b$. Note that $\frac{dA^r(x)}{dx} = 0$ in power utility and, therefore, decision makers using this utility function have *constant relative risk aversion* (CRRA).

[5] It should be noted that the mean-variance model allows an implicit way to measure attitudes to risk by using the *reliability ratio*, see Recker (2005) for details. This is defined as the ratio of the marginal utility of travel time variability to the marginal utility of mean travel time.

- Negative Exponential Utility Function: $u(x) = -e^{-ax}$, $u'(x) = ae^{-ax}$, and $u''(x) = a^2e^{-ax}$. Correspondingly, $\frac{dA^a(x)}{dx} = 0$ indicating *constant absolute risk aversion* (CARA).
- Hyperbolic Absolute Risk Aversion (HARA): $u(x) = \frac{(1-r)}{r}(\frac{ax}{1-r} + b)^r$, $u'(x) = a(\frac{ax}{1-r} + b)^{(r-1)}$, and $u''(x) = a^2(\frac{ax}{1-r} + b)^{(r-2)}$ where $b > 0$. This specification is the general class of utility function which is widely accepted due to its mathematical tractability. $A^a(x) = -a(\frac{ax}{1-r} + b)$ and $\frac{dA^a(x)}{dx} = \frac{-a^2}{1-r}$ imply *decreasing absolute risk aversion* (DARA) if $r < 1$, *increasing absolute risk aversion* (IARA) if $r > 1$, and CARA if $r$ approaches positive or negative infinity. In addition, HARA is often referred to as the general utility form which can be converted to the other structures. For instance, it becomes quadratic if $r = 2$, a negative exponential form if $r$ approaches negative infinity and $b = 1$, and a power function if $r < 1$ and $b = 0$.

### 3.3.2 RUM-EUT approach

There are only a few studies using a RUM-EUT framework (many works simply use expected value of travel time without risk attitudes, so are actually RUM-EVT models). One reason for this may be the difficulty of establishing such a mixed model incorporating EUT into a RUM structure. Empirical studies have explored the framework of this mixed model from a variety of perspectives. The initial attempts mainly account for EUT in a mean variance model, with Polak (1987) being the first to seek to account for a computationally tractable nonlinear utility form in transport. Based on Polak's findings, Senna (1994) elaborated the mean variance model specification using nonlinear utility, and explained the EUT approach for analysing travellers' attitude towards risk. In addition to the mean variance model, utility theory has also been applied to the other model frameworks, such as the scheduling model (Noland and Small, 1995). The common feature of these studies is that both attitude towards risk (nonlinear utility in this case) and probabilistic choice (random utility in this case) are included.

Batley and Daly (2004) take into account both the researcher's and the traveller's uncertainty in the context of departure time choice. They proposed the *random expected utility maximization* model by adding an error term of RUM into the expectation operator of EUT, i.e., $u^n = \sum_{k=1}^{K}(v_k^n + \varepsilon^n)p_k^n$. In this sense, the corresponding probability of choosing prospect *n* is expressed as $P^n = P(\sum_{k=1}^{K}(v_k^n + \varepsilon_k^n)p_k^n > \sum_{k=1}^{K}(v_k^m + \varepsilon_k^m)p_k^m)\ \forall\, m \neq n$. The outcome value $v_k^n$ is the weighted additive result of travel time, scheduled delay early, scheduled delay late, and a dummy variable of lateness. Although they merely employed a linear value $v_k^n$ rather than a nonlinear utility, this model specification is still flexible enough to allow other work to incorporate EUT and its alternative approaches.

Recker et al. (2005)compared EVT and EUT models in the risky route choice context, and observed significant difference of flow allocations between risk neutral travellers (modelled by EVT) and risk averse travellers (modelled by EUT). In their EUT specification, they employed a similar model to Batley and Daly's, but they also applied a nonlinear transformation of utility function, such that the prospect utility function was expressed as $u^n = \sum_{k=1}^{K}(g(v_k^n, \alpha) + \varepsilon_k^n)p_k^n$, where $g(v_k^n, \alpha)$is the nonlinear outcome utility function and $\alpha$ represents the parameter of risk attitude. Evidently both the probabilistic term represented by $\varepsilon^n$ and the risk attitude termed by $\alpha$ are successfully included in this single utility function. It should be noted that this specification is potentially highly complicated due to the expectation of the error term $\varepsilon^n$. Furthermore, the validity of this model specification it is still not clear, especially in respect to the expectation of error term.

Bates et al. (2001) established a more tractable model which only accounts for the expectation of observed utility, i.e., $u^n = \sum_{k=1}^{K} v_k^n p_k^n + \varepsilon^n$, where $\varepsilon^n$ is the additive random error applying to the whole prospect, and the outcome value function is $v_k^n = \beta_1 SDE + \beta_2 SDL + \beta_3 Fare + \beta_4 Headway$. Note that the expectation operator is only taken over travel time distribution, the prospect utility function is then expressed as:

$$u^n = \beta_1 E(SDE) + \beta_2 E(SDL) + \beta_3 Fare + \beta_4 Headway + \varepsilon^n \quad (3.13)$$

where $E(SDE)$ and $E(SDL)$ is the expected value of scheduled delay early and scheduled delay late, respectively. This specification is flexible enough to accommodate different RUM by changing the assumption of error term $\varepsilon^n$, e.g., GEV and MMNL. Note that $E(SDE)$ and $E(SDL)$ are generally linear functions implying risk neutrality.

Another extension of Bates et al. (2001)'s model, which is of interest to us, is to explore the RUM-EUT approach embedded with risk attitude in this model framework. Liu and Polak (2007) introduced an explicit approach to account for attitude towards risk. Based on Bates et al. (2001), they applied an exponential transformation to the outcome values, so that the prospect utility function turns out to be:

$$u_k^n = \sum_{k=1}^{K} \frac{1-e^{-\alpha v_k^n}}{\alpha} p_k^n + \varepsilon^n \tag{3.14}$$

where $v_k^n$ is the observed part of the outcome value as used in Bates et al. (2001), and $\alpha$ characterizes the traveller's attitude towards risk. Specifically, $\alpha < 0$ represents risk aversion, $\alpha > 0$ means risk proneness, and $\alpha = 0$ corresponds to risk neutrality. In a subsequent work Polak et al. (2008) investigated the possible heterogeneity in terms of travellers' risk attitudes. In this paper, they argue that the risk attitude parameter $\alpha$ is no longer a single value as shown in MNL model, but rather a specific distribution with mean and standard deviation. They conducted the new RUM-EUT model where the CARA utility function and the mixed multinomial logit (MMNL) approach is applied to the modelling of risk attitude and its unobserved heterogeneity respectively. The model fit is significantly improved if the observed and unobserved heterogeneity is taken into account, and the estimation shows that travellers, on average, are slightly risk averse, while it should also be noted that there exists substantial heterogeneity of risk attitude across the sample.

A simpler form of RUM-EUT was proposed by De Palma and Picard (2005) where risky outcome is exclusively associated with travel time, and the prospect utility is evaluated on the expectation of nonlinear transformation of travel time (i.e. it is evaluated from the utility domain rather than value domain). The De Palma and Picard model is expressed as $u^n = \sum_{k=1}^{K} g(t_k^n, \alpha)\, p_k^n + \varepsilon^n$, where $\varepsilon^n$ is the random error term, and $t_k^n$ is the travel time of the $k^{th}$ outcome. The nonlinear utility function $g(t_k^n, \alpha)$ serves as the methodology to measure attitude towards risk. In De Palma and Picard's research, both linear and nonlinear models are taken into account, namely mean-standard deviation model, mean-variance model and the CRRA and CARA models. The calibration results showed that 66% of people in the sample were risk averse or risk neutral, and around 33% were risk prone. While the main purpose of this paper is to explore the factors influencing attitude towards risk, the most essential contribution is their risky choice model specification, which merges RUM and EUT in a computationally tractable functional form.

This section reviews the studies that combine RUM and EUT in a proper functional form. Generally, the studies attempt to apply the expectation operator to different components of the utility function, namely random utility (the observed utility plus an error term), the observed utility, and a specific attribute. Moreover, the majority of these studies applied this integrated model in the context of risky route choice or departure time choice (Boyce et al., 1999, Liu et al., 2002, Tatineni et al., 1997). Given the possible computation problems of utility based specification, this thesis pays special attention to attribute based risky choice specifications.

## 3.4 Risky choice models in transport from non-EUT's perspective

.Over the past decade the field of transport studies has witnessed increased interest and growth in the use of non-EUT models. Possible reasons for the increasing research on non-EUT are twofold. Firstly, given the observed violations of EUT, it is interesting to identify whether non-EUT models with more complicated specifications may deliver better estimation and prediction results with respect to EUT in various transport contexts. Moreover, the intuitive appeal of non-EUT approaches also attracts researchers to establish alternative models for realism, and obtain new insights for project appraisal and transport policy making. This thesis is by no means an exception.

### 3.4.1 RUM-NEUT approach

Having discussed the theoretical development of non-EUT in section 2.5, this section reviewed the empirical studies applying non-EUT into RUM structure (RUM-NEUT). There are essentially two streams of transport research related to the RUM-NEUT approach. The first strand of research concentrates on the descriptive capability of non-EUT by synthesizing various non-EUT components into the RUM framework for realism. The earliest exploratory work on PT is concerned with a scheduling model. Jou and Kitamura (2002) conducted initial work to examine the applicability of a reference dependent approach to departure time choice. They assumed dual reference points: earliest acceptable arrival time and official work start time. This scheduling based PT model was further developed by Senbil and Kitamura (2004) who added another important reference point, preferred arrival time (termed as a pseudo-reference point), to the original dual reference point model of Jou and Kitamura (2002). The preferred arrival time is usually between the other two reference points, and an arrival time outcome between the earliest acceptable arrival time and the preferred arrival time is termed as a quasi-gain, which has a positive and a concave utility function. This so-called quasi-gain is also interpreted as scheduled delay early in a scheduling model, however, they have different estimated signs, i.e., whereas the quasi-gain utility is positive, the schedule delay early utility is negative in other literature (e.g. Bates et al., 2001).[6] This

[6] Another relevant study is that of Michea and Polak (2006) who proposed a CPT embedded scheduling model with only loss (both scheduled delay late and scheduled delay early are termed as loss, and the corresponding

contradiction potentially suggests that researchers should pay more attention to the relationship between preferred arrival time and reference point. van de Kaa (2008) conducted a meta-analysis of PT and other non-EUT assumptions in order to synthesize a generic theory for travel choice behaviour. This *meta theory of choice behaviour* enables him to propose a so called *extended prospect theory* which is capable of the description and prediction of travel choice behaviour. An extensive literature review on reference point of PT is presented in the Appendix A.

In the travel demand domain, research has almost exclusively focused on utility as the single scalar; a notable exception, however, is Chorus et al. (2008), who found that the basic framework of RT is compatible with probabilistic choice models like discrete choice models and, therefore, they proposed the so called *random regret minimization-approach*. They considered the most significant contribution of the random regret minimization-approach as being "*its ability to capture semi-compensatory behaviour as well as choice set-specific preferences, within a model that is as tractable and parsimonious as RUM's linear-additive MNL-model*". Chorus (2011) went on to develop another version of RT specification in a route choice context in which expected utility maximization is employed to measure preference. This modified version is similar to the theory of disappointment (TD) in which the non-decreasing disappointment function is replaced by the regret function.

The second stream of research aims to test non-EUT performances under advanced operational models. The generic purpose of these studies is to account for travellers' heterogeneities in taste parameters and even decision making paradigms. Hensher and Li (2012) evaluated the RDEU model within a mixed multinomial logit (MMNL) framework. In these methods, the unobserved heterogeneity of travel time parameter is addressed by using a random parameter following a constrained triangular distribution (Hensher and Greene, 2003, Revelt and Train, 1998). Hess et al. (2012) applied the latent class model to account for the heterogeneity in travellers' decision rules. In their PT case-study, the mixed model with three different reference points is simultaneously evaluated. They concluded that further research should take into account the heterogeneity in decision making paradigms.

parameters are negative). There is no explicit reference point which distinguishes gain and loss and, therefore, it is actually a general form of a RDEU model.

### 3.4.2 Overview of empirical results

Since the early 2000s, transport researchers have become increasingly interested in the adequacy of non-EUT to predict travellers' risky behaviour. While the importance of these relatively new methods has been properly recognized, empirical evidences on non-EUT models are surprisingly limited. One notable exception is the work conducted by Michea and Polak (2006) who present a systematic comparison between EUT and non-EUT models. Their model evaluation is based on the same SP dataset used by Bates et al. (2001) which describes a series of risky choices between hypothetical unreliable train services. Michea and Polak found that the inclusion of decreasing sensitivity parameters statistically significantly increased the goodness of fit. This finding is in line with the significant model fit improvement produced by EUT, which highlights the importance of accounting for nonlinear utility in the non-EUT models. While nonlinear utility was not included in Michea and Polak's non-EUT models (except CPT), the final calibration results still showed that all the non-EUT models led to a statistically significant increase of log-likelihood with respect to the basic model EUT.[7] This suggests a significant benefit of adopting non-EUT for improving model fit, and it especially indicates the potential benefits from using nonlinear weighting functions. As the only transport based study which has conducted systematic comparison of different non-EUT approaches, Michea and Polak (2006) have provided many empirical insights into how non-EUTs perform in a transport context.[8] More comparison and empirical tests, however, should be conducted to fill in this research gap.

Recent work is also interested in measuring the degree of loss aversion in transport context. The general method is to assess the ratio between scale parameters of gain and loss. van de Kaa (2008) summarized twenty trading experiments to elicit the loss aversion factor which turns out to be at the average level of 2.0. The similar method was applied by Hess et al. (2008) who obtained a higher loss aversion factor at 3.15 using SP data. Gao et al. (2010) compared EUT model and CPT model in the strategic route choice context, and it was found that CPT provides better model fit in terms of log-likelihood (-478.7 vs -720.3). They,

[7] The original Bates et al. (2001)'s SP data was collected on two corridors. The long corridor is between central London and the cities of Bristol and Cardiff, while the short corridor is between central London and Hayes. Michea and Polak (2006)'s work suggests that all four non-EUT models provide statistically significant improvement of model fit in the long distance corridor scenario, while only this is only the case for WUT in the short distance corridor scenario.

[8] To the author's knowledge, the other transport literature is Ramos et al. (2011) who also carried out a comparison of EUT and non-EUT models (prospect theory and regret theory). However, RUM is not taken into account and, therefore, the important parameters for non-EUT cannot be estimated. The prediction result is simply based on the number of correct predictions compared to experimental data.

meanwhile, identified that the loss aversion factor is around 2.09. The other potentially useful method is to measure the difference between WTA and WTP. This disparity reflects the individual's different response to owned goods and compared to acquiring the goods of others, from the valuation perspective (Rose and Masiero, 2010).

A recent work conducted by Koster and Verhoef (2010) proposed a ranked dependent scheduling model, and tested the cumulative probability weighting function within a linear value specification. Whilst attitude towards risk is not reflected by utility specification due to the adoption of a linear value function, they argued that travellers' risk perceptions should be primarily related to nonlinear decision weight rather than utility. Schwanen and Ettema (2009), meanwhile, explored the usage of CPT in the context of collecting children under congested networks. Using the same single-parameter weighting function for gain and loss, the estimation result shows an inverse S-shaped curve. They did not explain travellers' risk attitudes in their paper but, at least, the utility function implies risk aversion for loss and risk proneness for gain.[9] The sensitivity parameter turns out to be 1.09-1.10, suggesting convex utility for gain and concave for loss. Schwanen and Etterna argued that this is in line with intuition since lateness is more objectionable as travel time increases.

Given the fact that the weighting function and the associated estimates has a substantial influence on model fit and attitude towards risk, Table 3.2 shows a range of empirical evidence using five 'popular' weighting functions. It tells us the empirical range of weighting parameters estimated inside and outside the transport field. A relatively wide range of estimates can be observed. This suggests that the estimated weighting parameter may vary across different choice contexts. Based on our comprehensive literature review, future research could extend to meta-analysis of weighting functions.

---

9 Schwanen and Ettema (2009) demonstrated that travellers are increasingly sensitive to travel time, i.e., the outcome utility function is concave for loss and convex for gain. This finding contradicts the diminishing sensitivity derived from the original PT. Such a function with an increasing marginal utility implies risk aversion for loss and risk proneness for gain.

**Table 3.2: Overview of empirical estimates**

| Weighting function | γ | τ | Literature |
|---|---|---|---|
| TK $w(p)=\frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{(1/\gamma)}}$ | 0.55 | | Razo and Gao (2011) |
| | 5.31 | | Hensher and Li (2012) |
| | 0.76 | | Hensher et al. (2011) |
| | 0.69 | | Gao et al. (2010) |
| | 0.83 | | Gao et al. (2010) |
| | 0.56 | | Camerer and Ho (1994) |
| | 0.71 | | Wu and Gonzalez (1996) |
| | 0.60 | | Abdellaoui (2000) |
| | 0.96 | | Stott (2006) |
| | 0.67 | | Bleichrodt and Pinto (2000) |
| | 0.52 | | Camerer and Ho (1994) |
| | 1.15 | | Li and Hensher (2012a) |
| | 0.69 | | Gao et al. (2010) |
| | 0.83 | | Schwanen and Ettema (2009) |
| WG $w(p)=\frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{\tau}}$ | 0.72 | 1.57 | Wu and Gonzalez (1996) |
| | 0.93 | 0.89 | Stott (2006) |
| GE $w(p)=\frac{\tau p^{\gamma}}{\tau p^{\gamma}+(1-p)^{\gamma}}$ | 1.25 | 1.95 | Hensher et al. (2011) |
| | 6.73 | 0.26 | Roberts et al. (2008) |
| | 2.00 | 1.63 | Michea and Polak (2006) |
| | 0.77 | 0.69 | Tversky and Fox (1995) |
| | 0.68 | 0.84 | Wu and Gonzalez (1996) |
| | 1.59 | 0.31 | Birnbaum and Chavez (1997) |
| | 0.44 | 0.77 | Gonzalez and Wu (1999) |
| | 0.60 | 0.65 | Abdellaoui (2000) |
| | 0.55 | 0.82 | Bleichrodt and Pinto (2000) |
| | 0.96 | 1.40 | Stott (2006) |
| Prelec-I $w(p)=e^{-(-lnp)^{\gamma}}$ | 1.16 | | Hensher et al. (2011) |
| | 0.37 | | Razo and Gao (2011) |
| | 0.74 | | Wu and Gonzalez (1996) |
| | 0.94 | | Stott (2006) |
| | 0.53 | | Bleichrodt and Pinto (2000) |
| Prelec-II $w(p)=e^{-\tau(-lnp)^{\gamma}}$ | 1.42 | 0.73 | Hensher et al. (2011) |
| | 0.53 | 1.08 | Bleichrodt and Pinto (2000) |

# 3.5 Summary

This chapter has discussed the development of risky choice model in transport contexts. It has revealed a modelling strategy which is potentially capable of synthesizing risky choice theories and RUM into a realistic operational model utilising a risky choice framework. Taking a broad look at the main findings from the literature review, traditional methods such as mean-variance and scheduling apprpaches are still the dominant methods used in practice to address travel time risk. These standard practical methods consider travel time risk by incorporating the attribute of travel time variability into model. With the development of risky choice theories, substantial research has been conducted to identify the applicability of EUT and non-EUT in the domain of travel time risk. It could be concluded that EUT is the most widely used and practical method for modelling travel time risk, while it is still unknown whether the complexity of non-EUT can lead to better performance in practice. Evidently more research remains to be conducted in order to fill in some specific gaps in the research.

In particular, EUT has been referred to as the dominant theory for risky travel behaviour research, while non-EUT has attracted only very few studies. Furthermore, with only a few of exceptions,[10] most empirical studies have conducted non-EUT research outside a risky choice framework. That is, they did not specify risky outcomes and associated probabilities. The relative performances of non-EUT and EUT in terms of their estimation and prediction have not, therefore, been properly addressed in existing literature, and it is this gap that consequently forms the main focus of this thesis. In order to address this issue, systematic comparisons across EUT and its alternative models are required, but only Michea and Polak (2006) have provided the foundational work for this. More research is therefore required in order to gain insights into how these non-EUT models actually perform in the real world. Surprisingly, it was also found that almost all empirical studies of non-EUT approaches employed stated preference (SP) data rather than revealed preference (RP) data. This is another research gap which urgently needs to be filled. These shortcomings form the main focus of the Chapter 4 which explicitly demonstrates the theoretical and empirical challenges regarding applying non-EUT to travel choice behaviour. Moreover, the contributions of this thesis to the methodology employed in this area of research are also presented in the next chapter.

[10] Two notable exceptions are Michea and Polak (2006) and Hensher et al. (2011). Both applied SP survey data to show respondents various prospects with a series of risky outcomes and associated probabilities; they were consequently capable of establishing viable risky choice models.

# Chapter 4 MODEL AND DATA SPECIFICATION

## 4.1 Introduction

The discussion in Chapter 3 has highlighted the potential benefits of introducing several behavioural theories, and also clarified gaps from the perspective of model specifications and data usages. One of the evident complications is that components of non-EUT should be embedded into a risky choice framework which is flexible enough to accommodate different non-EUT approaches. We sought to explore various non-EUT approaches embedded into the risky choice framework, however, several theories were found to need extra information which cannot be achieved at the current stage, such as the prior probability in the prospective reference theory, and the certainty equivalent in the theory of disappointment aversion. Particular attention therefore is paid to the specifications of Subjective Expected Utility (SEU) Theory, Rank Dependent Expected Utility (RDEU), Prospect Theory (PT)Cumulative Prospect Theory (CPT). Another problem which has limited the application of risky choice models is that it is difficult to find a revealed preference (RP) data allowing sufficient variation and uncertainty of travel time.

Given the above problems, this chapter presents a methodology for establishing models and collecting RP data. The remainder of this chapter is organized as follows. Section 4.2 begins with the introduction of the notations and terminology of our risky choice framework, and a description of the specific modelling approaches for non-EUT. The methodology of RP data collection is discussed in section 4.3. Based on the methods introduced in this chapter, subsequent chapters will present a series of applications in an RP context. The chapter is summarized in section 4.4.

## 4.2 Model framework

As stated by McFadden (2000), “Even for routinized, ‘rational’ decisions such as work trip mode choice which may be consistent with the economists’ standard model, psychological elements are likely to be important in the construction and reinforcement of preferences…The cognitive psychology of choice should be required study for all travel demand analysts, even the die-hard RUM modellers.” Working from this basis, it is quite sensible to combine RUM with non-EUT approaches in an operational framework, and thereby to attempt to obtain insights into their respective performances. Here we present the general model framework used in this thesis (refer to subsection 2.2.1 for notation and terminology).

It has been found previously that decision makers would subjectively combine outcome value $v_m^n$ and the associated probability $p_m^n$ to produce a nonlinear utility $u_m^n$. This requires decision rules of EUT and non-EUT to address $v_m^n$ and $p_m^n$. Specifically, let the nonlinear transformation $u_m^n = g(v_m^n, \alpha)$ calculate outcome utility, and $w_m^n = \pi(p_m^n, \gamma)$ represent nonlinear decision weight. The extra parameter $\alpha$ characterizes an individual's attitude towards risk (refers to section 3.3.1) , and the parameter $\gamma$ in the weighting function $\pi(.)$ represents an individual's perceivable bias in terms of probability (refers to section 2.5.1). Thus, if there are $K$ outcomes, the expected utility of risky outcomes is generally expressed as:[11]

$$u_m^n = \sum_{k=1}^{K} g(v_{mk}^n, \alpha)\, \pi(p_{mk}^n, \gamma) \tag{4.1}$$

here, we can convert deterministic choice into a probabilistic choice formulation by introducing the RUM decision rule $f$ and an unobserved error term $\varepsilon^n$. As a result, the utility of prospect $n$ can be expressed as:

$$u^n = f(u_m^n, \beta) + \varepsilon^n \tag{4.2}$$

where $\beta$ corresponds to an individual's taste parameter that is to be estimated. Within this framework, the probability of an individual $i$ choosing the prospect $s^n$ is given by:

$$P(s^n|C) = Pr(u^n > u^t) \;\; \forall\, s^t \in C, t \neq n \tag{4.3}$$

Given this flexible model framework, a wide range of non-EUT approaches can be synthesized into the RUM-NEUT structure proposed here by making specific assumptions on the functional form of $g(.)$ and $\pi(.)$. The following subsections present the main methods for modelling non-EUT within this risky choice model framework.

[11] This is only the general functional form with only a single risky attribute $a_m^n$. Interactions with socio-demographic variables are also not shown in this form, for the sake of clarity. Another notable study has been conducted by Liu and Polak (2007) in which they proposed an EUT framework in which risky outcomes are characterized by multiple attributes including schedule delay early, schedule delay late, fare and headway.

### 4.2.1 Modelling reference dependence

A reference dependent approach, mainly PT and its advanced version CPT, has been applied fruitfully in finance and economics. In the domain of travel demand models, PT almost serves as the only non-EUT approach, and one that is employed by most researchers.

The original version of PT highlights two choice stages, namely an editing stage and an evaluation stage. The former applies different decision heuristics to simplify the choice context, while the latter aims at the evaluation of risky outcomes and prospects. Note that the reference point plays an essential role in the evaluation stage, with the outcome utility being measured differently according to its relative location to reference point. It is this asymmetrical measure of reference dependence that differentiates PT from other alternatives. This subsection, therefore, seeks to explore the method for modelling reference dependence and, in particular, for determining the reference point.

#### *4.2.1.1 Model specification*

In our PT specification, we define that the utility function of PT is based on the relative value of travel time rather than travel time *per se* (such as mean travel time). As such, it successfully addresses reference dependence by accounting for the difference of actual travel time and reference travel time. The utility function of PT should contain the following component:

$$u^n = \cdots + \beta_{gain} \sum_{g=1}^{K} \max(0, t_{ref}^n - t_g^n)\, \pi(p_g^n) + \beta_{loss} \sum_{l=1}^{K} \max(0, t_l^n - t_{ref}^n) \pi(p_l^n) + \cdots \tag{4.4}$$

where $t_{ref}^n$ gives the reference point for the travel time attribute. To measure outcome utilities asymmetrically, *K* travel time outcomes are divided into gain and loss according to their relative magnitudes with respect to $t_{ref}^n$. Thus, $t_g^n$ and $t_l^n$ corresponds to the travel time less and more than $t_{ref}^n$ respectively, suggesting travel time outcomes of gain and loss. Parameters $\beta_{gain}$ and $\beta_{loss}$ are expected to be positive and negative respectively, due to individuals' asymmetrical tastes towards gain and loss. In the most general case, outcome probability should be nonlinearly transformed to decision weight by using the weighting function $\pi(.)$.

If we admit the assumption of the reference dependence of PT, the corresponding prediction generated by this model must be highly reliant on the specific value of the reference point. Indeed, it has been found that determining the reference point has been regarded as the main problem for the application of PT. In experimental economics, take a gambling experiment for instance, risky outcomes are simply divided into gains and losses according to the natural reference point, £0. In travel choice behaviour, on the other hand, it is never easy to adopt the reference dependent approach since it is especially difficult to determine what a traveller's reference point is (refer to subsection A.2.1 for literature review). This problem encourages close attention to setting the reference point prior to any extensive application of PT.

#### *4.2.1.2 Methodology for determining reference point*

Based on the above discussions, we now introduce our methods for determining the reference point. The first issue is the definition of reference point. In our risky choice framework, the reference point can be applied to reference prospect, outcome and attribute. Given that, here, risk is only derived from unpredictable travel time, we assume that travellers may only account for reference travel time when they make a route choice. The second issue arises when it is attempted to determine this reference travel time. In this subsection three potential methods are presented.

The first method is the main method applied in existing transport literature associated with reference dependent choice behaviour. It assumes that a traveller's reference point may be a common or widely accepted travel time. For instance, decision makers might consider the mean or median travel time experienced by the target population as the reference travel time. Given the fact that travellers can only form the reference point from their travelling experiences, however, this proposal is arguably implausible since travellers cannot know these travel times. One solution is to give travellers the travel time information through a website or social network, but this still requires modelling the influence of information, e.g. dynamic models. Another possible solution is to track each individual's commuting history using advanced equipment, e.g. GPS and cell phone data, and then extract the average travel time of each respondent. In this way, the calculated travel time reflects individual respondents actual travelling experiences recorded in equipment. Despite the reliability and validity of this tracking data (e.g. GPS data), it is potentially time and resource intensive.

The second method attempts to generate the estimated reference point that best fits the data. It is computationally difficult to implement, however, since the calibration process is highly sensitive to the estimated value of the reference point. In fact, there is a kink of utility around the reference point which dramatically affects the estimation results. As such, model calibration is necessary to test different initial values of reference points in order to obtain the estimates under global maximization. Moreover, this endogenous estimation method potentially requires an iteration algorithm to find out the estimated reference point which satisfies the assumptions of PT.

The last method is closely related to survey design. In an RP survey exercise, respondents are required to provide extra information regarding their reference points, such as the ideal commuting time, average journey time of recent trips. It is very easy to obtain this information by explicitly presenting the questions like 'what is your most recent trip', 'when do you usually leave home in the morning peak', 'what is the journey time that you usually spend travelling from home to your office'. Unfortunately, however, to the author's knowledge, this kind of RP survey based method for collecting reference points has not been applied in existing transport literature.

In closing, whilst it is relatively arbitrary to assume a natural reference point, this method has been hitherto been regarded as a simple way to incorporate PT into risky route choice modelling. We should, however, pay careful attention to its validity for realism before applying it to our models. The benefit of the survey based method for valuing reference point is its simplicity and flexibility, but the drawback is that it cannot take account of respondents' perception errors and cognitive bias when they answer questionnaires. For instance, a traveller's perception of reference alternatives as shown in a questionnaire may not be the same as their real perception of reference alternatives in their real commuting experiences. In this case, the respondent considers the hypothetical reference alternative as being as common an alternative as others. In contrast, whilst the endogenous estimation method for valuing reference points is constrained by the sophisticated specification and complicated calibration required it does not need additional survey based information and, therefore, it avoids respondents' potential misperceptions due to inappropriate survey design. Consequently, endogenously estimated reference points and natural reference points are the selected methods in this thesis.

### 4.2.2 Modelling diminishing sensitivity

Another important component of PT is diminishing sensitivity, which leads to a nonlinear outcome utility function (refer to subsection 2.5.3.2 for the details). Two techniques are employed here to incorporate diminishing sensitivity into our PT and CPT models.

Firstly, a continuous method is introduced, which is in line with the proposal in the original version of PT. To model reference dependence and diminishing sensitivity jointly, the utility function is expressed as:

$$u^n = \cdots + \beta_{gain} \sum_{g=1}^{K} (\max(0, t_{ref}^n - t_g^n))^{\theta}\, \pi(p_g^n) + \beta_{loss} \sum_{l=1}^{K} (\max(0, t_l^n - t_{ref}^n))^{\gamma}\, \pi(p_l^n) + \cdots \tag{4.5}$$

where again $\beta_{gain}$ and $\beta_{loss}$ are the parameters for gain and loss respectively, and it is expected that $|\beta_{loss}| > |\beta_{gain}|$ if loss aversion holds. Note that the only difference between this functional form and the previous reference dependent functional form is the parameter $\theta$ and $\gamma$, which represent the different levels of diminishing sensitivity for gain and loss. In particular, travellers express diminishing sensitivity to travel time delay if $\gamma$ is statistically significantly smaller than one. The same also applies to diminishing sensitivity to time saving if $\theta$ is statistically significantly smaller than one.

One merit of this continuous method is that the linear value of gain and loss is converted to a nonlinear utility. That is, the diminishing sensitivity parameter is believed to have a similar utility distortion effect as EUT, i.e., convex utility for loss and concave utility for gain. We cannot, however, take decreasing sensitivity for granted without carefully estimating the models. One relatively convincing point is that the level of sensitivity should play an important role in behavioural models.

Unlike continuous methods using nonlinear parameters, discontinuous methods maintain the linear utility form, but separate outcomes into several cut-offs.[12] This piecewise linear approximation is capable of estimating different taste parameters for different ranges of a selected attribute. For instance, two attributes to represent the loss of travel time can be proposed, namely $t^{++}$ refers to the travel time increasing by 10 minutes with respect to the reference travel time, and $t^{+}$ refers to the travel time increasing by 5 minutes. Then parameters $\beta^{++}$ and $\beta^{+}$ are employed to characterize travellers' different tastes regarding different levels of travel time loss. It is expected that parameters for punctuality is $\beta^{++} < \beta^{+}$ <0, implicitly suggest diminishing sensitivity. This discontinuous method has a similar effect as a continuous method in embodying travellers' sensitivity to a specific attribute. Moreover, another benefit of using a discontinuous strategy is that this piecewise linear approximation approach maintains the functional form with linear parameters.

[12] Swait (2001) firstly proposed such a cut-off method which has been subsequently applied to freight transport by Danielis and Marucci (2007).

While the continuous approach is exactly in line with the assumption of the original version of PT, the discontinuous method also plays an important role in that the continuous method cannot provide statistically significant estimates of diminishing sensitivity parameters. Thus both methods are taken into account in this thesis.

### 4.2.3 Modelling nonlinear decision weight

A nonlinear weighting function, as an essential compensatory factor of nonlinear utility of EUT, has been increasingly applied in other domains, e.g. finance, environmental policy, marketing etc (refer to subsection 2.5.1 for the details). In transport, literature relating to decision weights is surprisingly scarce and more research urgently needs to be carried out to investigate whether travellers have biased perceptions of the likelihoods of travel time outcomes. Here we intend to combine a nonlinear utility function and a nonlinear weighting function into a generic function in order to jointly reflect travellers' attitudes towards risk.

#### *4.2.3.1 Model specifications*

For SEU and RDEU specifications, the functional form contains not only a nonlinear utility function $g(t_k^n, \alpha)$ as used in EUT, but also nonlinear weighting functions $\pi(p_k^n, \theta)$. The utility function of PT should contain the following component:

$$u^n = \cdots + \beta_{gain} \sum_{g=1}^{K} g(t_k^n, \alpha)\, \pi(p_g^n) + \cdots \tag{4.6}$$

The nonlinear weighting function also plays a crucial role in the applications of CPT, which is often referred to as the mixed model combining the main properties of PT and RDEU. As such, PT and RDEU generally serve as a special case of CPT. This allows different weighting functions to measure gains and losses respectively, and it is this which differentiates CPT from RDEU. The utility function of CPT should contain the following component:

$$u^n = \cdots + \beta_{loss} \sum_{l=1}^{j-1} (t_l^n - t_{ref}^n)^{\gamma}\, w(p_l^n) + \beta_{gain} \sum_{g=j-1}^{K} (t_{ref}^n - t_g^n)^{\theta}\, w(p_g^n) + \cdots \tag{4.7}$$

In this equation, outcomes are ranked from the worst to the best, so that $t_{loss}^n = \{t_l^n; 1 \le l \le j-1\}$ and $t_{gain}^n = \left\{t_g^n; j \le g \le K\right\}$. Again, reference dependence is represented by different taste parameters for gain and loss, while $\gamma$ and $\theta$ give different levels of diminishing sensitivity to loss and gain respectively. It should also be noted that the weighting function $\pi(.)$ is replaced by $w(.)$, in that the decision weight $w(.)$ used in CPT corresponds to the difference of cumulative probabilities distorted by weighting function $\pi(.)$ (refer to subsection 2.5.4 for the details). We differentiate $w(p_g^n)$ from $w(p_l^n)$ by using different weighting functions, since this allows for travellers' asymmetric tastes regarding decision weights for gain and loss.

If travellers actually react to risky outcomes depending on their decision weight rather than objective probability, it is the weighting function that should be applied in transport studies in order to reflect this perceived travel time distribution. It should be noted that an inherent connotation of nonlinear decision weight is that travellers' characteristics (optimism and pessimism) have a notable influence on their perception of risky prospects. More importantly, as shown in the following subsection, these characteristics are closely related to attitude towards risk.

#### *4.2.3.2 Extended discussion on risk attitude and weighting functions*

Research on travellers' risk attitudes has primarily been carried out from a nonlinear utility point of view.[13] Here, the alternative method to understanding attitudes towards risk from a decision weight perspective is now investigated. In subsection 2.5.2 it was concluded that people's characteristics are reflected by the shape of the weighting function curve. Specifically, pessimism is reflected by a convex weighting function, and optimism is reflected by a concave curve, with the counterpart to optimism and pessimism being believed to be risk proneness and risk aversion.

This is consistent with our intuition. For a convex weighting function, provided the probability of a good outcome is $p$ and the probability of a bad outcome is *1-p* respectively, the decision weight of a good outcome is $\pi(p)$, which is less than $p$ under a convex weighting function. Accordingly, the decision weight for the bad outcome is $\pi(1-p) = 1 - \pi(p) > 1 - p$. Evidently, therefore, the good outcome is under-weighted and the bad outcome is over-weighted, if the weighting function is convex. This distortion probability is directly related to pessimism. For a concave weighting function in RDEU for instance, provided the probability of a good outcome is $p$ and the probability of a bad outcome is *1-p* respectively, the decision weight of a good outcome is $\pi(p)$, which is more than $p$ under a concave weighting function. Accordingly, the decision weight for the bad outcome is $\pi(1-p) = 1 - \pi(p) < 1 - p$. Evidently, the good outcome here is over-weighted and the bad outcome is underweighted if the weighting function is concave. This suggests that an

[13] This is the main method used by EUT, for further detail refer to Von Neumann and Morgenstern (1947) and Marschak (1950).

optimistic traveller appears to pay more attention to better outcomes, and correspondingly overweight the likelihood of good outcomes. In this case, we can treat it as risk proneness.

It should be noted that, in some cases, the weighting function may display a mixed shape, such as an S-shape or an inverse S-shape. In these scenarios we cannot simply determine risk attitude from concavity or convexity. Instead, it is dependent on a specific distribution of outcomes. For instance, an inverse S-shape weighting function tends to over-weight small probabilities and under-weight large probability. In this case, if the probability of a good outcome is small enough to be inflated by a weighting function, it can be said that travellers appear to express risk proneness behaviour.

### 4.2.4 Modelling rank dependence

Thus far the method to analyse travellers' attitudes towards risk has been discussed from the point of view of the weighting function. It should be noted that this method only applies to RDEU and CPT with pre-processed rank ordering of outcomes (refer to subsection 2.5.2 for the details). As a result, how to determine the ranking of outcomes turns out to be an unavoidable problem for RDEU and CPT. This section, therefore, sets out a method for ranking the order of outcomes.

It is assumed that decision makers have conducted some pre-processing for the specific choice context that would account for the rank ordering of all possible outcomes. In a laboratory experiment, it is easy to rank monetary outcomes but deriving parallel behaviour data in a travel choice context is much more difficult since it contains multiple attributes and decision heuristics. Three plausible approaches for achieving this are now discussed.

The first approach is the single attribute-specific method, which concentrates on only one typical attribute. This usually applies to the context with only one influential attribute in terms of risky prospects, e.g. travel time, schedule delay, etc. For instance, if travel time is the only attribute associated with prospects, delayed travel time is naturally interpreted as a worse outcome with a lower rank. Therefore, we only need to present a series of possible travel times to respondents, and they would correspondingly form the ranking orders of travel time outcomes.

This attribute-specific method is especially useful for revealing outcome ranking orders from an RP dataset. Travel time outcomes can be elicited from a service dataset, and it can be assumed that ranking orders only depend on the magnitude of travel time. For instance, if observed travel time varies from 10 minutes to 20 minutes, we can assume that travellers would treat each time point between 10 to 20 minutes as a risky outcome, with the 10 minutes outcome being preferred to the 20 minutes outcome in that travellers' taste of travel time is negative (i.e. the more the worse). This assumption is in line with intuition, and it enables outcomes to be ranked without additional information from interview based surveys.

It should be noted that this attribute-specific method is unable to address a situation with multiple attributes of alternatives. In this case, however, an appropriate experimental design for an SP survey is capable of presenting ranking orders to respondents. This second method, therefore, is based on lexicographic techniques and framing effect. For instance, in a SP survey exercise, we can present travel time related attributes as 'arriving early', 'arriving late', and 'arriving on time', which infers that there are three potential outcomes for each trip. It is then assumed that respondents would naturally rank these outcomes from the worst to the best, i.e., late arrival<early arrival<on-time arrival.

The third potential method is based on the calibration and an iterative algorithm. It assumes that the ordering of outcomes is not known *a priori*, since it actually depends on the outcome utility and estimated parameters. It should be noted that the outcome mentioned in this method, is not an attribute-specific outcome (such as a travel time outcome), but rather a multi-attribute outcome. The premise of this multi-attribute outcome is that it is not only travel time that affects ranks, but also other factors, e.g. schedule delay. This iterative algorithm is expressed as follows:

- Fix taste parameters of EUT
- Rank outcomes according to the induced utilities
- RDEU model calibration
- Judgement. Iteration is required if the change of model fit is beyond a specific threshold, otherwise, RDEU model estimation finishes.

Given that the ranks and estimation have mutual impacts on each other, this kind of iteration algorithm is extremely useful in a situation where each risky outcome involves a trade-off among a number of attributes.

Since this study focuses on risky choice in an RP context, and since the second method can only be applied to an SP context it cannot be adopted here. The attribute-based method, meanwhile, is considered to be a simplification of iteration algorithm and represents an efficient way to determine the ranks as long as we ensure we are aware of travellers' preferences towards the single attribute, such as travel time.[14] As a result, in this research, the third method will be the primary method used, with the first method being employed in special conditions.

## 4.3 Data

This subsection will describe the data collection strategy, focusing on the data from travellers' actual behaviour, i.e. revealed preference data. This is an approach to researching travellers' choice behaviour by observing what they actually do. Given RP's high reliability and validity, it should be regarded as the natural data source for travel choice research. Indeed, as noted by Brownstone and Small (2005), RP results reflect what transportation planners need to know for transportation project evaluation, since they describes the real market data.

As has been repeatedly emphasized, however, risky choice behaviour contains an inherent connotation of repeatability, which potentially requires repeated observations of traveller's choices in order to reveal all the possible outcomes faced by travellers. It is extremely difficult to obtain this information from an RP exercise. In practice, therefore, RP studies are only rarely used for modelling travellers' risky choice behaviour. That said, while it might be difficult to conduct risky choice modelling using RP data, it is not impossible. Based on an thorough literature review, a methodology for RP data collection and usage is proposed here, serving as one of the significant contributions of this thesis.

### 4.3.1 Revealed versus stated preference

In an SP exercise, respondents are often asked to provide feedback on what they would choose under some hypothetical or virtual circumstances. This method is flexible as researchers are capable of address any requirements by designing questionnaires appropriately. Whilst SP is an efficient way to obtain substantial observations, there are issues regarding its validity and reliability. An appropriate questionnaire design is essential to ensure the internal validity of SP data, and external validity may be a problem if respondents behave differently in real life from their answers in the survey. This is true especially when

[14] For instance, if we constantly find that the taste parameter of travel time is negative, it is reasonable to believe the outcome with less travel time is ranked higher.

respondents deliberately give misleading answers, or omit important constraints that exist in the reality, e.g. time pressure for making decision, impatience in real traffic congestions.

In contrast, RP data, since it is obtained from individuals' actual behaviour, is less flexible but more reliable than SP data. Hence, this type of data should naturally be applied to risky choice model. Surprisingly, however, there are almost no convincing studies using RP data within a risky choice framework. The key problem associated with using RP data is that analyst has only very limited control over the covariance structure of the data and is therefore vulnerable to problems of collinearity between key attributes and limited overall variation in key dependent and explanatory variables. This limitation can in turn lead to difficulties in model estimation. By contrast, with SP data, the analyst in principle has completely control over the statistical properties of the choice contexts presented and can optimise these for efficiency in statistical estimation. Table 4.2 shows a comprehensive comparison of the potential challenges of applying RP within a risky choice framework.

| RP | SP |
|---|---|
| Hampered by the collinearity between travel time, travel time variability, and travel cost. | Controls relationships between attributes in order to avoid collinearity and maximise efficiency. |
| Can only account for the existing market. | New alternative can be presented in hypothetical choices, e.g. new tolled road. |
| Cannot address an attribute that does not exist in the market. | New attribute can be included, e.g. improved reliability of an existing road, new technology. |
| Time and resource intensive. | Relatively cheap and flexible. |
| Difficult to find situations with sufficient perceived variation in travel time. | Controls relationships between attributes allowing sufficient variations. |
| Requires additional assumptions on the possible outcomes faced by respondents.[15] | Risky outcome scenarios can be explicitly presented to respondents during interviews. |
| Has high reliability and validity since it is based on respondents' actual choices. This is also very important for the evaluation of transportation projects. | Behaviour reported in the context of hypothetical choices may not be replicated in real choices. |
| Travel time information is observed from respondents' actual travelling experiences. | The perception of travel time outcomes is affected by the design of the survey. |

**Table 4.2: Features of RP and SP within a risky choice framework**

[15] SP questionnaires directly present risky outcomes to respondents, while an RP dataset does not naturally show outcomes perceived by respondents. Hence, additional assumptions and techniques are required to reveal possible outcomes.

While there are several unavoidable drawbacks of RP, given an understanding of these drawbacks the application of advanced techniques in data collection methodology can help to mitigate them. The following subsection, therefore, moves on to introduce the methodology for RP data collection and usage which will be employed in this research.

## 4.3.2 RP data methodology and availability

In terms of data requirement, the RP dataset used in this thesis should consist of survey data and level-of-service data. The former reveals travellers' chosen alternatives, e.g. the chosen route, mode and departure time. Moreover, travellers' socio-economic and demographic data is also included in the survey data, e.g. age, gender, income level, flexibility, owned car type (important to estimate the toll level and cost in some cases), work starting time, etc. The level-of-service dataset records actual travel time information between origin and destination. This data can be applied to specify attributes that are not revealed in the survey data, such as travel time attributes, *inter alia*. In this way it is possible to obtain the true network performances of alternatives, and more importantly, the travel time extracted from the level-of-service dataset can be regarded as the attribute of travel time perceived by travellers under some assumptions. The premise of this technique is that it is assumed that travellers are experienced enough to be aware of the true travel time distribution (although non-EUT allows misperception of distribution, this distortion is still based on true travel time distribution).

Based on the above methodology, it is possible to collect the required data. Initially, however, existing studies were consulted since these represent the most efficient way to obtain qualified data. Additionally, even though collecting new RP data is time consuming for a PhD student, the new method proposed here for RP data collection for modelling risky choice behaviour is introduced and the availability of this data is discussed. Accordingly, , subsection 4.3.2.1 discusses the availability of existing datasets and briefly describes the selected dataset, i.e., SR91 data; while subsection 4.3.2.2 proposes several possible methods for new RP data collection, and outlines the new London Underground dataset used in this thesis.

### *4.3.2.1 Existing datasets*

**Travel diary data:** Senbil and Kitamura (2003) investigated the applicability of PT in the context of RP data. Their original data was collected in Japan in 2002 and they randomly mailed 1000 resident drivers who were asked to record departure and arrival time for three days, and answer whether they would change their departure time. A distinguishing feature of this RP exercise was that respondents were asked to supply their preferred arrival time (PAT), and thus PAT is regarded as the reference point in their PT model. It should be noted that this kind of mail survey corresponds to a diary survey, in that both methods are capable of recording respondents' daily travelling experiences. The travel time reported by respondents is the same as the travel time actually experienced by them. In the UK, a similar data set is the National Travel Survey (NTS) database. This survey consists of two parts – interview with respondents in their houses, and a seven day travel diary. Along with other traffic information, it is possible to apply NTS data to characterize individuals' travel behaviour. The lack of detailed origin and destination (OD) information is, however, the main drawback of the NTS database.

**GPS data:** With the development of technology, advanced equipment has been increasingly applied to RP surveys. This enables researchers to obtain detailed travelling information by automatically tracking respondents' actual travel choices. A recent study by Carrion and Levinson (2010) carried out an RP study to investigate drivers' willingness to pay for the improvement of reliability offered by a High Occupancy Toll road. They equipped Global Positioning System (GPS) devices to each respondent and traced their choice behaviour. It has been found that such devices are relatively mature and feasible, and the data from these loggers are more reliable in reporting accurate location and travel time than travel diary surveys. Whilst a GPS logger is bigger and heavier than the other alternatives such as a mobile phone, they can be installed into vehicles so that respondents would not consider its portability. Furthermore, the passive nature of the data collection reduces the load on respondents. Hence, vehicle-based GPS loggers are highly recommended for future route choice surveys.

**Floating car data:** Small et al. (2005b)'s study on the SR91 corridor in the US pays particular attention to travellers' route behaviour using both SP and RP data. In their RP sample, raw observations (438) for travel time were derived from field measurements on SR91 by students' repeated driving, i.e. floating car data. Each observation was of the route choice made by travellers between two routes – a tolled and an untolled route. The tolled route was assumed to be uncongested and to have a fixed and known travel time, whereas the untolled route was congested with an uncertain travel time. The magnitude of the variability in travel time depended on the demand, and hence varied according to the time of day. Travellers were assumed to have a fixed time of travel. Therefore any traveller in the RP sample was assumed to be choosing between a certain prospect (the tolled route) and an uncertain prospect (the untolled route), where the uncertainty in question is in travel time (see also Chapter 5).

In closing, Small et al. (2005b)'s SR91 dataset is used in this thesis due to three reasons. Firstly, a GPS survey exercise is extremely resource intensive and thus the corresponding sample size is usually quite small. Secondly, this research requires extra information from network data which is not available in the existing travel diary dataset. Finally, SR91 dataset offers a natural experiment for risky choice research with two parallel competitive routes, it is accessible, and its survey data and network data is suitable for our risky choice framework.

#### *4.3.2.2 Methodology for new RP data collection*

The feasibility of conducting a new RP exercise was also investigated. Whilst a risky choice framework appears to complicate RP data collection, as discussed above, this kind of data collection would still be feasible as long as tailored techniques and assumptions are adopted. The used in this thesis methodology of simultaneously collecting survey data and level-of-service data is illustrated in Figure 4.1. The former covers respondents' actual choices and their socio-economic information, while the latter is used to collect the data associated with alternatives, in particular travel time related data. Given that risky choice research has the feature of repeatability, the following discussion pays special attention to possible methods for addressing this issue.

Travellers' choices can be observed by either traditional roadside interviews or other advanced methods. The advantage of interviews is that the researcher can ask respondents for detailed information including basic demographic data and additional information. For instance, ideal journey time and preferred arrival time have been found to be possible reference points in travellers' reference dependent choice behaviour, and this information can be obtained from interview. Given the relatively large sample size needed in this research, it is difficult to employ this method in research which is of limited duration. Alternatively, new technological and commercial data, e.g. GPS, Automatic Number Plate Recognition (ANPR), cell phone data with customers' information can be used. These methods enable drivers' route choice behaviour to be conveniently observed, and greatly enlarge the sample size, however, it is extremely difficult to access drivers' socio-economic information.[16]

[16] In our feasibility studies, we contacted several councils and commercial companies which manage advanced data collection activities. We found that authorization is required to access drivers' information recorded in

In the studies presented in this thesis, survey data must be complemented by level-of-service data in order to estimate the attribute of travel time and this is another challenge for data collection, since it requires massive performance data and appropriate assumptions. First of all, we have to collect performance data and measure travel time. For road transport, travel time can be estimated using the data collected by floating cars, ANPR, GPS or loop detector. With the development of survey techniques, these methods are already capable of providing relatively reliable estimation of travel time. For rail transport, each train's performance is monitored by signalling systems, and therefore, the recorded travel time extracted from such performance data is much more accurate. Consequently, special attentions were paid to rail system, especially to London Underground.

In this thesis, the new RP data is based on the London Underground (LU) dataset. Survey data is from the Rolling Origin Destination Survey (RODS) dataset which records annual passenger survey results from a sample of underground stations. Train performance data is saved in the Network Management Information System (NetMIS) system through which we can retrieve historic data for each train (for details refer to Chapter 7). LU data is of special interest to us, in that it is the most efficient way to obtain qualified data for modelling risky choice behaviour. In particular, it avoids expensive and laborious field survey exercises, and the train running time system is much more accurate than the travel time estimated from network data.[17] It should be noted that this proposed method still faces an unavoidable drawback of collinearity among key variables, e.g. travel time, travel time variability and travel cost. This problem can be overcome in two ways, however: firstly, the travel time distribution can be estimated across days for a given narrow time-of-day interval and a given day-of-week. Secondly, the interaction between the travel time variable and the socio-economic variables can be specified, which provides additional variations.

It is also necessary to acknowledge that train running time and frequency information alone does not ncecessrily provide a complete representation of passengers' travel time experience, since it does not include representation of time spent in pedestrian access and circulation to and within stations nor account for the possibility that extreme train crowding

---

ANPR datasets, and companies managing cell phone and GPS data are not willing to share their customers' personal information.

[17] A series of comprehensive feasibility studies were conducted on travel time estimation, and research cases covering the M6 tolled road, the Maidstone area, and the Itchen bridge in the UK. It was found that it is extremely time consuming and expensive to install equipment and collect and analyse the data. Additionally, most of the loop detectors embedded in our target area are single loop, which provide only very limited information to estimate travel time. The associated possibility of inaccurate estimates of travel time was another important reason for not collecting data by ourselves. That said, we believe that the method proposed in this thesis will be useful for future study.

may extend platform waiting times. However, these limitataions are believed to be acceptable within the oveall context of the study.

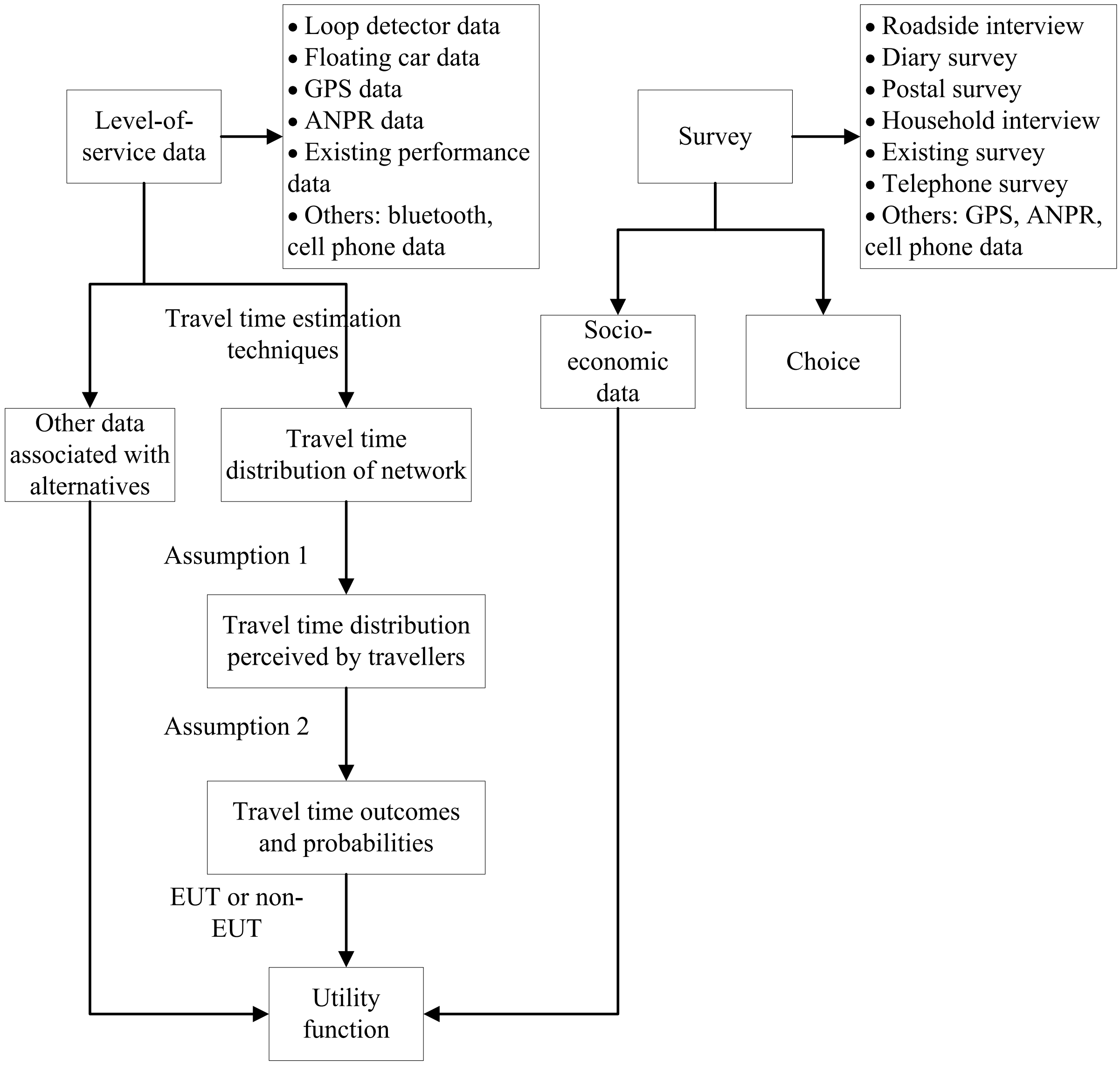



**Figure 4.1: Methodology for RP data collection in a risky choice context**

## 4.4 Summary

This chapter proposed the model forms and data collection strategy that will be used in the ensuing analysis of travellers' risky choice behaviour. Two principle methodological contributions have been proposed. The first is the generic risky choice framework incorporating various non-EUT approaches: SEU highlights the importance of nonlinear weighting function; rank-dependence is embedded in RDEU model on the basis of SEU; PT is the typical theory which is capable of modelling reference dependent behaviour; CPT synthesizes both reference-dependence and rank-dependence. Finally, all of these proposed non-EUT models can be evaluated using an RUM approach.

The second contribution is the RP data collection method tailored specifically for our risky choice framework. Given that SP studies have predominated in the existing literature, the exploratory work conducted here enables a better understanding of how non-EUT models perform in an RP context, which improves the validity and reliability of this research.

Based on comprehensive feasibility studies, it was decided to use two datasets: the first being an existing dataset, which was originally collected on the SR91 corridor in the US; the second was a dataset collected from the London Underground (LU) database. In the next chapter, we present the model result based on the SR91 data, while an extensive of their implementations is presented in Chapter 6. Chapter 7 presents a second case study using the LU data collected by us.

# Chapter 5 TOLL ROAD CASE STUDY

## 5.1 Introduction

Literature review in Chapter 2, 3 and methodology in Chapter 4 are referred to as the theoretical part in this thesis. This current chapter and the remainder of this thesis construct the applied part of this thesis. As illustrated by the discussion in the theoretical part, recent work on risky choice modelling has sought to address the shortcomings of expected utility theory (EUT) by using non-expected utility (non-EUT) approaches. However, to date these approaches have been merely tested on stated choice data. Moreover, travel demand area has witnessed the gap between the state-of-art, i.e. theoretical development of alternative approaches, and the state-of-practice, i.e. applications of non-EUT models to the real-world. The application presented in this chapter aims to fill in the gap at least in the area of route choice by empirically investigating the feasibility of non-EUT methods in a reveal preference context.

In the current case-study, we look at the binary route choice between free road with unreliable travel time and toll road with highly reliable service but monetary toll as the cost. We consider it as an ideal scenario for modelling risky choice in that drivers are assumed to treat the unpredictable travel time of free road as a risk, and some of them are willing to pay extra toll to alleviate such risk. To test the validity of this hypothesis, we established a series of candidate models, namely EUT, WUT, SEU, REDU and PT. And we subsequently present systematic comparisons between these models.

The remainder of this chapter is organized as follows. Section 5.2 provides a brief overview of toll road studies. Section 5.3 describes the RP dataset used in this empirical work. This is followed by the model specifications and the corresponding estimation results in Section 5.4 and 5.5. Section 5.6 aims at the systematic comparison of candidate models. Finally, this chapter close with a brief summary in Section 5.7.

## 5.2 Background

Toll roads offer a natural experiment for risky choice research due to their special characteristics which involves a trade-off between a free but very congested road versus a tolled (priced) but free-flow road. In the UK, there exist several toll facilities, such as the Dartford Crossing, M6 toll road, Severn bridges and Itchen bridge. According to our feasibility studies, some of them may potentially serve as suitable cases for modelling travellers' risky route choice behaviour. Whereas, to the author's knowledge, there exist no studies on driver's risky choice conducted on these sites, and it is almost impossible to obtain commercial data[18]. Given the availability of data collection, we cannot consider these local road-pricing cases at the current stage, but have to resort to existing dataset from other countries. Recently, a large body of toll road studies have been undertaken in Australia in 2008 (Hensher et al., 2011, Li and Hensher, 2012a). These empirical researches are based on a stated preference (SP) survey which presents three route choices across eight hypothetical scenarios. Toll levels vary across different route choice, which provides sufficient variation of toll. Moreover, each route choice is associated with three travel time outcomes, namely '*x* minutes earlier', 'arrival on time', '*x* minutes later'. Thus, such data is qualified for our risky choice framework. Such flexibility of SP exercise enables researchers to obtain substantive travel details, which RP survey can hardly achieve. Indeed, it is difficult to find any convincing RP studies which provide sufficient perceived variation in attributes. Whereas, a couple of exceptions have been found from toll road studies in California and San Diego in the US.

The first exploration is from the studies on the Interstate 15 (I15) facility which is a publicly funded toll road[19] (Brownstone et al., 2003, Ghosh, 2001). This project allows carpool drivers to use the express lane by paying tolls in order to avoid the eight mile congested segment. Loop detectors embedded in roads collect the data of time-of-day speed, and then travel time is estimated using this time-of-day speed. One distinguished feature of I15 is that the toll level is relating to the congestion level of the untolled road. In this case, drivers are capable of guessing the travel time according to the updated toll information showing at the entrance. For instance, if the toll at the specific time of day is extremely high, it suggests that the untolled road ahead is severely congested. It appears to be an attractive feature for drivers as the uncertainty of congestion is significantly reduced. However, it also leads to correlation between travel time attribute and toll. Furthermore, it still requires additional research on whether the prior information on toll level has influences on reducing travellers' perceived travel time risk. If the answer is yes, such RP data may be not suitable

[18] The Department for Transport (DfT) commissioned an SP-based analysis of travel time variability.
[19] The San Diego I-15 Congestion Pricing Project serves as one of the demonstration projects in the US to convert existing High Occupancy Vehicle (HOV) lane to High Occupancy Toll (HOT) lane.

for modelling risky choice behaviour since traveller's decision is under relatively certain situation.

The State Route 91 (SR91) toll road in Orange County of California is another notable road pricing project in the US. It has been found that some travellers on SR91 corridor prefer paying extra money to use toll road service in order to ensure little congestion, while the others who are not willing to pay such money can still use free infrastructure which is parallel with the toll facility. Given the fact that the quality of SR91 dataset is higher than I15 dataset as the former obtained better performance data (see Ghosh, 2001), we pay special attentions on the former. SR91 connects the residential area in riverside and San Bernadino Countries to job centres in Los Angeles and Orange Counties. It used to be the most congested corridors of California. The opening of new toll lanes in 1995 provides a new alternative operated by a private company[20]. It enables driver to pay tolls electronically via a transponder in order to use the *SR91 Express Lane*. Unlike dynamic pricing used in I-15 facility, time-varying toll structure is applied on SR91 to ensure profit maximization of the private operator. Thus drivers cannot predict the congestion ahead according to the scheduled toll levels.

Given the fact that RP study is extremely scarce, this project is capable of producing qualified data for several reasons: firstly, there exist only two route alternatives, i.e., it is binary choice problem. Given such limited choices, it is reasonable to assume that travellers are aware of travel time distributions on both routes. Secondly, as noted by Brownstone and Small (2005), the choice of choosing toll road is relatively independent since there exist little transit services along this corridor. Thus we can simply assume that travellers only face two options, namely free road and toll road. Thirdly, the variation of travel time and toll has been found to be relatively independent. Based on these special factors, a series of empirical studies have been undertaken on SR91 (Brownstone and Small, 2005, Lam and Small, 2001, Recker et al., 2005, Small et al., 2005b).

Among these studies is Lam and Small (2001) who initially collected RP data in 1997 to modelling traveller's various travel choice behaviour. In this research, travel time distribution is estimated from loop detector data, and survey data is based on two waves of telephone survey. They measured VTTS and value of reliability (VOR) by observing travellers' actual travel choice in the real road pricing context. Travel time variability is defined as the difference between 90$^{th}$ percentile and the median travel time. As such, their

[20] The project was originally based on a franchise agreement with the State of California before 2003. However, this franchise was then purchased by the Orange County Transportation Authority which serves as a public agency.

route choice model is based on the typical mean-variance specification. Whereas, noticed that loop data was collected one year before the survey, one drawback of their RP, as admitted by Lam and Small, is that the travel time extracted from such network data may be not accurate[21]. Given the relatively poor quality of network data, we have to resort to the other SR91 studies.

In 1999, Small et al undertook a combined RP-SP study of drivers travelling on SR91 corridor. In their research, SP data is applicable since local respondents are quite familiar with the SR91 project, and thus, they are able to understand the SP choice context. Given this thesis aims at modelling travellers' risky choice behaviour in RP context, only the RP data is applied to this current research, and the following section will provide detailed descriptions of this dataset.

## 5.3 Description of data

The data used in this case study is from the RP data collected on the SR91 corridor in the US, a survey that was originally undertaken to investigate the value of travel time and reliability (Small et al., 2005b, Small et al., 2005a). The original SR91 dataset consists of RP data and SP data collected in 1999-2000. Given the purpose of this current research, only the cross sectional RP data is employed, and a total of 438 observations are available. Respondents in this survey were asked about their most recent trips on the SR91 corridor during morning peak. The questions cover selected route (actual route choice), age, income level, trip distances, the flexibility of work arrival time, and various issues about personal characteristics et al. Each observation is of the route choice made by travellers as between two routes – a tolled and an untolled route. The toll road is assumed to be uncongested and have a fixed and known travel time whereas the untolled route is congested with uncertain travel time. The magnitude of the variability in travel time depends on the demand, and hence varies by time of day. Travellers are assumed to have a fixed time of travel. Therefore any traveller in the RP sample is assumed to be choosing between a certain prospect (the tolled route) and an uncertain prospect (the untolled route), where the uncertainty in question is in travel time.

In terms of network data, we have available floating car data that enables us to identify the time of day dependent distributions of travel time on the untolled route. It has

[21] Lam and Small (2001) realized the potential bias caused by the time difference. They measured the trend of growing congestion by analysing the loop detector data collected in the surveying year. It enables them to apply a growth factor to the travel time dataset.

been found that the traffic on the tolled road is observed to move freely at all time of day at the level of around 8 minutes. Therefore, the travel time on the tolled route is assumed to be constant at 8min. It should be noted that this dataset is also hampered by the low quality of network data[22]. Given the above basic information, two problems emerge when historic data is applied to reproducing travel time distributions: how to identify and define prospects and corresponding risky outcomes, and how to numerically draw the value of attributes and associated probabilities from the estimated distribution.

It is reasonable to assume that the travel time extracted from network data is the time experienced by travellers if travellers are experienced commuters. Consequently, we treat travel time from 8min to 20min[23] as discrete contingent outcome, and the range between two consecutive time outcomes is 1min.[24] Consequently, there are maximum 13 outcomes. Associated probabilities are valued according to the empirical frequency extracted from floating network dataset, for each time of day category. We tested two parametric distributions, viz., the normal distribution and the lognormal distribution, however the K-S test results show that these distributions do not fit the floating data well (partly due to the small sample sizes). We concluded that the limited observation from floating car data is not convincingly sufficient to obtain corresponding continuous travel time distribution. Consequently, the empirically observed discrete travel time distributions are used in this study.

## 5.4 Model specifications

### 5.4.1 Basic model

Issues in model specifications deserve further attentions. Given that this current study accounts for binary route choice, it is naturally to assume that travel time distribution has

[22] The floating car data was collected from 4:00 to 10:00 across 11 days, with only 210 observations. It was separately collected by California Department of Transportation and California Polytechnic State University at San Luis Obispo.

[23] This is the time taken to travel a 10 mile portion of the study corridor, and the observed range of travel time is from 8 minutes to 20 minutes.

[24] A natural question is how to determine the increment of consecutive travel time outcomes. Indeed, it is not necessary to assume 1 minute as the increment, and alternatively travellers may consider more aggregated travel time outcome set or even fuzzy outcome set, for instance, they just simply account for three outcomes, namely the best outcome, normal outcome, and the worst outcome. Whereas, it is impossible to identify the actual outcome set taken into account by travellers, and it is arbitrary to give a random number for the outcome increment. Moreover, as shown in section 5.5.3, the RDEV model with 2 minutes increment gives worse model fit than the model with 1 minute increment. Consequently, 1 minute is selected as the natural increment for travel time outcomes.

significant influence on travellers' route choices. Indeed, most relevant studies usually employ mean or median travel time as the variable reflecting the centrality of travel time distribution[25]. Noticed that expected value of travel time mathematically equals mean travel time, such linear specification is in line with Expected Value Theory (EVT) function. Thus, we treat EVT function as the basic model, and the utility function of choosing the toll road is expressed as:

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} p_k^n \, TT_k^n + \beta_{Toll}^n Toll^n + \beta_W^n H_i + \varepsilon^n \qquad (5.1)$$

where $TT_k^n$ is the travel time for the $k^{th}$ outcome of the free road alternative (the travel time for toll road is assumed to be constant at 8 minutes), and $p_k^n$ is the associated probability (e.g. a 10 minute travel time outcome with the associated probability). $H_i$ characterizes the $i^{th}$ agent's specific socio-demographic attributes such as income and gender. $\beta_{TT}^n$ , $\beta_{Toll}^n$ and $\beta_W^n$ are the parameters to be estimated. We also tested the inclusion of standard deviation (of travel time) into this model, though the estimated parameter for standard deviation turned out to be not statistically significant (for details refer to Table 5.1).

Existing literature on EUT exhibits a wide range of utility formulations incorporating decision makers' attitudes toward risk (see subsection 3.3.1 for details). It is arbitrary to select one of them unless we empirically test the actual performances of these utility specifications[26]. In this chapter, we employ constant relative risk aversion specification to nonlinearly transform utility function. The EUT model specification is expressed as follows:

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} p_k^n \frac{TT_k^{n(1-\alpha)}}{(1-\alpha)} + \beta_{Toll}^n Toll^n + \beta_W^n H_i + \varepsilon^n \qquad (5.2)$$

where the extra parameter $\alpha$ characterizes agent's attitude towards risk. As the traditional method for modelling risky choice behaviour, this EUT specification is capable of explaining whether individuals tend to nonlinearly distort the utility of travel time. In this thesis, EUT also serves as a basic model for the purpose of comparison. The following subsections concentrate on our alternative model specifications.

[25] We initially select the centrality-dispersion specification as the basic model structure (it is consistent with Small et al. (2005)). However, it is found that the parameter for standard deviation is not statistically significant at all.

[26] A preliminary analysis has been conducted to determine which utility functions better fit the dataset. Consequently, we found out that either the maximum number of iterations is reached in the exponential specification (e.g. CARA model), or the estimated parameter of travel time is not statistically significant in the power specification (e.g. Box-Cox and CRRA model). This undesirable result may be due to the relatively low quality of travel time data.

### 5.4.2 Weighted Utility Theory (WUT) Model

In WUT specification, risky outcome is not only weighted by associated probability, but also by the outcome value per se. Thus, we define the following functional form:

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} \frac{W(TT_k^n)}{\sum_{j=1}^{K} W(TT_k^n) p_j^n} TT_k^n p_k^n + \beta_{Toll}^n Toll^n + \beta_W^n H_i + \varepsilon^n \qquad (5.3)$$

where $W(TT_k^n)$ is a function assigning an additional weight to risky outcomes, and $TT_k^n$ is the value of the $k^{th}$ travel time outcome. Determining the function of the weight factor in WUT remains an open problem without general solutions. Sugden (2004) interpreted the real-value weight factor function as $W(TT_k^n) = (-TT_k^n)^\alpha$. We treat the above WUT function as a simple extension of EVT, i.e., each outcome is weighted by $\frac{W(TT_k^n)}{\sum_{j=1}^{K} W(TT_k^n) p_j^n}$. And EVT is a special case of WUT if $TT_k^n$ is identical or $\alpha = 0$. The basic assumption of WUT is that travellers tend to overweight the outcome with more travel time. Another promising function is the Box-Cox transformation of travel time shown as follows:

$$W(TT_k^n) = \begin{cases} \frac{(TT_k^n)^\alpha - 1}{\alpha} & if\ \alpha \neq 0 \\ \ln(TT_k^n) & if\ \alpha = 0 \end{cases} \qquad (5.4)$$

This is a more general class of power functions which has been widely used in transportation (Mandel, 1998; Lapparent, 2010). In this research, both Sugden's function and Box-Cox function will be tested.

### 5.4.3 Subjective Expected Utility Theory (SEU) Model

In this thesis, we adopt the following SEU model to embody the nonlinearity of probability.

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} \pi(p_k^n) \frac{TT_k^{n(1-\alpha)}}{(1-\alpha)} + \beta_{Toll}^n Toll^n + \beta_W^n H_i + \varepsilon^n \qquad (5.5)$$

where $\pi(p_k^n)$ is the weighting function of probability $p_k^n$. In line with our EUT model, CRRA utility function is also employed to address travellers' risk attitudes. As discussed in subsection 4.2.3, there exist a large body of functional forms for $\pi(.)$. Thus it is worth testing the influence of selecting different weighting functions on the final model fit. It should be noted that subjective expected value theory (SEV) theory is a special case of SEU when $\alpha = 0$, and EUT is a special case of SEU when $\pi(p_k^n) = p_k^n$.

### 5.4.4 Rank-Dependent Expected Utility Theory (RDEU)

The RDEU model is expressed by the following equation:

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} w(p_k^n) \frac{TT_k^{n(1-\alpha)}}{(1-\alpha)} + \beta_{Toll}^n Toll^n + \beta_W^n H_i + \varepsilon^n \quad (5.6)$$

where $w(p_k^n)$ characterizes individuals' decision weights towards risky outcomes, given the outcomes are defined in increasing order, i.e. the outcome with the least travel time (also the best outcome in this context) is ranked as $k$ = 13. We deliberately use $w(p_k^n)$ in order to discriminate it from $\pi(p_k^n)$ of SEU. In RDEU, $w(p_k^n)$ is numerically determined by the difference between two cumulative subjective probabilities, i.e. $w(p_k^n) = \sum_{l=k}^{13} \pi(p_k^n) - \sum_{l=k+1}^{13} \pi(p_k^n)$. Again, a number of functional forms of weighting function have been examined since the 1990s (see Scott, 2006 for details), thus, this research also aims to identify whether the performance of RDEU is affected by choosing weighting functions.

### 5.4.5 Prospect Theory (PT)

The proposed PT model specification is expressed as follows:

$$u_i^n = ASC + \beta_{TT(gain)}^n \sum_{k=1}^{j-1} p_k^n \left(TT_{rp}^n - TT_k^n\right) + \beta_{TT(loss)}^n \sum_{k=j}^{m-1} p_k^n \left(TT_k^n - TT_{rp}^n\right) +$$
$$\beta_{TT(loss^-)}^n \sum_{k=m}^{K} p_k^n \left(TT_k^n - TT_{rp}^n\right) + \beta_{Toll}^n Toll^n + \beta_W^n H_i + \varepsilon^n \quad (5.7)$$

where $TT_{rp}^n$ is the reference travel time which, in this case, is assumed to be the same across individuals. And the travel time parameter is divided into gain $\beta_{TT(gain)}^n$ , loss $\beta_{TT(loss)}^n$ , and diminishing loss $\beta_{TT(loss^-)}^n$ parameters respectively according to the relative location with respect to the reference point $TT_{rp}^n$. In the space of loss, the sensitivity towards travel time loss is diminishing insofar as travel time outcome turns out to exceed a travel time $TT_m^n$ to be estimated.

The PT model raises the question of how to endogenously estimate the reference point $TT_{rp}^n$. Surprisingly this issue has not been empirically studied very much despite its critical role in PT. In order to solve this nontrivial problem the algorithm shown in Figure 5.1 is applied.

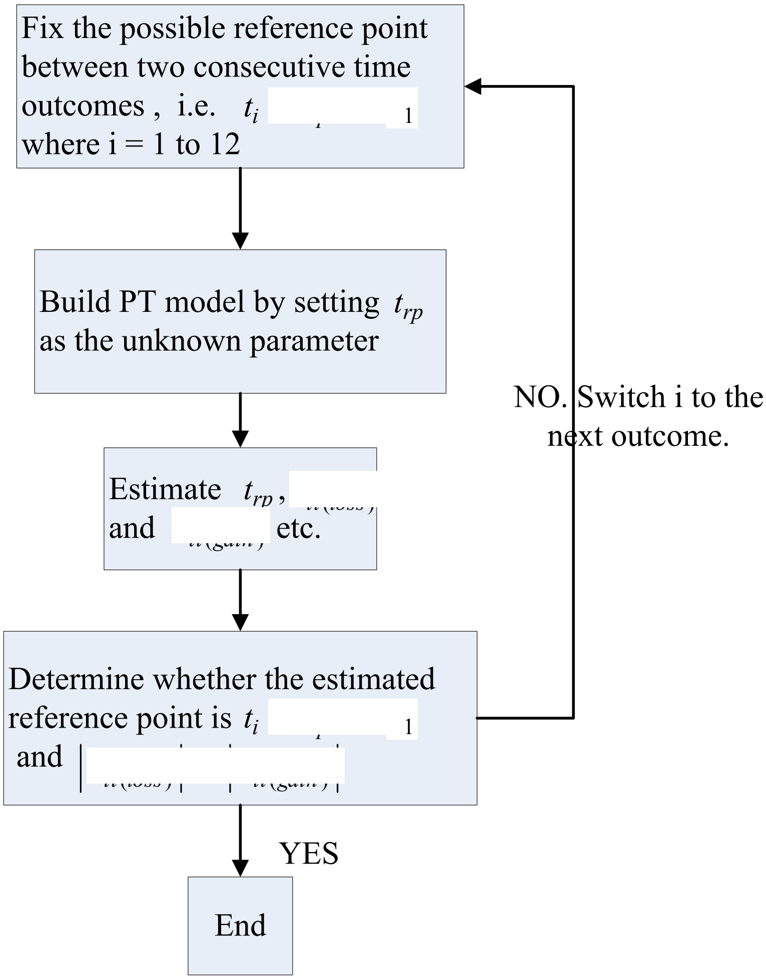



**Figure 5.1: Algorithm for the estimation of endogenous reference point**

To identify the value of $TT^n_{m,}$, the two-step profile likelihood approach is adopted. Step 1, build the PT models with fixed $TT^n_{rp}$ and candidate $TT^n_m$; Step 2, evaluate PT Model 1 with different$TT^n_m$, and select model which satisfies three criteria, namely better goodness of fit, statistically significant parameter estimates, and reasonable behavioural performances. Consequently, as shown in the next subsection, in this empirical context all the above criteria are satisfied only when $TT^n_{rp}$ =8.8 min and $TT^n_m$ =13 min.

## 5.5 Estimation results

The model estimation was carried out using BIOGEME 2.0 (Bierlaire, 2003). As the base models, EVT produced the log-likelihood of -230.414 and EUT delivered -230.151 as shown in Table 5.1.

| | EVT | | EVT_SD | | EUT | |
|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| ASC | -1.740 | -5.260 | -1.750 | -4.660 | -1.580 | -3.950 |
| $\beta_{toll\_low}$ | -0.450 | -4.170 | -0.451 | -4.190 | -0.459 | -4.240 |
| $\beta_{toll\_high}$ | -0.312 | -3.110 | -0.314 | -3.090 | -0.320 | -3.180 |
| $\beta_{TT}$ | -0.006 | -3.270 | -0.006 | -3.160 | -0.009 | -2.050 |
| SD | | | -0.002 | -0.010 | | |
| α | | | | | 0.102 | 0.920 |
| Age | 0.709 | 2.860 | 0.709 | 2.850 | 0.708 | 2.850 |
| Gender | 0.824 | 3.470 | 0.823 | 3.460 | 0.803 | 3.360 |
| Final LL(β) | -230.414 | | -230.395 | | -230.151 | |
| $\rho^2(0)$ | 0.241 | | 0.241 | | 0.242 | |
| Adj. $\rho^2(0)$ | 0.221 | | 0.218 | | 0.219 | |
| $\rho^2(ASC)$ | 0.086 | | 0.086 | | 0.087 | |
| Adj. $\rho^2(ASC)$ | 0.063 | | 0.058 | | 0.060 | |

**Table 5.1: Estimation results for EVT and EUT**

The EVT model results are clearly quite intuitive, with a negative constant on the toll road option, and statistically significant negative parameter on the travel cost and travel time. $\beta_{toll\_low}$ is the taste parameter of toll for the population with low income (≤$60000 per year), and $\beta_{toll\_high}$ is the taste parameter of toll for the high income population (>$60000 per year). The relationship $\beta_{toll\_low} < \beta_{toll\_high}$ suggests that low income drivers hold more negative attitudes towards toll. As a result, drivers' wiliness to pay toll in return for travel time savings varies across different income levels. Contrary to our expectation, EUT model fails to improve goodness of fit comparing to EVT. We also note that the CRRA utility parameter α turns out to be not statistically significant. These unappealing results may be partially due to the relatively low quality of network data used to generate the travel time distributions in this research.

### 5.5.1 WUT

The estimation results for the WUT model show that neither Sugden's power function nor Box-Cox transformation provide statistically significant estimators of the additional parameter $\alpha$. Furthermore, none of the candidate functions leads to increases in log-likelihood with respect to EUT. These estimation results are in contrast with the SP analysis of Michea & Polak (2006) in which $\alpha$ is statistically significantly different from zero and the model fit is improved with respect to EUT.

| | WUT 1<br>$W(TT_k^n) = (-TT_k^n)^\alpha$ | | WUT 2<br>$W(TT_k^n) = \frac{(TT_k^n)^\alpha - 1}{\alpha}$ | | WUT 3<br>$W(TT_k^n) = \ln(TT_k^n)$ | |
|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| ASC | -1.750 | -5.220 | -1.753 | -5.240 | -1.750 | -5.260 |
| $\beta_{toll\_low}$ | -0.450 | -4.170 | -0.440 | -4.070 | -0.439 | -4.190 |
| $\beta_{toll\_high}$ | -0.312 | -3.110 | -0.302 | -3.010 | -0.315 | -2.990 |
| $\beta_{TT}$ | -0.006 | -2.940 | -0.006 | -3.040 | -0.006 | -3.280 |
| α | 0.441 | 0.170 | 0.096 | 0.020 | | |
| Age | 0.709 | 2.860 | 0.709 | 2.860 | 0.709 | 2.860 |
| Gender | 0.823 | 3.470 | 0.823 | 3.470 | 0.823 | 3.470 |
| Final LL(β) | -230.401 | | -230.384 | | -230.380 | |
| $\rho^2(0)$ | 0.241 | | 0.241 | | 0.241 | |
| Adj. $\rho^2(0)$ | 0.218 | | 0.218 | | 0.221 | |
| $\rho^2(ASC)$ | 0.086 | | 0.086 | | 0.086 | |
| Adj. $\rho^2(ASC)$ | 0.059 | | 0.059 | | 0.063 | |

**Table 5.2: Estimation results for WUT**

Note that the model with $W(TT_k^n) = \ln(TT_k^n)$ provides a similar model fit as the EVT model without introducing any extra parameters. However, this specification mathematically addresses a couple of shortcomings of EUT by generating fanning out (refer to 2.5.5.4 for details). Specifically, the outcome with the travel time more than the average would result in a weight $\frac{\ln(TT_k^n)}{\sum_{j=1}^{K} \ln(TT_k^n)\, p_j^n} > 1$. This behaviour can be interpreted as pessimism or risk aversion as worse outcomes tend to be overweighted.

### 5.5.2 SEU

The results for SEU and SEV models, which employ various weighting functions, are shown in Table 5.3. We tested a range of weighting functions, however, none of these SEV models lead to significant improvement in LL over their EVT and EUT counterpart. Moreover, the results show that LL seems not to change across different SEV models, and they are almost constant at the range of -230.1 and -230.3. This is somewhat unexpected, given that weighting functions with binary or multiple weighting parameters are expected to be more flexible[27]. Furthermore, we face computational issues when we evaluate the models with complicated weighting functions due to their highly nonlinear specifications. Based on these observations we believe that, at least in this current context, the selection of weighting functions has little influence on model fit. Consequently, we decided to use Kahneman and Tversky's specification (SEV-TK), which provides the best LL, as the basic weighting function for the subsequent SEU model.

| | SEV-TK $\frac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{(1/\gamma)}}$ | | SEV-GE $\frac{\tau p^\gamma}{\tau p^\gamma+(1-p)^\gamma}$ | | SEV-WG $\frac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^\tau}$ | | SEU-TK | |
|---|---|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| ASC | -1.760 | -5.270 | -1.790 | -4.780 | -1.790 | -4.540 | -1.600 | -3.980 |
| $\beta_{toll\_low}$ | -0.448 | -4.150 | -0.447 | -4.130 | -0.446 | -4.120 | -0.460 | -4.250 |
| $\beta_{toll\_high}$ | -0.309 | -3.080 | -0.308 | -3.060 | -0.308 | -3.050 | -0.321 | -3.190 |
| $\beta_{TT}$ | -0.007 | -2.260 | -0.006 | -0.920 | -0.007 | -1.180 | -0.056 | -2.370 |
| α | | | | | | | 0.811 | 0.910 |
| γ | 1.150 | 9.970 | 1.070 | 4.920 | 1.040 | 7.910 | 1.110 | 18.060 |
| τ | | | 1.090 | 1.280 | 2.150 | 0.190 | | |
| Age | 0.711 | 2.860 | 0.711 | 2.860 | 0.711 | 2.860 | 0.710 | 2.860 |
| Gender | 0.830 | 3.490 | 0.834 | 3.490 | 0.834 | 3.490 | 0.809 | 3.380 |
| Final LL(β) | -230.111 | | -230.302 | | -230.301 | | -229.073 | |
| $\rho^2(0)$ | 0.242 | | 0.241 | | 0.241 | | 0.245 | |
| Adj.$\rho^2(0)$ | 0.219 | | 0.218 | | 0.218 | | 0.219 | |
| $\rho^2(ASC)$ | 0.087 | | 0.087 | | 0.087 | | 0.092 | |
| Adj.$\rho^2(ASC)$ | 0.060 | | 0.059 | | 0.059 | | 0.060 | |

**Table 5.3: Estimation results for WUT**

[27] Take SEV-GE model for instance, $\gamma$ identifies the degree of curvature, and $\tau$ represents the elevation of the weight function.

We observed several appealing results from SEU. First of all, LL of SEU is improved marginally w.r.t EVT and EUT. SEU even leads to the improvement of LL comparing to all SEV models. Furthermore, we also observed the improvement of SEU in terms of parameters' t-statistic, especially for weighting parameter $\gamma$. This finding supports that it is better to simultaneously apply nonlinear utility and decision weight into prospect function.

It should be noted that the form of weighting function seem to have little influence on the model goodness of fit. This raises a question on the necessity of choosing weighting functional forms for modelling travellers' risky choice behaviours. However, different weighting functions actually generate different estimates which potentially affect the valuation of travel time savings. Thus it is necessary to account for the role of various weighting functions when we calculate VTTS (refer to Chapter 6 for details).

### 5.5.3 RDEU

The estimation results of the EVT model shows that the estimated parameter for travel time is negative, which enables us to rank risky outcomes according to travel time, i.e. the outcome with 8min is ranked as number 13, and correspondingly the outcome with 20min is number 1. Given the above construct, the Tversky & Kahneman (T-K) specification and Prelec specifications (refer to Table 3.2 for details) are tested. Moreover, the influence of travel time difference between two consecutive outcomes on the estimation results is also of interest to this study. As shown in Table 5.4, the estimator $\gamma$ of the T-K model improves when the travel time increment switches from 1min to 2min, implying that the estimated travel time distribution actually affects the estimation of the decision weight. However, it seems that model fit does not vary across different models.

In terms of RDEV models, all the estimated parameters are statistically significant at the 1% level except $\gamma$ which is not significantly different from 1. There could be three possible reasons for this, viz., poor dataset quality, inappropriately estimated travel time distributions, and inappropriate weighting function. The first reason is highly possible since there are merely 210 observations in the floating dataset. And the second reason is also quite likely as indicated by the improvement of the estimated $\gamma$ from RDEV-TK with 1min to RDEV-TK with 2min. We also tested several other weighting functions, such as the weighting functions with two parameters (Gonzalez and Wu, 1999). However, the negative sign of weighting parameter is in contrary to our expectation. Therefore, the two-parameter nonlinear probability transformation specifications were not accepted.[28] Based on the above analysis, the parsimonious T-K weighting function (2min) is selected as the recommended structure for subsequent RDEU model in this context. However, RDEU fails to improve model fit even it introduces nonlinear utility function. In terms estimated parameter, the utility parameter α is not statistically significant, while the other parameters seem relatively constant across different model specifications.

| | RDEV-TK 1min $\frac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{(1/\gamma)}}$ | | RDEV-TK 2min $\frac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{(1/\gamma)}}$ | | RDEV-Prelec $e^{-(-lnp)^\gamma}$ | | RDEU-TK $\frac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{(1/\gamma)}}$ | |
|---|---|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| ASC | -1.760 | -5.220 | -1.760 | -5.220 | -1.800 | -5.290 | -1.570 | -3.940 |
| $\beta_{toll\_low}$ | -0.450 | -4.180 | -0.449 | -4.160 | -0.450 | -4.160 | -0.462 | 1.130 |
| $\beta_{toll\_high}$ | -0.313 | -3.120 | -0.309 | -3.080 | -0.308 | -3.080 | -0.324 | -4.260 |
| $\beta_{TT}$ | -0.006 | -3.180 | -0.006 | -2.810 | -0.006 | -3.080 | -0.009 | -3.210 |
| α | | | | | | | 0.124 | 1.030 |
| γ | 0.850 | 2.140 | 0.676 | 2.420 | 0.587 | 1.520 | 0.701 | 4.620 |
| Age | 0.709 | 2.850 | 0.707 | 2.850 | 0.715 | 2.880 | 0.707 | 2.850 |
| Gender | 0.825 | 3.480 | 0.825 | 3.480 | 0.856 | 3.610 | 0.800 | 3.350 |
| Final LL(β) | -230.367 | | -230.478 | | -230.464 | | -229.99 | |
| $\rho^2(0)$ | 0.241 | | 0.241 | | 0.241 | | 0.242 | |
| Adj.$\rho^2(0)$ | 0.218 | | 0.218 | | 0.218 | | 0.216 | |
| $\rho^2(ASC)$ | 0.086 | | 0.086 | | 0.086 | | 0.088 | |
| Adj.$\rho^2(ASC)$ | 0.059 | | 0.058 | | 0.058 | | 0.056 | |

**Table 5.4: Estimation results for RDEU**

The shape of RDEU weighting function contains information on individuals' attitudes toward risk, hence, it is necessary to analyse the shape of the weighting function and the associated decision weight. Figure 5.2 shows inverted-S shaped weight curves for both RDEV and RDEU models. Given the fact that the decision weight for the best outcome is equal to the value of the nonlinear weighting function (*w(p)=π(p))*, the inverse S shaped curve implies that individuals tend to concavely overweight the best outcome with a low probability, while convexly underweighting the best outcome with a high probability. Another observation from Figure 5.2 is that different weighting functions display different distortion capacity in terms

[28] It should be noted that this finding conflicts with some SP studies, such as Hensher et al. (2011) who found the two-parameter weighting function to be behaviourally better than the one-parameter specification.

of probability, though all of them keep the inverse S shaped curves. RDEV-Prelec largely overweight the probability that is less than 0.75. And RDEV-TK exhibits relatively less capacity of transformation, with the cross point with probability line at around 0.4. RDEU-TK generates the curve of weighting function which is similar as the one of RDEV-TK, while its distortion of probability is even less significant than RDEV-TK.

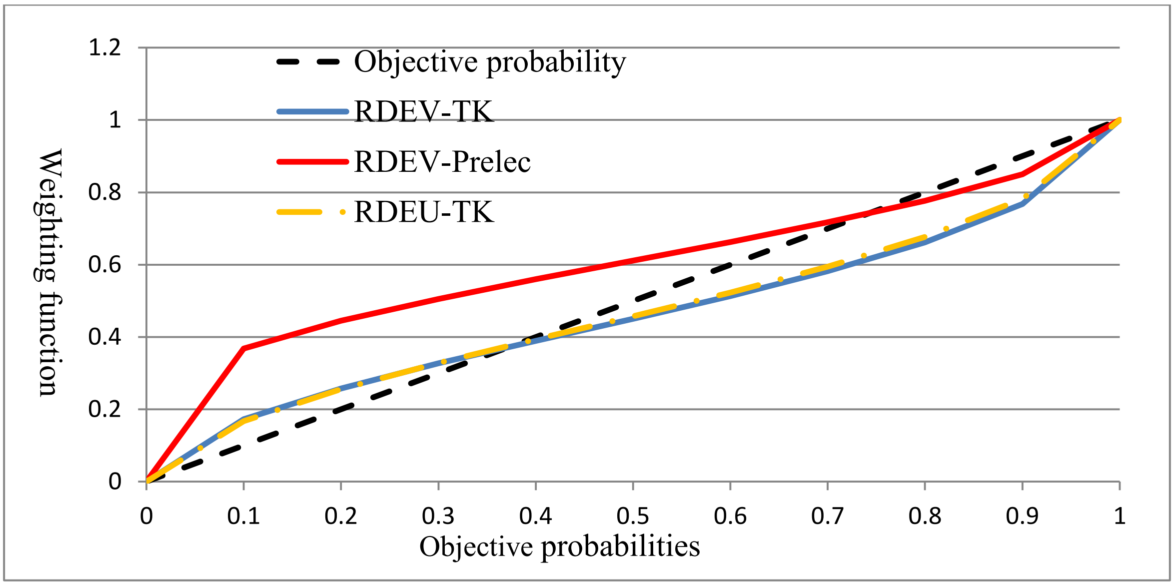



**Figure 5.2: Probability weighting function with one parameter**

Figure 5.2 does not show the final decision weight attached to each risky outcome. Based on the aggregated probability of outcomes, we are able to demonstrate the features of the decision weights in Figure 5.3. It should be noted that the objective probability is considered as the observed probability of trips.

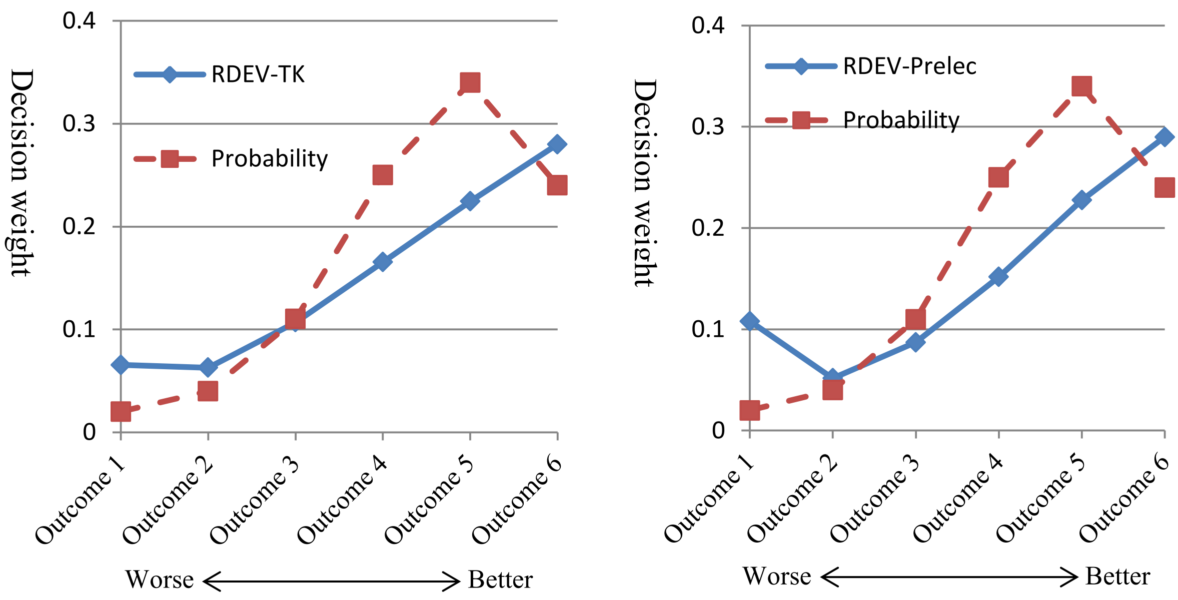



**Figure 5.3: Decision weight for RDEV models**

Despite the observed difference in terms of weighting functions, Figure 5.3 shows similar shape of decision weights estimated from the RDEV-TK and RDEV-Prelec models. In particular, the decision weights of outcome 3, 4 and 5 (that is from 10min to 15min) are always less than the corresponding objective probabilities, while the other outcomes are overweighted to different extents, in particular for outcome 1 (that is from 18min to 20min) which is referred to as the worst travel time outcome.

Three steps are proposed to identify attitude towards risk from the decision weight perspective only. Firstly, we should reveal the possible ranking order of outcomes. Then the perceived outcome distribution should be identified based on an estimated weighting function. Finally, the perceived distribution should be compared with the actual distribution to see whether the good and bad outcomes are over-weighted or under-weighted. In this research, the inverse S-shape of decision weights suggests that the probabilities of extremely bad outcomes are inflated at the range 2% to 7%, while the probabilities of the normal outcomes (outcome 4 and 5) are deflated at the range 4% to 13%. Evidently, drivers pay more attentions to extreme outcomes, especially the extremely bad outcome, whilst underweight the likelihood of normal situations. This finding suggests the presence of pessimism or risk aversion behaviour in the sampled respondents.

### 5.5.4 PT

Table 5.5 shows the estimation results for the PT models. The endogenously estimated reference points is around 8.8 minutes, and the other critical time point from which diminishing sensitivity applies is estimated to be $TT_m^n$=13 minutes[29]. PT 1 model only addresses reference dependence, while PT 2 model also captures diminishing sensitivity. Both models lead to marginal gain of model fit over the basic EUT model.

| | PT 1 | | PT 2 | |
|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. |
| ASC | -1.830 | -4.760 | -1.900 | -5.460 |
| $\beta_{toll\_low}$ | -0.437 | -4.010 | -0.434 | -3.990 |
| $\beta_{toll\_high}$ | -0.292 | -2.870 | -0.289 | -2.840 |
| $\beta_{TT}$ | | | | |
| $\beta_{TT(gain)}$ | 0.004 | 1.550 | 0.006 | 2.210 |
| $\beta_{TT(loss)}$ | -0.007 | -2.430 | -0.009 | -3.030 |
| $\beta_{TT(loss^-)}$ | | | -0.005 | -2.070 |
| Ref | 8.830 | 1.890 | 8.800 | 1.903 |
| Age | 0.709 | 2.840 | 0.710 | 2.840 |
| Gender | 0.820 | 3.410 | 0.840 | 3.500 |
| Final LL(β) | -228.358 | | -227.916 | |
| $\rho^2(0)$ | 0.248 | | 0.249 | |
| Adj. $\rho^2(0)$ | 0.221 | | 0.223 | |

[29] We initially used the method of setting reference point to estimate $TT_m^n$. Whereas, we found out that PT model is very sensitive to the value of $TT_{ref}^n$ and $TT_m^n$ (there is actually a kink around such point), and it is extremely difficult to evaluate the model with two estimated points. Thus, we tested a series of candidate model with different $TT_m^n$, and found out that the model with $TT_m^n$=13 min fits the data well.

| | | |
|---|---|---|
| $\rho^2(ASC)$ | 0.094 | 0.096 |
| Adj. $\rho^2(ASC)$ | 0.063 | 0.064 |

**Table 5.5: Estimation results for PT**

Consistent with our expectation, Table 5.5 shows that the estimated travel time parameter is positive when it is framed as a gain, while travel time is valued negatively when it is framed as a loss. Furthermore, the ratio between the absolute values is $\left|\beta_{TT(loss)}/\beta_{TT(gain)}\right| = 1.57 > 1$, and the t-ratios for the difference between $\beta_{TT(loss)}$and $\beta_{TT(gain)}$ is relatively high at 2.64, which empirically supports the validity of loss aversion. From a behavioural point of view, individuals show a significantly asymmetrical response to travel time decreases and increases from 8.8min. This asymmetrical behaviour also exists in the loss space where the outcomes with travel time more than 13min is valued less than the outcomes with less travel time. This is empirically demonstrated by the fact that $\left|\beta_{TT(loss)}\right| > \left|\beta_{TT(loss^-)}\right|$with high t-ratios at 10.55. Note that the taste parameter of travel time is associated with a probability, thus $\beta_{TT(loss)}$ and $\beta_{TT(loss^-)}$ represents the sensitivity to a risky outcome with weighted travel time. Hence, we believe that the smaller value of $\left|\beta_{TT(loss^-)}\right|$ is in line with intuition since the sampled respondents are not sensitive to the extremely bad situation with a small probability (the probability for the travel time outcome between 13min and 20min is only 0.16). We also tested nonlinear expressions characterizing diminishing sensitivity, however the parameter estimates turned out to be insignificant. Similar results are also reported in Masiero & Hensher (2010) in which the parameter of punctuality decreasing 0 to 2 percent is bigger than the parameter of punctuality decreasing 3 to 4 percent.

Figure 5.4 explicitly demonstrates the changes in utility with respect to the changes in the weighted travel time of gain and the weighted travel time of loss. A strong asymmetric response in travel time is clearly shown in Figure 5.4. In particular, an decrease of 1min in gain space results in an increase in utility of 0.21 units, an increase of 1min in loss space results in an increase in disutility of 0.31 units. Consequently, such asymmetrical utility function also supports the existence of loss aversion behaviour.

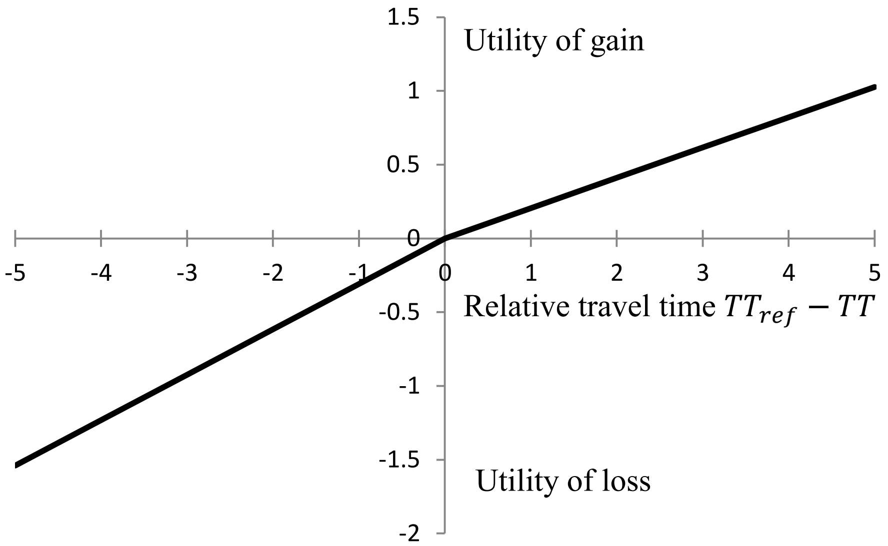



**Figure 5.4: Illustration of asymmetrical preference towards gain and loss**

## 5.6 Comparison of substantive estimation results

Non-EUT specifications proposed in this thesis gain more insights into travellers' behaviour, with the trade-off of more complicated model structures. However, whether complexity leads to superior models is clearly questionable. In fact, each risky choice theory/model explains travellers' behaviour from different perspectives, and their applicability may be dependent on specific choice contexts. It seems sensible to conduct a meta-analysis in order to identify whether the estimates of taste parameters and attitude parameters significantly vary across contexts in future research. In this research, we merely explore alternative specifications to determine which theoretical models actually provide the best fit.

In this section, we pay special attentions to the comparison of model goodness of fit and their behavioural implications. It should be noted that the comparison of estimated coefficients across models are also of interest to us. In particular, the marginal utility of travel time and monetary toll plays an extremely important role in implications. This is usually often referred to as the value of travel time savings (VTTS), or the willingness to pay (WTP) extra toll in return for decreases in travel time. Such a trade-off between travel time and toll is discussed in greater detail in Chapter 6.

### 5.6.1 Goodness of fit

In terms of goodness of fit, we notice that our models are actually divided into nested model and non-nested model. The former is referred to as the model that can be a parametric generalization of the others. In this research, it consists of all the models except PT. And the non-nested model, i.e. PT model in this research, applies to the situation in which comparative models are not the parametric generalisation of the given model itself.

To make comprehensive analysis of goodness-of-fit for nested models, we employ a series of indexes as shown in Table 5.6.

| | |
|---|---|
| $\rho^2$ test | $1 - LL(\beta)/LL(ASC)$ |
| Adjusted $\rho^2$ test | $1 - (LL(\beta) - K)/LL(ASC)$ |
| Akaike Information Criterion (AIC) | $-2LL + K * 2$ |
| Bayesian Information Criterion (BIC) | $-2LL + K * ln(N)$ |
| Consistent AIC | $-2LL + K * (1 + ln(N))$ |
| Corrected AIC | $-2LL + K * (2 + 2(K + 1)(K + 2)/(N - K - 2))$ |
| Likelihood ratio | $LR = -2(LL_R - LL_G)$ |

**Table 5.6: The test criterion of model fit for nested models**

The decision rule of $\rho^2$ statistic and adjusted $\rho^2$statistic is that a higher $\rho^2$ indicates a better fit to the data. Additionally, Hurvich and Tsai (1989) also suggested a variety of information criteria statistics to test model fit. It is generally specified as $-2LL + K * \delta$, where LL is the estimated value of log-likelihood, K is the number of parameters and $\delta$ is a penalty constant. Akaike Information Criterion (AIC) developed by Akaike (1974) is a simple way to compare models where $\delta = 2$. Gupta and Chintagunta (1994) used an alternative method called Bayesian Information Criterion (BIC) where N is the number of observations. Other more complex statistics are also adopted in this paper, e.g. consistent AIC and corrected AIC. To compare the EUT and its alternative models, we employ the likelihood ratio test where $LL_R(\beta_R)$ is the final log-likelihood value of the restricted model, and $LL_G(\beta_G)$ is the final log likelihood value of the general model. Here we regard EVT and EUT as the restricted model, and the other nested models are general models which become restricted model if several restrictions are imposed on parameters. Given that null hypothesis is formulated as 'restricted model is true', we reject the null hypothesis if $LR > \chi^2_\alpha$, where $\chi^2_\alpha$ is the critical value from the $\chi^2$ distribution. Noticed that there is no consensus in the literature suggesting the best test statistic, we use all the test statistics to comprehensively compare these nested models, namely EVT, EUT, WUT, SEV, SEU, RDEV, and RDEU (shown in Table 5.7).

| Models | EVT | EUT | WUT | SEV | SEU | RDEV | RDEU |
|---|---|---|---|---|---|---|---|
| Parameters | 6 | 7 | 7 | 7 | 8 | 7 | 8 |
| Final LL(β) | -230.414 | -230.151 | -230.401 | -230.111 | **-229.073** | -230.478 | -229.990 |
| $\rho^2(0)$ | 0.241 | 0.242 | 0.241 | 0.242 | **0.245** | 0.241 | 0.242 |
| Adj. $\rho^2(0)$ | **0.221** | 0.219 | 0.218 | 0.219 | 0.219 | 0.218 | 0.216 |
| $\rho^2(ASC)$ | 0.086 | 0.087 | 0.086 | 0.087 | **0.092** | 0.086 | 0.088 |
| Adj. $\rho^2(ASC)$ | **0.063** | 0.060 | 0.059 | 0.060 | 0.060 | 0.058 | 0.056 |
| AIC | **472.828** | 474.302 | 474.802 | 474.222 | 474.146 | 474.956 | 475.980 |
| BIC | **497.321** | 502.878 | 503.378 | 502.798 | 506.804 | 503.532 | 508.638 |
| Consistent AIC | **503.321** | 509.878 | 510.378 | 509.798 | 514.804 | 510.532 | 516.638 |
| Corrected AIC | **475.301** | 476.652 | 477.152 | 476.572 | 477.510 | 477.306 | 479.344 |
| Likelihood ratio w.r.t EVT | | 0.526 | 0.026 | 0.606 | 2.682 | -0.128 | 0.848 |
| Likelihood ratio w.r.t EUT | | | -0.500 | 0.080 | 2.156 | -0.654 | 0.322 |

**Table 5.7: Comparison of goodness-of-fit**

The first observation that can be made from Table 5.7 is that the difference between models in terms of LL is modest, though all the proposed non-EUT models turn out to provide slightly better model fit to the data. Among these non-EUT models is SEU which produces the best model fit from the $\rho^2$-statistic perspective. However, the parsimonious EVT model even outperforms SEU when we account for adjusted $\rho^2$-statistic. The same result can be found from AIC and BIC. If we look at likelihood ratio test results, SEU leads to marginal improvement of model fit at 0.25 of p-value w.r.t EVT model, and at 0.20 level w.r.t EUT model.

Moreover, we also notice that the model improvement is very limited if we only use either nonlinear utility or nonlinear weighting function, e.g. EUT, SEV, and RDEV. Whereas, the models jointly using nonlinear utility and weighting function can not only give higher LL but also generate more reliable estimates in terms of t-static. Plus, we also observed that different weighting functions of RDEV and SEV seem not to affect log-likelihood in this specific study. And different weighting functions consistently exhibit similar distortion of probability, e.g. the inverse S shaped decision weight for RDEV, and the S shaped decision weight for SEV.

Finally, despite the intuitive appeals of weighted utility and rank dependence, most comparison criterions surprisingly suggest that WUT and RDEV appear to generate the worst model fit to data (likelihood ratio w.r.t EUT is even negative). There exist two possible reasons for such relatively low LL of WUT and RDEV. Firstly, the model structure is not suitable for this specific research. For instance, though we tested three candidate utility weight $W(TT_k^n)$ for our WUT model, $W(TT_k^n)$ is still arbitrarily determined in this research. Also, model performance may be highly sensitive to the order ranking of outcomes, whilst we can only make assumptions which may be not consistent with the reality. In addition to model perspective, the other essential factor affecting estimation results is the data, in particular the network dataset. As discussed earlier, RP analysis requires reliable and high quality travel time distributions extracted from the network, whereas one of the shortcomings of our first case study is the floating car dataset which has very limited observations.

We also calculated the non-nested model fit comparison criteria for the PT models, showing a notable improvement of log-likelihood w.r.t EVT at 0.05 confidence level, and at 0.03 confidence level w.r.t EUT. However, as discussed at the beginning of this section, PT is not nested within our EUT model structures, in that it cannot be the parametric generalisation of any of the other models structures explored in this case study[30]. Therefore, we employ non-nested tests to compare the PT model with the other models.

There exist a large body of literature describing different methods to test non-nested hypotheses (Horowitz, 1983). One strategy is to merge the basic model and alternative model together, and then apply traditional tests to compare them (Davidson and MacKinnon, 1981). In our research, we follow another method used by Ben-Akiva and Lerman (1985) as well as Bhat and Pulugurta (1998), which determines whether the adjusted likelihood ratio index difference between two non-nested models is statistically significantly different. It is expressed as:

$$pr(\bar{\rho}_1^2 - \bar{\rho}_2^2 > \tau) \leq \Phi\{-[-2\tau LL(ASC) + (N_2 - N_1)]^{1/2}\} \quad (5.8)$$

where $\bar{\rho}_l^2$ is the adjusted likelihood ratio index for model *l,* and $N_l$ is the number of parameters in model *l.* $\Phi\{.\}$ is the cumulative standard normal distribution function. A small probability $pr$ indicates that the difference of adjusted likelihood ratio is statistically significant, and the model with higher $\bar{\rho}^2$ is preferred. Based on this test technique, we can compare PT with all the other models, and the test result is shown in Table 5.8.

| PT | Test statistics | *P*-values |
|---|---|---|

[30] PT model uses parameters $\beta_{TT(gain)}$, $\beta_{TT(loss)}$ and $\beta_{TT(loss^-)}$ to characterize travellers' tastes to travel time, while the others only use $\beta_{TT}$. Consequently, they represent different extensions of EVT, and all the other models cannot be a special case of PT models.

| | | |
|---|---|---|
| *vs* EVT | -1.650 | 0.3 |
| *vs* EUT | -1.863 | 0.2 |
| *vs* WUT | -1.992 | 0.18 |
| *vs* SEV | -1.841 | 0.2 |
| *vs* RDEV | -2.031 | 0.15 |

**Table 5.8: Non-nested test results for PT**

The non-nested test results indicate that if we take into account the cost of extra parameters the PT model does not fit the data better than the other models. Therefore, despite the improvement of LL from PT, we cannot, at least in this current stage, conclude that PT produces a statistically significant gain of model fit to the data.

## 5.6.2 Behavioural implications

The non-EUT models tested in this case study have behavioural appeal which enables modellers to more accurately model travellers' risky choice behaviours. For instance, the empirical estimation of the SEU model supports the hypothesis that individuals potentially use subjectively transformed probabilities as decision weights. It should be noted that such nonlinear weights not only depend on the objective probability but also on the utility per se (WUT) and the rank of outcome utilities (RDEU). The WUT model estimated with the RP data assigns increasing weights to the outcomes with increasing disutility, whilst RDEV and RDEU are capable of generating more flexible decision weights with an inverted-S shape. The evidence also reinforces the importance of empirically selecting a preferred weighting functional form. Two-parameter forms were abandoned in this case study due to the non-intuitive negative sign of the estimated parameter, while the estimation results do not support the one-parameter function (as the travellers' weighting pattern) in that log-likelihood of the RDEU-TK model is not improved with respect to our basic models. However, we are still able to extract essential information on risky attitudes from the estimated decision weights.

Travellers' attitudes toward risk were traditionally captured via the nonlinear transformation of the utility function. The risk parameter of EUT, SEU, and RDEU is found to be consistently less than 1. This is usually interpreted as risk aversion in the traditional gambling experiment, as the concave utility of monetary income implies that pessimistic gamblers consider the utility of risky outcome is less than the utility of certain outcome although both outcomes actually have the same utility. It should be noted that the utility of monetary income (in the gambling experiment) is positive, while travel time actually generates negative disutility as demonstrated in our case study. Thus, the induced risk parameter in our models actually produces a convex utility function, which suggests that drivers exhibit risk prone behaviours in the context of route choice. Another interpretation of the risk parameter is the diminishing marginal utility of travel time. In this research, we treat it as drivers' diminishing sensitivity to travel time, for instance, an extremely large travel time is perceived as being much smaller than its real magnitude. This suggests, for example, that drivers who are used to congestion and serious delay do not weight a 60 minute journey time as being twice as bad as a 30 minute journey time.

Interesting conclusions can be drawn if we jointly account for nonlinear utility and nonlinear weighting function. Such is the case with the SEU and RDEU models where we tried to explore risky attitudes from the deviation between nonlinear decision weights and objective probabilities. The inverse S shaped decision weight of RDEU shows that small probability tends to be overweighted, while large probability is underweighted. For instance, when the probability for a bad travel time outcome is very low, say 0.1, drivers appear to overweight it, $w(0.1) = 0.168$ according to our RDEU model. And correspondingly the usual outcomes are underweighted, $w(0.9) = 1 - 0.168 = 0.832$. This nonlinear distortion of probability suggests pessimistic behaviour. This is in line with consumer choice of purchasing life insurance to avoid unacceptable loss even though the associated likelihood is very small.

Only the PT model shows a marginally improved model fit over EUT. This improvement is mainly derived from reference dependence and diminishing sensitivity to travel time loss. Instead of accounting for travel time per se, individuals compare all possible travel time outcomes against 8.8min which is treated as the reference point. As expected, extremely good outcomes are interpreted as gain ($TT$<8.8$min$), whilst diminishing sensitivity occurs when travel time exceeds a threshold ($TT$>13$min$). Free-lane drivers appear to enjoy the trip when they perceive similar travel time as the time on toll road, especially when such smooth trip turns out to be free. It is the comparison that leads drivers to consider relative travel time as joy and gain. When the actual travel time exceed the reference travel time, drivers became cautious about the cost of travel time which is increasing but in a diminishing trend.

## 5.7 Summary

This chapter serves as the introduction of our empirical applications, with the first case study being a toll road analysis using a range of non-EUT specifications.

Taking a simple formulation of EVT and EUT as a basis for comparison, in this chapter we have presented results on the performance of non-EUT models based on weighted utility theory (WUT), subjective expected utility theory (SEU), rank-dependent expected utility theory (RDEU) and prospect theory (PT), using revealed preference data of real-life travel choices. In the context of this case study, none of the non-EUT models significantly outperforms the EUT and EVT models in terms of goodness of fit etc. However, we found that the quality of network data significantly affects the estimated travel time distribution and the induced estimation of decision weights. Hence, this result may be due, at least in part, to the shortcomings of RP data which cannot provide enough variation of travel time and the induced risky outcomes in real transportation context. In summary, the findings presented in this case study reinforce the importance of exploring non-EUT models within a revealed preference context before they can be applied reliably to modelling risky choices in the real world.

There exists another strand of research that applies more complicated model structures, such as the MMNL (Mixed MNL) model, to improve goodness-of-fit for non-EUT models. For instance, Hensher and Li (2012) demonstrate that the heterogeneity of the risk attitude parameter has significant influence on goodness-of-fit and estimates of their RDEU model. In future research, it would be interesting to incorporate heterogeneity in to the risk parameter, weighting parameter, diminishing sensitivity parameter etc. in our non-EUT models.

Furthermore, it is clear that a couple of problems associated with non-EUT models should be addressed before these advanced models are applied into practice. Specifically, reference points may not be fixed in reality, rather we expect them to context dependent and individual specific, i.e. there may be more than one reference point. One possible technique to endogenously estimate the possible reference points is to treat them as missing data or latent variables, and estimate them using the Expectation Maximization algorithm (as discussed in Appendix A). Moreover, in terms of applications of non-EUT approaches, an important aspect is the value of travel time savings (VTTS). Chapter 6 extends this first cae study by examining the implications to VTTS of the various non-EUT specifications. Finally, given the fact that the model fit improvement is modest and risk parameter is insignificant in this case-study, we raise the question of the quality of network data. To determine whether network data has a significant influence on the final results, we carried out a second case-study using qualified network data in Chapter 7.

# Chapter 6 VTTS CASE-STUDY

## 6.1 Introduction

Chapter 5 presents an RP case-study concentrating on a binary route choice with different travel time distributions. A range of risky choice models were established to understand drivers' behaviour better when unpredictable travel time is treated as a source of risk. It was found that drivers on the SR91 corridor tend to exhibit bounded rational choice behaviour, which is difficult to explain purely through economic intuition.

This chapter presents a more detailed model estimation using the same dataset. Particular attention is paid to the estimated parameters and their interactions. One of the most important applications of these parameters is calculation of the value of travel time savings (VTTS), which is an essential input in travel demand modelling and appraisal. In fact, the literature on VTTS for both freight transportation and passenger transportation is extensive and well developed. Valuation of travel time is traditionally determined by the marginal rates of substitution of travel time and travel cost in the mode choice context. However, existing studies usually overlook the possible influence of decision makers' attitude towards risk, and their behavioural attributes, in estimating VTTS. This chapter, therefore, presents updated methods for measuring VTTS in a risky choice framework, and highlights the importance of accounting for risk attitudes, nonlinear decision weights, rank dependence and loss aversion. Given the importance of VTTS for policy and appraisal, this research concludes that its usage should be properly identified prior to implementation.

The remainder of this chapter is organized as follows. The ensuing section thoroughly reviews theories of VTTS, and conducts a meta-analysis based on existing empirical studies. Section 6.3 demonstrates the research gap, and this is followed by a description of our methodology in section 6.4. Substantive estimation results are presented in section 6.5, in which the effect of model structures on VTTS is discussed. Finally, section 6.6 summarizes the key findings of this case-study.

## 6.2 Overview of VTTS studies

Value of travel time savings (VTTS) plays a vital role in measuring travellers' willingness to pay for travel time savings. Researchers are increasingly interested in VTTS for several reasons. Firstly, it is a critical measure in transportation policy decision making and transport infrastructure appraisal. Take transportation investment for instance, the most important objective of this investment is to reduce travel time and improve reliability. Indeed, as observed by Small and Verhoef (2007) in the US, around 45% of social variable cost is the cost of travel time and reliability, compared to vehicle capital costs (19%), operating costs (16%) and incident costs (16%). In the UK, roughly 80% of the benefits associated with new transport infrastructure investment are related to travel time savings. This is not surprising since travel demand models have consistently found that travel time is one of the most significant variables, with even more explanatory power than travel costs. Accurate estimates of VTTS therefore have great influence on the appraisal of transportation investment (such as cost-benefit analysis). Moreover, it is also considered to be a crucial index in travel behaviour models and travel demand models. Take the traditional four-step model for instance: trip assignment is determined by the sum of travel time and travel cost rolled into a generalised cost, in which the cost of travel time equals the product of travel time and assumed VTTS, with the traveller being assigned to the route with the lowest generalized cost. This method is still applied to most toll road demand models. Given the importance of VTTS, this section aims to present our findings from basic theories and empirical evidence.

### 6.2.1 Theories: what we know

Microeconomics literature initially applied the allocation of time between different activities to consumer behaviour analysis, and then the shadow price of time savings became an essential factor in the individual choice framework. The basic assumption is that time can be transferred freely between different activities, for instance, the time saved from leisure activity can be used to increase labour income. In this case, it is sensible to determine the trade-off between money and time, of course including travel time.

Travel time is commonly treated as an added cost to travel choice, thus it is natural to measure travellers' willingness to pay (WTP) for travel time savings. Indeed, a large body of literature has sought to develop the model specification for VTTS, with the common starting point being the allocation framework of Becker (1965). Becker stated that consumers' satisfaction is not only subject to income constraints but also time constraints, which are divided into work, travel and leisure. As such, any marginal savings of time can be freely applied to increase income, with its value being simply the marginal trade-off between the monetary factor and time. The theory of time allocation and evaluation has been elaborated over the intervening decades by a number of studies from various perspectives (e.g., Johnson (1966); Oort (1969); Evans (1972) and DeSerpa (1971)). Jara-Díaz (2000) provided a detailed overview of the development of theories regarding the allocation of travel time savings, and more recently, Small (2012) carried out a selective review of a series of essential conceptual issues in VTTS. In fact, VTTS has long been an essential application of discrete choice modelling in transport. In line with consumer choice theories, its calculation under a discrete choice model is quite straightforward:

$$VTTS = \frac{\partial u/\partial time}{\partial u/\partial cost} \tag{6.1}$$

where $u$ gives the observed utility, and time and cost corresponds to travel time and travel cost respectively. The above equation is usually referred to as the marginal rate of substitution between travel time and travel cost. In the simplest case, when the utility specification is linear, i.e., $VTTS = \beta_{time}/\beta_{cost}$, VTTS corresponds to the ratio between the travel time parameter $\beta_{time}$ and the travel cost parameter $\beta_{cost}$. Estimates of these parameters are obtained from the calibration of discrete choice models. It should be noted that estimated VTTS may vary across populations with different gender, age, income level, trip purpose, and other characteristics. For instance, it has been found that the driver with a work commute trip purpose may reflect a more negative attitude towards travel time delay (higher absolute value of $\beta_{time}$), and correspondingly have a higher VTTS according to $VTTS = \beta_{time}/\beta_{cost}$. Another example is the observed difference between drivers with various incomes, given that high income drivers are less sensitive to travel cost (lower absolute value of $\beta_{cost}$), and tend to have higher VTTS. Additionally, the valuation of travel time should be differentiated when travel time is divided into several components, e.g. in vehicle time, platform waiting time, congested time and uncongested time. For instance, it has been found that passengers tend to hold less negative attitudes to in train time than platform waiting time. One possible reason for this difference is that passengers may prefer sitting in a moving train.

Since travel time variability also serves as an extra cost to travel choice, it is also natural to account for the value of reliability (VOR). Indeed, recent transport studies have increasingly recognized the fact that travellers are also willing to pay for an improvement in terms of travel time reliability, i.e., $VOR = \beta_{variability}/\beta_{cost}$ (insofar as the utility function is linear), where $\beta_{variability}$ is the parameter for travel time variability. Several empirical studies have even demonstrated higher estimated VOR than VTTS (Asensio and Matas, 2008, Batley and Ibáñez, 2009). In the traditional mean-variance model, VOR corresponds to the marginal trade-off between mean variance and travel cost. Interesting findings can be found if we jointly account for VTTS and VOR. Black et al. (1993) initially defined an index as VOR/VTTS. They called it a reliability ratio (RR) which can be simplified as $RR = \beta_{variability}/\beta_{time}$. This has been regarded as the traditional way to measure attitude towards risk, since it describes the extent to which travel time variability is more undesirable or desirable relative to travel time. Another benefit of using the RR index is to avoid the computation problems of monetary exchange rate and changes in the consumer price index when we compare different empirical studies.

Similar approaches can be extended to scheduling models where schedule delay early (SDE) and schedule delay late (SDL) deserve valuations as well. Again, the value of SDE and SDL is defined as the marginal rate of substitution between attributes: $VSDE = \frac{\partial u/\partial SDE}{\partial u/\partial cost}$ and $VSDL = \frac{\partial u/\partial SDL}{\partial u/\partial cost}$, where SDE and SDL represents schedule delay early and late respectively. Similar to the reliability ratio, we can also use the ratios VSDE/VTTS and VSDL/VTTS to represent travellers' asymmetric tastes to the cost of schedule delay and travel time.

With the increasing use of non-EUT models in transport studies, researchers have realized the explanatory power of VTTS in demand models and its applicable power in practice. The full theory of VTTS specifications in non-EUT cases is still not straightforward, however. Indeed, the most widely used functional form $VTTS = \beta_{time}/\beta_{cost}$ can only be applied to an EVT model in this research, since it requires the condition of linear utility specifications. More insights therefore need to be obtained to address complicated situations in a realistic way.

The brief sketch of VTTS presented in this subsection enables us to understand the basic theory of VTTS, and the following subsection will describe our comprehensive analysis based on estimates of VTTS from existing studies with the main objective of finding out how big VTTS is and which possible factors influence the estimation of VTTS.

### 6.2.2 Empirical evidence: what we can learn

#### *6.2.2.1 Valuation of travel time savings*

MVA et al. (1987) and Accent Marketing & Research (1996) provided an early summarization of the UK evidence on VTTS. Both market studies, using cross sectional data, proposed a more general model form for the allocation of monetary and time resources, rejecting a simple proportionality between VTTS and income, and indicating that VTTS would grow over travel time. Additionally, there is a large body of literature which attempts to estimate VTTS in various contexts by using revealed or stated preference methods. To better compare VTTS estimates in the existing literature, a sample of relevant studies are listed in Table 6.1.[31] It is evident that empirical estimates of VTTS vary significantly across the literature. This is not in itself surprising, given that respondents from various regions are likely to exhibit different acceptable levels of the valuation of travel time savings. However, even on the same corridor, notable differences were observed in terms of VTTS.

If we pay particular attention to a series of studies on the SR91 road, given that this dataset is of special interest to us, there are three crucial studies Small et al. (1999); Lam and Small (2001) and Small et al. (2005b), which here are referred to simply as SR1, SR2 and SR3. Each of these studies published their estimated result for the value of time based on data collected from the SR91 corridor. The first observation is that SR3 and SR2 generate a much higher VTTS than SR1. Since the main difference between SR1 and SR2 is that the former used SP data whereas the latter used RP data, it would be sensible to conclude that the type of dataset has a direct influence on VTTS. Indeed, evidence for a similar effect can be found in other literature, for instance, Ghosh (2001) demonstrated that the estimated VTTS from RP data is almost four times higher than the VTTS from their SP dataset ($40/h vs $13/h). This suggestion can be convincingly supported by setting out the VTTS estimated across a large body of literature, as in Figure 6.1.

[31] It should be noted that some of the EUT and non-EUT studies illustrated in Table 6.1 do not explicitly demonstrate VTTS, such as Michea and Polak (2006), De palma and Picard (2005), and De Lapparent (2010). We estimated their VTTS based on their estimation results.

| Study | Data | Year | Mode | VTTS | VOR | RR | VSDE | VSDL | Risky choice model | Utility | Decision weight |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bates et al. (2001) | SP | 1999 | Rail | NA | NA | NA | $52.42/h | $106.39/h | EVT | NA | NA |
| Small et al. (1999) (SR1) | SP | 1995 | Car | $3.9/h | $12.6/h | 3.23 | NA | $18.6/h | NA | NA | NA |
| Hensher (2001) | SP | 1999 | Car | $8.7/h | $5/h | 0.57 | NA | NA | NA | NA | NA |
| Tilahun and Levinson (2010) | SP | NR | Multi modes | $7.82/h | $6.93/h | 0.89 | $0.41/h | $7.11/h | NA | NA | NA |
| Polak et al. (2008) | SP | 1999 | Car & rail | $8.81/h | NA | NA | $2.15/h | $17.61/h | EVT | NA | NA |
| Lam and Small (2001) (SR2) | RP | 1997-1998 | Car | $24/h | $12/h (male) $30/h (female) | 0.5-1.3 | NA | NA | NA | NA | NA |
| Liu et al. (2004) | RP/SP | 1999-2000 | Car | $12.8/h | $20.6/h | 1.61 | NA | NA | NA | NA | NA |
| Small et al. (2005) (SR3) | RP | 1999-2000 | Car | $27.44/h | $24.31/h | 0.89 | NA | NA | NA | NA | NA |
| Small et al. (2005) | SP | 1999-2000 | Car | $11.99/h | $5.54/incident | NA | NA | NA | NA | NA | NA |
| Brownstone and Small (2005) | RP/SP | 1999-2000 | Car | $12.55/h | $5.02/h | 0.4 | NA | NA | NA | NA | NA |
| Noland et al. (1998) | SP | 1994 | NR | NR | NR | 1.27 | NA | NA | EVT | NR | NR |
| Black et al. (1993) | SP | NR | Multi modes | $7.24/h | $5.08/h | 0.7 | NA | NA | NA | NA | NA |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bhat and Sardesai (2006) | RP/SP | 2003 | Multi modes | $12.2/h | $3.3/h (flex); $6.1/h (infle) | 0.2-0.5 | NA | NA | NA | NA | NA |
| Ghosh (2001) | RP | 1998-1999 | Route | $40/h | NA | NA | NA | NA | NA | NA | NA |
| Ghosh (2001) | SP | 1998-1999 | Route | $13/h | NA | NA | NA | NA | NA | NA | NA |
| Carrion and Levinson (2010) | RP | 2008-2009 | Route | $9.15/h | $5.99 /h | 0.91 | NA | NA | NA | NA | NA |
| Asensio and Matas (2008) | SP | 2005 | Car | $18.89/h | NA | NA | $12.06/h | $28.27-$68.47 /h | NA | NA | NA |
| Hollander (2006) | SP | 2004 | Bus | $6.47/h | $0.62/h | 0.1 | $4.81/h | $13.48/h | NA | NA | NA |
| Senna (1994) | SP | NR | Car & bus | $0.22/h-$1.23/h | $1.08/h-$2.00/h | 0.8-1.6 | NA | NA | EUT | Mixed[c] | NA |
| Liu and Polak (2007) | SP | 1999 | Rail | NA | NA | NA | £8.25/h | £22.3/h | EUT | Risk aversion | NA |
| Polak et al. (2008) | SP | 1999 | Car & rail | $3.14/h | NA | NA | $1.17/h | $6.49/h | EUT | Risk aversion | NA |
| Hensher and Li (2012) | SP | 2008 | Car | $21.13/h | NA | NA | NA | NA | RDEU | Risk proneness | Risk aversion |
| Hensher et al. (2011) | SP | 2008 | Car | $6.30/h-$9.65/h | NA | NA | NA | NA | SEU | Risk proneness | NR |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Li et al. (2009) | SP | 2008 | Car | $19.1/h | NA | NA | NA | NA | SEU | Risk aversion | NR |
| Michea and Polak (2006) | SP | 1999 | Rail | NA | NA | NA | $34.96/h | $70.53/h | RDEV | NA | Mixed |
| De Palma and Picard (2005) | SP | 2000 | Car | NA | NA | NA | NA | NA | EUT | Mixed[c] | NA |
| Hess et al. (2008) | SP | 2004 | Car | WTA= $4.5/h, WTP= $7.3/h [a] | NA | NA | NA | NA | PT | NA | NA |
| Rose and Masiero (2010) | SP | 2004 | Car | WTP= $1.6/h, WTA= $5.8/h [b] | NA | NA | NA | NA | PT | NA | NA |
| Masiero and Hensher (2010) | SP | 2008 | Multi modes | WTP= $9.04/h, WTA= $24.53/h | WTP= $78.37/h, WTA= $227.37/h | NA | NA | NA | PT | NA | NA |
| De Lapparent (2010) | RP | 2002 | Flight | $59.50/h | NA | NA | NA | NA | RDEU | Risk neutrality | Risk proneness |

NA: not applicable; NR: not reported
For the purposes of comparison, all the values have been converted into US dollars based on the exchange rates current at 20th February 2013.
a Here only commuters' WTP and WTA w.s.t free flow travel time is recorded. Further details are provided in Hess, Ross and Hensher (2008)
b Here only commuters' WTP and WTA w.s.t free flow travel time in WTP/WTA space is considered. Preference space is provided.
c Risk prone for commuters with a fixed arrival time, and risk aversion for the others
d 66% risk aversion and risk neutrality; 33% risk proneness

**Table 6.1: Summary of empirical studies on VTTS**

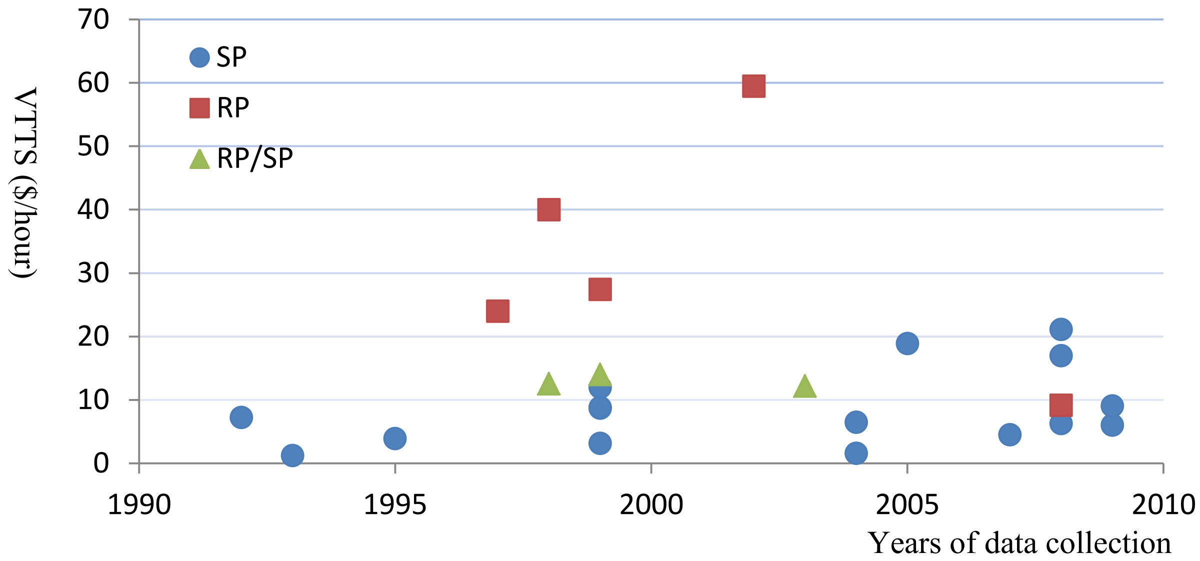



**Figure 6.1: Illustration of estimated VTTS from selected studies**

Based on the VTTS shown in Figure 6.1, it seems that the estimated VTTS from RP studies is slightly higher than the estimate from SP studies, which is most likely due to the systematic effects of SP and RP methodologies. Alhough we cannot conclude which method is more suitable for VTTS studies, there are at least two reasons why SP studies appear to underestimate VTTS. Firstly, as noted by Brownstone and Small (2005), respondents on SR91 reported perceived travel time savings were almost twice that of the real travel time savings using a toll facility. This perception bias magnifies the benefits of using a toll road,[32] and may potentially lead to drivers having higher estimates of VTTS in the RP context. Secondly, respondents usually fail to take into account the actual situations in real travelling experience when they answer hypothetical questions in SP surveys, and therefore, they tend to display inconsistency compared to actual behaviour. For instance, in a real situation, drivers may be impatient about serious congestion, which result in an intensified feeling of regret or loss. This negative attitude towards travel time delay may also serve to enlarge VTTS, whereas respondents in an SP survey, which is more remote from the event, are unlikely to feel the same as they do in real congestion. However, it should be noted that most studies  used in this comparison are from the US, and the empirical studies in the UK are different. For instance, Abrantes and Wardman (2011) conducted a meta-analysis on VTTS studies undertaken in the UK, and concluded that the valuation of SP is not greatly different from RP values. Consequently, given the regional variation, empirical practices should be conducted to accurately estimate VTTS in differnt cases.

[32] Ghosh (2001) found out that drivers with bigger perception errors tend to be more likely to use toll roads.

To mitigate the effect of methodological differences between SP and RP, we compare SR2 and SR3 using an RP sample, and SR1 and SR3 using an SP sample. Given the same type of data, VTTS of SR3 is still higher than the estimates of SR1 ($11.99/h vs $3.9/h) and SR2 (($27.44/h vs $24/h). There are at least two reasons for this continuing difference. The first explanation is from a model perspective, given the fact that SR1, SR2 and SR3 employ MNL, Nested Logit, and MMNL model structure respectively. Similar findings are obtained by Hensher and Li (2012) who obtained higher VTTS from their MMNL model ($21.13/h vs $17.92/h). This empirical evidence would suggest that the MMNL model specification tends to deliver higher estimates of VTTS when respondents' heterogeneity is taken into account.

The second reason is from a data perspective. For RP studies, SR2 extracted travel time distribution is from loop detector data, while SR3's travel time data is based on their floating car dataset. For SP studies, SR1 verbally presents five arrival time scenarios for each alternative to represent travel time variability, whilst the SP survey in SR3 describes travel time variability as the frequency of 10 minutes or more delay for each alternative. It has been found that the different formats and presentations of SP questionnaires may have an impact on respondents' judgements, in particular on travel time and travel time variability. According to Tseng et al. (2009), it is easier for respondents to understand travel time distribution when it is described verbally. From his point of view, SR1 is preferred.

Another argument may be that different data collection periods appear to have an influence on VTTS, given that the data was collected in 1995, 1997 and 1999 for SR1, SR2 and SR3 respectively. It is almost certainly the case that more recent data is capable of generating higher VTTS if the other factors are constant, due to the increase due to inflation. This is also supported by the increasing trend shown in Figure 6.1.

Additionally, it should be noted that other factors potentially play significant roles in determining VTTS, such as choice type, trip purpose, gender and regional difference, etc. For instance, the highest VTTS shown in Table 6.1 is obtained by De Lapparent (2010) focusing on air route choice, with the value of almost 60 dollars per hour. It is very likely that flight passengers tend to spend more money on saving journey time than the other passengers, say train passengers, given that a flight ticket is much higher than a train ticket. Finally, having understood the possible impact of survey and discrete choice models on VTTS, further research should attempt to address whether EUT and non-EUT approaches affect the estimates of VTTS.

*6.2.2.2 Valuation of travel time reliability*

We have also found out different estimates of VOR, in a range from \$0.62/h to \$30/h. One possible explanation for such a wide variation is that different representations of travel time variability are related to final VOR. For instance, SR1, SR2 and SR3 employ standard deviation, dmp90 and dmp80[33] of travel time to characterize travel time variability, and correspondingly they obtained different VOR. Similar to the comparison of VTTS, comparing VOR across different time periods may also be affected by inflation and exchange rate. To eliminate this impact, it is sensible to adopt the reliability ratio (RR) as the index for comparison. It should be noted that another merit of using RR is that it reflects attitude to risk, since this ratio indicates whether the cost of travel time variability is more or less undesirable than the cost of travel time for decision making. Specifically, travellers are risk averse if RR>1 (travel time variability is more undesirable), risk prone if RR<1(travel time is more undesirable), and risk neutral if RR=1 (indifferent between travel time and travel time variability). In fact, as shown in Figure 6.2, our observations from 15 empirical studies are mixed, with the mean estimated RR=0.95.

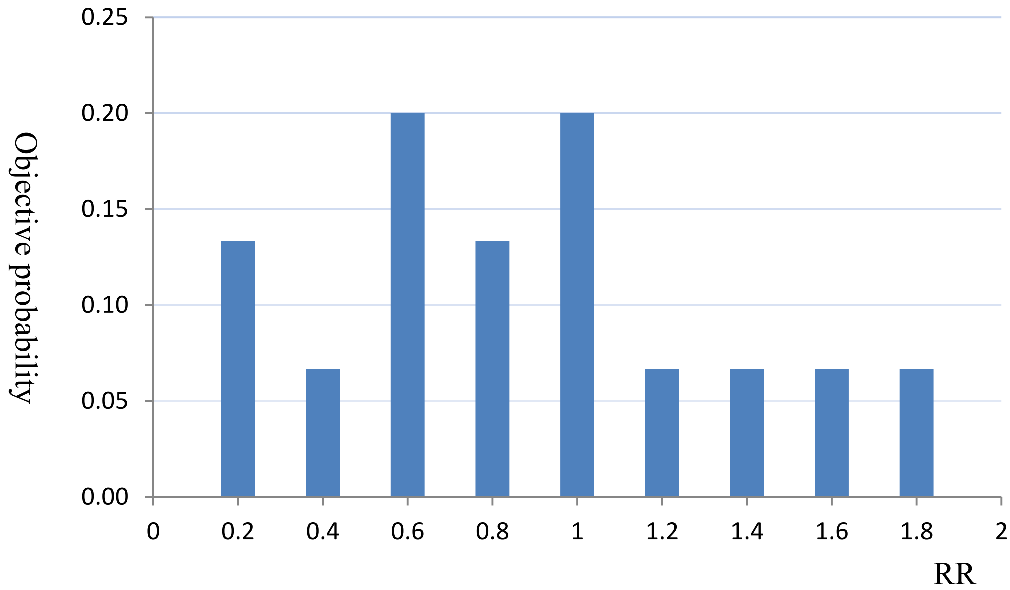



**Figure 6.2: Distribution of reliability ratio (RR) based on selected studies**

In scheduling models, the relationship between different valuations appears to be relatively stable. It has consistently been found that $VSDL > VSDE$, which reflects travellers more

[33] Dmp90 means the 90$^{th}$ percentile of travel time minus median travel time, and dmp80 is the 80$^{th}$ percentile of travel time minus median travel time.

negative taste to SDL than SDE. This is intuitively plausible, since there may be an additional penalty for late arrival, especially for workers. Indeed, as noted by Bates et al. (2001), punctuality is highly valued by respondents, and VSDL is almost twice that of VSDE. Noland et al. (1998) compared the performances between the mean-variance model and the scheduling model, concluding that the scheduling model provides the best goodness of fit and acceptable estimators, with the relationship being $VSDL > VTTS > VSDE$. Hollander (2006) conducted a similar study showing $VSDL > VTTS > VOR > VSDE$.

In summary, a large body of literature has applied a single index to valuate travel time variability such as VOR, VSDL and VSDE. It should be noted that all the indexes are constrained to specific models (mean-variance model and scheduling model), and they cannot deal with different definitions of travel time variability. In this current research, however, we focus on the travel time distribution *per se* rather than its moment. As a result, we can apply EUT and non-EUT models to modelling the distribution perceived by travellers, and then estimate the VTTS which is embedded within the information on travellers' attitudes toward risk.

## 6.3 Gaps in existing work

It is evident that most empirical studies on the valuation of travel time merely apply a linear-utility approach, and simply calculate VTTS as $\beta_{time}/\beta_{cost}$. This linear-utility specification, however, overlooks three essential components in real decision making situations.

First, unpredictable travel time implies risks, whilst traditional VTTS specification omits travellers' attitude towards risk. Thus, it can only address the valuation of travel time for the specific population who are risk neutral. Secondly, it has consistently been found that reported travel time by respondents to questionnaires is different from actual travel time, and therefore, subjective travel time distribution should be taken into account. Thirdly, it is still unclear whether elements of non-EUT models, such as reference dependence and rank dependence, have an influence on the estimation of VTTS.

In the remainder of this chapter, we demonstrate our methods and estimation results in order to bridge the gaps evident in VTTS. We present a formulation in which the VTTS from EUT and non-EUT models is no longer just the ratio between the travel time parameter and the travel cost parameter. Instead, different VTTS specifications are established for each model. More importantly, unlike the previous section which focuses on the comparison of VTTS across studies, the application presented in subsequent sections aims to investigate the impact of model structures on VTTS using the same dataset, which hopefully provides more valid insights into the improvement of future transport services and project appraisal.

## 6.4 Methodology

### 6.4.1 Data and Model specifications

The study presented here employs the same data as used in Chapter 5. To analyse travellers' trade-off between travel time and travel cost, we establish a range of utility functions which characterize travellers' risky choice behaviour. After model calibrations, we can obtain estimates on how travellers evaluate travel time with respect to travel cost. To express it in a flexible way, the utility function is given by:

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} w(p_k^n)\, g(TT_k^n) + \beta_{Toll}^n Toll^n + \beta_W^n W_i + \varepsilon^n \qquad (6.2)$$

where $g(TT_k^n)$ is the utility of travel time, and $w(p_k^n)$ corresponds to the decision weight of probability $p_k^n$. Socio-demographic attributes $W_i$ are also included in this model, since we are capable of investigating the impacts of respondents' characteristics on VTTS. Model estimation is partly based on the results shown in Chapter 5, but we also include other alternative specifications which, although they fail to improve goodness-of-fit, do provide different estimates of VTTS.

As a starting point, we consider both $g(TT_k^n)$ and $w(p_k^n)$ as a linear specification, i.e., $g(TT_k^n) = TT_k^n$ and $w(p_k^n) = p_k^n$. Evidently, this gives an EVT model specification with the traditional valuation of travel time as $VTTS = \beta_{time}/\beta_{toll}$. In our EUT model, attitude towards risk is taken into account. In particular, we tested two popular utility specifications, i.e., constant relative risk aversion (CRRA) $g(TT_k^n) = \beta_{time} \sum_{k=1}^{K} \frac{(TT_k^n)^{1-\alpha}}{1-\alpha}$, and constant absolute risk aversion (CARA) $g(TT_k^n) = \sum_{k=1}^{K} \frac{(1-e^{-\alpha TT_k^n})}{\alpha}$.

The SEV model gives a nonlinear weighting function $w(p_k^n)$, and the SEU model jointly accounts for the nonlinear weighting function and utility function. It assumes that travellers make decisions based on subjective travel time distribution rather than objective distribution. Decision maker's subjective distortion of probability is embodied by nonlinear $w(p_k^n)$. Additionally, it has been found not only that nonlinear weighting functions affect decision making, but also tht the rank ordering of outcomes matters. To investigate whether rank dependence has an influence on VTTS, we establish RDEV and RDEU models. In this current research, we test VTTS of SEV, SEU, RDEV and RDEU based on five popular weighting functions:

$$\text{TK: } w(p) = \frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{(1/\gamma)}}$$

$$\text{WG: } w(p) = \frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{\tau}}$$

$$\text{GE: } w(p) = \frac{\tau p^{\gamma}}{\tau p^{\gamma}+(1-p)^{\gamma}}$$

$$\text{Prelec-I: } w(p) = e^{-(-lnp)^{\gamma}}$$

$$\text{Prelec-II: } w(p) = e^{-\tau(-lnp)^{\gamma}} \tag{6.3}$$

Finally, reference dependence is also of interest to us, given that asymmetrical preference to gain and loss might also apply to the measurement of VTTS. This is evaluated in our PT model which divided travel time outcomes into gain and loss according to their relative location to the reference point.

### 6.4.2 VTTS specifications

Within an EVT model, the linearity of the model form indicates that the value of travel time savings can be straightforwardly obtained from the ratio between estimates, i.e. $VTTS = \beta_{time}/\beta_{toll}$. In the case of EUT and non-EUT models, the derivations of valuation measures are quite different. The following discussion aims to present the general VTTS specifications which are capable of incorporating essential components omitted by traditional measurements.

Within an EUT model, VTTS is not only dependent on the parameters of travel time and travel cost, but is also dependent on utility functional forms and the specific travel time distribution. For the CRRA approach where $u_i^n = \cdots + \beta_{time} \sum_{k=1}^{K} p_k^n \frac{(TT_k^n)^{1-\alpha}}{1-\alpha} + \cdots$, VTTS is expressed as follows:

$$VTTS = \frac{\partial u}{\partial time} \Bigg/ \frac{\partial u}{\partial toll} = \frac{\beta_{time} \sum_{k=1}^{K} p_k^n (TT_k^n)^{-\alpha}}{\beta_{cost}} \tag{6.4}$$

For the CARA model where $u_i^n = \cdots + \beta_{time} \sum_{k=1}^{K} p_k^n \frac{(1-e^{-\alpha TT_k^n})}{\alpha} + \cdots$, VTTS is given by:

$$VTTS = \frac{\partial u}{\partial time} \Bigg/ \frac{\partial u}{\partial toll} = \frac{\beta_{time} \sum_{k=1}^{K} p_k^n e^{-\alpha TT_k^n}}{\beta_{cost}} \tag{6.5}$$

where $\alpha$ is the risk parameter which characterizes the traveller's attitude towards risk. As shown in the above equation, the VTTS functional form varies across different utility functions, while the function $VTTS = \frac{\partial u/\partial time}{\partial u/\partial toll}$ is a general form which can be applied into any other models. For instance, if we consider a more complicated model, say a hyperbolic absolute risk aversion model where $u_i^n = \cdots + \beta_{time} \sum_{k=1}^{K} p_k^n \frac{1}{\alpha-1} (\gamma + \alpha TT_k^n)^{\frac{(\alpha-1)}{\alpha}} + \cdots$, the corresponding VTTS is expressed as:

$$VTTS = = \frac{\beta_{time} \sum_{k=1}^{K} p_k^n (\gamma + \alpha TT_k^n)^{(\frac{-1}{\alpha})}}{\beta_{toll}} \tag{6.6}$$

Basically, the SEV and RDEV approaches do not change the functional form of VTTS if we merely account for the model with nonlinear decision weight or rank dependence. Interestingly, the VTTS specification is restructured if we consider utility and the weighting function jointly, such as in SEU and RDEU. In this research, both SEU and RDEU models employ CRRA utility, given that the other utility functions either fail to provide significant estimates or cannot give an improvement in model fit. Thus, the corresponding VTTS function for SEU and RDEU is expressed as:

$$VTTS = \frac{\beta_{time} \sum_{k=1}^{K} w(p_k^n)(TT_k^n)^{-\alpha}}{\beta_{toll}} \tag{6.7}$$

As shown in the above equation, in addition to parameter $\beta_{time}$ and $\beta_{cost}$, both risk parameter $\alpha$ and decision travel time distribution have some impact on the final estimates of VTTS. It is of special interest to us to reveal what the extent of this impact is.

Until now, all the VTTS specifications demonstrated in this section belong to symmetric measurement, i.e., decision makers are indifferent between gain and loss. These models produce identical estimates of the willingness to pay for travel time savings (WTP) and the willingness to accept an increase in travel time in return for toll savings (WTA). In fact, these asymmetric models assume that VTTS corresponds to the marginal ratio between the travel time parameter and the travel cost parameter. Nonetheless, within the PT model, loss aversion implies that indifference curves are kinked at a reference point. As a result, the previous measurement of VTTS is invalidated, WTP turns out to be different from WTA due to reference dependence, and WTA is believed to be larger than WTP if loss aversion holds. Expressions of WTP and WTA for the PT model are not as straightforward since the parameter of travel time is divided into loss space, diminishing loss space, and gain space. Specifically, WTP is expressed as:

$$WTP = \frac{\beta_{time(gain)}}{\beta_{toll}} \tag{6.8}$$

And WTA is given by:

$$WTA = \frac{\beta_{time(loss)}}{\beta_{toll}} \tag{6.9}$$

Based on WTP and WTA, it is possible to seek evidence of drivers' asymmetric response to travel time, which would support the reference dependence of PT from VTTS perspective.

## 6.5 Estimation results

The results of our estimations are summarized in Table 6.2 for EVT and EUT models. In this table, estimation results have been shown for several essential parameters including the travel time parameter, the toll parameter for both a low income population (≤$60000 per year) and a high income population (>$60000 per year). The segmentation of income is crucial since previous studies have consistently found that the negative taste to the cost of the toll varies across different income levels. This distinction in terms of toll cost actually leads to a significant difference in VTTS. Thus, the first observation is that the decision makers with high incomes appear to over-weight their VTTS compared to decision makers with low incomes. This finding arises repeatedly in each model estimation result through this section. In fact, we believe that it is in line with intuition, given that drivers with a high income are very likely to suffer a more serious penalty for their late arrivals and, therefore, they are willing to pay higher toll fees in order to avoid possible travel time delay.

### 6.5.1 VTTS estimation

Examining the results in more detail, the EVT model is treated as the basic model, with a linear utility functional form suggesting risk neutrality. That is, attributes are riskless and travellers have no attitude towards risk in their decision makings. This risk free model delivers a VTTS with a relatively low value at roughly $27.70/h for low income drivers and $40.00/h for high income drivers. The estimates of VTTS turn out to be quite different, however, if attitude towards risk is taken into account. In this research, we account for CRRA and CARA utility functional forms. The former is a power function and the latter an exponential function. The estimates of risk parameters for both CRRA and CARA are less than one (suggesting risk proneness), and both EUT models give larger estimates of VTTS compared to the EVT models. This signifies that failure to take into account attitude towards risk would result in severe underestimations of VTTS. In terms of goodness-of-fit, neither CRRA nor CARA can significantly improve model fit. In particular, CARA function even gives inferior estimation results if we adopt LL measures. It should be noted, however, that Li and Hensher (2012b) found opposite results in which linear models tended to overestimate VTTS. This is an indication that attitude towards risk can act in either direction.[34]

| | EVT | | EUT-CRRA | | EUT-CARA | |
|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| $\beta_{toll\_low}$ | -0.450 | -4.170 | -0.459 | -4.240 | -0.409 | -5.240 |
| $\beta_{toll\_high}$ | -0.312 | -3.110 | -0.320 | -3.180 | -0.297 | -1.660 |
| $\beta_{TT}$ | 0.006 | -3.270 | -0.009 | -2.050 | -0.012 | -1.130 |
| Final LL(β) | -230.414 | | -230.151 | | -230.551 | |
| Adj. $\rho^2(0)$ | 0.221 | | 0.219 | | 0.218 | |
| *Low income* VTTS ($/hour) | 27.703 | | 31.072 | | 30.072 | |
| *High income* VTTS ($/hour) | 39.957 | | 44.568 | | 41.222 | |

**Table 6.2: VTTS results based on EVT and EUT models**

What effect could nonlinear sensitivity have on a travel demand model and, in particular, on measures of VTTS? The answer has been illustrated in Table 6.3 where the estimation results for SEV and SEU are compared using different weighting functions. For SEV models, despite almost identical goodness-of-fit across models, variations of VTTS are still observed

[34] It should be noted that Li and Hensher (2012) employ SP data and a mean-variance approach which is different from this current study.

in a range of \$29.5/h - \$33.1/h for low income drivers and \$42.9/h-\$48.0/h for high income drivers. This suggests that different weighting functions may have little impact on model fit, but a significant impact on the implied VTTS.

Moreover, all three candidate SEV models produce higher estimates of VTTS than EVT. Among these SEV models is the SEV-TK model, which delivers the highest valuation of travel time, at roughly \$5 - \$8 more than the VTTS estimated by EVT. The estimated value of travel time is even higher when we evaluate SEU models. Although the increase in VTTS is relatively modest with respect to SEV models, this still supports the contention that a nonlinear utility function can enlarge the estimated value of travel time savings, especially when nonlinear decision weight and utility functions are considered jointly.

| | SEV-TK | | SEV-GE | | SEV-WG | |
|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| $\beta_{toll\_low}$ | -0.448 | -4.150 | -0.447 | -4.130 | -0.446 | -4.120 |
| $\beta_{toll\_high}$ | -0.309 | -3.080 | -0.308 | -3.060 | -0.308 | -3.050 |
| $\beta_{TT}$ | -0.007 | -2.260 | -0.006 | -0.920 | -0.007 | -1.180 |
| Final LL(β) | -230.111 | | -230.302 | | -230.301 | |
| Adj. $\rho^2(0)$ | 0.219 | | 0.218 | | 0.218 | |
| *Low income* VTTS (\$/hour) | 33.099 | | 29.543 | | 30.347 | |
| *High income* VTTS (\$/hour) | 47.988 | | 42.876 | | 43.943 | |

| | SEU-TK | | SEU-GE | | SEU-WG | |
|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| $\beta_{toll\_low}$ | -0.460 | -4.250 | -0.462 | -4.260 | -0.456 | -4.210 |
| $\beta_{toll\_high}$ | -0.321 | -3.190 | -0.323 | -3.200 | -0.319 | -3.170 |
| $\beta_{TT}$ | -0.056 | -2.370 | -0.030 | -1.030 | -0.002 | -1.200 |
| Final LL(β) | -229.073 | | -229.335 | | -230.06 | |
| Adj. $\rho^2(0)$ | 0.219 | | 0.215 | | 0.213 | |
| *Low income* VTTS (\$/hour) | 35.314 | | 30.784 | | 30.686 | |
| *High income* VTTS (\$/hour) | 50.605 | | 44.728 | | 44.221 | |

**Table 6.3: VTTS results based on SEV and SEU models**

As shown in Table 6.4, interesting findings can be obtained if rank dependence is taken into account. It should be noted that the only difference between the RDEV and SEV model is that the RDEV specification requires, along with nonlinear sensitivity, prior processing regarding rank ordering of risky outcomes. Since relatively higher estimates of VTTS have been consistently obtained from SEV models with respect to the EVT model, it is surprising to observe that our RDEV model actually delivers the lowest VTTS of all alternative models. Furthermore, it seems that the selection of various weighting functions has little influence on the final LL and VTTS. This is also opposite to the results of SEV estimations in which the range of VTTS across different weighting functions is relatively larger. If a nonlinear utility function is incorporated into RDEV, however, (i.e. an RDEU-TK model),[35] the estimated VTTS turns out to be even larger than the estimates of SEU-TK. It is difficult to explain the increase of VTTS from RDEV to RDEU, although we should note that the utility function and weighting function in RDEU-TK give different interpretations of risk attitudes. Specifically, in RDEU-TK, the risk parameter of the utility function is less than 1, which implies risk prone behaviour, whereas the corresponding weighting function tends to over-weight bad outcomes and under-weight good outcomes, i.e., it is an inversed-S shaped weighting function, which suggests pessimism and risk aversion (refer to section 5.5.3 for a detailed discussion).

| | RDEV-TK | | RDEV-GE | | RDEV-PrelecI | | RDEV-PrelecII | | RDEU-TK | |
|---|---|---|---|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| $\beta_{toll\_low}$ | -0.449 | -4.160 | -0.447 | -4.100 | -0.450 | -4.160 | -0.451 | -4.180 | -0.462 | 1.130 |
| $\beta_{toll\_high}$ | -0.309 | -3.080 | -0.297 | -2.930 | -0.308 | -3.080 | -0.312 | -3.100 | -0.324 | -4.260 |
| $\beta_{TT}$ | -0.006 | -2.810 | -0.006 | -3.640 | -0.006 | -3.080 | -0.009 | -2.920 | -0.009 | -3.210 |
| Final LL(β) | -230.478 | | -230.141 | | -230.464 | | -229.653 | | -229.99 | |
| Adj. $\rho^2(0)$ | 0.218 | | 0.219 | | 0.218 | | 0.217 | | 0.216 | |
| *Low income* VTTS ($/hour) | 25.295 | | 25.316 | | 26.562 | | 26.222 | | 36.243 | |
| *High income* VTTS ($/hour) | 36.756 | | 38.102 | | 38.809 | | 38.002 | | 51.680 | |

**Table 6.4: VTTS results based on RDEV and RDEU models**

We now look at the estimation results for PT models, as shown in Table 6.5. Note that travel time is evaluated differently according to the position relative to the reference travel time, The PT1 model addresses reference dependence by using two different parameters for travel time gain and loss, respectively. As a result, respondents demonstrate negative taste to loss,

[35] Here we merely present RDEU-TK model, as the other weighting functions either deliver insignificant estimates or result in computation problems.

and positive taste to gain. Moreover, the asymmetric nature of the PT model leads to a significant disparity of travel time, with $\beta_{TT(loss)} = -0.007$ and $\beta_{TT(gain)} = 0.004$. This asymmetric measurement has been found in the valuation of travel time, with the ratio WTA/WTP equalling 1.75. For PT2, diminishing sensitivity is taken into account by including an extra parameter $\beta_{TT(loss^-)}$ for an extremely bad travel time outcome ($TT > 13min$). Consequently, $|\beta_{TT(loss^-)}| < |\beta_{TT(loss)}|$ suggests that the sampled respondents are not sensitive to the extremely bad situation with a small probability. Furthermore, in comparison with EVT, the WTP of the PT model decreases slightly by 5% for low income individuals, while it remains roughly the same at about 40 $/hour for high income individuals. Both population segments express much higher value of travel time loss than travel time gain (41.3 $/hour and 62.1 $/hour respectively). Specifically, the ratio WTA/WTP is 1.56 for low income individuals and 1.57 for high income individuals. It should be noted that our estimated WTA/WTP ratio is relatively lower than the estimates from studies on markets. For instance, Horowitz and McConnell (2002) reported that the mean WTA/WTP is roughly 2.9 from a comparison of 59 ratios from nine studies on market goods. It is easy to understand the big gap between WTP and WTA in consumer choice, however, given that an decision maker may deliberately elevate the selling price or transaction cost in order to obtain higher profits.

Within transport studies, a larger WTP-WTA gap is found in an SP context. For instance, Masiero and Hensher (2010) demonstrated that WTA/WTP equalled 2.7 using their SP data in a freight transport context. It is still not clear why SP studies seem to overestimate WTA/WTP, although it has been found that the size of WTP-WTA gap may be related to the respondent's familiarity with the choice environment and choice task. Thus it is probable that the valuation of travel time is misperceived by respondents who fail to understand the presentation of the questionnaire correctly. In the hypothetical context of SP, respondents could also omit several essential factors that they would take into account for real choices, such as scheduling pressure and the impatience in congestion. As a result, they tend to overestimate the compensation for increasing travel time (WTA) and even underestimate the money which they are willing to pay for reducing travel time (WTP), which leads to a larger gap between WTA and WTP. Special attention should also be paid to determining the reference point and travellers' perceived travel time distribution.

| | PT1 | | PT2 | |
|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. |
| $\beta_{toll\_low}$ | -0.437 | -4.010 | -0.434 | -3.990 |
| $\beta_{toll\_high}$ | -0.292 | -2.870 | -0.289 | -2.840 |
| $\beta_{TT(gain)}$ | 0.004 | 1.550 | 0.006 | 2.210 |
| $\beta_{TT(loss)}$ | -0.007 | -2.430 | -0.009 | -3.030 |
| $\beta_{TT(loss^-)}$ | | | -0.005 | -2.070 |
| Final LL(β) | -228.358 | | -227.916 | |
| Adj. $\rho^2(0)$ | 0.221 | | 0.223 | |
| *Low income* | | | | |
| WTP | 19.786 | | 26.406 | |
| WTA | 31.488 | | 41.360 | |
| WTA_dim | | | 24.513 | |
| *High income* | | | | |
| WTP | 29.611 | | 39.655 | |
| WTA | 47.125 | | 62.111 | |
| WTA_dim | | | 36.812 | |

**Table 6.5: VTTS results based on PT models**

### 6.5.2 Merging results of model estimation and VTTS

Recall the model estimation presented in Chapter 5, the results of the goodness-of-fit tests suggest that none of the non-EUT models produce a statistically significant increase of log-likelihood with respect to the EVT and EUT model. Though such a relatively ‘bad’ performance of the non-EUT models is not what we would expect, this cannot be regarded as conclusive proof due to the relatively poor network data available for this case study. This shortcoming is addressed in the second case study presented in Chapter 8, which uses a more rigorous dataset. Although it seems that various modelling approaches have little impact on the final model fit, VTTS turns out to significantly vary across different models.

As shown in Figure 6.3, the population with high income, as would be expected intuitively, values travel time savings higher than the population with low income, due to their different wealth levels. Despite the variety of VTTS estimated by these models, the VTTS difference between the low-income group and the high-income group is almost identical, at \$10. An exception is WTA estimated by PT2 in which WTA for high-income drivers is 50% higher than the WTA for low-income drivers.

Another observation is the variation across alternative models, and even across different weighting functions. We found that the average VTTS across 19 models is \$29.7/h for low income drivers and \$42.9/h for high income drivers, and the standard deviation is \$3.5 and \$4.9 for low income and high income drivers respectively.[36] EVT, RDEV and PT models deliver lower estimates of VTTS than the mean value (for PT model, WTP and WTA- is lower). The other models give higher valuations of travel time, with the highest estimates from WTA of PT2. Given drivers' asymmetric measure of travel time due to loss aversion, this relatively high WTA is in line with expectations. Based on the observations on the estimation of VTTS, it can be concluded that model structures have influences on the valuation of travel time. Specifically, nonlinear utility and weighting functions appear to enlarge VTTS compared to the linear EVT model, while loss aversion leads to a significant disparity of WTP and WTA. These findings suggest that future research should pay special attention to components such as risk attitudes, rank dependence, nonlinear decision weight, and loss aversion before VTTS is implemented in practice.

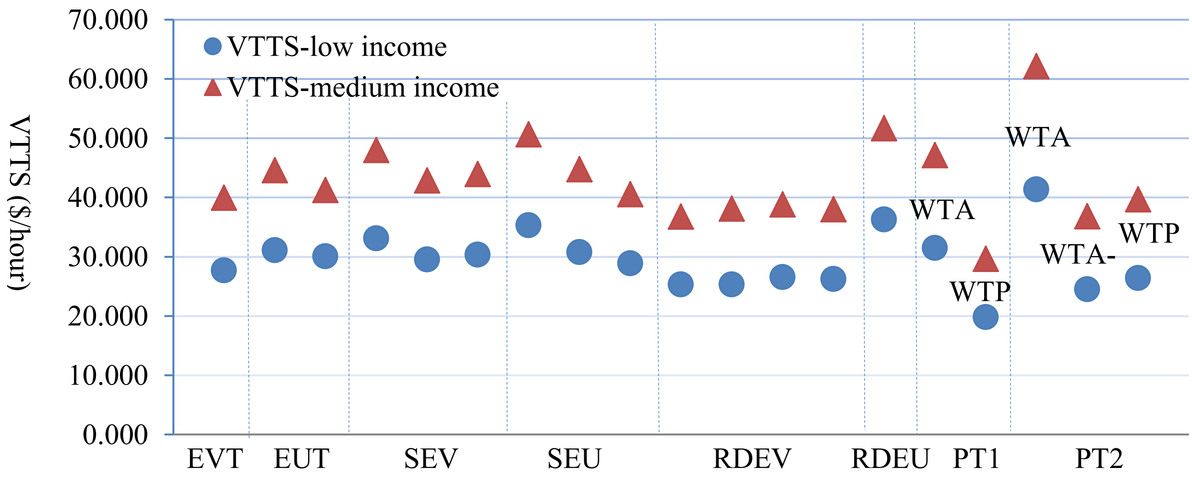



**Figure 6.3: The influence of model structures on the estimation of VTTS**

Traditional estimation methods for VTTS rely on the marginal trade-off of travel time and travel cost parameters in the calibration of the discrete choice model. It has been found that this linear utility approach overlooks several crucial components of real decision making under risk. This research proposes a series of alternative modelling approaches for valuing travel time savings which are able to take into account the behavioural factors omitted by

[36] To calculate a comparable VTTS based on PT models, we use the geometric average of the WTP and the WTA to estimate the mean VTTS for reference-free preference (De Borger and Fosgerau, 2008). As a result, the estimated VTTS is \$24.941/h and \$49.922/h for PT1, \$29.651/h and \$42.922/h for PT2.

traditional models, such as attitude towards risk, subjective probability distribution, rank dependence and reference dependence. It should be noted that the value of reliability (VOR) can be also addressed if we incorporate the variable of travel time variability (e.g. standard deviation or mean variance of travel time) into the utility function[37]. This research explores a new way to incorporate travel time risk into VTTS by using nonlinear utility function, although mean variance specification is not adopted.

## 6.6 Summary

The value of travel time savings is an essential factor for travel demand modelling and toll road project appraisal. In this chapter, we not only thoroughly reviewed relevant literature and compared empirical estimates across studies, but also discussed our own empirical findings of VTTS using the same RP dataset. The analysis presented in this chapter has revealed the significant variation of VTTS across different model structures, in particular EUT and non-EUT approaches, although model structures appear to have little impact on model fit.

One critical purpose of testing different models within the same choice context was to compare the influence of modelling approaches on the estimated VTTS. It was found that nonlinear utility and weighting function tend to overestimate VTTS, while rank-dependence seems to underestimate it. In this present context the most interesting observations arise from the PT model. Despite the evidence of loss aversion that has been found in a large body of literature using laboratory experiments and SP surveys, to our knowledge, no transport research has attempted to reveal whether these hypothetical observations in an SP context apply to actual travel behaviour. In fact, existing literature cannot even explain whether loss aversion in travel choice is an artefact created in the hypothetical SP context which cannot be applied to real choices. To address this question, this current research demonstrated a disparity between WTP and WTA which reinforces the validity of loss aversion in an RP context, although the observed WTA/WTP ratio is relatively lower compared to the SP context. Hence, the finding from the PT models explains the puzzle as to why the penalty for late arrival may be higher than the benefits of reduced travel time. In summary, this empirical evidence suggests that VTTS should be carefully identified before it is applied into toll road

[37] We cannot obtain well-determined estimate of standard deviation in this case study, and VOR is therefore not calculated. However, future research is worth comparing it with standard VOR estimate.

project appraisal. Inappropriate methods could lead to improper VTTS and thus induce unnecessary losses and risks.

As has been repeatedly discussed in Chapter 5 and Chapter 6, the modelling results of EUT and non-EUT may be sensitive to the travel time data, and may even be context dependent. To obtain more insights into the validity of these behavioural models, it is worth conducting another empirical study based on a new dataset, preferably in a different choice context. The following chapter, therefore, aims to carry out a new case-study based on the London underground data, which hopefully will reinforce our findings regarding the explanatory power and predictive effect of our non-EUT models.

# Chapter 7 LONDON UNDERGROUND CASE-STUDY

## 7.1 Introduction

The two case-studies presented in Chapters 5 and 6 have discussed the estimation and application of non-EUT models using an RP dataset for road travel in the California SR91 corridor. They have shown that in the SR91 case study, non-EUT approaches, such as SEU, RDEU and PT, do not significantly outperform EVT. This finding, to some extent, indicates that different model structures have limited impact on the final model fit, although they still contribute to the application of non-EUT approaches in transport. This conclusion cannot be drawn based only on a single experiment/observation. Further research is, therefore, urgently required to provide an extensive analysis using more reliable data to compare the proposed non-EUT models. This chapter seeks to meet this need by analysing route choice in the London Underground (LU) network. In this case study, we will investigate not only the calibration of candidate models, but also their predictive performance for the purpose of model comparison.

The remainder of this chapter is organized as follows. Section 7.2 introduces the basic situation of London Underground, and is followed in Section 7.3 by a brief description of the choice context. Section 7.4 presents the data used in this research, and model specifications are discussed in Section 7.5. Section 7.6 presents the calibration and the corresponding comparison based on the calibration performance of the candidate models. Model validation is presented in Section 7.7, and the chapter is concluded in Section 7.8.

## 7.2 Risky Choice Context

The London Underground is one of the largest and most complex underground networks in the world. It is composed of 11 lines, with 402 kilometres of track and 270 underground stations. Although it is called “underground”, the proportion of the route in tunnels is only 45 per cent, with the other 55 per cent being above ground. More than one billion passengers travel by LU each year, which is as many as the entire number of passengers travelling by National Rail in the UK in 2010. This makes the LU the fourth largest underground system in the world in terms of track kilometres. This large and complex transport system is operated by London Underground Limited (LUL) which serves as a wholly owned subsidiary of Transport for London (TfL). It is the complexity of the LU network that provides substantive route choice for passengers. The layout of LU lines in central London (Zone 1 and Zone 2 for illustration) is presented in Figure 7.1, which additionally illustrates the corridors of interest to us (the selection of these corridors is explained in section 7.3), namely, Waterloo (WL) to Baker Street (BS), Finsbury Park (FP) to King's Cross St. Pancras (KS), Finsbury Park to Green Park (GP), and King's Cross St. Pancras to Green Park.

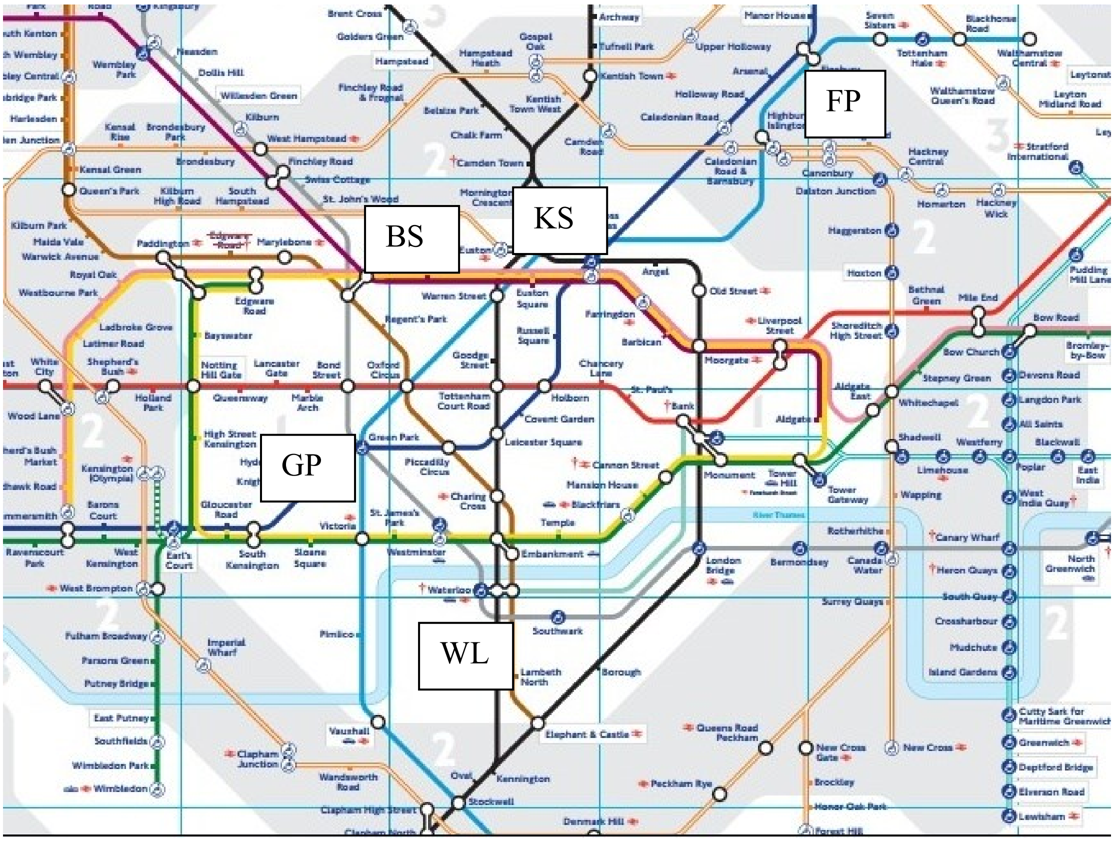



**Figure 7.1: Map of London Underground network with study corridors**

More lines serve north London than the south due to unfavourable geology and historical competition with surface railways in the south. Specifically, all the lines pass through the central London except the Waterloo and City line, East London line and Circle line. These underground lines are distributed in different ways. Some lines horizontally separate central London, e.g. the Central line and District line; or they are vertically distributed, e.g. the Northern line and Bakerloo line; or they even provide a loop service, e.g. the Circle line, and part of the Hammersmith & City line. Such a highly connected network 'unavoidably' results in a large number of interchange stations between different lines inside/outside central London. It is easy to observe that these lines tend to be linked with the major National Rail termini, such as King's Cross & St. Pancras, Victoria, Liverpool Street and Waterloo stations. This type of network enables passengers to have more route options if they depart from these stations connected with multiple lines. In fact, underground lines frequently cross each other, with the range being from seven interchange stations (on the Bakerloo line) to 18 interchange stations (on the District line), which suggests that each line cross each of the others almost twice on average. Indeed, there are a number of underground line pairs sharing more than one interchange stations along their routes, which naturally offers options for passengers in terms of route, such as South Kensington station and Ealing Broadway station, which are connected by both the Piccadilly line and District line.

### 7.2.1 Heterogeneous underground lines

As shown in Table 7.1, the 11 underground lines differ from each other in several aspects, such as infrastructure, location, rolling stock type, operation and service quality, etc Even though LU has been designed as a highly integrated and connected system, different underground lines actually provide passengers diverse travelling experiences.

| **Underground lines** | **Length (km)** | **Stations** | **Train types** | **Number of cars** | **Capacity per train** | **AWP** | **PWT (min)** | **ITT (min)** |
|---|---|---|---|---|---|---|---|---|
| Bakerloo | 23 | 25 | 1972 | 7 | 730 | 302869 | 2.55 | 11.97 |
| Central | 74 | 49 | 1992 stock | 8 | 892 | 589734 | 2.66 | 17 |
| Circle | 27 | 36 | C stock | 6 | 739 | 218136 | 4.69 | 15.78 |
| H & C | 25.5 | 29 | C stock | 6 | 739 | 149405 | 4.12 | 15.01 |
| District | 64 | 60 | C/D stock | 6 | 739/827 | 556252 | 3.16 | 18.94 |
| Jubilee | 36 | 27 | 1996 stock | 7 | 817 | 405878 | 2.54 | 15.75 |
| Metropolitan | 67 | 34 | A stock | 8 | 1045 | 186271 | 4.19 | 26.89 |
| Northern | 58 | 50 | 1995 stock | 6 | 665 | 660395 | 3.51 | 17.05 |
| Piccadilly | 71 | 53 | 1973 | 6 | 684 | 529550 | 2.44 | 19.07 |
| Victoria | 21 | 16 | 2009 stock | 8 | 864 | 511714 | 2.23 | 12.99 |
| W & C | 2.4 | 2 | 1992 stock | 8 | 892 | 37173 | 2.75 | 7.27 |

H & C: Hammersmith and City line; W & C: Waterloo and City line.

AWP, PWT and ITT correspond to average weekday passengers, passenger waiting time and in train time respectively. All time data is based on Journey Time Metric 2007 from TfL.

Train information is based on Transport for London (2012) and Transport for London (2013 a).

**Table 7.1: Basic characteristics of LU lines**

From an infrastructural perspective, the length of lines varies from 2.4 kilometres (the Waterloo and City line) to 74 kilometres (the Central line), and the number of stations also varies across lines, ranging from 2 (the Waterloo and City line) to 60 (the District line). It should be noted that the data shown in Table 7.1 merely describes the general characteristics of each line, while the situation may be quite different if we turn to look at a specific segment of each line. For instance, though the Metropolitan line has 31 more stations than the Jubilee line, the Jubilee line has 5 more stations than the Metropolitan line when we focus on the segment between Wembley Park station and Finchley road station. A path with many inter-stations between a specific OD appears to be less appealing than a path with fewer stations, given that frequent braking and accelerating results in more journey time and travel time variability. We should also note the other important attributes which potentially lead to varied travelling experiences for passengers. For instance, the LU infrastructure can be classified as subsurface (e.g. the Circle line and the District line) and deep-tunnel (e.g. the Bakerloo line and the Jubilee line). The former was constructed by the cut-and-cover method, with the average level being just 5 metres below the surface, while the latter applied a tunnelling shield running approximately 20 metres below the surface (this is the reason for the name of 'the tube'). Some parts of the underground lines are even above surface, such as the Central line section between White City and Ealing Broadway.

In terms of train type, current lines employ a variety of rolling stock, produced from 1969 to the present. This rolling stock has different sizes and capacities: specifically, the A, C and D stock is designed for subsurface lines with more interior spaces and capacity, whilst the other stock is designed for deep-tunnels and has fewer spaces due to the constraint of tunnel size. A consequence of the limited interior space of cars is the high level of in-train congestion. In regard to the Jubilee line, for example, passengers in the East London area used to complain that the 1996 stock was too small to feed the vast demand. Furthermore, the mode of train also directly affects the size of platform; for instance, Central line platforms are normally 120 metres long, while the station size of other stations is mostly 105 metres. Self-evidently, different sizes of platform have an influence on the level of platform congestion.

Travel time always serves as a vital attribute for route choice, and this is unlikely to be an exception for the decision making of LU passengers. Transport for London (TfL) developed the Journey Time Metric (JTM) to measure each component of journey time, namely the access and egress time between platform and station entrance, ticket queuing and purchase time, platform waiting time, in-train time, etc. Every component is given both a scheduled time and an actual time, with the difference being referred to as the excess journey time, which measures the average level of lateness. The values are annually estimated for 13 periods since 1998. In-train travel time and platform waiting time vary significantly across LU lines. Based on JTM, the most frequent service is the Victoria line which provides an average platform waiting time of 2.23 minutes which is almost 2.5 minutes less than the Circle line. Although the length of the Central line is longer the Metropolitan line, passengers on the Metropolitan line spend almost 10 minutes more on train time, on average, than do Central line users.[38]

In summary, the complexity of the LU network results in a number of underground stations from which passengers have multiple route options to reach their destinations. Furthermore, the LU network also provides materially different underground services which potentially lead to risky choices for travellers' decision making. Consequently, the LU can be considered an ideal context in which to examine travellers' risky choice behaviour.

### 7.2.2 Corridor selection

Having discussed the complexity and heterogeneity of LU network, we now look at the issues related to the identification of choice scenarios in the LU network. This enables us to understand whether LU passengers face risky route choices in their travel decisions, and which alternatives they may take into account.

In this research, 15 corridors were selected, which provide multiple, and potentially competitive, underground services.[39] Based on our selection criteria, as shown in Table 7.2, it is possible to determine the final choice scenarios. First, there must be at least two comparable underground services which are really competitive without dominating any alternative. This is regarded as an essential condition, since passengers would not take the other options into account if the dominating underground line provides overwhelmingly better services in terms of journey time, reliability, train delay and congestion.[40] Secondly, the comparable lines must not share the same track or platform, otherwise it is highly likely that passengers would choose the first coming train without taking into account the attributes

---

[38] JTM defines on train time as the time elapsed from wheel start of a train boarded to door opening of train alighted, and platform wait time as the time from customer arrival at midpoint of a platform to wheel start of boarded train.

[39] In our feasibility study, we investigated alternative proposals, but concluded that many of them have problems with data collection. For instance, many Docklands Light Railway (DLR) stations are not fully gated, and there are often not even any barriers to collect tickets or touch smartcards in/out. It has been found that many passengers do not touch out when they leave DLR stations. As a result, we are not able to obtain these passengers' route information from TfL's system, such as the Oyster smartcard data.

[40] To determine whether there is dominating service, we measured all the important attributes, such as journey time, standard deviation of travel time, and headways etc. Afterwards, we discussed these candidate routes with the staffs in TfL for further insights.

associated with different services. Thirdly, these comparable underground lines should have the same (or very closely spaced) station entrances and exits. This criterion is to exclude the influence of station location on passengers' route choices. Finally, the scenario selection is also constrained by data availability, in that the survey data used in this research was only selectively collected on limited stations annually, as described in greater detail in section 7.3.

Based on the above analysis, four route choice scenarios satisfying the above criteria were selected from the initial 15 listed in Table 7.2, and these four scenarios are shown in Table 7.3.

| Scenarios | Route choice | Excess Time *(min)* | Delay number (>15min) | D_ Time | D_ Delay | Mean Time | Different track or platform | The same station | Competitive choices | Long enough | Data availability |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Paddington --Baker Street | Bakerloo | 4.41 | 212 | 0.45 | 33 | 5 | Y | Y | Y | N | Y |
| | Circle | 4.86 | 179 | | | | | | | | |
| Edgware Road --Embankment | Bakerloo | 4.41 | 212 | 0.45 | 33 | 12 | Y | N | Y | Y | N |
| | Circle | 4.86 | 179 | | | | | | | | |
| Waterloo-- Baker Street | Bakerloo | 4.41 | 212 | 0.01 | 160 | 9 | Y | Y | Y | Y | Y |
| | Jubilee | 4.39 | 372 | | | | | | | | |
| Charing Cross --Waterloo | Bakerloo | 4.41 | 212 | 0.48 | 30 | 3 | Y | Y | Y | N | Y |
| | Northern | 3.93 | 182 | | | | | | | | |
| Ealing Broadway --Mile End | Central | 5.75 | 270 | 2.24 | 173 | 42 | Y | Y | N | Y | Y |
| | District | 3.52 | 443 | | | | | | | | |
| Liverpool street --White City | Central | 5.75 | 270 | 0.89 | 91 | 25 | Y | Y | N | Y | Y |
| | Circle | 4.86 | 179 | | | | | | | | |
| Stratford--Bond Street | Central | 5.75 | 270 | 1.36 | 102 | 21 | Y | Y | N | Y | Y |
| | Jubilee | 4.39 | 372 | | | | | | | | |
| Baker Street--Aldgate | Circle | 4.86 | 179 | 1.27 | 260 | 17 | N | Y | Y | Y | Y |
| | Metropolitan | 6.12 | 439 | | | | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ealing Common -- South Kensington | District | 3.52 | 443 | 1.22 | 118 | 20 | Y | Y | Y | Y | N |
| | Piccadilly | 4.73 | 325 | | | | | | | | |
| West Ham -- Westminster | Jubilee | 4.39 | 372 | 0.88 | 71 | 22 | Y | Y | N | Y | N |
| | District | 3.52 | 443 | | | | | | | | |
| Wembley Park --Baker Street | Jubilee | 4.39 | 372 | 1.73 | 67 | 14 | Y | Y | N | Y | Y |
| | Metropolitan | 6.12 | 439 | | | | | | | | |
| Uxbridge --Rayners Lane | Metropolitan | 6.12 | 439 | 1.39 | 114 | 15 | N | Y | Y | Y | Y |
| | Piccadilly | 4.73 | 325 | | | | | | | | |
| Finsbury Park --Green park | Piccadilly | 4.73 | 325 | 1.15 | 142 | 16 | Y | Y | Y | Y | Y |
| | Victoria | 5.89 | 183 | | | | | | | | |
| Finsbury Park --King's Cross | Piccadilly | 4.73 | 325 | 1.15 | 142 | 16 | Y | Y | Y | Y | Y |
| | Victoria | 5.89 | 183 | | | | | | | | |
| King's Cross --Green Park | Piccadilly | 4.73 | 325 | 1.15 | 142 | 16 | Y | Y | Y | Y | Y |
| | Victoria | 5.89 | 183 | | | | | | | | |

a All the travel time distribution data is extracted from the *Journey Time Metric dataset* which recorded the assembled tube performance data from 11/2010 to 12/2012.
Excess Time: the difference between measured journey time and scheduled journey time.
D_ Time: the difference of excess journey time between two alternatives
D_Delay: the difference of delays between two alternatives
Y: Yes; N: No

**Table 7.2: Competitive tube lines with different travel time distribution [a]**

| Origin station | Tube line 1 | Tube line 2 | Destination station |
|---|---|---|---|
| Waterloo | Bakerloo | Jubilee | Baker Street |
| King's Cross St. Pancras | Piccadilly | Victoria | Green Park |
| Finsbury Park | Piccadilly | Victoria | Green Park |
| Finsbury Park | Piccadilly | Victoria | King's Cross St. Pancras |

**Table 7.3: Choice scenarios for the London Underground (LU) study**

In each choice scenario, passengers have binary choices associated with different characteristics. All the study stations are in Zone 1 of central London, except Finsbury Park station, which is in Zone 2. In the first choice scenario, passengers at Waterloo station are assumed to either choose the Bakerloo line or the Jubilee line to arrive at Baker Street station. The origin, Waterloo station, serves as a main transportation terminus for both mainline railway and London Underground, with almost 100 million passenger entries and exits per year. Baker Street station, meanwhile, is a historic station which used to serve the world's first underground line. All the other three scenarios describe a binary choice problem between the Piccadilly line and the Victoria line. As the biggest underground interchange station in London, King's Cross St. Pancras station serves six underground lines as well as several railway lines. The lines running through King's Cross St. Pancras serve Finsbury Park to the north, which is also an important interchange station for railway and underground services, and Green Park to the south.

## 7.3 Description of data

### 7.3.1 Data requirement

The data required in the research needs to include the information on passengers' choice behaviour and the level-of-service for underground lines. As a result, the qualified data are twofold: the choice dataset reveals passengers' route choice and their socio-demographic information; the train performance dataset indicates the basic characteristics of each London underground line.

For the train performance data, we assume that the travel time extracted from the train performance dataset is the same as the time experienced by passengers. This is based on the condition that travellers have an idea of the actual travel time distributions of underground services they have used[41]. To make this statement more plausible, we also require that the observed travel time distribution must be specific to each passenger's particular departure time and objective origin-destination (OD). This is done by calculating the observed travel time distribution for every specific observation in the sample. Here, we use all the real journey time data one year before the observed trip.[42] This one-year performance data then enables us to calculate the actual travel time distribution which may be perceived by the specific respondent.

This data processing method leads, therefore, to significant variation of on-train time and headway[43] across observations in the sample, while it relies heavily on the quality and reliability of the train performance dataset. LU mainly applies Journey Time Metrics (JTM) to assess performance, and journey time is divided into five aggregate components from a passenger's perspective.[44] This data provides a general view of train performance, but it cannot be employed in this research because the travel time data is aggregated but not specific to time-of-day. As a result, we have to resort to detailed performance data which describes each train's movement, and, for this, the Network Management Information System (NetMIS) dataset is the only qualified data source.

For passenger choice data, there are two candidates which contain information on passengers actual route choices, namely Oyster Card data and the Rolling Origin Destination Survey (RODS) data. Oyster Card data records almost 80% of transactions at the origin station and destination station in the LU network. It relies on an automatic fare collection (AFC) system which records passengers' touch-in and touch-out details if they use an Oyster card as the payment medium. At the beginning of our research, we carried out a feasibility study on Oyster Card data. Although it records large amount of journey information and can potentially greatly enlarge our sample size, there are a number of limitations. Firstly, the most important problem is that there is no route choice information between OD pairs. Oyster Card data only captures the entry/exit time and station, but is incapable of identifying the chosen route if there are multiple routes between a specific OD (which is exactly the situation of interest to us). Secondly, the current system suffers from a technical constraint that entry and

[41] Future research should measure the potential impact of visitors/tourists.
[42] For instance, if the actual route choice is from Finsbury Park station to Green Park station using the Victoria line at 10:00 on 31/12/2009, we should extract all the real disaggregate journey time data for the Victoria line trains which also travel between Finsbury Park station to Green Park station and depart at around 10:00 from 01/01/2009 to 30/12/2009
[43] Headway is defined as the measure of time gap between two trains departing from the same station.
[44] It consists of access, egress, and interchange time; platform wait time; on-train time; ticket purchase time; and Closures.

exit timestamps are truncated at the minute level, and, indeed, the recorded timestamps are sometimes even incorrect. Finally, Oyster Card data is only capable of capturing card holder's travel information, while all non-Oyster card passengers are ignored. Consequently, this data cannot represent the whole population. Hopefully, future research could aim to solve the problems posed in the use of the otherwise very powerful Oyster card data for disaggregate choice modelling.

For the present analysis, the data on passengers' choice behaviour and characteristics is obtained from the Rolling Origin Destination Survey (RODS) data, which is currently collected and maintained by TfL. This is the only reliable data source that reveals passengers' actual route choices in the London underground network.

### 7.3.2 Rolling Origin Destination Survey (RODS) data

RODS is an annual rolling survey programme which was launched in 1998. For this research, RODS is of interest to us because it reveals respondents' selected path as well as their characteristics, which is vital for us to implement discrete choice model. Its main purpose for LU is to generate the OD matrix for daily underground services, and to estimate the flows between each OD pair. This is done by conducting passenger surveys at a random sample at selected underground stations (usually 30-40 stations subject to budget constraints) over continuous years. The selection of underground stations is determined by the expected changes in service provision and/or station ridership. During the survey, RODS questionnaires are randomly distributed to passengers who enter the station. The sample size is determined by using the hourly control totals for each underground station adjusted by the expected response rate. The assigned questionnaires are expected to be returned by mail. As shown in Appendix B, passengers are asked to provide detailed information on their journeys. Specifically, the key questions include, but are not limited to:

- The origin and destination for this particular trip (address and postcode)
- Selected underground service
- Departure time of this journey
- All the other transport modes that are used in this journey
- Trip purpose
- Ticket type
- All underground stations used in this journey

- Socio-demographic data including age, gender, travelling frequency and physical ability

This information is considered to be an essential input for several underground performance models. For instance, the Train Service Model applies RODS data as the key input for the analysis of demand. The Journey Time Metric (JTM) also uses RODS to calculate the weight of delay at a particular node in the whole journey. Moreover, the Pedroute Strategic Model (PEDS), which assesses the congestion and delay of underground lines, estimates flows at each entrance and exit on the basis of the passenger flows recorded in RODS.

While RODS represents a valuable data source to reveal route choice, we should be aware of several factors that make it less than perfect, and which have potential impacts on this research. Firstly, the sample size is still relatively small due to its limited sampling and low response rate (between 20% and 30% in recent years). This is the main obstacle for large-scale modelling since limited observations may lead to inaccurate estimation. Secondly, RODS is incapable of capturing the annual changes of passengers' travel patterns since most stations are surveyed every 8-10 years. Consequently, some old samples cannot be employed since it is impossible to obtain corresponding train performance data (as explained in 7.4.3, performance data prior to 2006 cannot be retrieved). Last but not the least, RODS surveys are merely distributed to passengers from 7:00 to midnight on weekdays, and thus it does not record the actual travel pattern at weekends.

To address these weaknesses, we only look at weekday travel pattern, and attempt to enlarge sample size by incorporating more corridors. From the RODS data, we extracted a subsample of respondents who made a journey along one of the four corridors that we identified in Table 7.3. This original sample contained 702 passengers, but the final sample was reduced to 661 passengers after data-cleaning (missing data, and compatibility between RODS and the level-of-service dataset).[45] This current sample compares favourably to the SR91 sample, which has only 438 observations. We then split the sample into two for the purpose of calibration and prediction respectively. For this present analysis, a 75% subsample (497 observations) was used as a calibration sample, and the remainder (164 observations) retained for model validation.

The descriptive statistics of the calibration sample is shown in Table 7.4. In terms of subsamples on each study corridor, the most observations were collected from the Waterloo

[45] The original observations of the RODS dataset (from 1998 to 2011) are much larger than the sample used in our final sample. However, we found out that many surveys were conducted before 2008, while the train performance data collected before 2008 is not available in the current TfL system (NetMIS data). Therefore, we had to abandon relatively old RODS data.

station — Baker Street station (WB) scenario, with a total of 210 observations collected in 2006 and 2009. The King's Cross St. Pancras station — Green Park station (KG) scenario also provides a relatively large sample with a total of 134 observations in 2008 and 2010, whilst the Finsbury Park station — King's Cross St. Pancras station (FK) scenario merely contributes 67 observations collected in 2009. Finally, we also retrieved 86 more observations from the RODS dataset for our Finsbury Park station — Green Park station (FG) scenario, which was surveyed in 2009. The observed statistics of the chosen route favour the Jubilee line and Victoria line, with an overall sample proportion of 25% and 40% respectively.

We are also interested in the segmentations of this population. This led to a subsample of 263 passengers travelling at peak hours, 234 passengers travelling at off-peak hours, 313 work trips, and 164 non-work trips. If we split the sample into two subsamples according to journey frequency, we find out that the proportion of frequent travellers varies across each choice scenario. Specifically, most respondents in the WB and FG samples reported that they normally have five or more trips per week, whilst only 38% and 39% respondents in the KG and FK samples, respectively, make frequent journeys using the particular underground service.

| | Fraction of Samples | | | | |
|---|---|---|---|---|---|
| **Choice** | **WB** | **KG** | **FG** | **FK** | **Total Sample** |
| Bakerloo line | 0.41 | | | | 0.17 |
| Jubilee line | 0.59 | | | | 0.25 |
| Piccadilly line | | 0.40 | 0.19 | 0.28 | 0.18 |
| Victoria line | | 0.60 | 0.81 | 0.72 | 0.40 |
| **Time periods** | | | | | |
| (6:00-10:00) [46] | 0.55 | 0.51 | 0.53 | 0.55 | 0.53 |
| **Age** | | | | | |
| (<35 years old) | 0.44 | 0.26 | 0.38 | 0.37 | 0.37 |
| **Female** | 0.54 | 0.34 | 0.57 | 0.52 | 0.49 |
| **Work** | 0.61 | 0.63 | 0.64 | 0.70 | 0.63 |
| **Journey Frequency** (≥5 per week) | 0.94 | 0.38 | 0.79 | 0.39 | 0.69 |
| **Observations** (All day) | 210 | 134 | 86 | 67 | 497 |

[46] Peak hour is defined as between 7:00 and 10:00 by RODS. If we look at the demand trend from 2000 to 2010, however, we found out that passenger growth in travel before 7am is approximately 100% which is much more than the growth in the other periods (normally 20% growth). Given the considerable change of demand trend in the early morning period, we decided to define the morning peak period as 6:00-10:00.

**Table 7.4: Descriptive statistics of calibration sample**

The RODS data presented in this section plays an essential role in this current research, given that it reveals passengers' route choices and their socio-demographic information. It is expected that future research could employ an extra dataset to enlarge sample size, such as the London Underground Oyster Card data. For the current analysis, however, only the RODS dataset is adopted but we conclude that RODS is good enough to provide sufficient observations for calibration and validation. Based on this survey data, we can now proceed in the exploration of level-of-service data.

### 7.3.3 NetMIS data

NetMIS, as an event-driven log, records the performance of each London underground train in the network. This data is transferred through several sections before it arrives at the NetMIS database. First, the signalling system is divided into a number of track circuits which indicates whether this section of track is occupied by a train. This track circuit data is then transferred to the signalling computers which process the data and send it to the TrackerNet database. It is the TrackerNet that enables NetMIS to extract 'train event' information by using logic. For instance, the event of the observed train arrival time at a platform is measured by adding an offset time to the track occupancy record.

While the automatically collected data of the current NetMIS is an improvement over the manually collected data of the previous NetMIS system, there are still several deficiencies. The most important problem is that data is missing from some stations due to issues with the signalling systems (Hickey, 2011). In addition, there are instances where train identifications are sometimes erroneously changed during a trip. These problems may affect the accuracy of estimates, especially headways between trains. Finally, some lines do not have NetMIS data since the signalling computers do not collect any track circuit data on these lines. The lines without NetMIS data are the Hammersmith & City, Circle, District and Metropolitan lines. For this research, however, all the lines of interest to us do have NetMIS data, and we found that the missing data for these stations is very limited. To avoid possible errors, we still conducted data cleaning to remove trip data with either missing performance data or erroneous train ID.

To retrieve the performance data, we can simply input the study periods, OD, and the specific underground line of interest to us. For instance, Figure 7.2 shows the user page of the NetMIS database through which we attempt to retrieve all the Piccadilly line trains' movement data in 2009. The corresponding output includes the unique identification number of each train, train number and trip number, the arrival/departure time at the origin station and the arrival/departure time at the destination station. Then we can calculate the actual in-train journey time by the difference between the train departure time at the origin station and the train arrival time at the destination station, and also calculate the dwell time by the difference between train arrival time and departure time at the origin station, and the observed headway by the departure time difference between two successive trains.

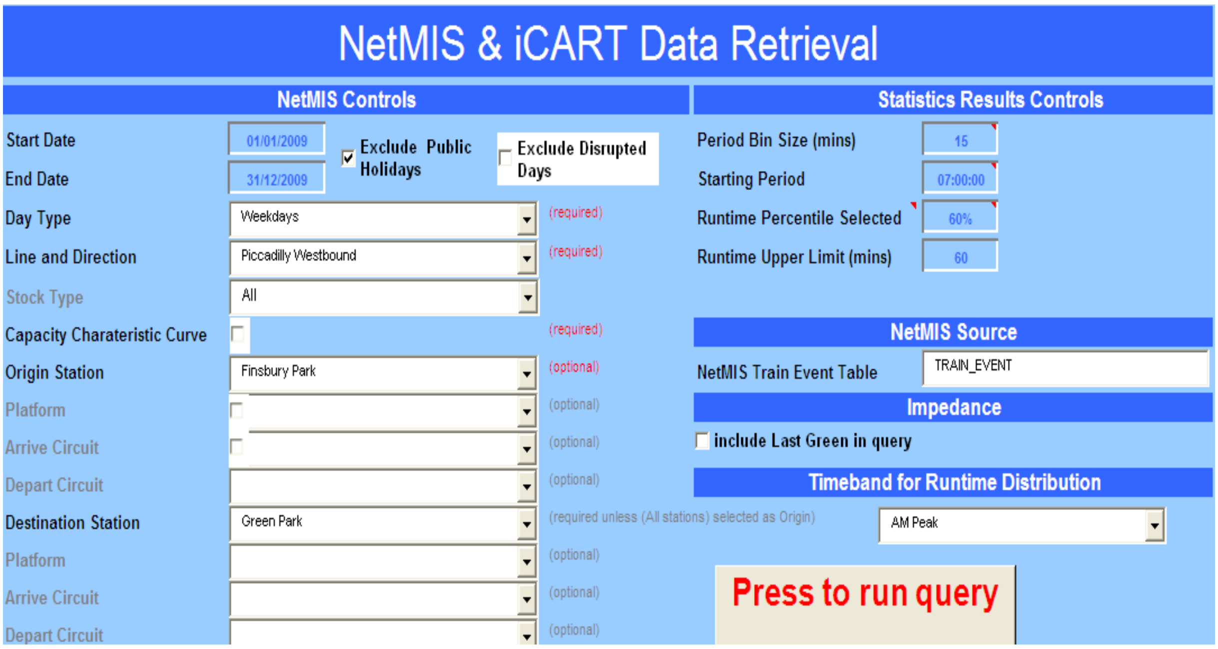



**Figure 7.2: The user page for the retrieval of NetMIS data**

We also retrieved other performance data saved in Journey Time Metric (JTM), such as on train time and platform waiting time. Crowding level is also indirectly assessed by the excess time for station walking and ticketing recorded in JTM .[47] JTM, however, only records the aggregate data, which is not specific to the time of day. Thus, as shown in Table 7.5, we only extract the performance data from JTM to illustrate the heterogeneity of London underground lines. Here, to have an overview of our study corridors, we take into account the aggregated information including mean travel time, standard deviation of travel time, headway, and distance between OD.

[47] In JTM the excess time is defined as the difference between ideal time and actual time.

| Scenarios | Lines | Mean travel time (min) | SD (min) | Headway (min) | Distance (miles) |
|---|---|---|---|---|---|
| KG_2008 | Piccadilly | 9.40 | 2.73 | 3.10 | 2.20 |
| | Victoria | 7.78 | 2.46 | 3.30 | 2.19 |
| KG_2010 | Piccadilly | 9.27 | 2.52 | 3.00 | 2.20 |
| | Victoria | 7.50 | 2.33 | 3.30 | 2.19 |
| FG | Piccadilly | 18.22 | 1.85 | 3.00 | 4.68 |
| | Victoria | 14.38 | 1.48 | 3.20 | 4.91 |
| FK | Piccadilly | 8.23 | 1.28 | 3.00 | 2.48 |
| | Victoria | 6.24 | 0.80 | 3.00 | 2.72 |
| WB_2009 | Bakerloo | 10.49 | 0.80 | 3.30 | 2.55 |
| | Jubilee | 8.90 | 0.94 | 3.50 | 3.41 |
| WB_2006 | Bakerloo | 10.80 | 1.58 | 3.30 | 2.55 |
| | Jubilee | 8.84 | 1.77 | 3.50 | 3.41 |

**Table 7.5: Performance statistics for choice scenarios (based on JTM)**

## 7.4 Model structures and specifications

The description of the model specification is split into two subsections. The first part analyses the explanatory variables used in our models, and their interactions with socio-demographic characteristics. We then look at the candidate modelling approaches presented in this research. Specifically, this current research established and compared the Expected Value Theory model (EVT), Expected Utility Theory model (EUT), Subjective Expected Value Theory model (SEV), Subjective Expected Utility Theory model (SEU), Rank-Dependence Expected Value model (RDEV), Rank-Dependence Expected Utility model (RDEU), Prospect Theory model (PT), and Cumulative Prospect Theory model (CPT).

### 7.4.1 Explanatory variables and interactions

In the exploration of the initial specification, a large number of variables were taken into account, e.g., standard deviation of travel time, mean variance, mean headway, standard deviation of headway, dwell time, crowding index, and distances, etc. Most of these turned out not to be statistically significant in model calibrations but the main explanatory variables influencing the sampled passengers' route choice were in-train journey time (hereafter termed as travel time) and headway.[48] Moreover, passengers' tastes regarding headway were found to interact with trip purpose. To address this interaction, we apply two different headway parameters to represent the segmentation in terms of work trips and non-work trips.

At the beginning of this exploration, we pay particular attention to measuring travel time distribution, accounting not only for the central tendency but also the dispersion. This is in line with the traditional method used in mean-variance studies. The estimated coefficient of standard deviation is not statistically significant in any of the candidate models (for details refer to Table 7.6), however. A large body of RP literature has concluded that the insignificant parameter of standard deviation is due to its correlation with travel time (Brownstone and Small, 2005). This may be true in this research as well, given that the observed standard deviation seems to be correlated with mean travel time (shown in Figure 7.3). Indeed, mean travel time has a similar pattern as standard deviation, while mean headway appears to be more independent.

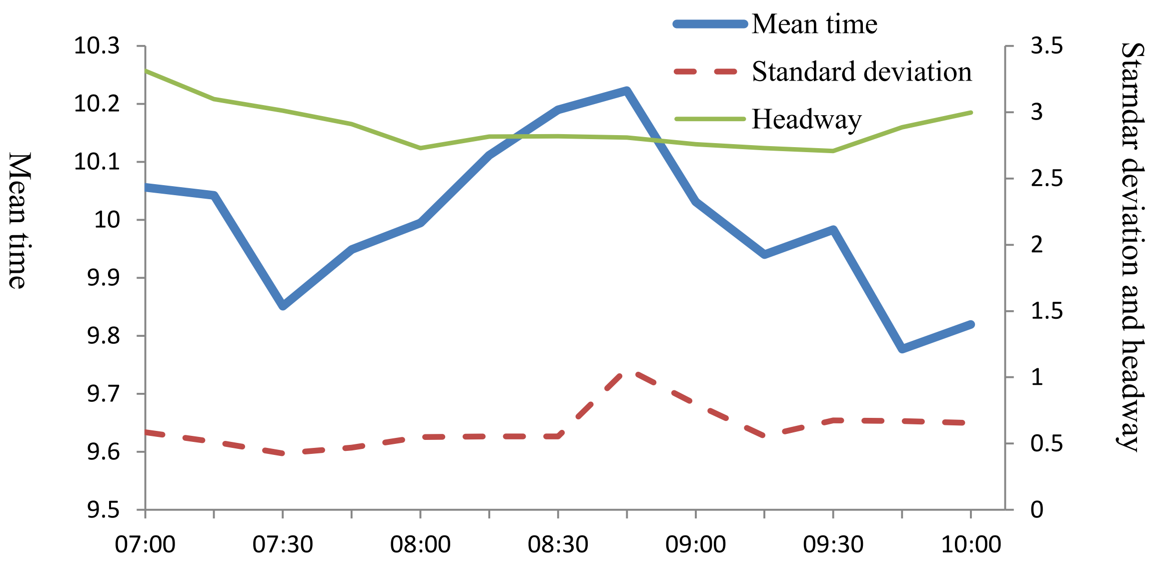



**Figure 7.3: Illustration of correlation among mean travel time, mean headway and standard deviation of travel time (the Bakerloo line)**

48 We cannot interpret insignificance since there is no explanatory power to LU route choice, given they are context dependent. For instance, access and egress time have been applied to the demand model in many studies, however, these attributes have little effect in that we have selected choice scenarios with identical access and egress distance and largely excluded their influence.

We have proposed a series of risky choice models which are capable of embodying passengers' attitude towards risk as well as perceptual conditioning. This is done by introducing extra parameters to be estimated for travel time utility and associated probability (e.g. SEU and RDEU), or even by dividing the travel time variable into two variables according to the reference point (e.g. PT and CPT). Based on our calibration results, we found that passengers' attitudes towards risk vary across populations with different travelling frequencies. We therefore creatively interact the risk attitude parameter of the travel time utility with a dummy variable reflecting passengers' travelling frequency.[49] Thus, two different parameters of risk attitude are applied to address the segmentation in terms of travelling frequency. Finally, it should be noted that travel cost information is deliberately omitted here, since LU ticket price is the same between the same zones. We therefore cannot analyse valuation of travel time savings in this case study.

### 7.4.2 Candidate model specifications

We now look at the specific modelling approaches proposed in this research. Risky outcomes correspond to the observed travel time extracted from the NetMIS database, while the time difference between two successive outcomes is 1 minute. We noticed that the observed travel time distribution varies across different times-of-day and underground lines, thus offering additional variation to the sample. More importantly, we can apply these disaggregated travel time outcomes to the risky choice framework.

Here, EVT and EUT are considered as the basic models. Both specifications are based on the assumption of instrumental rationality as discussed in Chapter 2. The simplest model of EVT maintains a linear functional form by computing explanatory variables weighted by taste parameters. If we incorporate the nonlinear utility function $g(TT_k^n, \alpha)$ into the EVT model, this is then transferred to the EUT model with an extra parameter $\alpha$ to embody risk attitude. Thus, the EVT and EUT model can be generally expressed as following:

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} p_k^n \, g(TT_k^n, \alpha) + \beta_{Hway_{work}}^n I(work)Hway^n + \beta_{Hway_{nonwork}}^n I(nonwork)Hway^n + \varepsilon^n \qquad (7.1)$$

[49] It is defined as 1 if the specific passenger use the same service five or more times per week, and otherwise it is defined as zero.

where $Hway^n$ represents the headway of the $n^{\text{th}}$ alternative, and the dummy variable $I(.)$ indicates whether the trip purpose is work or non-work. $\beta_{TT}^n$, $\beta_{Hway\_work}^n$ and $\beta_{Hway\_nonwork}^n$ are the parameters to be estimated. If $g(TT_k^n, \alpha)$ is linear, the model is referred to as EVT; if $g(TT_k^n, \alpha)$ is nonlinear, the corresponding model is EUT. We now look at the non-EUT models, starting with SEV and SEU. SEV applies a nonlinear distortion of probability rather than utility, while SEU jointly incorporates nonlinear utility and a probability weighting function. The SEV and SEU models are generally expressed as:

$$u_i^n = ASC + \beta_{TT}^n \textstyle\sum_{k=1}^{K} \pi(p_k^n, \gamma)\, g(TT_k^n, \alpha) + \beta_{Hway_{work}}^n I(work) Hway^n + \beta_{Hway_{nonwork}}^n I(nonwork) Hway^n + \varepsilon^n \tag{7.2}$$

where $\pi(p_k^n, \gamma)$ is the weighting function of probability, and the parameter $\gamma$ captures passengers' perceptual conditioning. Here, $g(TT_k^n, \alpha)$ also serves as the utility function and, if it is linear, the model is referred to as SEV; if it is nonlinear, the model is SEU. In this research, the only difference between SEU and RDEU is the use of rank-dependence. This is the idea that individuals evaluate risky alternatives not on the basis of probabilities but rather using (nonlinear) decision weights that reflect, *inter alia*, the preference ordering of potential outcomes. As a result, the decision weight of RDEU/RDEV is termed as $w(p_k^n, \gamma)$ in order to differentiate it from the decision weight $\pi(p_k^n, \gamma)$ of SEU/SEV.

$$u_i^n = ASC + \beta_{TT}^n \sum_{k=1}^{K} w(p_k^n, \gamma)\, g(TT_k^n, \alpha) + \beta_{Hway_{work}}^n I(work) Hway^n + \beta_{Hway_{nonwork}}^n I(nonwork) Hway^n + \varepsilon^n \tag{7.3}$$

In our PT model, travel time outcomes are divided into a group of gains and a group of losses. Consequently, we can apply the difference between travel time outcome and the reference point to represent the idea of reference dependence. Unlike the discontinuity method used in the SR91 case study, this current study employs the continuity method to embody passengers' diminishing sensitivity to travel time. The functional form of PT is expressed as following:

$$u_i^n = ASC + \beta_{TT(gain)}^n \sum_{k=1}^{K} p_k^n \max(0, t_{ref}^n - t_g^n)\left(TT_{ref}^n - TT_k^n\right)^{\alpha} + \beta_{TT(loss)}^n$$

$$\sum_{k=1}^{K} p_k^n \max(0, t_g^n - t_{ref}^n)(TT_k^n - TT_{ref}^n)^\alpha + \beta_{Hway_{work}}^n I(work)Hway^n$$

$$+\beta_{Hway_{nonwork}}^n I(nonwork)Hway^n + \varepsilon^n \quad (7.4)$$

where $\beta_{TT(gain)}^n$ is the parameter for the travel time $TT_k^n$ that is less than the reference point $TT_{ref}^n$, and $\beta_{TT(loss)}^n$ is the parameter for the travel time $TT_k^n$ that is more than the reference point $TT_{ref}^n$. The parameter $\alpha$ serves as the diminishing sensitivity parameter. If we incorporate the decision weight of RDEU into the PT model function, it is then transferred to the CPT model as follows:

$$u_i^n = ASC + \beta_{TT(gain)}^n \sum_{k=1}^{K} w(p_k^n, \gamma) \max(0, t_{ref}^n - t_g^n)\left(TT_{ref}^n - TT_k^n\right)^\alpha + \beta_{TT(loss)}^n$$

$$\sum_{k=1}^{K} w(p_k^n, \gamma) \max(0, t_g^n - t_{ref}^n)(TT_k^n - TT_{ref}^n)^\alpha + \beta_{Hway_{work}}^n I(work)Hway^n +$$

$$\beta_{Hway_{nonwork}}^n I(nonwork)Hway^n + \varepsilon^n \quad (7.5)$$

# 7.5 Model estimation results

This section discusses the findings of the modelling analysis, starting with the demonstration of estimation results for the EVT and EUT models. This is followed by the presentation of results for the proposed non-EUT models including SEV, SEU, RDEV, RDEU, PT and CPT. The section concludes with a comparison of calibration results across these candidate models. In particular, it aims to investigate whether alternative models can actually improve goodness-of-fit compared to EVT and EUT, and which specific method results in an improvement if the non-EUT model does provide a better fit to the data. All the models are coded in MATLAB which enables us to process calibration and the following validation.

## 7.5.1 Findings from the basic model

In this research, EVT and EUT models are considered to be the basic model, with their corresponding estimation results being shown in Table 7.6. For the purpose of comparison, we pay special attention to the goodness-of-fit and the explanatory effect of the estimates.

Both parameters of travel time and headway turned out to be negative, which was consistent with our expectations, and indicates that passengers have a more negative attitude to headway than to travel time. Moreover, we found out that $\beta_{Hway\_work} < \beta_{Hway\_nonwork}$ which indicates that passengers with a work purpose tend to over-weight the cost of headway delay. This is in line with intuition, since passengers with a work purpose may suffer a greater penalty for their late arrival. We also found out that passengers in this sample tend to put more cost on headway than travel time, given that $\frac{\beta_{Hway_{work}}}{\beta_{TT}} = 2.25$ and $\frac{\beta_{Hway_{nonwork}}}{\beta_{TT}} = 1.14$ as implied from EVT estimation result.

A lower estimate of headway suggests that LU passengers would rather stay in train where they can sit down, read a newspaper, or at least feel like they are on their way, rather than waiting. It should be noted that EUT gives a different estimate of $\beta_{TT}$ which is even more negative than $\beta_{Hway\_nonwork}$, although we cannot compare this with the travel time parameter estimated by EVT since travel time in EUT has been nonlinearly distorted by the CRRA function. This utility transformation embodies passengers' attitude towards risk by using an extra parameter $\alpha$ which nonlinearly transforms the curvature of utility. As a result, $\alpha > 0$ indicates risk proneness in terms of travel time risk.[50] It should be noted that travel time risk is different from money risk as presented in economics because travel time utility is usually negative but the monetary utility is positive. Moreover, risk proneness here is only specific to travel time, and it indicates the travel time utility is convex, which is consistent to the concept of diminishing marginal utility and diminishing sensitivity toward travel time. This suggests, for example, that drivers who are used to congestion and serious delay do not weight a 60 minute journey time as being twice as bad as a 30 minute journey time.

We also found the observed heterogeneity of risk attitude parameter $\alpha$. That is, the passengers who usually travel through the chosen underground line five or more times per week (frequent travellers) are less risk prone than the others, given that $\alpha_{infrequent} < \alpha_{frequent}$. It should be noted that the higher estimate of $\alpha$ implies a weaker utility distortion power from the CRRA perspective. Therefore, the higher estimate of $\alpha_{frequent}$means that frequent travellers are more objective to travel time.

[50] It should be noted that travel time risk is different from the money risk emphasized in the economic literature. A formal discussion on the treatment of travel time risk is available in Bates and Whelan (2001) and Batley and Ibáñez (2012).

| | EVT | | EVT_SD | | EUT | |
|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| ASC | -0.319 | -1.490 | -0.288 | -1.510 | -0.262 | -1.207 |
| $\beta_{TT}$ | -0.396 | -7.844 | -0.233 | -5.221 | -0.711 | -1.299 |
| $\beta_{Hway\_work}$ | -0.891 | -2.729 | -0.773 | -3.232 | -0.794 | -2.378 |
| $\beta_{Hway\_nonwork}$ | -0.560 | -1.929 | -0.461 | -1.989 | -0.476 | -1.678 |
| SD | | | -0.051 | -0.141 | | |
| $\alpha_{frequent}$ | | | | | 0.339 | 0.965 |
| $\alpha_{infrequent}$ | | | | | 0.178 | 1.848 |
| Final LL (β) | -301.04 | | -300.051 | | -298.705 | |
| $\rho^2(0)$ | 0.126 | | 0.129 | | 0.133 | |
| Adj. $\rho^2(0)$ | 0.115 | | 0.114 | | 0.115 | |
| $\rho^2(ASC)$ | 0.026 | | 0.030 | | 0.034 | |
| Adj. $\rho^2(ASC)$ | 0.014 | | 0.013 | | 0.015 | |

**Table 7.6: EVT and EUT estimation results for LU data**

## 7.5.2 SEV and SEU models

In this section, we look at the explanatory power of the model specification using subjective probability rather than objective probability. A nonlinear probability weighting function is applied in order to account for passengers' perceptual conditioning of occurrence probabilities attached to travel time outcomes.[51] This nonlinear weighting function is capable of explaining several violations of EUT, e.g. the Allais paradox (Allais, 1953), meaning that SEV and SEU are more appropriate for behavioural realism. It is still necessary to reveal whether nonlinear probability weighting functions can actually improve model fit in a real choice context, however, and which weighting function performs better. Hence, the following discussion will primarily focus on the findings in terms of weighting functions and the explanatory power of their parameters.

[51] The subjective probability used in this current research is, strictly speaking, different from the so-called subjective probability obtained by asking respondents during surveys. The latter usually suffers serious problems with data validity due to respondents' perception biases.

### *7.5.2.1 SEV model*

The results for the SEV models, which consist of five candidate weighting functions, are shown in Table 7.7. Although nonlinear utility is not applied to SEV models, all of these models deliver improvements in LL over their counterpart EVT model by 4.73, 4.72, 4.62, 4.27 and 4.62 units respectively. Note that since the only difference between SEV and EVT is the interpretation of likelihood, we can conclude that it is the subjective probability that leads to the improvement in LL.

We now look at the explanatory power of different weighting functions. There are merely marginal changes in LL across different SEV models, with the biggest difference being of 0.461 unit. This finding suggests that the selection of weighting function may not have a significant influence on final LL. Specifically, SEV_GE provides the best fit in terms of LL, at the cost of two additional parameters for the GE weighting function. Out of all the SEV models with a single weighting parameter SEV_TK gives the biggest LL, which is even better than SEV_WG and SEV_Pr2 with double weighting parameters. Given that a two-parameter weighing function is more flexible they may be behaviourally better than their counterpart single-parameter weighting function.[52] Our empirical findings, however, suggest that complicated weighting functions do not necessarily deliver a better model fit.

Once again, the parameters for travel time and headway consistently remain negative, which shows the disutility associated with the costs of these attributes, and the ratio between $\beta_{Hway\_work}$ and $\beta_{TT}$ is approximately 1.35-1.76, which is similar to the ratio estimated from EVT. It should be noted that the weighting parameter $\gamma$ determines the curvature of the weighting function, while $\tau$ determines the elevation in two-parameter functions. In this case, we observed an S-shaped weighting function in all the SEV models. Beginning with SEV_TK, $\gamma$ is statistically significantly different from 1, with a t-ratio of 2.6. Likely, we also observed statistically significant $\gamma$ for all the other models except SEV_GE and SEV_Pr2. In terms of estimated $\tau$, it is significant but only at the 10% level (with a t-ratio of 1.8 for SEV_GE), suggesting some evidence of an elevation of the weighting function. SEV_WG is often referred to as a general form of SEV_TK by substituting $\tau$ for $1/\gamma$, whilst $\tau$ turns out to be not significant in the estimation results. Moreover, an interesting finding is that the estimated value of $\tau$ is similar to the value of $1/\gamma$, and the LL for both models is also identical, suggesting that it is not necessary to adopt SEV_WG compared with SEV_TK, despite its flexibility.

---

[52] Gonzalez and Wu (1999) claimed that curvature and elevation are logically independent, and both factors should be jointly considered by a two-parameter weighting function.

| | SEV_TK | | SEV_WG | | SEV_GE | | SEV_Pr1 | | SEV_Pr2 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| ASC | -0.141 | -0.612 | -0.142 | -0.606 | -0.116 | -0.489 | -0.237 | -1.087 | -0.160 | -0.680 |
| $\beta_{TT}$ | -0.567 | -5.711 | -0.566 | -3.891 | -0.593 | -4.599 | -0.476 | -7.230 | -0.578 | -4.029 |
| $\beta_{Hway\_work}$ | -0.772 | -2.335 | -0.772 | -2.334 | -0.762 | -2.301 | -0.825 | -2.512 | -0.784 | -2.366 |
| $\beta_{Hway\_nonwork}$ | -0.504 | -1.710 | -0.504 | -1.707 | -0.495 | -1.696 | -0.540 | -1.791 | -0.509 | -1.726 |
| γ | 1.212 | 14.663 | 1.211 | 14.301 | 0.905 | 4.883 | 1.161 | 15.111 | 1.129 | 7.779 |
| τ | | | 0.838 | 0.758 | 1.178 | 11.909 | | | 1.073 | 8.971 |
| Final LL (β) | -296.310 | | -296.311 | | -296.211 | | -296.772 | | -296.43 | |
| $\rho^2(0)$ | 0.140 | | 0.140 | | 0.140 | | 0.139 | | 0.140 | |
| Adj. $\rho^2(0)$ | 0.125 | | 0.122 | | 0.123 | | 0.124 | | 0.122 | |
| $\rho^2(ASC)$ | 0.042 | | 0.042 | | 0.042 | | 0.040 | | 0.041 | |
| Adj. $\rho^2(ASC)$ | 0.026 | | 0.022 | | 0.023 | | 0.024 | | 0.022 | |

**Table 7.7: SEV estimation results using five weighting functions**

### *7.5.2.2 SEU models*

It is sensible jointly to take into account individuals' perception biases on travel time outcome and its associated probability. The natural way is to combine a subjective probability weighting function with a nonlinear utility function in a SEU specification. By incorporating a utility component into SEV, our SEU models enable us simultaneously to analyse attitude towards risk and subjective probability in a single model. The interesting results from this estimation are shown in Table 7.8.

The common observation across models is that the structural influence of weighting functions on the final LL of SEU is quite limited, while all SEU models improve LL compared with their counterpart SEV models, with an average increase of 1 unit. Specifically, all the SEU models give a similar LL at a level of approximately -295.5, which is 3.21 units more than the LL of EUT. Those models with a single-parameter weighting function, especially SEU_TK, again provide a better model fit if we take the parameter number into account, with an $\chi^2$ $p$-value of 0.01 compared with EUT.

| | SEU_TK | | SEU_WG | | SEU_GE | | SEU_Pr1 | | SEU_Pr2 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. | est. | t-stat. |
| ASC | -0.127 | -0.557 | -0.120 | -0.522 | -0.103 | -0.440 | -0.198 | -0.902 | -0.134 | -0.575 |
| $\beta_{TT}$ | -0.977 | -1.273 | -1.030 | -1.205 | -1.018 | -1.240 | -0.918 | -1.268 | -1.088 | -1.278 |
| $\beta_{Hway\_work}$ | -0.689 | -2.026 | -0.686 | -2.016 | -0.681 | -2.003 | -0.728 | -2.149 | -0.693 | -2.035 |
| $\beta_{Hway\_nonwork}$ | -0.429 | -1.467 | -0.426 | -1.465 | -0.422 | -1.461 | -0.455 | -1.525 | -0.430 | -1.474 |
| γ | 1.144 | 12.023 | 1.144 | 12.133 | 0.924 | 5.899 | 1.111 | 15.300 | 1.088 | 9.586 |
| τ | | | 0.666 | 0.564 | 1.119 | 11.175 | | | 1.047 | 11.222 |
| $\alpha_{frequent}$ | 0.329 | 0.858 | 0.341 | 0.882 | 0.324 | 0.834 | 0.369 | 0.990 | 0.369 | 0.989 |
| $\alpha_{infrequent}$ | 0.210 | 1.595 | 0.222 | 1.620 | 0.207 | 1.578 | 0.237 | 1.694 | 0.246 | 1.714 |
| Final LL (β) | -295.458 | | -295.444 | | -295.374 | | -295.764 | | -295.463 | |
| $\rho^2(0)$ | 0.142 | | 0.142 | | 0.143 | | 0.141 | | 0.142 | |
| Adj. $\rho^2(0)$ | 0.122 | | 0.119 | | 0.119 | | 0.121 | | 0.119 | |
| $\rho^2(ASC)$ | 0.045 | | 0.045 | | 0.045 | | 0.044 | | 0.044 | |
| Adj. $\rho^2(ASC)$ | 0.022 | | 0.019 | | 0.019 | | 0.021 | | 0.019 | |

**Table 7.8: SEU estimation results using five weighting functions**

A consistently lower value of $\alpha_{infrequent}$ is obtained for all SEU models, showing a higher utility distortion capability. From a behaviourial perspective, this implies that passengers who frequently travel through their chosen underground lines are less risk prone than the other passengers. This is in line with the findings from EUT. It should be noted, however, that all the estimates of $\alpha_{frequent}$ are insignificant, although this segmentation does improve model fit. One possible reason for the undesirable $\alpha_{frequent}$ is that passengers with high travelling frequency are more likely aware of the true travel time distribution, and thus are capable of making decisions which are hardly affected by risk attitudes and perceptual errors.

SEU models seem to have less power to distort probability compared with SEV, since the estimates of weighting parameter γ are consistently closed to one. This suggests that the curvature of the weighting function is closer to the line of objective probability, although we can still observe the deviation between objective probability and subjective probability when we look at the aggregate travel time distribution (shown in Figure 7.4). Evidently, subjective probability tends to be under-weighted compared to objective probability[53]. This shift to some extent reflects the fact that the travel time probability we observed actually varies compared to the probability perceived by passengers.

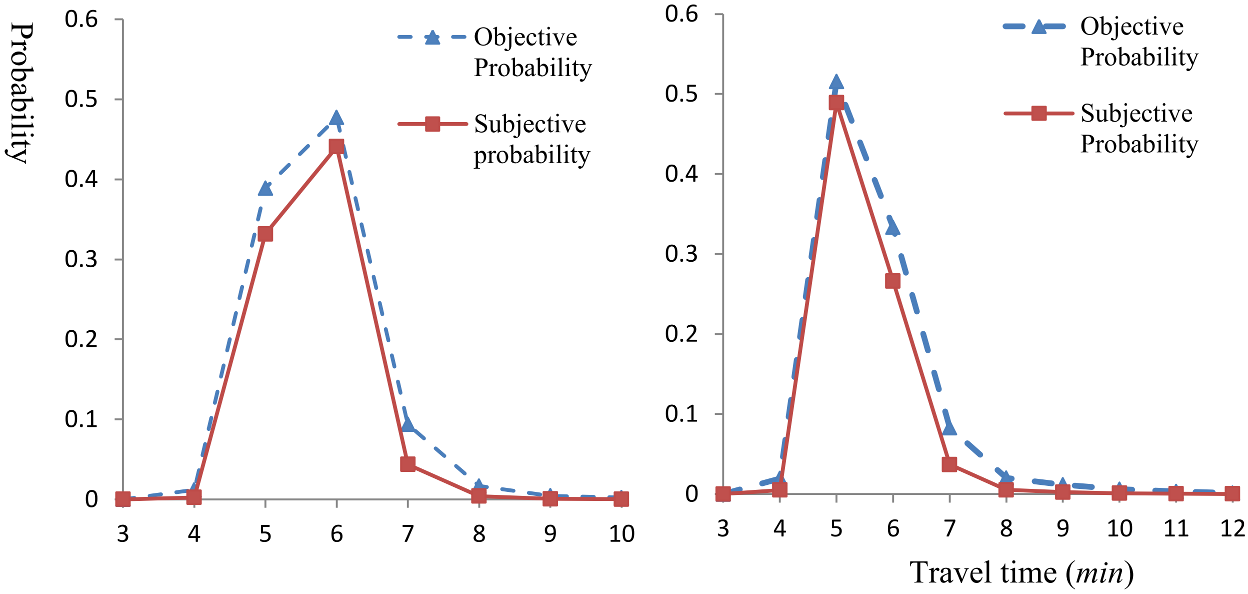



**Figure 7.4: Deviation of objective probability and subjective probability for the Piccadilly line (left) and the Victoria line (right) in the KG subsample**

### 7.5.3 RDEV and RDEU models

On the basis of SEV and SEU, we now consider the situation where rank dependence is also included into the model. The corresponding estimation results for RDEV and RDEU are shown in Table 7.9. Initially, we explored a set of probability weighting functions for RDEV, and found out that the TK function outperformed the other candidates. Hence, the estimation results illustrated in this subsection are based purely on the RDEV_TK and RDEU_TK models. Particular attention will be paid to the influence of rank dependence on model fit and induced estimates.

[53] It should be noted that the subjective probability used in this research is not necessarily summed up to 1.

| | RDEV | | RDEU | |
|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. |
| ASC | -0.316 | -1.467 | -0.256 | -1.168 |
| $\beta_{TT}$ | -0.396 | -7.838 | -1.197 | -1.829 |
| $\beta_{Hway\_work}$ | -0.900 | -2.753 | -0.788 | -2.330 |
| $\beta_{Hway\_nonwork}$ | -0.565 | -1.941 | -0.477 | -1.674 |
| γ | 0.812 | 2.718 | 0.875 | 2.705 |
| $\alpha_{frequent}$ | | | 0.574 | 0.669 |
| $\alpha_{infrequent}$ | | | 0.394 | 1.956 |
| Final LL (β) | -298.547 | | -296.824 | |
| $\rho^2(0)$ | 0.133 | | 0.138 | |
| Adj. $\rho^2(0)$ | 0.119 | | 0.118 | |
| $\rho^2(ASC)$ | 0.035 | | 0.040 | |
| Adj. $\rho^2(ASC)$ | 0.018 | | 0.017 | |

**Table 7.9: Estimation results for RDEV and RDEU**

Despite the potential behavioural appeal of rank dependence, we surprisingly found that RDEV and RDEU failed to outperform SEV and SEU respectively. Given that we differentiate the RDEV model from the SEV model by ordering the ranks of outcomes, the less good model fit for the RDEV model raises a question as to whether rank dependence actually exists in the RP context. This finding is opposite to the general conclusions from SP studies where RDEV or RDEU normally serve as preferred models with better goodness-of-fit (Hensher and Li, 2012, Koster and Verhoef, 2010, Razo and Gao, 2011). In SP surveys, risky outcomes are normally presented to respondents with a specific order in questionnaires, which is relatively easy for them to rank these outcomes in mind. RP studies, without a clear presentation of risky outcomes, can only rely on setting up assumptions on the ranking orders. Consequently, it is too arbitrary to conclude that rank dependence is not applicable to RP before we properly understand whether passengers account for ranks of travel time outcomes. And if they do rank travel time outcomes, the next question is how they rank these outcomes in reality. In this research, we simply assume that the less travel time the better the outcome, whilst the real decision making procedure seems to be more complex, and more research needs to be done in transportation like psychology and behavioural science.

The curvature of the weighting function, with the estimate showing γ equalling 0.812 in RDEV and 0.875 in RDEU, exhibits an inverse S-shape. This is different from the shape observed from SEV and SEU in which γ is larger than one and the weighting function is S-shaped. As a result, we observed a different probability transformation for RDEV and RDEU. To illustrate the observed difference straightforwardly, the decision weight estimated from RDEV is shown in Figure 7.5 (for comparability, this is also based on the KG subsample).[54]

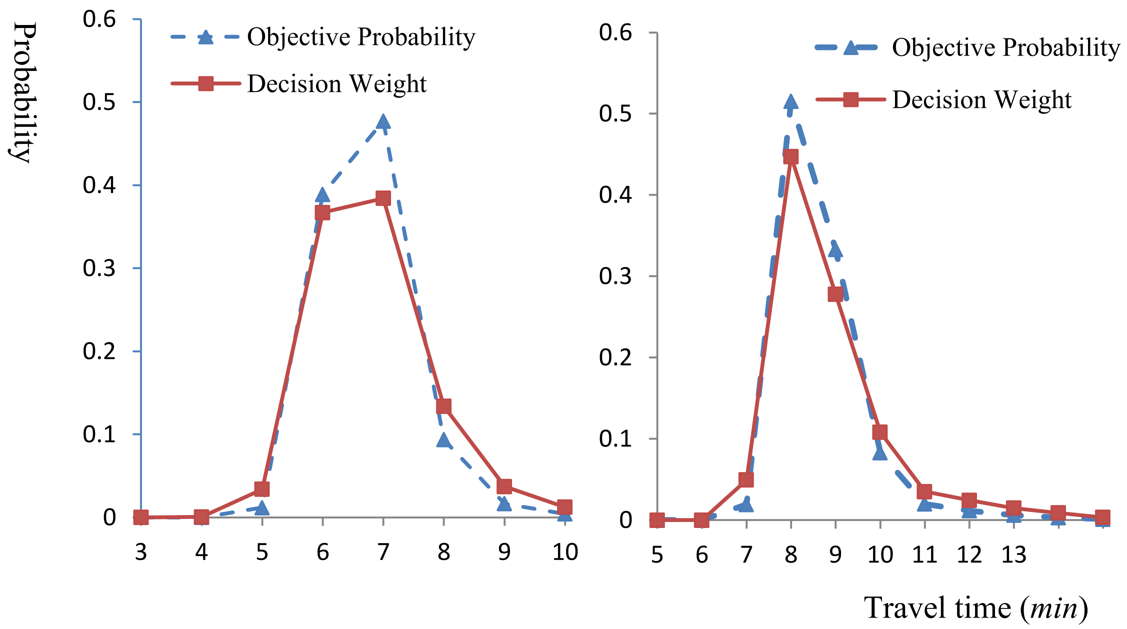



**Figure 7.5: Deviation of objective probability and decision weight for the Piccadilly line (left) and the Victoria line (right) in the KG subsample**

In this research, risky outcomes are ranked in an increasing order in terms of travel time. All the decision weights are calculated as the difference of cumulative probability transformed by the probability weighing function. Only the decision weight of the best outcome (i.e. the outcome with the least travel time) is simply the same as the weighting function value. Hence, a convex and concave weighting function can lead to under-weighting or over-weighting of the best outcome, suggesting individuals' attitude towards risk. The curvature observed in this research is mixed (inverse S-shaped), but we can still find out clues on risk attitude from Figure 7.4. This shows that the mean value of travel time is significantly under-weighted, while the right tail of distribution turns out to be slightly over-weighted. This suggests that passengers tend to under-weight the normal situation subjectively, whilst over-weighting the

[54] It should be noted that decision weight in RDEV and RDEU is the counterpart of subjective probability in SEV and SEU for the purpose of distinction.

extremely bad situation, which implies a slight pessimism and risk aversion from a decision weight perspective.

### 7.5.4 PT and CPT models

We turn our attention next to the estimation of PT and CPT models, with both these models highlighting the importance of reference dependence, and the latter also adopting a nonlinear decision weight. In the present context, we explore a set of possible reference points in terms of travel time, and finally select mean travel time for the current PT analysis since it gives the best model fit.[55] This means that we no longer simply use a single variable for travel time but that the travel time variable is divided into gain and loss according to its relative difference to the reference point. Moreover, passengers' sensitivity to travel time is expected to be diminished, which is expressed by an extra parameter $\alpha$. Consequently, travel time utility is given by:

$$u(time) = \begin{cases} \beta_{TT\_gain}\left(tt_{ref} - tt\right)^{\alpha} & if\ tt < tt_{ref} \\ \beta_{TT\_loss}\left(tt - tt_{ref}\right)^{\alpha} & if\ tt \geq tt_{ref} \end{cases} \tag{7.6}$$

where parameters $\beta_{TT\_gain}$ and $\beta_{TT\_loss}$ characterize the travel time, interpreted as gain and loss respectively. And the parameter $\alpha$ embodies the diminishing sensitivity of travel time. It should be noted that $\alpha$ varies across the population ($\alpha_{frequent}$ and $\alpha_{infrequent}$), but is constant between loss and gain. According to the calibration results, the weighting parameter $\gamma$ of CPT is also found to be identical between loss and gain.[56] The estimation results of the PT and CPT models are summarized in Table 7.10.

| | PT | | CPT | |
|---|---|---|---|---|
| | est. | t-stat. | est. | t-stat. |
| ASC | -0.249 | -1.225 | -0.381 | -1.625 |
| $\beta_{TT\_gain}$ | 3.222 | 1.939 | 4.900 | 2.057 |
| $\beta_{TT\_loss}$ | -4.556 | -4.164 | -5.541 | -4.716 |

[55] We also managed to estimate the reference point endogenously as we have done in Chapter 5, however, we cannot obtain a well-determined estimate of reference point as the statndard deviation is large. Given that there are four different choice scenarios, it is impossible to identify a uniform reference point for all respondents. Further research should take into account this heterogeneity for the endogenous estimation of reference points.
[56] The original version of CPT allowed different $\alpha$ and $\gamma$ for loss and gain. This is more flexible, but our estimation results show that such segmentation cannot improve model fit.

| | | | | |
|---|---|---|---|---|
| $\beta_{Hway\_work}$ | -0.771 | -2.326 | -0.630 | -1.834 |
| $\beta_{Hway\_nonwork}$ | -0.444 | -1.621 | -0.393 | -1.502 |
| γ | | | 0.859 | 11.201 |
| $\alpha_{frequent}$ | 0.687 | 5.065 | 0.687 | 7.342 |
| $\alpha_{infrequent}$ | 0.836 | 5.635 | 0.888 | 7.045 |
| Final LL (β) | -294.209 | | -293.041 | |
| $\rho^2(0)$ | 0.146 | | 0.149 | |
| Adj. $\rho^2(0)$ | 0.126 | | 0.126 | |
| $\rho^2(ASC)$ | 0.049 | | 0.052 | |
| Adj. $\rho^2(ASC)$ | 0.026 | | 0.026 | |

**Table 7.10: PT and CPT estimation results for LU data**

It is of interest to look first at the difference observed from $\beta_{TT\_gain}$ and $\beta_{TT\_loss}$. Consistent with our expectations, the observation is the correct sign for both parameters. Positive $\beta_{TT\_gain}$ shows that passengers attach positive attitudes to a travel time which is less than their reference travel time, while a negative $\beta_{TT\_loss}$ means that passengers attach negative attitudes to a travel time which is greater than their reference travel time. The second notable observation is the asymmetrical weights to gain and loss, i.e. $|\beta_{TT\_loss}| > |\beta_{TT\_gain}|$, suggesting passengers' behaviour relating to loss aversion. As shown in Figure 7.5, travel time utility is clearly kinked at the reference point. Evidently, the slope of loss is steeper, and looms larger than gain. Specifically, the ratio $|\beta_{TT\_loss}|/|\beta_{TT\_gain}|$ equals 1.414 for PT but only 1.13 for CPT, which means that loss aversion is more significant for PT. Although the ratio of PT is relatively high, it is still much smaller than in the other PT literature using SP data, such as Hess et al. (2008)'s 3.15 and Gao et al. (2010)'s 2.09. It is still unknown why RP data produces a lower ratio between loss and gain, but this current study at least contributes evidence as to what the level of loss aversion is in reality.

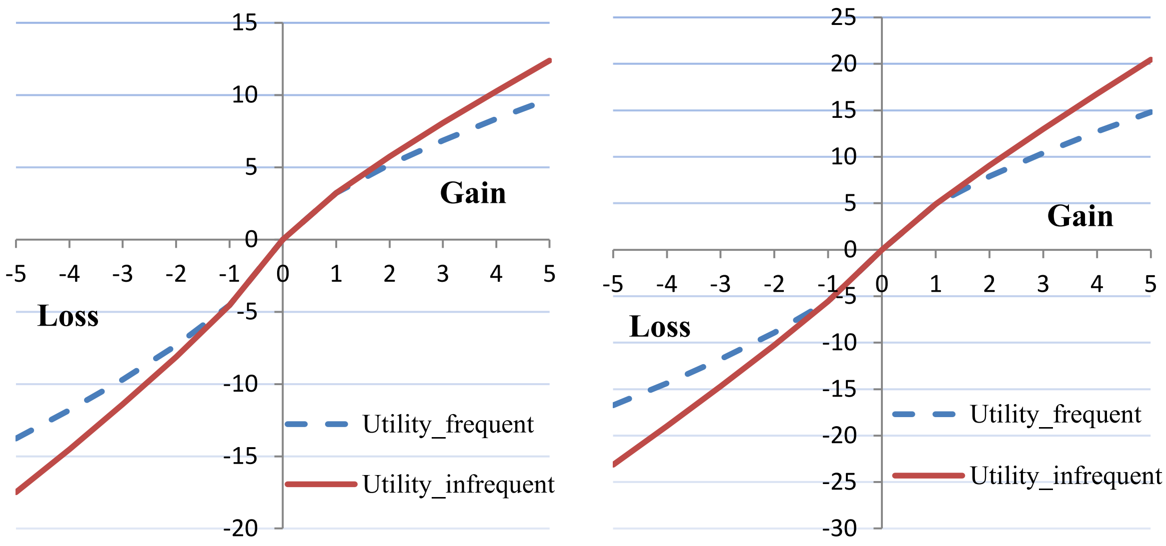



**Figure 7.6: Illustration of loss aversion and diminishing sensitivity for PT (left) and CPT (right)**

The last finding, from a utility perspective, is the concavity of gain and convexity of loss due to $\alpha < 1$. Such nonlinearity of utility is usually interpreted as diminishing sensitivity, i.e. the psychological impact of the marginal disutility of travel time decreases when actual travel time is increasingly further from reference point. In terms of attitude towards risk, the convexity of utility can be interpreted as risk proneness, while concavity implies risk aversion. In this case, LU passengers have different risk attitudes towards travel time outcomes, with risk aversion in gain and risk proneness in loss.

A preliminary analysis showed that the selection of weighting functional forms has little influence on the final model fit of CPT. For the current research relating to CPT, we apply a TK weighting function, as used in RDEU, since this best fits the data. As a result, we found that the estimate gave $\gamma = 0.859$, which is similar to the $\gamma$ estimated by RDEU. This leads to an inverse S-shaped weighting function, as shown in Figure 7.6. Due to the distortion of the nonlinear weighting function, a probability of less than 0.38 is over-weighted, while a probability of more than 0.38 is under-weighted.

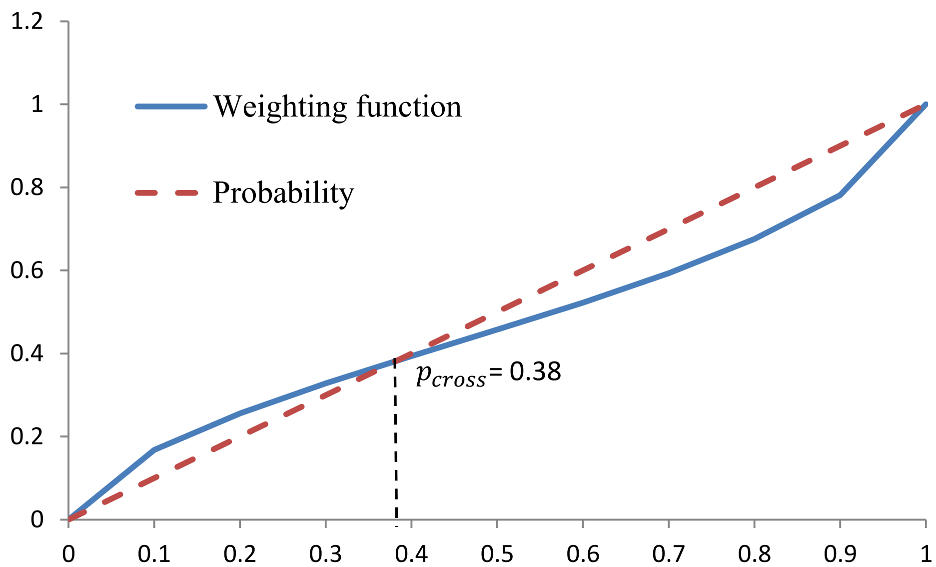



**Figure 7.7: Weighting function for CPT**

In terms of model performance, it is clear that both PT and CPT give a better fit to the data, with 4.496 and 5.664 units improvement of LL compared to the EUT model. This is also supported by the likelihood ratio (LR) test, with $\chi^2$ p-values of less than 0.005. It should be noted, however, that PT and CPT belong to a non-nested model which is not a generalization of an EUT functional form; consequently, we cannot compare PT and CPT simply using the method for nested models. The next subsection will discuss model comparison in more detail.

### 7.5.5 Model comparison

The final stage of the analysis, before model validation, is to compare substantive estimation results across models. It is of interest to reveal whether the non-EUT models proposed in this research actually outperform EVT and EUT models, and which modelling techniques lead to an improvement in model fit. The answer to these questions relies on a series of formal statistical tests to determine the structure that best fits the data. It is essential to realize that these candidate models have different parameters and even structures. Specifically, PT and CPT cannot be treated as the parametrical generalization of any other models. Thus, we applied both nested and non-nested tests to assess their empirical performances.

Table 7.11 shows the measures of fit in the estimation sample. The adjusted likelihood ratio index favours PT and CPT models, with the highest value $\rho^2(0) = 0.126$ and $\rho^2(ASC) = 0.026$. Only the SEV model achieves the same level of $\rho^2(ASC)$ as PT and CPT, while CPT appears slightly to outperform PT from an AIC point of view, with an improvement of just 0.336 unit. This finding highlights the importance of reference dependence and diminishing sensitivity for modelling route choice behaviour, but does not necessarily support the complexity of CPT.

| | **EVT** | **EUT** | **SEV** | **SEU** | **RDEV** | **RDEU** | **PT** | **CPT** |
|---|---|---|---|---|---|---|---|---|
| Parameter | 4 | 6 | 5 | 7 | 5 | 7 | 7 | 8 |
| LL(β) | -301.040 | -298.705 | -296.311 | -295.458 | -298.547 | -296.824 | -294.209 | **-293.041** |
| $\bar{\rho}^2(0)$ | 0.115 | 0.115 | 0.125 | 0.122 | 0.119 | 0.118 | **0.126** | **0.126** |
| $\bar{\rho}^2(ASC)$ | 0.014 | 0.015 | **0.026** | 0.022 | 0.018 | 0.017 | **0.026** | **0.026** |
| AIC | 610.080 | 609.410 | 602.622 | 604.916 | 607.094 | 607.648 | 602.418 | **602.082** |
| BIC | 626.409 | 633.903 | **623.033** | 633.492 | 627.505 | 636.224 | 630.994 | 634.740 |
| ConAIC | 630.409 | 639.903 | **628.033** | 640.492 | 632.505 | 643.224 | 637.994 | 642.740 |
| CorAIC | 610.636 | 610.973 | **603.596** | 607.266 | 608.068 | 609.998 | 604.768 | 605.446 |
| LR_EVT | | 4.670 | 9.458 | 11.164 | 4.986 | 8.432 | NA | NA |
| LR_EUT | | | 4.788 | 6.494 | 0.316 | 3.762 | NA | NA |

LL(β): The final log-likelihood based on the calibration sample;

ConAIC: Consistent AIC;

CorAIC: Corrected AIC;

LR_EVT: Likelihood ratio w.r.t EVT;

LR_EUT: Likelihood ratio w.r.t EUT;

NA: Not applicable.

**Table 7.11: Measures of goodness-of-fit**

When the other measures of fit are used, we surprisingly obtain a rather different conclusion. That is, the SEV model turns out to be consistently favoured by BIC, consistent AIC and corrected AIC. Given the relatively simple model structure of SEV, we can conclude that its good performance is due to the nonlinear probability weighting function. This finding suggests that ignoring the weighting function may lead to an incorrect model for analysing risky choice behaviour, and it is therefore essential to account for subjective probability in future research. SEU provides an even better LL than SEV, but it fails to compete with SEV if the number of parameters is taken into account.

We cannot find evidence in favour of rank dependence using any of the measures of fit. In fact, although RDEV and RDEU outperform EVT and EUT, they provide worse model fit compared to SEV and SEU. Given that RDEV and RDEU incorporate rank dependence on the basis of SEV and SEU, their relatively unexpected performance raises a question as to whether rank dependence actually matters in reality. If it is not a factor of concerned to passengers, future research should turn more attention to the other models like PT and SEV. If it does matter, we should concentrate on the techniques for determining the real rank orders of outcomes perceived by passengers in an RP context.

To test if the model fits between EUT and non-EUT are significantly different, we use both nested and non-nested test. The latter applies to PT and CPT, while the former employs an LR test to assess the other nested models. The LR statistics for EVT, EUT, SEV, SEU, RDEV and RDEU are set out at the bottom of Table 7.11. In terms of nested models, SEV significantly improves model fit compared to EVT, with a $\chi^2$ p-value of less than 0.005. Although nonlinear utility is not included in SEV, it still significantly outperforms the EUT model, which again reinforces the benefit of using the nonlinear probability weighting function. SEU also gives a statistically significant improvement in terms of LR statistics, with a $\chi^2$ p-value of 0.01. RDEV provides similar LL as EUT, whilst RDEV is still preferred, given it has one less parameter than EUT. Although the LL of RDEU is almost 2 units more than EUT, it only delivers a $\chi^2$ p-value of 0.05 compared to EUT.

We also applied a non-nested test to compare PT and CPT with all the other models (refer to Chapter 5 for the non-nested test method), with the results shown in Table 7.12.

| **PT** | **Test statistics** | ***P*-values** |
|---|---|---|
| *vs* EVT | -3.265 | 0.001 |
| *vs* EUT | -2.827 | 0.002 |
| *vs* SEV | -1.485 | 0.069 |
| *vs* SEU | -1.581 | 0.057 |
| *vs* RDEV | -2.584 | 0.005 |
| *vs* RDEU | -2.287 | 0.012 |

| **CPT** | **Test statistics** | ***P*-values** |
|---|---|---|
| *vs* EVT | -3.464 | 0.000 |
| *vs* EUT | -3.054 | 0.001 |
| *vs* SEV | -1.881 | 0.030 |
| *vs* SEU | -1.958 | 0.026 |
| *vs* RDEV | -2.831 | 0.002 |
| *vs* RDEU | -2.562 | 0.005 |
| *vs* PT | -1.156 | 0.124 |

**Table 7.12: Non-nested test results for PT and CPT**

It should be noted that the *p*-value corresponds to the upper bound of probability, although the other models could provide higher adjusted likelihood ratios than PT or CPT models by chance. The results clearly favour the PT and CPT models, which outperform most models with a very low p-value, except the SEV and SEU models. Specifically, both PT and CPT give a statistically significantly better model fit compared to EVT and EUT models, with a p-value of only 0 to 0.002. These are also the models which merit more consideration compared to the RDEV and RDEU models for this data, given that the highest probability is only 0.012 for PT versus RDEU. We cannot be entirely sure, however, whether the PT and CPT models provide better performances than the SEV and SEU models, in particular PT versus SEV, where the probability is approximately 0.07.

To conclude, the comparison results based on the calibration sample offer the same evidence that non-EUT models are preferred to EVT and EUT models. CPT provides the highest LL, while, according to the non-nested test results, the statistical benefit of data fit does not seem to overcome the penalty of having more parameters. SEV turns out to be an efficient model specification with a fair model fit and a relatively simple model structure. These findings highlight the particular importance of three modelling techniques, namely nonlinear utility (and diminishing sensitivity), nonlinear probability weighting function, and reference dependence.

## 7.6 Model validation

The final part of the analysis is concerned with model validation. This is done by applying all the candidate models with the estimated parameters into the validation sample. A great deal of literature also provides more complicated methods, such as cross validation and bootstrap (Breiman and Spector, 1992, Kohavi, 1995, Wassenaar et al., 2005). In this research, model estimation is based on the sample randomly drawn from 75% of the total sample, with the remainder of the sample constituting the hold-out sample. It should be noted that, although all the non-EUT models turn out to outperform the EVT and EUT models in terms of data fit, the difference between the non-EUT models themselves is very modest. Moreover, it is still uncertain whether these alternative models also perform well in prediction. To address this issue, both aggregate measures and disaggregate measures are employed to compare their predictive performances.

For the disaggregate test method, we adopt two measures. The first is a predictive adjusted likelihood ratio index, which is based on the trade-off between predicted likelihood and number of parameters. This test describes the model fit to the validation data, but it cannot provide the probability that each model correctly predicts respondents' chosen alternatives. This, however, can be done by the second method, i.e. average probability of correct prediction (APCP). Here, we first calculate the probability of a specific decision maker's chosen alternative, and then average all the observations in the validation sample to obtain the final index of average probability of correct prediction. This method is expressed by the following function:

$$APCP = \frac{\sum_i \sum_k \delta_{ik} p_{ik}}{N} \tag{7.7}$$

where N corresponds to the number of total observations, $\delta_{ik}$ is a dummy variable indicating whether the individual $i$ actually chooses the alternative $k$, and $p_{ik}$ is the predicted probability of individual $i$ choosing the alternative $k$. At the aggregate level, we employ the root mean square error (RMSE) and the mean absolute percentage error (MAPE) test.

The results of model validation are summarized in Table 7.13. Considering first the results of the disaggregate measure, a general observation is that the rank ordering of the predicted LL is not consistent with the estimated LL. In the validation results, SEU gives the best performance in terms of LL, but PT and CPT fail to fit the validation data well. This is also supported by the predicted likelihood ratio test, with a highest value of 0.097 from SEU. If we take number of parameters into account, however, the adjusted predicted likelihood ratio index favours SEV instead, with a relatively high p-value of 0.04. In terms of average probability of correct prediction, it is SEU that provides the most accurate prediction of respondents' preferences. If we consider whether the superiority of SEV and SEU is also found in aggregate measures of model predictions it is found that both RMSE and MAPE demonstrate the index in favour of SEV, although the differences between SEV, SEU and CPT are quite trivial.

To conclude, all the non-EUT models provide better predictive performances than EUV and EUT, from both an aggregate and a disaggregate perspective. This finding is in line with the calibration results. We also found that there are slight deviations between the aggregate and disaggregate tests: specifically, APCP appears to favour SEU, while the RMSE and MAPE indexes support the superiority of SEV.

| | EVT | EUT | SEV | SEU | RDEV | RDEU | PT | CPT |
|---|---|---|---|---|---|---|---|---|
| Parameters | 4 | 6 | 5 | 7 | 5 | 7 | 7 | 8 |
| Predicted LL(β) | -100.560 | -99.813 | -96.053 | **-95.127** | -99.427 | -98.914 | -96.765 | -96.399 |
| Predicted LR | 0.045 | 0.056 | 0.088 | **0.097** | 0.052 | 0.061 | 0.078 | 0.084 |
| Adj. predicted LR | 0.007 | -0.001 | **0.040** | 0.030 | 0.004 | -0.006 | 0.012 | 0.008 |
| APCP | 0.619 | 0.635 | 0.695 | **0.706** | 0.669 | 0.676 | 0.680 | 0.682 |
| RMSE | 12.820 | 12.733 | **11.986** | 11.987 | 12.399 | 12.235 | 12.040 | 11.998 |
| MAPE | 9.883 | 9.513 | **8.968** | 8.972 | 9.258 | 9.112 | 9.040 | 9.017 |

**Table 7.13: Prediction test results on validation sample**

## 7.7 Summary and conclusions

This chapter has described an analysis of passengers' risky choices on the London Underground network. The main purpose of this research is systematically to compare the performances of EUT models and their counterpart non-EUT models which have been developed for this research. The non-EUT models of interest to us consist of SEV, SEU, RDEV, RDEU, PT and CPT models, with the comparison among these candidate models focusing on their estimation performances as well as their predictive performances. While most of the relevant literature is dominated by SP studies, this research is based on RP data collected from the London Underground system and involves the choice between alternative competitive underground services linking pairs of stations. The estimation sample was randomly extracted from 75% of the total sample, with the remainder of the sample being used for model validation.

In terms of model performance, all the non-EUT models lead to a modest improvement in model fit. The nested test of fit indicates the superiority of the SEV model over the other alternative model specifications. According to the non-nested test result, PT and CPT also show an improvement in model fit over the other models, except for SEV and SEU. In addition to the statistical test of the estimation sample, the predictive test using aggregate and disaggregate methods also reveals that non-EUT models actually provide better predictive performances. The results from calibration and validation jointly reinforce the importance of adopting a critical and empirically driven approach to evaluating the merits of non-EUT models, especially taking into account the much greater complexity involved in the estimation and application of these models. Specifically, the empirical findings especially highlight the importance of nonlinear utility, weighting function and reference dependence.

Moreover, it should be noted that the good performances of non-EUT models in this research is in contrast to the relatively poor performances of similar non-EUT approaches in Chapter 5. We can now conclude that the estimation of non-EUT is potentially sensitive to the quality of RP data, given that the level-of-service data used in the LU case study is much more plausible than the data used in the SR91 case study. Hence, future RP research should pay particular attention to the statistical properties of the underlying network performance data for extracting travel time distribution, especially the network data. Furthermore, an essential avenue for future research is the analysis of the potential heterogeneity across parameters, and even different non-EUT models, using a large-scale sample.

# Chapter 8 CONCLUSIONS AND FUTURE RESEARCH RECOMMENDATIONS

This chapter provides a conclusion on the research undertaken in this thesis and recommends possible directions for future research based on the results developed in this thesis.

## 8.1 Substantive results

The objectives and motivations for this research were developed in Chapter 1, with the general aim of establishing and comparing EUT and non-EUT approaches in the RP context. This section revisits these objectives to summarize the main findings and contributions.

### 8.1.1 Identify opportunities and challenges in modelling traveller's risky choice using non-EUT approaches

This thesis is based on a comprehensive literature review (in Chapter 2) which offers a foundation for subsequent research from both a micro-economics and behavioural economics perspective. In particular, Chapter 2 identified appropriate choice theories and discussed the reasons for using these selected theories.

Sections 2.4 and 2.5 reviewed the main risky choice theories and especially highlighted the motivations for using non-EUT methods. It was found that the prevailing approach for modelling travellers' risky choice behaviour is EUT, which is based on the assumption of utility maximization and instrumental rationality. The behavioural assumptions underlying EUT are frequently claimed to be unrealistic and overly simplistic, however, and therefore more recent work on risky choice has sought to address these perceived shortcomings by using various, more general, non-EUT approaches.

The work presented in Chapter 3, meanwhile, was concerned with the applications of EUT and non-EUT in the existing transport modelling literature, especially in the context of travel time variability, which results in risks for travellers' decision making. Section 3.3 and 3.4 discussed the opportunities and methods to incorporate these models into a risky choice framework. The key finding was that we can combine RUM with EUT and non-EUT to model travellers' risky choice behaviour.

To date, little attention has been given to the empirical evaluation of these non-EUT approaches – the case for their use has been rhetorical rather than empirical. Despite their current popularity in some academic circles, we have very little evidence regarding whether they actually produce materially different and better results, and such evidence as we have is almost exclusively based on stated rather than revealed preference data. These hypothetical data collection strategies are flexible and economical, but have significant weaknesses in terms of the external validity and generalizability of their results. This state of affairs provides a strong motivation for this thesis.

### 8.1.2 Develop a novel RUM-NEUT framework

In terms of model specification, Chapter 4 determined a possible modelling strategy for synthesizing non-EUT and RUM into an operational and flexible structure for a realistic model. Following this strategy, section 4.2 set out our risky choice framework, which is capable of characterizing and predicting travellers' decision making under risky conditions, such as unpredictable travel time. This model framework successfully addresses travellers' uncertainty regarding travel time and modellers' uncertainty regarding real choice context, and, more importantly, incorporates several behavioural factors, such as reference dependence, diminishing sensitivity, nonlinear decision weight and rank dependence.

### 8.1.3 Develop the method of RP data collection for modelling risky choice behaviour

In terms of data collection, problems may arise with RP data because it is difficult to acquire sufficient detail to model risky choice behaviour. This was the main obstacle in data collection and analysis for this research, since we could not afford the cost of observing detailed travel time distributions specific to each respondent. To solve this critical issue, we simplified the research by using several assumptions and techniques. Firstly, we acquired travel time distribution information from a network performance dataset, and assumed that this travel time, in terms of a specific corridor, was the same as the time perceived by travellers who were used to travelling through this corridor. Consequently, we could consider the travel time extracted from the performance data to equate to the variable of travel time in the models.

Secondly, we assumed that travellers make decisions by measuring a set of possible travel time outcomes rather than a uniform travel time such as mean travel time. Specifically, observed travel time was divided into a set of discrete contingent outcomes, while the likelihood of occurrence recorded in the performance data served as the associated probability of travel time outcome. In this way, the set of travel time and associated probability could be converted into the input for our risky choice framework.

### 8.1.4 Compare EUT and non-EUT models

This thesis presents the empirical findings of two case studies using the RP data collected on the SR91 corridor and the London Underground network, respectively. This applied work was undertaken by comparing the calibration performances of each candidate model, making ours the first study to compare EUT and non-EUT systematically in an RP context. We will now briefly summarise the results of the two studies along with the overall findings in relation to model comparison.

The research presented in Chapter 5 was our first attempt to estimate and compare the proposed models empirically. This data describes a binary choice between a tolled route with a reliable travel time and an untolled route with high travel time variability. A set of techniques were proposed to embody the behavioural theories. In particular, section 5.4.4 introduced the algorithm for endogenously estimating the reference point, and section 5.5.3 demonstrated the method for determining attitude towards risk by analysing the nonlinear weighting probability. The estimation result suggested that the main attributes affecting route choice in this RP context are the cost of travel time and cost of the toll, while the segmentation by age and gender also had a significant impact on decision making. The estimation results showed the behavioural appeal of non-EUT, and explained drivers' route choice behaviours which turned out to vary from economic intuition. Although the difference of mode fit between the models was insignificant, we concluded that this 'undesirable' result may be due to the shortcomings of the RP data used in this research, given that there were merely 210 observations in the floating car dataset.

The analysis using the RP data collected from the London Underground was presented in Chapter 7. Travel time was still found to have a major impact on risky choice behaviour, and headway also served as an essential factor. Journey cost, however, had no impact in this research since the ticket price is uniform between tube lines. Here we randomly used 75% of the total sample as the estimation sample, and treated the remainder as a validation sample. Both survey data and level-of-service data were originally collected by us and sufficient observations from the network data enable us to obtain accurate travel time distributions and investigate whether the quality of the network data has an impact on final model fit in comparison with the SR91 case-study. We evaluated each candidate model separately, and analysed which method could lead to an improvement in model fit. It was found that all non-EUT models offer a modest gain in goodness-of-fit compared to EVT and EUT models. We concluded that the superiority of non-EUT models highlighted the importance of subjective probability, reference dependence and diminishing sensitivity.

The empirical results of the two case studies are not directly comparable due to the difference of choice context and the year of data collection. Several general conclusions can still be reached from both a model estimation and a behavioural interpretation point of view, however.

An improvement in model fit relative to EVT and EUT was found in both case studies, and especially in the London underground case study. These results from calibration reinforce the benefit of using non-EUT approaches for modelling travellers' risky choice behaviour, especially taking into account the much greater complexity involved in model specifications, such as nonlinear weighting function, reference dependence, and diminishing sensitivity. We also observed that the improvement in non-EUT model fit compared to the EUT models was larger in the London Underground case study than in the SR91 case study. The better performance of non-EUT models, especially SEV and SEU, serves as strong evidence in favour of the high-quality network data used in the London Underground case study, since this data was able to generate relatively accurate travel time distributions which may affect the estimation of subjective probability and nonlinear utility.

Although the improvements in model fit in both case studies are relatively modest, the advanced models still offer more insights into risky choice behaviour. For instance, the RDEU calibration results show that respondents tend to over-weight extremely bad outcomes and under-weight the likelihood of normal situations, which implies the presence of pessimism and risk aversion in the sample. In addition, a strong loss aversion was also observed in both studies, namely that $|\beta_{TT\_loss}|/|\beta_{TT\_gain}|$ equals 1.57 in the SR91 case study and 1.414 in the LU case study. This finding indicates that loss aversion may be a perceptual conditioning that generally exists in travellers' decision making under risk.

To conclude, the main contribution of the model comparison is that we analysed how non-EUT approaches perform in the real world, rather than in laboratory experiments or hypothetical choice contexts. Moreover, we step-by-step diagnosed which non-EUT method outperforms the EUT methods and explained the possible reasons for this improvement.

### 8.1.5 Implementation with respect to the valuation of travel time savings

Chapter 6 not only conducted a comprehensive analysis of VTTS from existing literature, but also proposed our own specifications for measuring VTTS based on non-EUT approaches. By testing different models with the same data, we compared the influence of different modelling approaches on the estimated VTTS. Various weighting functions were tested, with the results showing that the selection of weighting function has little impact on the final model fit and VTTS. Although goodness-of-fit did not vary across models in the SR91 case-study, model structures do have significant impact on VTTS. In this research, RDEV tended to underestimate VTTS compared to linear EVT models, whilst EUT and all the other non-EUT models seemed to overestimate it. Section 6.5 demonstrated a disparity between WTP and WTA which reinforces the validity of loss aversion in an RP context. Moreover, it was found that the observed WTA/WTP ratio in this RP context was much lower than the estimated ratio from SP studies. This suggests that EVT in an SP context may result in misleading VTTS if the impacts of non-EUT components and RP data are omitted.

### 8.1.6 Implementation with respect to prediction

Having compared the model performances in terms of calibration in Chapter 7, research was then undertaken to identify the implementation with respect to prediction. Both aggregate measures and disaggregate methods were applied to measure the correct predictions of route choice. It was found that all the proposed non-EUT models delivered better predictive performances in the validation sample relative to EUT and EVT, although the improvement was trivial. The improvement of predictive performance was too small to arrive at a convincing conclusion favouring non-EUT models.

## 8.2 Conclusion

In the theoretical part of this thesis (Chapter 2, 3 and 4), the research demonstrated the development of risky choice theories, and it seems that an increasing amount of research have been attracted to develop ever more complex non-EUT approaches. This prosperity of theories actually offers the opportunity to explore the role of risky choice theory for characterizing travellers' risky choice behaviours more fully. It should be noted that these theories are derived from different assumptions which cannot be simultaneously true in the same choice context. Empirical tests on these theories, therefore, are urgently required to properly discriminate them. However, it was concluded that there exist a research gap between the state-of-the-art and the state-of-the-practice, especially in the model evaluation and data usage. Indeed, more resources should be allocated to the studies on the empirical performances of risky choice models.

The applied part (Chapter 5, 6 and 7) has shown that non-EUT models do have the potential to deliver better model performance, especially in the LU case study. This result is encouraging, and especially indicates that non-EUT is worth receiving extensive research on its usage and validity in transport. Furthermore, it also serves an important empirical evidence for experimental economics to explore the performance of non-EUT in the real world. However, it is still arguable about whether the complexity of models is capable of outperforming EUT, given that there is no any major difference of prediction effect between models. Furthermore, these advanced models usually require extra information from data collection, such as travellers' reference points, which again restrict their applications in the real-world. These problems are considered as the main obstruct to transfer non-EUT from the theory to the practice in large scale analysis. Given these difficulties, it is still too early to claim that non-EUT is conclusively ready to replace EUT for modelling risky choice behaviour in transport.

Although the improvements of model performance are modest in this thesis, non-EUT models still show their crucial roles in the implication of VTTS. Researchers should be careful about the variation of VTTS since a significant difference of estimate (say 10%) could lead to the adjustment of policy making. Therefore, it was concluded that misleading estimates of VTTS may be obtained if we omit the impact of attitude towards risk and non-EUT components.

Having realized the power of model comparison, we conclude that more empirical tests should be carried out before non-EUT is applied into large-scale practice, in particular by paying special attention to both model specifications and data collection. It would be worth identifying how other advanced approaches perform in a transportation context, such as regret theory. Moreover, more testing is necessary to identify whether the results produced in this thesis can be extended to other datasets and choice scenarios, such as mode choice and departure time choice. One important issue that this current thesis has not addressed is the identification of respondent heterogeneity at an individual level. Chapter 5 found that drivers' tastes of preference varied across different income levels, ages and genders, and Chapter 7 also demonstrated the observed heterogeneity in risk attitude parameters. It would be natural to consider whether unobserved heterogeneity also exists in taste parameters, such as the travel time parameter and cost parameter, and attitude parameters, such as the risk attitude parameter and the weighting function parameter. This can be done by adopting a mixed multinomial logit (MMNL) model with a nonlinear utility function. This allows these parameters to be random, and then assumes different distributions for them. Note that even MMNL still assumes that all individuals use the same utility function with the homogenous decision rule, but future research could actually allow the heterogeneity in decision rules across individuals by employing such advanced statistical methods as latent class (LC) models or Expectation Maximization (EM) models.

One main limitation of the empirical studies in this thesis is the relatively limited sample size and unreliable network performance data used in the SR 91 case-study. Chapter 7 also highlights the importance of employing high-quality network data to improve the validity of RP studies. GPS data is potentially the most promising alternative to obtain highly reliable data for risky choice research. GPS probes are capable of recording drivers' actual trips, therefore, allowing the estimation of the real journey time experienced by drivers. Such travel time data is specific to each individual's actual journey experience, and is potentially more reliable than the travel time extracted from level-of-service data as used in this current research. Similar to GPS data, there are a number of alternative methods, such as mobile phone data, Bluetooth data and ANPR data. The main problems of such data collection methods are issues with privacy protection and high costs. To enlarge the sample size, however, future research could turn to several large-scale survey databases, such as NTS and Oyster data in the case of the UK. Finally, it is also potentially worth combining RP and SP data in future research.

# Appendix A: AN EM APPROACH TO REFERENCE-DEPENDENT RISKY CHOICE

## A.1 Introduction

Recent work on risky choice modelling has sought to address the perceived shortcomings of expected utility theory (EUT) by using non-expected utility (non-EUT) approaches. One of the most popular non-EUT approaches is prospect theory (PT). A key feature of PT (and its many derivatives) is the idea of reference dependence. Despite its popularity in the recent literature, however, the empirical evidence for reference dependence is entirely limited to stated choice data. Moreover, no credible theory for the definition of reference points has been proposed, resulting in most empirical studies resorting to ad hoc and essentially arbitrary definitions of reference points.

Here we address these two weaknesses, first by evaluating PT approaches using revealed preference data and second, by extending existing methods to advanced PT models incorporating estimated and natural reference points. The data for this study describes a simple route choice context involving drivers choosing between a free flowing and reliable tolled facility and a congested and unreliable untolled facility.

PT models are estimated within a random utility framework. In the proposed PT models, the segment of reference travel time which each individual belongs to was treated as missing data and estimated using an expectation-maximisation (EM) approach. This EM algorithm is the general iteration method of finding the maximum log-likelihood and parameter estimates when data is incomplete or has missing values. Specifically, three model specifications with six scenarios were tested, and all models produced intuitively plausible estimation results with the EM embedded PT model providing the best overall goodness of fit. The findings presented in this study reinforce the importance of exploring elaborate PT models within a revealed preference context.

The structure of this appendix is organized as follows. The next section outlines reference dependence in Prospect Theory and the specification of the Expectation-Maximization algorithm. This is followed by a description of econometric model forms. The subsequent section sets out the empirical application and interprets the model estimation results. The appendix concludes with suggestions for further research.

## A.2 Methodology

### A.2.1 What is a traveller's reference point?

Reference point, in travel behaviour modelling, is commonly regarded as some threshold value that distinguishes gains and losses (Avineri and Bovy, 2008). The question is what such a threshold value is. There is no consistent definition of reference point or any proper models to estimate such a threshold value. In fact, the lack of consensus regarding the reference point value has been considered as the main obstruction for the adoption of PT.

The original PT suggested that the current wealth condition (or status quo) is a promising reference point. Hence, a number of subsequent studies employed this proposition for the reference point and obtained empirical evidence on reference dependence, e.g. current job (Tversky and Kahneman, 1991) and current dwelling (Habib and Miller, 2009), etc. The situation in regard to the reference point is much more complicated in the context of travel behaviour, however, and there are therefore various representations of the reference point. A simple assumption is that a traveller's reference point is related to his/her recent past travelling experiences (Avineri and Prashker, 2003). De Borger and Fosgerau (2008), meanwhile, argued that the current trip is the most plausible reference point in a car-commuter survey. This is in line with the definition of status quo proposed in the original version of PT. More importantly, such a definition of the reference point enables researchers to obtain the value of the reference point by observing individuals' actual travelling practice, such as the average journey time of the latest ten trips. An alternative definition of the reference point is proposed by Kőszegi and Rabin (2006) who assumed that the reference is the individual's rational expectation determined by their personal equilibrium. Additionally, empirical evidence has shown that the future goal can also serve as the reference point (Heath et al., 1999).

The above literature provides a potentially useful set of alternatives regarding reference points. However, we should also be aware of several other practical factors influencing the value of the reference point. In the real world, travellers might not have sufficient cognitive capacity and, therefore, find it difficult to form a reference point based on their experienced trips. Additionally, some travellers have no or merely a few travelling experiences on the objective trip, and thus have insufficient information to form an adequate reference point. Furthermore, it has been found that the reference point is influenced by explicit and implicit information, such as *a priori* travelling information from a travel planning website. All of these factors serve to distinguish the reference point applied in modelling from the one perceived in reality. Given the complexity of the reference point in travel behaviour, it seems necessary to make plausible assumptions regarding reference points that are specific to different choice contexts.

Jou et al. (2008) observed travellers' asymmetric reactions to gain and loss, and their estimated parameters of losses and gains provided insights into travellers' tastes to different arrival times. Specifically, an arrival time outcome between the preferred arrival time and the work start time is most preferred in that it generated maximum positive estimates of 0.0423, while an arrival time outcome later than the work start time turned out to be most undesirable with negative estimates of -0.1089. While this three reference point model enjoys intuitive strength, it is difficult to collect sufficient information regarding all the three reference points.

In addition to departure time choice, route choice behaviour has also been extended to PT. Note that route choice is primarily affected by travel time, so this stream of studies consistently focuses on the reference point in the attribute of travel time. It is assumed that route choice is influenced by the difference of actual travel time and reference travel time. The simplest way to elicit reference travel time is hypothetically to set some natural travel time. For instance, Gao et al. (2010) assumed that travellers' route choice behaviour is related to their reference travel time. Regardless of the heterogeneity across travellers, free flow travel time is assumed to be the reference travel time in their CPT route choice model. Rose and Masiero (2010) considered the free flow travel time and slowed down travel time of recent trips as the reference points. In this case, the parameters of gains and losses are statistically significant, whereas the fit for the PT model is worse than the fit of the normative model without a reference point. Avineri and Prashker (2005) employed expected travel time as the reference point in their risky route choice model. They concluded that the CPT prediction results are highly sensitive to such a reference point. Avineri (2009) subsequently extended this concept of reference point to the fuzzy set, which is a range of reference points, for instance, a fuzzy set of reference travel time is (25min, 30min, 35min). In addition to the above methods for valuing reference points, the concept of reference dependence has been applied in other transport literature (Páez and Whalen, 2010). Some recent studies on reference dependence models are summarized in Table A.1.

| Literature | Choice | Mode | Definition of reference point(s) | Loss aversion | Diminishing sensitivity |
|---|---|---|---|---|---|
| Michea and Polak (2006) | Departure time | Rail | Preferred arrival time | Yes | Yes |
| Senbil and Kitamura (2004)* | Departure time | Car | Earliest permissible arrival time; work start time; preferred arrival time | Yes | Yes |
| Jou et al. (2008) | Departure time | Public transport | Earliest permissible arrival time; work start time | Yes | NR* |
| Hess et al. (2008) | Route | Car | Current route | Yes | Yes |
| Rose and Masiero (2010) | Route | Car | Current route | Yes | Yes |
| Masiero and Hensher (2010) | Route | Multi modes | Current route | Yes | Yes |
| De Borger and Fosgerau (2008) | Route | Car | Current trip | Yes | Yes |
| Gao et al. (2010) | Routing policy | Car | Free flow-travel time | Yes | NA* |
| Avineri and Prashker (2005) | Route | Car | Expected travel time | Yes | NR |
| Avineri (2009) | Route | NA | A set of travel times | Yes | NA |
| Schwanen and Ettema (2009) | Arrival time | Car | Closing time; Event-based time; preferred arrival time | Yes | NR |

NA: not applicable
NR: not reported
Senbil and Kitamura (2004) suggested that the earliest permissible arrival time< preferred arrival time< work start time

**Table A.1: Recent transport studies using reference dependent effect**

### A.2.2 Expectation-Maximization (EM)

The Expectation-Maximization (EM) algorithm is a method for maximizing a likelihood function when direct maximization is difficult and/or there is missing data (Bilmes, 1998). It has been applied in various fields to address choice models with multiple unknown population segments (Aitkin and Aitkin, 1996, Horsky et al., 2006). In transport, Bhat (1997) applied the EM algorithm to a latent class model with up to four classes, and Train (2008) subsequently extended Bhat's work by specifying the EM algorithm for discrete mixing distribution with a large number of classes. This current study proposes a new reference dependence specification which incorporates the EM algorithm into PT models with multiple reference points. The missing data in this research is the category of the reference travel time of the decision maker.

Given the values of reference points as segments, the missing data in the PT model is the reference travel time segment which each decision maker belongs to. Consequently, the reference travel times $r \in 1, \dots, R$ are treated as the latent classes, and $r = k$ means the sample was generated by the segment with the $k^{th}$ reference point. Similar to mixture-density parameter estimation, the log-likelihood turns out to be:

$$LL(\theta) = \sum_{i=1}^{N} logp(x_i;\theta) = \sum_{i=1}^{N} log \sum_{r=1}^{R} \alpha_r p(x_i, r; \theta_r) \quad \text{(A.1)}$$

where $\alpha_r = p(r)$ is the prior probability of segment $r$, and $i^{th}$ individual's probability of choosing the alternative with attribute $x_i$ is $p(x_i, r; \theta_r) = \frac{exp(U_{x_i,r})}{\sum_{k=1}^{K} exp(U_{k,r})}$ which is in line with the logit formula. To maximize the likelihood function, note that it is difficult to calculate this function due to the log of the sum. Therefore, Jensen's inequality (Dempster et al., 1977) is applied to calculate $LL(\theta)$, i.e.,

$$LL(\theta) = \sum_{i=1}^{N} log \sum_{r=1}^{R} \alpha_r^T p(x_i, r; \theta_r)$$

$$= \sum_{i=1}^{N} log \sum_{r=1}^{R} \alpha_r^T p(x_i, r; \theta_r) \frac{p(r|x_i;\theta_r^T)}{p(r|x_i;\theta_r^T)}$$

$$\geq \sum_{i=1}^{N} \sum_{r=1}^{R} p(r|x_i;\theta_r^T) log \frac{\alpha_r^T p(x_i,r;\theta_r)}{p(r|x_i;\theta_r^T)} = E(\theta) \quad \text{(A.2)}$$

where $p(r|x_i;\theta_r^T)$ is the posterior probability of each segment r, given the current estimate $\theta_r^T$ and $\alpha_r^T$ after the $T^{th}$ iteration, and $p(r|x_i;\theta_r^T) = \frac{\alpha_r^T p(x_i,r;\theta_r)}{\sum_{j=1}^{R} \alpha_j^T p(x_i,r;\theta_j)}$ from Bayes's rule. This last step is exactly the E-step which constructs the lower-bound on $LL(\theta)$. The subsequent M-step aims to optimize the lower-bound $E(\theta)$. To incorporate a Random Utility Model (RUM) with the EM algorithm, note that maximization of expectation $E(\theta)$ corresponds to a separate maximization of each parameter. Hence, we should elaborate $E(\theta)$ as:

$$\mathrm{E}(\theta) = \sum_{i=1}^{N} \sum_{r=1}^{R} p(r|x_i;\theta_r^T) log\alpha_r^T + \sum_{i=1}^{N} \sum_{r=1}^{R} p(r|x_i;\theta_r^T) log\, p(x_i, r; \theta_r) \quad \text{(A.3)}$$ [57]

Given the initial value $\theta_r^T$, M-step is divided into two simple maximization problems, that is,

$$\alpha_r^{T+1} = argmax \sum_{i=1}^{N} \sum_{r=1}^{R} p(r|x_i;\theta_r^T) log\alpha_r \quad \text{(A.4)}$$

[57] The $E(\theta)$ formula should include $-\sum_{i=1}^{N} \sum_{r=1}^{R} p(r|x_i;\theta_r^T) log(p(r|x_i;\theta_r^T))$ as well, but the weights $p(r|x_i;\theta_r^T)$ does not depend on $\theta_r$, so we only consider the first and second sum.

$$\theta^{T+1} = \text{argmax} \sum_{i=1}^{N} \sum_{r=1}^{R} p(r|x_i; \theta_r^T) \log p(x_i, r; \theta_r) \quad \text{(A.5)}$$

The estimated parameters are again applied into the input of the E-step if convergence is not achieved. The convergence of the EM recursion is usually defined as a sufficiently small change in log-likelihood (Weeks and Lange, 1989).

# A.3 Model structure

In this research we implement three candidate models in six risky choice scenarios, and each model is followed by estimations using our proposed methods. The compressive comparisons of estimation results are subsequently conducted according to appropriate criteria. Specifically, the following models are tested in this research:

- A basic PT model which only consists of a single reference travel time. The assumption behind this model is that all individuals hold the same reference point when they make a choice.
- A hybrid model with a number of fixed points where the share of each point is estimated by repeatedly maximizing the weighted log-likelihood function.
- A PT-EM model where all coefficients and shares of each point are estimated by repeated estimation of the embedded MNL model.

In A.3.1 the basic PT specifications are demonstrated in detail, while the hybrid model is set out in A.3.2, and the proposed PT-EM specification in A.3.3.

## A.3.1 PT model

In this model, individuals are assumed to take into account the difference between actual travel time and reference travel time when they choose routes. Hence, the utility function for individual *i* is given by:

$$U_i = \beta_{xi} \sum_{k=1}^{K} p_{ki}(x_{ki} - x_{rp}) + \beta_{yi} y_i + \beta_i W_i \quad \text{(A.6)}$$

where, $\beta$ is the coefficient to be estimated, $x_i$ is the variable of travel time which is the source of risk in this case, $p_{ki}$ is the associated probability of travel time in the $k^{th}$ risky outcome, $y_i$ is the other variables associated with the level of road service, such as travel cost and travel time variability, and $W_i$ represents individual *i*'s socio-economic variables such as age, gender and income. It can easily be seen that the above model is equivalent to the expected value model. As such, an adapted formulation is required to incorporate the features of PT:

$$U_i = \beta_{gain} \sum_{k=1}^{m-1} p_{ki} |x_{ki} - x_{rp}| + \beta_{loss} \sum_{k=m}^{n-1} p_{ki} |x_{ki} - x_{rp}| + \beta_{loss^-} \sum_{k=n}^{K} p_{ki} |x_{ki} - x_{rp}|$$

$$+ \beta_{yi} y_i + \beta_i W_i \qquad \text{(A.7)}$$

The risky outcome is interpreted as gain if $x_{ki} - x_{rp} \leq 0$, and loss if $x_{ki} - x_{rp} \leq 0$, and diminishing loss if $x_{ki} - x_{rp} > \delta$, given that δ is a specific threshold value beyond which individuals express diminishing sensitivity towards time loss. We expect that the coefficient of gain $\beta_{gain}$ is positive, and the coefficient of loss $\beta_{loss}$ and $\beta_{loss^-}$ is negative. It should be noted that this function enables us to investigate the loss aversion effect as well as the diminishing sensitivity effect by using three different coefficients associated with travel time. The specification is slightly different from the original PT formulation, where power function is adopted to address diminishing sensitivity. The reason for the different interpretation of diminishing sensitivity is that we found that the coefficient for a continuous nonlinear specification is not statistically significantly different from unit.

## A.3.2 Hybrid model with fixed coefficients

The model described in this section does not required logit estimation as shown in 3.1. Instead, the coefficients in the utility functions are fixed (fixed points), and only the share of each point should be estimated. Again, we assume there are *R* reference points, and each reference point is referred to as a specific segment. Hence, it is a restricted specification of the EM algorithm with *R* classes, except that the estimated parameters do not include $\beta_r$ for $r = 1., \ldots, R$. By increasing the number of segments, the estimations of shares are better. In this research, the number of segments (reference points) should be specified from estimation results of PT in advance, which limits the size of the grid. In this way, the hybrid model simply becomes:

$$\alpha_r^{T+1} = argmax \sum_{i=1}^{N} \sum_{r=1}^{R} p(r|x_i; \theta_r^*) log\alpha_r \qquad \text{(A.8)}$$

where $\alpha_r^{T+1}$ is the estimates of share *r* in the *(T+1)$^{th}$* iteration, given the known values of parameters $\theta_r^*$. Iterations enable us to estimate the shares as follows:

$$\alpha_r^{T+1} = \frac{1}{N}\sum_{i=1}^{N} p(r|x_i;\theta_r^*) \tag{A.9}$$

The algorithm described in this section is implemented by the following steps:

1. Determine the fixed points including the coefficients $\theta_r^*$ and reference points $\mathrm{x_r}$. One solution to select the fixed points is to specify the range of coefficient values according to the estimation results of the PT model.
2. Guess the initial share $\alpha^T$ .
3. Compute the probability of choosing alternatives for each individual at each fixed point, i.e. $p(x_i, r; \theta_r)$. The MNL model is adopted at this step.
4. Calculate the posterior segment probabilities $p(r|x_i;\theta_r^*)$ for $r = 1., \dots, R$.

$$p(r|x_i;\theta_r^*) = \frac{\alpha_r^T p(x_i,r;\theta_r^*)}{\sum_{j=1}^{R} \alpha_j^T p\left(x_i,r;\theta_j^*\right)} \tag{A.10}$$

5. Update the share of each segment.
6. Repeat steps 4 and 5 until convergence. $\alpha_r^*$

This method is illustrated in Figure A.1.

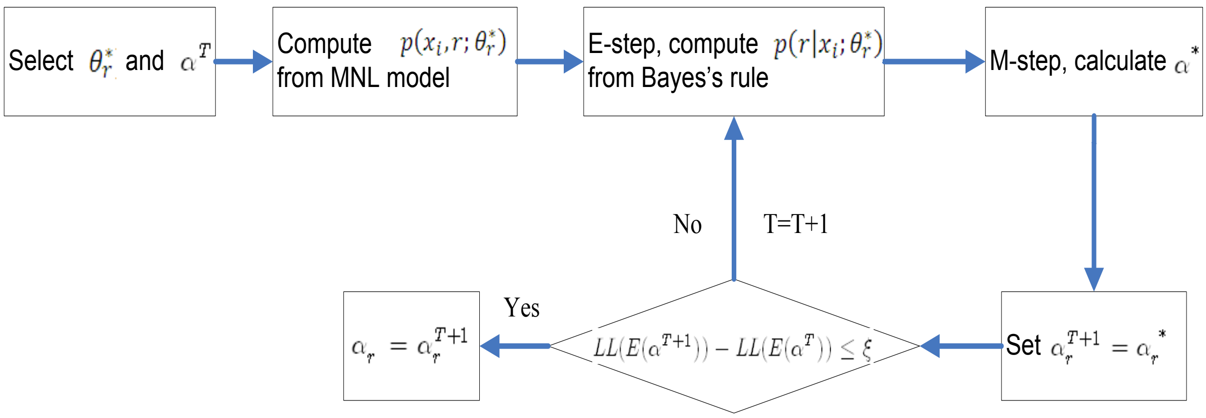



**Figure A.1: Flowchart of the hybrid approach**

### A.3.3 PT-EM model with all coefficients to be estimated

In this section, we modify the previous model by treating both coefficients $\beta_r$ and shares $\alpha_r^i$as parameters to be estimated homogeneously. Again, reference travel time corresponds to segments in the population, and the missing data is the specific segment which each decision maker belongs to. As a result, we do not arbitrarily assign fixed coefficients to each segment, but estimate all the parameters within the EM framework. The PT-EM algorithm consists of the following steps:

1. Specify the initial values of the coefficients $\theta_r^T$ and corresponding shares $\alpha_r^T$ in every segment of the reference travel time.
2. Calculate the probability of choosing alternatives for each individual in the $T^{th}$ iteration, i.e. $p(x_i, r; \theta_r^T)$.
3. Calculate the posterior segment probabilities $p(r|x_i; \theta_r^T)$ for $r = 1., \ldots, R$.

$$p(r|x_i; \theta_r^T) = \frac{\alpha_r^T p(x_i,r;\theta_r^T)}{\sum_{j=1}^R \alpha_j^T p\left(x_i,r;\theta_j^T\right)} \tag{A.11}$$

4. Update the share of each segment.

$$\alpha_r^{T+1} = \frac{1}{N}\sum_{i=1}^N p(r|x_i; \theta_r^*) \tag{A.12}$$

5. Use the MNL model to substitute updated parameter $\theta^{T+1}$for $\theta_r^T$, note that the log-likelihood function is weighted by $p(r|x_i; \theta_r^T)$ as follows:

$$\theta^{T+1} = \ argmax \sum_{i=1}^N \sum_{r=1}^R p(r|x_i; \theta_r^T) \log p(x_i, r; \theta_r) \tag{A.13}$$

6. Repeat steps 3 to 5 using updated $\alpha_r^{T+1}$ and $\theta^{T+1}$ until convergence.

This procedure is illustrated in Figure A.2.

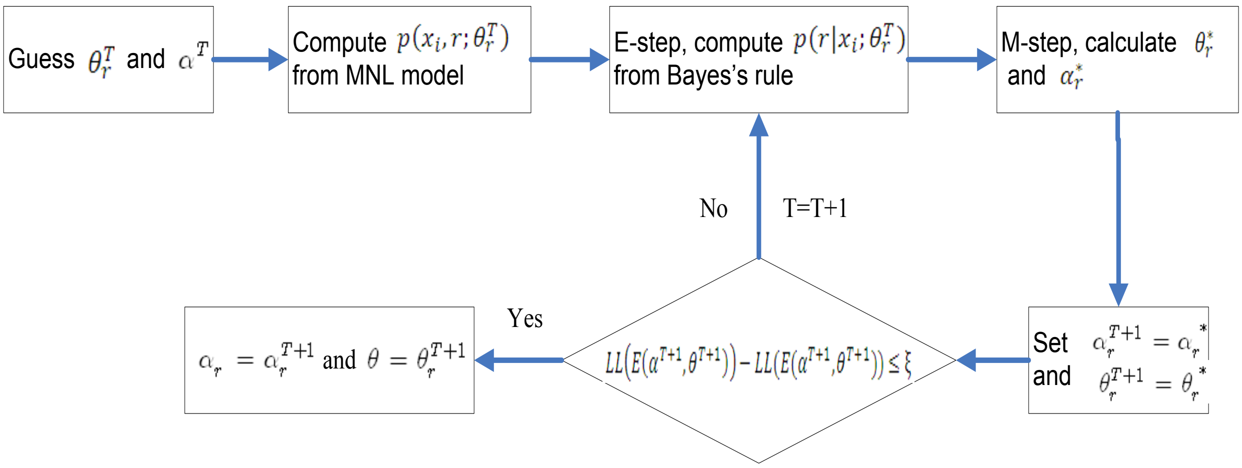



**Figure A.2: Flowchart of the PT-EM approach**

## A.4 Application to route choice modelling

In this section, we present the empirical results obtained from applying the proposed PT-EM models, and other alternative models, to an analysis of route choice behaviour under unreliable travel time.

### A.4.1 Data

Almost all existing studies have used stated preference (SP) data to develop risky choice models. Revealed preference (RP) data, obtained from individuals' actual behaviour, is less flexible but more reliable than SP data, however. Consequently, more reliable RP data is urgently required in risky choice studies. The data adopted in this research is from the RP data collected on the California State Route 91 (SR91) corridor in the USA, a survey that was originally undertaken to investigate the value of travel time and travel time reliability (Lam and Small, 2001, Small et al., 2005a, Small et al., 2005b).

The route of interest is a 10 mile portion of SR 91 which consists of one tolled road and a parallel freeway road. A sample of 438 observations is included in this study. Level of service data was generated based on floating car data with 210 trip observations.

### A.4.2 Variables

Consistent with Small et al. (2005b), we also assume that the travel time for the untolled route is constant at 8 minutes, since the traffic on the tolled road is observed to move freely at all time of day. Therefore, only a travel time distribution for the toll route is required. The limited number of observations from floating car data is not convincingly sufficient to obtain a corresponding continuous travel time distribution, however.[58] Consequently, empirical frequency of travel time is employed, i.e., we treat travel time from 8 minutes to 20 minutes as discrete contingent outcomes.

### A.4.3 PT model estimation results

Both PT models M1 and M2 are regarded as the base model. M1 is the simplest model with only one natural reference travel time which is assumed to be the free-flow travel time at 8 minutes. Notice that all travel time outcomes in M1 are interpreted as a loss, since all actual travel time is greater than free flow travel time $x_{rp} = 8\ min$. M2 adopted the estimated travel

[58] We tested a couple of distributions for the travel time, viz., normal distribution and lognormal distribution, however the K-S test results show that these distributions do not fit the floating data well.

time $x_{rp} = 8.8\, min$ which is treated as an extra parameter and endogenously estimated. As a result, travel time outcome is considered as a gain when $x_k \leq 8.8\, min$, a loss when $8.8\, min < x_k \leq 13\, min$, and a diminished loss when $13\, min < x_k \leq 20\, min$ (details are available in (Hu et al., 2012b)).

| | M1 | | M2 | |
|---|---|---|---|---|
| | Value | (*t*-ratio) | Value | (*t*-ratio) |
| ASC | -1.740 | (-5.26) | -1.9 | (-5.46) |
| $\boldsymbol{\beta_{toll\_low}}$ | -0.450 | (-4.17) | -0.434 | (-3.99) |
| $\boldsymbol{\beta_{toll\_medium}}$ | -0.312 | (-3.11) | -0.289 | (-2.84) |
| $\boldsymbol{\beta_{TT(gain)}}$ | | | 0.0056 | (2.21) |
| $\boldsymbol{\beta_{TT(loss)}}$ | -0.006 | (-3.27) | -0.0087 | (-3.03) |
| $\boldsymbol{\beta_{TT(loss^-)}}$ | | | -0.0052 | (-2.07) |
| Age | 0.710 | (2.86) | 0.71 | (2.84) |
| Gender | 0.82 | (3.47) | 0.84 | (3.5) |
| LL(0) | -303.598 | | -303.598 | |
| LL(ASC) | -252.163 | | -252.163 | |
| Final LL(β) | -230.414 | | -227.916 | |
| $\boldsymbol{\rho^2}$(0) | 0.241 | | 0.249 | |
| Adj. $\boldsymbol{\rho^2}$(0) | 0.221 | | 0.223 | |
| $\boldsymbol{\rho^2}$(ASC) | 0.086 | | 0.096 | |
| Adj. $\boldsymbol{\rho^2}$(ASC) | 0.063 | | 0.064 | |

**Table A.2: The estimation results of M1 and M2**

The negative sign of $\beta_{TT(loss)}$ and $\beta_{TT(loss^-)}$ shows that travellers dislike the extra travel time if they take reference dependence into account. $\beta_{TT(gain)}$ is positive in that travellers benefit from the shorter journey time compared to the reference travel time. We also found that the ratio between the absolute values is $\frac{|\beta_{TT(loss)}|}{|\beta_{TT(gain)}|} = 1.57 > 1$, and the t-ratio for the difference between $\beta_{TT(loss)}$and $\beta_{TT(gain)}$is very high at 9.68, which empirically supports the validity of loss aversion. $\beta_{TT(loss^-)}$ is statistically significantly smaller than $\beta_{TT(loss)}$. We also found that $|\beta_{TT(loss^-)}| < |\beta_{TT(loss)}|$, while the t-ratio of difference is relatively low at 0.94. In terms of goodness of fit, M2 performs better than M1. This is supported by the $\rho^2$ measure. However, the likelihood ratio test value of 5 gives a $p$-value of 0.08 on the $\chi^2$distribution.

## A.4.4 Hybrid model estimation results

This hybrid model is simply the combination of all candidate models with different reference points. In fact, the estimation procedure *per se* is computationally easy, although the key elements in this algorithm are the determination of candidate fixed coefficients and reference points. Train (2008) recommended the complete grids specification to find the fixed points. Specifically, he determined the maximum and minimum of $n$ coefficients, and arbitrarily assigned five even spaced values to each coefficient, such that there are $5^n$ different set of coefficients in this grid. In this research, we found that the estimation of the PT model is extremely sensitive to the location of the reference point, and it is therefore not appropriate arbitrarily to assume a grid with a number of assumed reference points. As a result, the estimated coefficients and reference points from M1 and M2 are employed as the fixed points. Hence, we assume that there are two segments of population who take into account the reference dependence, and the possible reference travel time is 8 minutes and 8.8 minutes as shown in M1 and M2.[59]

Given fixed coefficients, only the weight of each segment is required to be estimated. The iteration specification shown in Figure A.1 is adopted. Consequently, the share of the first class is $S_1 = 0.7$ when reference travel time is 8 minutes, and the share of the other class is $S_2 = 0.3$ when reference travel time is 8.8 minutes. It should be noted that the starting value of shares are 0.5 respectively. Although it advisable to check whether the local maximum is global by using a different starting value, in this research, we found that the estimation result is not sensitive to the initial value of shares since the estimated shares are roughly stable at 0.7 and 0.3 respectively with different starting values. The final log-likelihood is relatively stable at -228.445, which is between the fit of M1 and M2.

## A.4.5 PT-EM model estimation results

The third model can be treated as a combination of a logit model and the EM algorithm. Likewise, we assume there are two segments of population with different reference travel times. It should be noted that logit models of two segments use the same observations but different weights in each iteration. Moreover, the maximum likelihood of a logit model is weighted by the posterior probability $p(r|x_i;\theta_r^*)$ of each segment *r*. Again, we should be also cautious about assessing convergence in that likelihood can move slowly near

[59] We also arbitrarily used 30 and more reference travel time to test the model, however, the log-likelihood only rose by 0.05.

convergence. 70 iterations are used to maximize likelihood, which rose less than 0.005 at the end of the iterations. Each iteration requires much more running time due to the homogenous estimation of coefficients and shares. To speed up computation, starting values are based on the estimated parameters of M1 and M2. As introduced in 3.1, we established three model specifications in different scenarios. M4 and M5 maintain the locations and shares of the coefficients of M1 and M2 respectively, and estimate the coefficients and shares of the second class model. M6 treats all the coefficients and shares as parameters to be estimated.

| | M4 | | M5 | | M6 | |
|---|---|---|---|---|---|---|
| | Class 1 | Class 2 | Class 1 | Class 2 | Class 1 | Class 2 |
| ASC | -1.8620<br>(-4.38) | | | -2.0835<br>(-4.00) | -1.9550<br>(-4.40) | -2.7719<br>(-3.67) |
| $\boldsymbol{\beta_{toll\_low}}$ | -0.4558<br>(-3.36) | | | -0.4261<br>(-2.66) | -0.2551<br>(-1.80) | -1.2115<br>(-4.45) |
| $\boldsymbol{\beta_{toll\_medium}}$ | -0.2872<br>(-2.27) | | | -0.2601<br>(-1.75) | -0.2636<br>(-1.87) | -0.4495<br>(-2.44) |
| $\boldsymbol{\beta_{TT(gain)}}$ | | | | 0.0041<br>(1.10) | | 0.0049<br>(1.01) |
| $\boldsymbol{\beta_{TT(loss)}}$ | -0.0075<br>(-3.16) | | | -0.0122<br>(-2.82) | -0.0046<br>(-1.85) | -0.0201<br>(-3.3) |
| $\boldsymbol{\beta_{TT(loss^-)}}$ | | | | -0.0048<br>(-1.31) | | -0.0163<br>(-3.19) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Age | 0.7203<br>(2.30) | | | 0.7490<br>(2.01) | 0.5036<br>(1.56) | 1.8167<br>(3.21) |
| Gender | 0.8903<br>(2.93) | | | 0.7692<br>(2.18) | 1.9114<br>(6.05) | -1.8866<br>(-3.14) |
| Shares | 0.62 (4.66) | 0.38 (7.78) | 0.52 (9.06) | 0.48 (8.45) | 0.58 (3.28) | 0.42 (2.36) |
| LL(0) | -303.598 | | -303.598 | | -303.598 | |
| LL(ASC) | -252.163 | | -252.163 | | -252.163 | |
| Final LL(β) | -228.150 | | -227.821 | | -225.216 | |
| $\boldsymbol{\rho}^2(0)$ | 0.249 | | 0.250 | | 0.258 | |
| Adj. $\boldsymbol{\rho}^2(0)$ | 0.193 | | 0.194 | | 0.202 | |
| $\boldsymbol{\rho}^2$(ASC) | 0.095 | | 0.100 | | 0.107 | |
| Adj.$\boldsymbol{\rho}^2$(ASC) | 0.028 | | 0.029 | | 0.039 | |

**Table A.3: The estimation results of M4, M5 and M6**

Table A.3 gives the estimation results of PT-EM models. All the estimates are statistically significant except $\beta_{TT(gain)}$. As with the estimates of M2, significant loss aversion is also found in M5 and M6, where the asymmetric preference $\frac{|\beta_{TT(loss)}|}{|\beta_{TT(gain)}|}$ is 2.98 and 4.10 respectively. In both M5 and M6, the *t*-ratio for the difference between $\beta_{TT(gain)}$ and $\beta_{TT(loss)}$ is 2.86 and 3.16 respectively. Given the initial starting share as 0.5, we consistently obtained a larger estimated share of the first segment which is from 0.52 to 0.62. One possible interpretation is that more of the population would consider 8 minutes to be the reference travel time. The estimates in M4 are similar to the estimates in M1 except $\beta_{toll\_medium}$ and $\beta_{TT(loss)}$. Likewise, M5 provides similar estimates as M2 except the estimated coefficients of travel time. It should be noted that estimates of travel time and toll of class 1 in M6 are smaller than the corresponding values in the other models, while these estimates of class 2 are significantly above the average level.

## A.5 Empirical results analysis

In terms of goodness of fit, the likelihood ratio test value of M6 and the base model M1 is 10.4 which give a p-value of 0.1 on $\chi^2$ distribution, suggesting a marginal improvement of model fit. M1 provides the lowest likelihood but also a parsimonious structure. It is difficult, therefore, to measure the performances of each model simply by the value of log-likelihood. Instead, appropriate comparison criteria are required to analyse these candidate models.

Hence, as shown in Table A.4, we use a series of test statistics to compare our candidate models.

| Models | TEST CRITERIONS | | | | | |
|---|---|---|---|---|---|---|
| | $\boldsymbol{\rho^2(ASC)}$ | $\boldsymbol{\bar{\rho}^2(ASC)}$ | *AIC* | *BIC* | *Consistent AIC* | *Corrected AIC* |
| M1 | 0.086 | 0.058 | 474.828 | *503.404* | *510.404* | 477.178 |
| M2 | 0.096 | *0.064* | *471.832* | 504.490 | 512.490 | *475.196* |
| M3 | 0.095 | 0.027 | 490.512 | 559.910 | 576.910 | 518.264 |
| M4 | 0.095 | 0.028 | 490.300 | 559.698 | 576.698 | 518.052 |
| M5 | 0.097 | 0.029 | 489.642 | 559.040 | 576.040 | 517.394 |
| M6 | *0.107* | 0.039 | 484.432 | 553.830 | 570.830 | 512.184 |

**Table A.4: The model test results**

Only M6 provides significant improvement in model fit with respect to the base model M1. The parsimonious models of M1 and M2, as shown in $\bar{\rho}^2(ASC)$ and all the information criteria statistics, are superior if the number of parameters is incorporated into the test criteria. Consequently, from a model fit point of view, M6 is better, albeit with the cost of introducing more coefficients.

We now compare the computational speed of these candidate models. M1 and M2 were done using the BIOGEME, and all the EM recursion embedded models were conducted by using Matlab. As a result, the speed to convergence is dramatically different depending on the complexity of model specifications. M1 requires the least running time, while M6 takes more than 30 minutes. Figure A.3 gives the estimation results of EM recursion embedded PT models. Each point in the figure represents an iteration of the proposed approaches, and the vertical axis represents the values of log-likelihood function. In line with our expectations, all the algorithms turn out to monotonically increase log-likelihood, although it moves very slowly near the likelihood maximum.

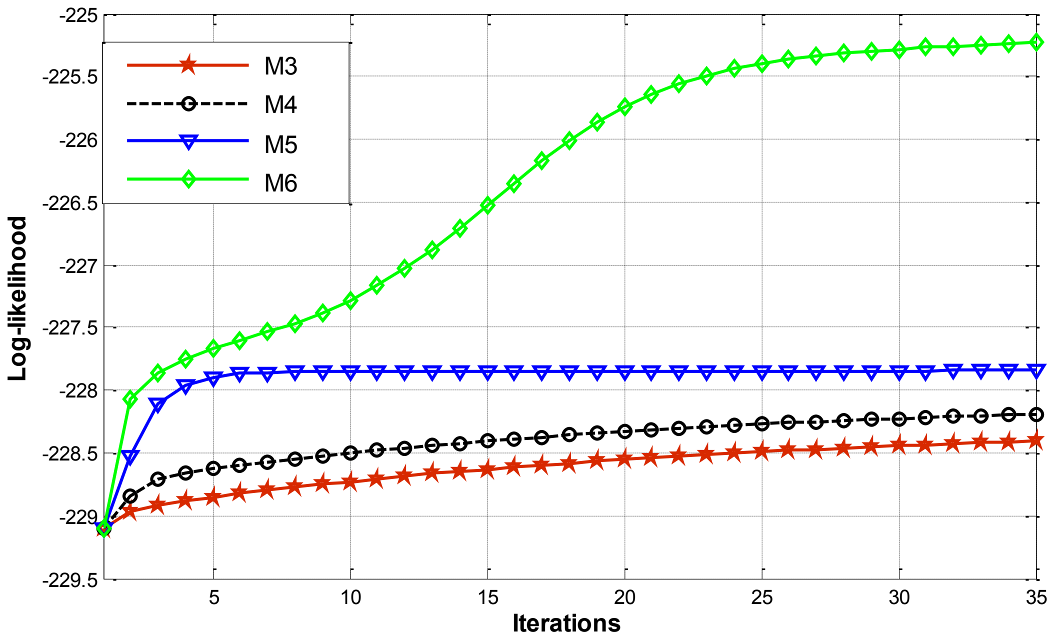



**Figure A.3: The log-likelihood of EM-embedded models at each iteration**

M3, M4 and M5 show a concave pattern, i.e. these algorithms have the largest increase of log-likelihood before the 5th iteration, and the rate of increase decreases thereafter. It is easy to see that M5 takes the least iterations to convergence. In the contrast, M6, after the leap in the first iteration, has a stable S-shaped pattern which is consistent with the findings of Bhat (1997). It should be noted that all the algorithms used 0.01 difference between successive lower-bound values of log-likelihood as the criterion of convergence. Less running time is required if we increase the convergence criterion, although the S-shaped pattern of M6 suggests that we should be extremely cautious about selecting a bigger value of convergence criterion. For instance, the M6 algorithm might mistakenly stop at the first concave segment of the S curve if an inappropriate convergence criterion is adopted. The issue regarding the convergence of the EM algorithm is an urgent area for future research.

Estimation results consistently show individuals' asymmetric preference towards travel time. The ratio between the travel time estimates is $\frac{|\beta_{TT(loss)}|}{|\beta_{TT(gain)}|} > 1$ which suggests that the influence of travel time loss is much more significant than the influence of gain on the utility of alternative. As an illustration of the asymmetries, Figure A.4 shows the impact of travel time changes on utility.

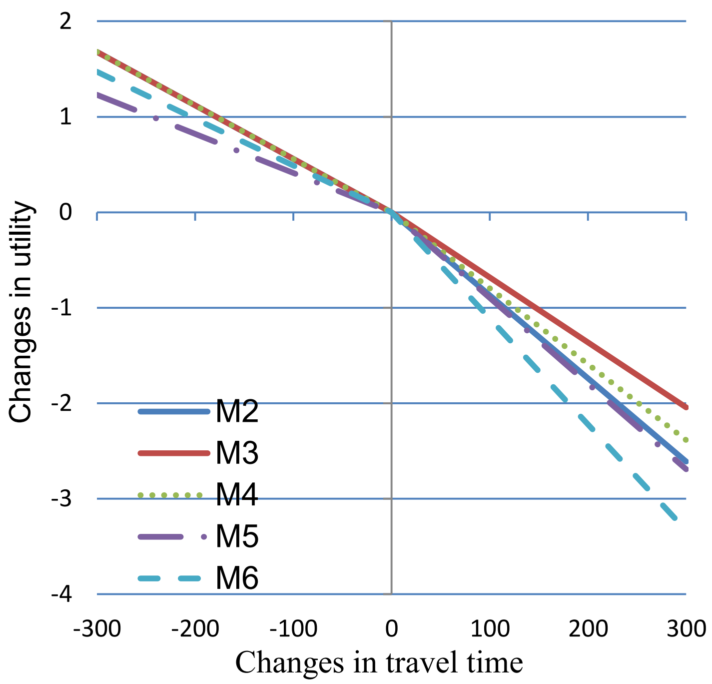



**Figure A.4 Evidence of asymmetrical response to changes of travel time**

## A.6 Conclusion

This current study has developed PT models with an EM algorithm to capture the value and distribution of travellers' reference travel time in the revealed preference context of a binary route choice between a tolled route and an untolled route. The proposed EM embedded PT model is capable of incorporating multiple reference points, namely natural reference travel time and estimated reference travel time.

Within this research, the basic assumption of decision making under risk is that travellers are aware of the travel time distribution. This is reasonable if decision makers are experienced commuters in the revealed preference context. One could argue, however, (as Henn and Ottomanelli (2006) have done) that it is more realistic to assume that the distribution is not completely known to all travellers, which is called 'decision making under uncertainty'. It is therefore of interest to extend the risky choice model to choice under uncertainty and apply it to transport. Moreover, it could also be interesting to compare the performance of the same risky choice model using different types of data, e.g. SP data, RP data and simulation data.

The empirical evidence explored here reinforces the advantage of the EM algorithm as a theoretically feasible and computationally attractive procedure for the estimation of a PT model with multiple reference points. This approach would be a possible breakthrough for the investigation of heterogeneity in terms of reference points. After systematic model comparison, we found that the PT-EM model with estimated coefficients and shares provided the best goodness of fit. Researchers should be cautious about using reference dependence, however, in that the reference point is likely to be context-dependent as well as individual-specific. Furthermore, the reference point is dynamic in some cases, for instance, the last property condition might be the reference point for the dwelling choice of the current house, while the current property condition might be the next reference point to select the future house.

Finally, instead of assuming decision makers as 'utility machines' who tend to maximize their utility, it is of interest to investigate alternative models with different decision rules. For instance, Bettman et al. (1991) introduced seven types of decision heuristics which have been empirically identified in consumer choice studies. Sugden (2004) and Starmer (2000) also summarized a number of non-EUT models, including regret theory, rank-dependent utility theory, theory of disappointment and prospective reference theory.

# Appendix B: RODS QUESTIONNAIRE

Transport for London

**London Underground**

## London Travel Survey

### 5 minutes of your time will help us to improve your journey

This survey is about the journey you were actually making when you were handed this questionnaire. For example, if you were on your way home from work we want to know about where you started your journey (work) and where you ended your journey (home).

The information you provide will tell us how public transport is being used in London and will help us to improve our service for you.

The questionnaire will only take **5 minutes to complete**. When complete, please return it in the FREEPOST envelope provided.

**Or if you prefer visit www.london-travel-survey.co.uk and enter your unique serial number (as shown above).**

**Thank you very much for your help.**

---

Please answer the questions below about the journey you were making **WHEN YOU WERE HANDED THIS QUESTIONNAIRE.**

Please either tick the relevant box ☑ or write in the appropriate answer e.g. 9 . 3 7 AM

### Section 1: Your journey to the Underground

**1** Which of the following options below best describes where you have just come from?

**Please tick ONE box only**

- Home ☐ 1
- Normal workplace ☐ 2
- Other workplace/business meeting ☐ 3
- Visiting friends/relatives/on holiday ☐ 4
- Theatre/cinema/concert/sporting activity/ event/other social (e.g. restaurant, pub) ☐ 5
- Museum/exhibition ☐ 6
- Shopping ☐ 7
- School/college/university (as student) ☐ 8
- School/college (accompanying pupil) ☐ 9
- Taking someone to airport, station, hotel etc. ☐ 10
- Meeting someone at airport, station, hotel etc ☐ 11
- Personal business (e.g. doctor, hospital, bank) ☐ 12
- Sightseeing ☐ 13
- Hotel/guest house etc ☐ 14
- Other (please tick and write in) ☐ 15

**2** It would help us if you were willing to enter the address of the place where you **started this journey**. This information is used to plan station entrances and exits and will not be used for marketing purposes.

**Please give us as much information as possible**

Name of shop/hotel etc. (if appropriate)

Street & number

District/Town Postcode

**3** At what time did you set out on this journey?

AM PM

**MAYOR OF LONDON**

**4** At what time did you actually reach Whitechapel station?

AM PM

**5** How did you get to Whitechapel station from the place mentioned in Q2?
**Please complete ONE box only**
**If you used more than one type of transport please complete the MAIN method used**

National Rail (⇌) - Please give origin station 1

Docklands Light Railway - Please give origin station 2

Bus - Please give bus route number 3

Tram - Please give origin station 4

Another Underground train (⊖) - Please give origin station 5

| | | | |
|---|---|---|---|
| Car/van - parked at/near station | ☐ 6 | Air | ☐ 11 |
| Car/van - dropped off | ☐ 7 | Taxi/minicab | ☐ 12 |
| Coach/workbus | ☐ 8 | Walked all the way from the start | ☐ 13 |
| Motorcycle | ☐ 9 | Boat | ☐ 14 |
| Bicycle | ☐ 10 | Other (please tick and write in) | ☐ 15 |

## Section 2: Your journey from Whitechapel station

**6** On this journey which train service did you use from Whitechapel station?
**Please tick ONE box only and continue to Q7 unless otherwise indicated**

District Line ☐
Hammersmith & City Line ☐
London Overground ☐
None ☐ go to Q9

**7** At which National Rail, DLR or Underground station did you finish your journey?
**Please write in the name of the National Rail (⇌), DLR or Underground (⊖) station where you ended your journey**

**8** Please write in the name(s) of all the Underground (⊖), DLR and National Rail (⇌) station(s) where you changed trains during your journey from Whitechapel station.
**Please leave blank if you did not make any changes**

| First change at | Second change at | Third change at | Fourth change at |
|---|---|---|---|
| | | | |

**9** When you arrived at your destination station, how did you complete the journey to your destination address?
**Please complete ONE box only**
**If you used more than one type of transport please complete the MAIN method used**

Bus - Please give bus route number 3

Trams - Please give destination station 4

| | | | |
|---|---|---|---|
| Car/van - parked at /near station | ☐ 6 | Air | ☐ 11 |
| Car/van - picked up | ☐ 7 | Taxi/minicab | ☐ 12 |
| Coach | ☐ 8 | Walked all the way from the station | ☐ 13 |
| Motorcycle | ☐ 9 | Boat | ☐ 14 |
| Bicycle | ☐ 10 | Other (please tick and write in) | ☐ 15 |

**10** Why were you travelling to this place/destination? **Please tick ONE box only**

- Going home ☐ 1
- Going to normal workplace ☐ 2
- Going to other workplace/business meeting ☐ 3
- Visiting friends/relatives/on holiday ☐ 4
- Theatre/cinema/concert/sporting activity/ event/other social (e.g. restaurant, pub) ☐ 5
- Going to a museum/exhibition ☐ 6
- Going shopping ☐ 7
- Going to school/college/university (as student) ☐ 8
- Accompanying pupil to/from school/college ☐ 9
- Taking someone to airport/station/hotel etc ☐ 10
- Meeting someone at airport, station, hotel etc ☐ 11
- Personal business (e.g. doctor, hospital, bank) ☐ 12
- Going sightseeing ☐ 13
- Going to hotel/guest house etc ☐ 14
- Other (please tick and write in) ☐ 15

**11** Were you carrying anything during this journey?

- None ☐ 1
- Walking aids (wheelchair, walking stick/frame) ☐ 2
- Briefcase/laptop case/shopping ☐ 3
- Suitcase or rucksack ☐ 4
- Pushchair/buggy/pram ☐ 5
- Sports bag, sports kit etc ☐ 6
- Tools or other work equipment ☐ 7
- Other bulky or heavy item ☐ 8

**12** It would help us if you were willing to enter the address of the place you were **travelling to**. This information is used to plan station entrances and exits and will not be used for marketing purposes.
**Please give us as much information as possible**

Name of shop/hotel etc. (if appropriate)

Street & number

District/Town

Postcode

## Section 3: Ticket type

**13** **Did you use an Oyster card?**

Yes ☐ No ☐

If yes, please enter your twelve digit Oyster Card number below (please note that this number can be found on the reverse of your card, in the top right hand corner).

**14** What type of ticket, Oyster card or pass did you use for the Underground part of this journey?
**Please tick ONE box only**

**Oyster/Travelcards**

- Oyster PAYG ☐ 1
- One Day Travelcard (anytime) ☐ 2
- One Day Travelcard (off-peak) ☐ 3
- Weekly Oyster/Travelcard ☐ 4
- Monthly Oyster/Travelcard ☐ 5
- Annual Oyster/Travelcard ☐ 6
- Other length Travelcard ☐ 7
- 16+ Zip Oystercard ☐ 8
- 18+ Zip Oystercard ☐ 9

**Paper Tickets**

- Single ticket ☐ 10
- Return ticket ☐ 11
- Extension ticket ☐ 12
- Privilege pass ☐ 13

**Passes/Permits**

- Elderly Freedom Pass ☐ 14
- Disabled Freedom Pass ☐ 15
- TfL/NR Staff/nominee staff pass ☐ 16
- Police pass ☐ 17
- **Travelling using NR only** ☐ 18
- **Other** (please write in) 19

**15** **Travelcard Users Only**, otherwise go to **Q16**
**Please tick ALL zones covered by your Travelcard**

| Zone 1 | Zone 2 | Zone 3 | Zone 4 | Zone 5 | Zone 6 | Outer zone station |
|---|---|---|---|---|---|---|
| | | | | | | |

## Section 4: Background information

**We need to finish this questionnaire by collecting some background information about you.**

**16** How often do you make this particular journey? **Please tick ONE box only**

| 5 or more days a week | 1-4 days days a week | Once a fortnight | Once a month | Less than once a month | First time ever |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |

**17** And are you Male .......................................................... ☐ Female .......................................................... ☐

**18** What age were you on your last birthday? **Please tick ONE box only**

| Under 16 | 16 - 19 | 20 - 24 | 25 - 34 | 35 - 44 | 45 - 59 | 60 - 64 | 65 - 70 | 71 - 74 | Over 74 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

**19** Do you have any long-term physical or mental disability which limits your daily activities or the work you can do, including problems due to age?

| Mobility impairment | Hearing impairment | Mental health condition | None |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| **Visual impairment** | **Learning disability** | **Serious long-term illness** | **Other** |
| 5 | 6 | 7 | 8 |

**20** **If the journey you are answering about does not start or finish at home, please write in where you normally live or the postcode.** This information will not be used for marketing purposes.

Street & number

District/Town

If outside UK, please give country Postcode

**Thank you for taking the time to complete this questionnaire.**

The questionnaire is to be returned in the envelope provided. If you lose your envelope please send the completed questionnaire to:

**Freepost Plus RSJS-GCLK-ZYCU**
London Travel Survey (RODS), Kantar Operations Data Centre
Ealing Gateway, 26-30 Uxbridge Road, LONDON W5 2AU

No stamp is required.

Please be assured that your responses will be treated as strictly confidential and used by London Underground for planning purposes only.

Yours Faithfully,

Richard Parry,
Director of Strategy and Commercial
London Underground Limited

Yours Faithfully,

Peter Field,
Director of London Rail Development
London Rail

References


Abdel-Aty, M. A., Kitamura, R. and Jovanis, P. P. (1995), 'Investigating effect of travel time variability on route choice using repeated-measurement stated preference data', *Transportation Research Record***,** 39-45.

Abdellaoui, M. (2000), 'Parameter-free elicitation of utility and probability weighting functions', *Management Science,* **46,** 1497-1512.

Abeler, J., Falk, A., Goette, L. and Huffman, D. (2011), 'Reference points and effort provision', *The American Economic Review,* **101,** 470-492.

Abrantes, P. A. and Wardman, M. R. (2011), 'Meta-analysis of UK values of travel time: An update', *Transportation Research Part A: Policy and Practice,* **45,** 1-17.

Accent Marketing & Research, H. C. G. (1996), 'The Value of Travel Time on UK Roads-1994', *Final Report prepared for Department of Transport, UK*.

Aitkin, M. and Aitkin, I. (1996), 'A hybrid EM/Gauss-Newton algorithm for maximum likelihood in mixture distributions', *Statistics and Computing,* **6,** 127-130.

Akaike, H. (1974), 'A new look at the statistical model identification', *Automatic Control, IEEE Transactions on,* **19,** 716-723.

Aliev, R., Pedrycz, W., Fazlollahi, B., Huseynov, O. H., Alizadeh, A. and Guirimov, B. (2012), 'Fuzzy logic-based generalized decision theory with imperfect information', *Information Sciences,* **189,** 18-42.

Allais, M. (1953), 'Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine', *Econometrica: Journal of the Econometric Society,* **21,** 503-546.

Allais, M. (1979), 'The so-called Allais paradox and rational decisions under uncertainty', *Expected utility hypotheses and the Allais Paradox: contemporary discussions of decisions under uncertainty with Allais' rejoinder***,** 437.

Arrow, K. J. (1965) *Aspects of the theory of risk-bearing,* Yrjö Jahnssonin Säätiö.

Asensio, J. and Matas, A. (2008), 'Commuters' valuation of travel time variability', *Transportation Research Part E: Logistics and Transportation Review,* **44,** 1074-1085.

ATOC (2005) *Passenger Demand Forecasting Handbook,* Association of Train Operating Companies, London.

Avineri, E. (2009), 'Fuzzy Meaning of Reference-Based Perceptions in Travel Choice Modeling', *Transportation Research Board 88th Annual Meeting*.

Avineri, E. and Bovy, P. H. L. (2008), 'Identification of parameters for a prospect theory model for travel choice analysis', *Transportation Research Record: Journal of the Transportation Research Board,* **2082,** 141-147.

Avineri, E. and Prashker, J. N. (2003), 'Sensitivity to uncertainty: Need for a paradigm shift', *Transportation Research Record: Journal of the Transportation Research Board,* **1854,** 90-98.

Avineri, E. and Prashker, J. N. (2005), 'Sensitivity to travel time variability: Travelers' learning perspective', *Transportation Research Part C: Emerging Technologies,* **13,** 157-183.

Bates, J., Polak, J., Jones, P. and Cook, A. (2001), 'The valuation of reliability for personal travel', *Transportation Research Part E: Logistics and Transportation Review,* **37,** 191-229.

Bates, J. and Whelan, G. (2001) *Size and Sign of Time Savings,* Institute of Transport Studies, University of Leeds , Leeds, UK.

Batley, R. (2007), 'Marginal valuations of travel time and scheduling, and the reliability premium', *Transportation Research Part E: Logistics and Transportation Review,* **43,** 387-408.
Batley, R. (2008), 'On ordinal utility, cardinal utility and random utility', *Theory and decision,* **64,** 37-63.
Batley, R. and Daly, A. (2004), 'Departure time choice under uncertainty: a Model of random utility maximisation for wagers', *European Transport Conference*.
Batley, R., Grant-Muller, S., Nellthorp, J., de Jong, G., Watling, D., HESS, S. and POLAK, J. (2008), 'Multimodal travel time variability-', *Final Report, Department of Transport,*.
Batley, R. and Ibáñez, J. N. (2012), 'Randomness in preference orderings, outcomes and attribute tastes: An application to journey time risk', *Journal of Choice Modelling,* **5,** 157-175.
Batley, R. and Ibáñez, N. (2009) In *International Choice Modelling Conference*.
Becker, G. S. (1965), 'A Theory of the Allocation of Time', *The Economic Journal,* **75,** 493-517.
Becker, J. L. and Sarin, R. K. (1987), 'Gamble dependent utility', *Management Science,* **33,** 1367-1382.
Bell, D. E. (1982), 'Regret in decision making under uncertainty', *Operations research,* **30,** 961-981.
Bell, D. E. (1985), 'Disappointment in decision making under uncertainty', *Operations research,* **33,** 1-27.
Ben-Akiva, M. and Bierlaire, M. (1999) In *Handbook of Transportation Science*Kluwer Academic Publishers, Norwell, Massachusetts.
Ben-Akiva, M. and Lerman, S. (1985) *Discrete choice analysis: theory and application to travel demand,* MIT press.
Benartzi, S. and Thaler, R. H. (1993), 'Myopic loss aversion and the equity premium puzzle', *National Bureau of Economic Research*.
Bettman, J. R., Johnson, E. J. and Payne, J. W. (1991), 'Consumer decision making', *Handbook of consumer behavior***,** 50-84.
Bhat, C. R. (1997), 'An endogenous segmentation mode choice model with an application to intercity travel', *Transportation science,* **31,** 34-48.
Bhat, C. R. and Pulugurta, V. (1998), 'A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions', *Transportation Research Part B: Methodological,* **32,** 61-75.
Bhat, C. R. and Sardesai, R. (2006), 'The impact of stop-making and travel time reliability on commute mode choice', *Transportation Research Part B: Methodological,* **40,** 709-730.
Bierlaire, M. (2003), 'BIOGEME: a free package for the estimation of discrete choice models', *Presented at the 3rd Swiss Transportation Research Conference, Ascona.* .
Bilmes, J. A. (1998), 'A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models', *International Computer Science Institute,* **4,** 126.
Birnbaum, M. H. and Chavez, A. (1997), 'Tests of theories of decision making: Violations of branch independence and distribution independence', *Organizational Behavior and Human Decision Processes,* **71,** 161-194.
Black, I. G., Towriss, J. G., Britain, G., Cranfield Institute of Technology.Centre for, L., Transportation and Limited, P. A. S. (1993) *Demand effects of travel time reliability,* Dept. of Transport.
Block, H. D. and Marschak, J. (1960), 'Random orderings and stochastic theories of responses', *Contributions to probability and statistics,* **2,** 97-132.

Booij, A. S. and Van de Kuilen, G. (2009), 'A parameter-free analysis of the utility of money for the general population under prospect theory', *Journal of Economic Psychology,* **30,** 651-666.
Boyce, D. E., Ran, B. and Li, I. Y. (1999), 'CONSIDERING TRAVELERS'RISK-TAKING BEHAVIOR IN DYNAMIC TRAFFIC ASSIGNMENT', *Transportation Networks: Recent Methodological Advances. Selected Proceedings of the 4th EURO Transportation Meeting*.
Boyd, J. H. and Mellman, R. E. (1980), 'The effect of fuel economy standards on the US automotive market: an hedonic demand analysis', *Transportation Research Part A: General,* **14,** 367-378.
Breiman, L. and Spector, P. (1992), 'Submodel selection and evaluation in regression. the x-random case', *International statistical review/revue internationale de Statistique***,** 291-319.
Brooks, P. and Zank, H. (2005), 'Loss averse behavior', *Journal of Risk and Uncertainty,* **31,** 301-325.
Brown, T. C. (2005), 'Loss aversion without the endowment effect, and other explanations for the WTA–WTP disparity', *Journal of Economic Behavior & Organization,* **57,** 367-379.
Brownstone, D., Ghosh, A., Golob, T. F., Kazimi, C. and Van Amelsfort, D. (2003), 'Drivers' willingness-to-pay to reduce travel time: evidence from the San Diego I-15 congestion pricing project', *Transportation Research Part A: Policy and Practice,* **37,** 373-387.
Brownstone, D. and Small, K. A. (2005), 'Valuing time and reliability: assessing the evidence from road pricing demonstrations', *Transportation Research Part A: Policy and Practice,* **39,** 279-293.
Camerer, C. F. and Ho, T. H. (1994), 'Violations of the betweenness axiom and nonlinearity in probability', *Journal of Risk and Uncertainty,* **8,** 167-196.
Camerer, C. F., Loewenstein, G. and Rabin, M. (2011) *Advances in behavioral economics,* Princeton University Press.
Carrion, M. and Levinson (2010), 'Value of Reliability: High Occupancy Toll Lanes, General Purpose Lanes, and Arterials', *Presented at the 4th International Symposium on Transportation Network Reliability, July 2010, Minneapolis, MN.*
Cavagnaro, D. R., Gonzalez, R., Myung, J. I. and Pitt, M. A. (2013), 'Optimal decision stimuli for risky choice experiments: An adaptive approach', *Management Science,* **59,** 358-375.
Chew, S. H. (1983), 'A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox', *Econometrica: Journal of the Econometric Society***,** 1065-1092.
Chew, S. H. (1989), 'Axiomatic utility theories with the betweenness property', *Annals of Operations Research,* **19,** 273-298.
Chew, S. H., Epstein, L. G. and Segal, U. (1991), 'Mixture symmetry and quadratic utility', *Econometrica: Journal of the Econometric Society,* **59,** 139-163.
Chew, S. H. and MacCrimmon, K. (1979), 'Alpha-nu Choice Theory: a Generalisation of Expected Utility Theory', *working paper 669 U.British Columbia*.
Chorus, C. G. (2011), 'Regret theory-based route choices and traffic equilibria', *Transportmetrica***,** 1-15.
Chorus, C. G., Arentze, T. A., Molin, E. J. E., Timmermans, H. J. P. and Van Wee, B. (2006a), 'The value of travel information: Decision strategy-specific conceptualizations and numerical examples', *Transportation Research Part B: Methodological,* **40,** 504-519.

Chorus, C. G., Arentze, T. A. and Timmermans, H. J. P. (2008), 'A random regret-minimization model of travel choice', *Transportation Research Part B: Methodological,* **42,** 1-18.

Chorus, C. G., Molin, E. J. E., Arentze, T. A., Hoogendoorn, S. P., Timmermans, H. J. P. and Van Wee, B. (2007), 'Validation of a multimodal travel simulator with travel information provision', *Transportation Research Part C: Emerging Technologies,* **15,** 191-207.

Chorus, C. G., Molin, E. J. E., Van Wee, B., Arentze, T. A. and Timmermans, H. J. P. (2006b), 'Responses to transit information among car-drivers: regret-based models and simulations', *Transportation Planning and Technology,* **29,** 249-271.

Chorus, C. G., Walker, J. L. and Ben-Akiva, M. E. (2010), 'The value of travel information: A search-theoretic approach', *Journal of Intelligent Transportation Systems,* **14,** 154-165.

Conlisk, J. (1996), 'Why bounded rationality?', *Journal of Economic Literature***,** 669-700.

Cox, J. C., Sadiraj, V. and Schmidt, U. (2011), 'Paradoxes and mechanisms for choice under risk', *Kiel Working Papers*.

Crawford, M. T., McConnell, A. R., Lewis, A. C. and Sherman, S. J. (2002), 'Reactance, compliance, and anticipated regret', *Journal of Experimental Social Psychology,* **38,** 56-63.

Cyert, R. M. and March, J. G. (1963), 'A behavioral theory of the firm', *Englewood Cliffs, NJ,* **2**.

Daganzo, C. F. (1997), 'Fundamentals of transportation and traffic operations', *New York: Pergamon*.

Danielis, R. and Marcucci, E. (2007), 'Attribute cut-offs in freight service selection', *Transportation Research Part E: Logistics and Transportation Review,* **43,** 506-515.

Davidson, R. and MacKinnon, J. G. (1981), 'Several tests for model specification in the presence of alternative hypotheses', *Econometrica: Journal of the Econometric Society***,** 781-793.

De Borger, B. and Fosgerau, M. (2008), 'The trade-off between money and travel time: A test of the theory of reference-dependent preferences', *Journal of urban economics,* **64,** 101-115.

De Jong, G., Kroes, E., Plasmeijer, R., Sanders, P. and Warffemius, P. (2004) In *Proceedings of the European Transport Conference*.

De Lapparent, M. (2010), 'Attitude toward risk of time loss in travel activity and air route choices', *Journal of Intelligent Transportation Systems,* **14,** 166-178.

De Palma, A., Ben-Akiva, M., Brownstone, D., Holt, C., Magnac, T., McFadden, D., Moffatt, P., Picard, N., Train, K. and Wakker, P. (2008), 'Risk, uncertainty and discrete choice models', *Marketing Letters,* **19,** 269-285.

De Palma, A. and Picard, N. (2005), 'Route choice decision under travel time uncertainty', *Transportation Research Part A: Policy and Practice,* **39,** 295-324.

Dekel, E. (1986), 'An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom', *Journal of Economic Theory,* **40,** 304-318.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), 'Maximum likelihood from incomplete data via the EM algorithm', *Journal of the Royal Statistical Society. Series B (Methodological)***,** 1-38.

DeSerpa, A. C. (1971), 'A theory of the economics of time', *The Economic Journal,* **81,** 828-846.

Diecidue, E. and Wakker, P. P. (2001), 'On the intuition of rank-dependent utility', *Journal of Risk and Uncertainty,* **23,** 281-298.

Domenich, T. and McFadden, D. (1975), 'Urban Travel Demand: A Behavioral Approach', *North-Holland, Amsterdam*.
Douglas, M. (2013) *Risk and Acceptability,* Routledge.
Edelman, M. (1985) *The symbolic uses of politics,* University of Illinois Press.
Evans, A. W. (1972), 'On the theory of the valuation and allocation of time*', *Scottish Journal of Political Economy,* **19,** 1-17.
Farquharson, R. (1969) *Theory of voting,* Yale University Press New Haven, CT.
Fennema, H. and Wakker, P. P. (1997), 'Original and cumulative prospect theory: A discussion of empirical differences', *Journal of Behavioral Decision Making,* **10,** 53-64.
Fishburn, P. C. (1970), 'Utility theory for decision making', *New York: John Wiley & Sons*.
Fishburn, P. C. (1978), 'On Handa's" New theory of cardinal utility" and the maximization of expected return', *The Journal of Political Economy,* **86,** 321-324.
Fishburn, P. C. (1982), 'Nontransitive measurable utility', *Journal of Mathematical Psychology,* **26,** 31-67.
Gachter, S., Johnson, E. and Herrmann, A. (2007), 'Individual-level loss aversion in riskless and risky choices'.
Gao, S., Frejinger, E. and Ben-Akiva, M. (2010), 'Adaptive route choices in risky traffic networks: A prospect theory approach', *Transportation Research Part C: Emerging Technologies,* **18,** 727-740.
Gaver, D. P. (1968), 'Headstart strategies for combating congestion', *Transportation science,* **2,** 172.
Ghosh, A. (2001), 'Valuing time and reliability: commuters' mode choice from a real time congestion pricing experiment', *Ph.D. Dissertation, Department of Economics, University of California at Irvine*.
Gigerenzer, G. and Todd, P. M. (1999), 'Fast and frugal heuristics: The adaptive toolbox', *In G. Gigerenzer, P.M. Todd, and the ABC Research Group, Simple heuristics that make us smart(pp. 3–34). New York:Oxford University Press.*
Gonzalez, R. and Wu, G. (1999), 'On the shape of the probability weighting function', *Cogn Psychol,* **38,** 129-66.
Gul, F. (1991), 'A theory of disappointment aversion', *Econometrica: Journal of the Econometric Society,* **59,** 667-686.
Gupta, S. and Chintagunta, P. K. (1994), 'On using demographic variables to determine segment membership in logit mixture models', *Journal of Marketing Research***,** 128-136.
Habib, M. A. and Miller, E. J. (2009), 'Reference-dependent residential location choice model within a relocation context', *Transportation Research Record: Journal of the Transportation Research Board,* **2133,** 92-99.
Hagen, O. (1979), 'Towards a Positive Theory of Preferences Under Risk', *Expected utility hypotheses and the Allais Paradox: contemporary discussions of decisions under uncertainty with Allais' rejoinder***,** 271.
Handa, J. (1977), 'Risk, probabilities, and a new theory of cardinal utility', *The Journal of Political Economy,* **85,** 97-122.
Hargreaves Heap, S. (1992) *The theory of choice: a critical guide,* Blackwell.
Harrison, G. W. and Rutström, E. E. (2008), 'Risk aversion in the laboratory', *In: Cox, J.C., Harrison, G.W. (Eds.), Risk Aversion in Experiments, Research in Experimental Economics, vol. 12. Emerald, Bingley, UK*.
Harrison, G. W. and Rutström, E. E. (2009), 'Expected utility theory and prospect theory: One wedding and a decent funeral', *Experimental Economics,* **12,** 133-158.

Heath, C., Larrick, R. P. and Wu, G. (1999), 'Goals as reference points', *Cognitive Psychology,* **38,** 79-109.
Henn, V. and Ottomanelli, M. (2006), 'Handling uncertainty in route choice models: from probabilistic to possibilistic approaches', *European journal of operational research,* **175,** 1526-1538.
Hensher, D. A. (2001), 'Measurement of the valuation of travel time savings', *Journal of Transport Economics and Policy (JTEP),* **35,** 71-98.
Hensher, D. A. and Greene, W. H. (2003), 'The mixed logit model: the state of practice', *Transportation,* **30,** 133-176.
Hensher, D. A., Greene, W. H. and 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,** 954-972.
Hensher, D. A. and Li, Z. (2012), 'Valuing Travel Time Variability within a Rank-Dependent Utility Framework and an Investigation of Unobserved Taste Heterogeneity', *Journal of Transport Economics and Policy (JTEP),* **Forthcoming**.
Herstein, I. N. and Milnor, J. (1953), 'An axiomatic approach to measurable utility', *Econometrica, Journal of the Econometric Society***,** 291-297.
Hess, S., Bierlaire, M. and Polak, J. W. (2005), 'Estimation of value of travel-time savings using mixed logit models', *Transportation Research Part A: Policy and Practice,* **39,** 221-236.
Hess, S., Rose, J. M. and Hensher, D. A. (2008), 'Asymmetric preference formation in willingness to pay estimates in discrete choice models', *Transportation Research Part E: Logistics and Transportation Review,* **44,** 847-863.
Hess, S., Stathopoulos, A. and Daly, A. (2012), 'Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies', *Transportation,* **39,** 565-591.
Hickey, S. W. (2011) In *Dept. of Civil and Environmental Engineering*, Vol. S.M. in Transportation Massachusetts Institute of Technology. .
Hodgson, G. M. (2012), 'On the limits of rational choice theory', *Economic Thought,* **1**.
Hollander, Y. (2006), 'Direct versus indirect models for the effects of unreliability', *Transportation Research Part A: Policy and Practice,* **40,** 699-711.
Horowitz, J. K. and McConnell, K. E. (2002), 'A review of WTA/WTP studies', *Journal of Environmental Economics and Management,* **44,** 426-447.
Horowitz, J. L. (1983), 'Statistical comparison of non-nested probabilistic discrete choice models', *Transportation science,* **17,** 319-350.
Horsky, D., Misra, S. and Nelson, P. (2006), 'Observed and unobserved preference heterogeneity in brand-choice models', *Marketing Science***,** 322-335.
Hu, G., Sivakumar, A. and Polak, J. W. (2012a) In *The 87th Annual Meeting of Transportation Research Board*Washington D.C., USA.
Hu, G., Sivakumar, A. and Polak, J. W. (2012b), 'Modelling travellers' risky choice in a revealed preference context: a comparison of EUT and non-EUT approaches', *Transportation***,** 1-17.
Hurvich, C. M. and Tsai, C. L. (1989), 'Regression and time series model selection in small samples', *Biometrika,* **76,** 297-307.
Jackson, W. B. and Jucker, J. V. (1982), 'An empirical study of travel time variability and travel choice behavior', *Transportation science,* **16,** 460.
Jara-Díaz, S. R. (2000), 'Allocation and valuation of travel time savings', *Handbooks in Transport,* **1,** 303-319.
Johnson, E. J., Gächter, S. and Herrmann, A. (2006), 'Exploring the nature of loss aversion', *IZA Discussion Papers No. 2015*.

Johnson, M. B. (1966), 'Travel time and the price of leisure', *Economic Inquiry,* **4,** 135-145.
Johnson, N. L. and Kotz, S. (1969) *Distribution in statistics: Discrete distribution,* Wiley.
Jou, R.-C., Kitamura, R., Weng, M.-C. and Chen, C.-C. (2008), 'Dynamic commuter departure time choice under uncertainty', *Transportation Research Part A: Policy and Practice,* **42,** 774-783.
Jou, R. and Kitamura, R. (2002), 'Commuter Departure Time Choice: A Reference Point Approach', *Mimeograph*.
Kahneman, D. (2011) *Thinking, fast and slow,* Macmillan.
Kahneman, D. and Frederick, S. (2002), 'Representativeness revisited: Attribute substitution in intuitive judgment', *Heuristics and biases: The psychology of intuitive judgment***,** 49-81.
Kahneman, D., Knetsch, J. L. and Thaler, R. H. (1991), 'Anomalies: The endowment effect, loss aversion, and status quo bias', *The Journal of Economic Perspectives***,** 193-206.
Kahneman, D. and Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica: Journal of the Econometric Society,* **47,** 263-291.
Knetsch, J. L. (1989), 'The endowment effect and evidence of nonreversible indifference curves', *The American Economic Review,* **79,** 1277-1284.
Knight, T. E. (1974), 'An approach to the evaluation of changes in travel unreliability: A "Safety margin" hypothesis', *Transportation,* **3,** 393-408.
Kohavi, R. (1995), 'A study of cross-validation and bootstrap for accuracy estimation and model selection', *In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (pp. 1137–1143). San Francisco, CA: Morgan Kaufmann,* **14,** 1137-1145.
Koster, P. and Verhoef, E. T. (2010), 'A Rank Dependent Scheduling Model', *Tinbergen Institute Discussion Papers*.
Kőszegi, B. and Rabin, M. (2006), 'A model of reference-dependent preferences', *The Quarterly Journal of Economics,* **121,** 1133-1165.
Lam, T. C. and Small, K. A. (2001), 'The value of time and reliability: measurement from a value pricing experiment', *Transportation Research Part E: Logistics and Transportation Review,* **37,** 231-251.
Lancaster, K. J. (1966), 'A new approach to consumer theory', *The Journal of Political Economy,* **74,** 132-157.
Li, Z. and Hensher, D. A. (2012a), 'Behavioural implications of preferences, risk attitudes and beliefs in modelling risky travel choice with travel time variability', *Transportation***,** 1-19.
Li, Z. and Hensher, D. A. (2012b), 'Estimating values of travel time savings for toll roads: Avoiding a common error', *Transport Policy,* **24,** 60-66.
Li, Z., Hensher, D. A. and Rose, J. M. (2009), 'Valuation of Travel Time Reliability in an Extended Expected Utility Theory Framework', *12th International Conference on Travel Behaviour Research, Jaipur, India*.
Lieberman, M. D., Gaunt, R., Gilbert, D. T. and Trope, Y. (2002), 'Reflexion and reflection: A social cognitive neuroscience approach to attributional inference', *Advances in experimental social psychology,* **34,** 199-249.
Lintner, J. (1965), 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *The review of economics and statistics,* **47,** 13-37.
List, J. A. (2005), 'Scientific numerology, preference anomalies, and environmental policymaking', *Environmental and Resource Economics,* **32,** 35-53.
Liu, H. X., Ban, X., Ran, B. and Mirchandani, P. (2002), 'Analytical dynamic traffic assignment model with probabilistic travel times and perceptions', *Transportation Research Record: Journal of the Transportation Research Board,* **1783,** 125-133.

Liu, H. X., Recker, W. and Chen, A. (2004), 'Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data', *Transportation Research Part A: Policy and Practice,* **38,** 435-453.
Liu, X. and Polak, J. W. (2007), 'Nonlinearity and specification of attitudes toward risk in discrete choice models', *Transportation Research Record: Journal of the Transportation Research Board,* **2014,** 27-31.
Long, N. E. (1958), 'The local community as an ecology of games', *American Journal of Sociology***,** 251-261.
Loomes, G. (2010), 'Modeling choice and valuation in decision experiments', *Psychological Review,* **117,** 902.
Loomes, G., Orr, S. and Sugden, R. (2009), 'Taste uncertainty and status quo effects in consumer choice', *Journal of Risk and Uncertainty,* **39,** 113-135.
Loomes, G., Starmer, C. and Sugden, R. (1991), 'Observing violations of transitivity by experimental methods', *Econometrica: Journal of the Econometric Society***,** 425-439.
Loomes, G., Starmer, C. and Sugden, R. (2003), 'Do Anomalies Disappear in Repeated Markets?*', *The Economic Journal,* **113,** C153-C166.
Loomes, G. and Sugden, R. (1982), 'Regret theory: An alternative theory of rational choice under uncertainty', *The Economic Journal,* **92,** 805-824.
Loomes, G. and Sugden, R. (1986), 'Disappointment and dynamic consistency in choice under uncertainty', *The Review of Economic Studies,* **53,** 271-282.
Lopes, L. L. (1993), 'Reasons and resources: The human side of risk taking', *In N. J. Bell & R. W. Bell (Eds.), Adolescent Risk Taking. Newbury Park, CA, US: Sage Publications.*
Luce, R. D. (1959), 'Individual choice behavior', *John Wiley and Sons, New York.*
Luce, R. D. and Suppes, P. (1965), 'Preference, utility, and subjective probability', *Handbook of mathematical psychology,* **3,** 249-410.
Manktelow, K. (2012) *Thinking and reasoning: An introduction to the psychology of reason, judgment and decision making,* Psychology Press.
Manski, C. F. (1977), 'The structure of random utility models', *Theory and decision,* **8,** 229-254.
March, J. G. and Simon, H. A. (1958), 'Organizations', *New York: John Wiley & Sons*.
Markowitz, H. (1952), 'The utility of wealth', *The Journal of Political Economy***,** 151-158.
Marschak, J. (1950), 'Rational behavior, uncertain prospects, and measurable utility', *Econometrica: Journal of the Econometric Society,* **18,** 111-141.
Marschak, J. (1960), 'Binary-choice constraints and random utility indicators', *Proceedings of a Symposium on Mathematical Methods in the Social Sciences*.
Masiero, L. and Hensher, D. A. (2010), 'Analyzing loss aversion and diminishing sensitivity in a freight transport stated choice experiment', *Transportation Research Part A: Policy and Practice,* **44,** 349-358.
McFadden, D. (1972) *Conditional logic analysis of qualitative choice behavior,* Institute of Urban & Regional Development, University of California.
McFadden, D. (1978) *Modelling the choice of residential location,* Institute of Transportation Studies, University of California.
McFadden, D. (1999), 'Rationality for economists?', *Journal of Risk and Uncertainty,* **19,** 73-105.
McFadden, D. (2000), 'Disaggregate behavioral travel demand's RUM side', *Travel Behaviour Research***,** 17-63.
Michea, A. and Polak, J. (2006), 'Modelling rsky choice behaviour: evaluating alternatives to expected utility theory', *11th International Conference on Travel Behaviour ResearchKyoto*.

MVA, C., Leeds, I. o. and Oxford, T. S. U. U. o. (1987), 'The value of travel time savings', *Policy Journals*.
Myerson, R. B. (2013) *Game theory: analysis of conflict,* Harvard university press.
Noland, R. B. and Polak, J. W. (2002), 'Travel time variability: a review of theoretical and empirical issues', *Transport Reviews,* **22,** 39-54.
Noland, R. B. and Small, K. A. (1995) *Travel-time uncertainty, departure time choice, and the cost of the morning commute,* Transportation Research Board.
Noland, R. B., Small, K. A., Koskenoja, P. M. and Chu, X. (1998), 'Simulating travel reliability', *Regional Science and Urban Economics,* **28,** 535-564.
Oort, C. J. (1969), 'The evaluation of travelling time', *Journal of Transport Economics and Policy***,** 279-286.
Páez, A. and Whalen, K. (2010), 'Enjoyment of commute: A comparison of different transportation modes', *Transportation Research Part A: Policy and Practice,* **44,** 537-549.
Pells, S. (1987), 'The evaluation of reductions in the variability of travel times on the journey to work', *Transport Policy, Proceedings of Seminar A, PTRC, 313-325*.
Petty, R. E., Cacioppo, J. T. and Schumann, D. (1983), 'Central and peripheral routes to advertising effectiveness: The moderating role of involvement', *Journal of consumer research***,** 135-146.
Pidgeon, N., Hood, C., Jones, D., Turner, B. and Gibson, R. (1992), 'Risk perception', *Risk: analysis, perception and management***,** 89-134.
Polak, J. W. (1987), 'Travel time variability and departure time choice: a utility theoretic approach', *Transport Studies Group, Discussion Paper,* **15**.
Polak, J. W., Hess, S. and Liu, X. (2008) In *Transportation Research Board 87th Annual Meeting*.
Pratt, J. W. (1964), 'Risk Aversion in the Small and in the Large', *Econometrica: Journal of the Econometric Society,* **32,** 122-136.
Quiggin, J. (1982), 'A theory of anticipated utility', *Journal of Economic Behavior & Organization,* **3,** 323-343.
Ramos, G. M., Daamen, W. and Hoogendoorn, S. (2011) In *Transportation Research Board 90th Annual Meeting*.
Razo, M. and Gao, S. (2011), 'A rank-dependent expected utility model for strategic route choice with stated preference data', *Transportation Research Part C: Emerging Technologies*.
Recker, W., Chung, Y., Park, J., Wang, L., Chen, A., Ji, Z., Liu, H., Horrocks, M. and Oh, J.-S. (2005), 'Considering Risk-Taking Behavior in Travel Time Reliability', *California Partners for Advanced Transit and Highways (PATH) Research Report. Institute of Transportation Studies. University of California, Berkeley.*
Revelt, D. and Train, K. (1998), 'Mixed logit with repeated choices: households' choices of appliance efficiency level', *Review of Economics and Statistics,* **80,** 647-657.
Roberts, D. C., Boyer, T. A. and Lusk, J. L. (2008), 'Preferences for environmental quality under uncertainty', *Ecological Economics,* **66,** 584-593.
Rose, J. M. and Masiero, L. (2010), 'A Comparison of the Impacts of Aspects of Prospect Theory on WTP/WTAEstimated in Preference and WTP/WTA Space', *EJTIR,* **4**.
Rosenberg, M. J. and Hovland, C. I. (1960), 'Cognitive, affective, and behavioral components of attitudes', *Attitude organization and change: An analysis of consistency among attitude components,* **3,** 1-14.
Salant, Y. (2011), 'Procedural analysis of choice rules with applications to bounded rationality', *The American Economic Review,* **101,** 724-748.

Sargent, T. J. (2011), 'Bounded rationality in macroeconomics: The Arne Ryde memorial lectures', *Oxford: Clarendon Press*.
Schmidt, U. and Traub, S. (2002), 'An experimental test of loss aversion', *Journal of Risk and Uncertainty,* **25,** 233-249.
Schmidt, U. and Zank, H. (2008), 'Risk aversion in cumulative prospect theory', *Management Science,* **54,** 208-216.
Schwanen, T. and Ettema, D. (2009), 'Coping with unreliable transportation when collecting children: Examining parents' behavior with cumulative prospect theory', *Transportation Research Part A: Policy and Practice,* **43,** 511-525.
Segal, U. (1990), 'Two-stage lotteries without the reduction axiom', *Econometrica: Journal of the Econometric Society,* **58,** 349-377.
Senbil, M. and Kitamura, R. (2003) In *Uncertainty Modeling and Analysis, 2003. ISUMA 2003. Fourth International Symposium on*, pp. 15-20.
Senbil, M. and Kitamura, R. (2004), 'Reference points in commuter departure time choice: A prospect theoretic test of alternative decision frames', *Journal of Intelligent Transportation Systems,* **8,** 19-31.
Senbil, M. and Kitamura, R. (2008), 'Policy effects on decisions under uncertain conditions: Simulation with mixed logit models of toll expressway use', *Transportation Research Record: Journal of the Transportation Research Board,* **2076,** 1-9.
Senna, L. A. D. S. (1994), 'The influence of travel time variability on the value of time', *Transportation,* **21,** 203-228.
Shogren, J. F., Shin, S. Y., Hayes, D. J. and Kliebenstein, J. B. (1994), 'Resolving differences in willingness to pay and willingness to accept', *The American Economic Review***,** 255-270.
Simon, H. A. (1955), 'A behavioral model of rational choice', *The Quarterly Journal of Economics,* **69,** 99-118.
Simon, H. A. (1979), 'Rational Decision-Making in Business Organizations', *American Economic Review,* **69,** 493-513.
Sinn, H. W. (1983) *Economic decisions under uncertainty,* North-Holland.
Small, K. and Verhoef, E. T. (2007) *The economics of urban transportation,* Routledge.
Small, K. A. (1982), 'The scheduling of consumer activities: work trips', *The American Economic Review,* **72,** 467-479.
Small, K. A. (2012), 'Valuation of travel time', *Economics of Transportation*.
Small, K. A., Noland, R. B., Chu, X. and Lewis, D. (1999) *Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation,* Transportation Research Board.
Small, K. A., Winston, C. and Yan, J. (2005a), 'Supplement to "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability."', *Econometrica Supplementary Material,* **73**.
Small, K. A., Winston, C. and Yan, J. (2005b), 'Uncovering the distribution of motorists' preferences for travel time and reliability', *Econometrica,* **73,** 1367-1382.
Smith, R. D. (1996), 'Is Regret Theory an alternative basis for estimating the value of healthcare interventions?', *Health Policy,* **37,** 105-115.
Starmer, C. (2000), 'Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk', *Journal of Economic Literature,* **38,** 332-382.
Starmer, C. and Sugden, R. (1993), 'Testing for juxtaposition and event-splitting effects', *Journal of Risk and Uncertainty,* **6,** 235-254.
Stigler, G. J. (1961), 'The economics of information', *The Journal of Political Economy,* **69,** 213-225.

Stoltz, G. and Lugosi, G. (2005), 'Internal regret in on-line portfolio selection', *Machine Learning,* **59,** 125-159.
Stott, H. P. (2006), 'Cumulative prospect theory's functional menagerie', *Journal of Risk and Uncertainty,* **32,** 101-130.
Sugden, R. (2004), 'Alternatives to expected utility: Foundations', *Handbook of utility theory,* **2,** 685–756.
Swait, J. (2001), 'A non-compensatory choice model incorporating attribute cutoffs', *Transportation Research Part B: Methodological,* **35,** 903-928.
Tatineni, M., Boyce, D. E. and Mirchandani, P. (1997), 'Comparisons of deterministic and stochastic traffic loading models', *Transportation Research Record: Journal of the Transportation Research Board,* **1607,** 16-23.
TfL (2010), 'London Underground Demand Trends', *Tranport for London,*
TfL (2012), 'Rolling Stock', *Transport for London,* http://www.tfl.gov.uk/corporate/modesoftransport/londonunderground/rollingstock/1614.aspx.
TfL (2013a), 'Key facts', *Transport for London,* http://www.tfl.gov.uk/corporate/modesoftransport/londonunderground/1608.aspx.
TfL (2013b), 'Tube reliability stays on track as demand rises', *Transport for London,* http://www.tfl.gov.uk/corporate/media/newscentre/metro/27397.aspx.
Thomson, J. (1968), 'The value of traffic management', *Journal of Transport Economics and Policy***,** 3-32.
Thurstone, L. L. (1927), 'A law of comparative judgment', *Psychological Review; Psychological Review,* **34,** 273.
Tilahun, N. Y. and Levinson, D. M. (2010), 'A moment of time: Reliability in route choice using stated preference', *Journal of Intelligent Transportation Systems,* **14,** 179-187.
Train, K. (2003) *Discrete choice methods with simulation,* Cambridge Univ Pr.
Train, K. E. (2008), 'EM algorithms for nonparametric estimation of mixing distributions', *Journal of Choice Modelling,* **1,** 40-69.
Tsalatsanis, A., Hozo, I., Vickers, A. and Djulbegovic, B. (2010), 'A regret theory approach to decision curve analysis: a novel method for eliciting decision makers' preferences and decision-making', *BMC medical informatics and decision making,* **10,** 51.
Tseng, Y.-Y., Verhoef, E., de Jong, G., Kouwenhoven, M. and van der Hoorn, T. (2009), 'A pilot study into the perception of unreliability of travel times using in-depth interviews', *Journal of Choice Modelling,* **2,** 8-28.
Tversky, A. and Fox, C. R. (1995), 'Weighing risk and uncertainty', *Psychological Review,* **102,** 269.
Tversky, A. and Kahneman, D. (1986), 'Rational choice and the framing of decisions', *Journal of business,* **59,** 251-278.
Tversky, A. and Kahneman, D. (1991), 'Loss aversion in riskless choice: A reference-dependent model', *The Quarterly Journal of Economics,* **106,** 1039-1061.
Tversky, A. and Kahneman, D. (1992), 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and Uncertainty,* **5,** 297-323.
van de Kaa, E. J. (2008), 'Extended prospect theory', *TRAIL Research School, Delft, The Netherlands*.
Van Lint, J., van Zuylen, H. J. and Tu, H. (2008), 'Travel time unreliability on freeways: Why measures based on variance tell only half the story', *Transportation Research Part A: Policy and Practice,* **42,** 258-277.
Varian, H. R. (1992) *Microeconomic analysis,* Norton New York.
Vickrey, W. S. (1969), 'Congestion theory and transport investment', *The American Economic Review,* **59,** 251-260.

Viscusi, W. K. (1989), 'Prospective reference theory: Toward an explanation of the paradoxes', *Journal of Risk and Uncertainty,* **2,** 235-263.
Von Neumann, J. and Morgenstern, O. (1947), 'Theory of games and economic behavior. Princeton', *NJ: Princeton University*.
Wakker, P. (1994), 'Separating marginal utility and probabilistic risk aversion', *Theory and decision,* **36,** 1-44.
Wakker, P. P. and Zank, H. (2002), 'A simple preference foundation of cumulative prospect theory with power utility', *European Economic Review,* **46,** 1253-1271.
Wassenaar, H. J., Chen, W., Cheng, J. and Sudjianto, A. (2005), 'Enhancing discrete choice demand modeling for decision-based design', *Journal of Mechanical Design,* **127,** 514.
Weeks, D. E. and Lange, K. (1989), 'Trials, tribulations, and triumphs of the EM algorithm in pedigree analysis', *Mathematical Medicine and Biology,* **6,** 209-232.
Winter, S. G. (1965), 'Economic" natural selection" and the theory of the firm', *Yale Economic Essay, 4, 225-72*.
Wu, G. and Gonzalez, R. (1996), 'Curvature of the probability weighting function', *Management Science***,** 1676-1690.
Yaari, M. E. (1987), 'The dual theory of choice under risk', *Econometrica: Journal of the Econometric Society,* **55,** 95-115.
Zajonc, R. B. (1980), 'Feeling and thinking: Preferences need no inferences', *American psychologist,* **35,** 151.