Assessing Reference Dependence in Travel Choice Behaviour

Huang, B.

DOI
10.4233/uuid:179c7c74-7a84-4622-9c7f-4bd56b32eb91

Publication date
2022

Document Version
Final published version

Citation (APA)
Huang, B. (2022). *Assessing Reference Dependence in Travel Choice Behaviour*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:179c7c74-7a84-4622-9c7f-4bd56b32eb91

Assessing Reference Dependence in Travel Choice Behaviour

Bing Huang 黄冰

Assessing Reference Dependence in Travel Choice Behaviour

Bing Huang

Delft University of Technology

This research was funded by the China Scholar Council

Assessing Reference Dependence in Travel Choice Behaviour

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op 19 mei 2022 om 10:00 uur

door

Bing HUANG

Master of Science in Transport Planning and Management,
University of Shanghai for Science and Technology, China
born in Zhengzhou, China

Dit proefschrift is goedgekeurd door de:
promotor: Prof.dr.ir. C.G. Chorus
copromotor: Dr.ir. S. van Cranenburgh

Samenstelling van de promotiecommissie:

Rector Magnificus	Voorzitter
Prof.dr.ir. C.G. Chorus	Technische Universiteit Delft, promotor
Dr.ir. S. van Cranenburgh	Technische Universiteit Delft, Copromotor

Onafhankelijke leden:

Prof. dr.ir. G.P. Van Wee	Technische Universiteit Delft
Prof. dr.ir. T.A. Arentze	Technische Universiteit Eindhoven
Prof. dr.ir. M. Vandebroek	Katholieke Universiteit Leuven
Dr.ir. O. Cats	Technische Universiteit Delft

Reserve leden:

Prof. dr.ir. L. Tavasszy	Technische Universiteit Delft

TRAIL Thesis Series no. T2020/06, the Netherlands Research School TRAIL

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-309-1

Printed in the Netherlands

Dedicated to

My parents and Zongchen

for their endless love, support and encouragement

Preface

Life is a matter of choices, and every choice you make makes you.

— John C. Maxwell

Choosing to pursue a PhD was probably one of the biggest choices that I have made in my life. A risk-averse person may never make such a decision because the journey of pursuing a PhD is packed with many challenges, unknowns, uncertainties, stress and struggles. Unfortunately (or fortunately), I am a risk-seeking person by nature. At this moment, standing right at the finish line and looking back on my PhD journey, I do not regret the choice that I made five years ago! I appreciate all the experience and self-growth that I have gained during these years. The process of doing a PhD does a lot more than equip me with a creative mind and analytical skills. It helps me build resilience to tackle complex problems and perseverance in the face of difficulties. More importantly, I have received a lot of help, support, care, friendship and love during my journey. Therefore, I would like to take this opportunity to express my gratitude and thanks to the people who have always been there for me and helped me along my journey.

First and foremost, I would like to say a big thank you to my supervisors, Caspar Chorus and Sander van Cranenburgh, who have both been very supportive during my PhD, especially during my first two years, who put a lot of extra time and energy into me. Caspar, thank you for accepting me to be your student. It is such a privilege to work with you. I have benefited a lot from every meeting and discussion that I had with you. Your quick mind, broad knowledge, and great presentation skills have made a great impact on me. Sander, I appreciate that your door is always open to me for a quick question or discussion. You helped me a lot not only in choice modelling knowledge but also in paper writing. I still remember how you guided me to write my gonogo proposal section by section and also other papers and chapters of this thesis. Sometimes I relied on you too much, which was a burden to you. But thank you for your patience with me.

I would also like to express my gratitude to Cees Timmers. Cees, you encouraged me a lot when I applied for the PhD position. Without you, I would never start my PhD journey at TU Delft.

I still remember the first time we met in Shanghai in 2015. The two-hour conversation gave me a lot of encouragement and strength. When I started my PhD, you continuously helped me. Thank you for showing up at my gonogo meeting rehearsal and providing much good advice for my presentation. I hope I will have a chance to see you again and have some catch-ups when I get back to the Netherlands.

I would like to thank my committee members. Bert van Wee, Martina Vandebroek, Theo Arentze, Lori Tavasszy and Oded Cats, thank you very much for the time and efforts that you devoted to reviewing and commenting on my draft thesis. I am looking forward to our discussions in my defence.

Thank you, all my colleagues at the TLO section. We had so many wonderful times together in the past years, especially before the Covid-19 outbreak. I enjoyed our lunchtime, section outings, and every year's Christmas dinner. Special thanks go to Jan Anne Annema. Jan Anne, I am so lucky and happy to be a colleague with you. You are the person who always brings positive energy and happiness to others. You made lunchtime more fun and lively. Moreover, thank you for offering me an opportunity to supervise a master's student with you. I was not very confident at first, but you gave me a lot of encouragement and help. Kartika, you are not just a colleague or friend to me, but more of a sister. Thank you for always being by my side, listening to my complaints and giving me support and encouragement. Lizet, we always have a lot to talk about, sharing not only funny things or gossip but also sadness and sorrows. Without you, my PhD journey would be a lot more boring. Kailan, thank you for all the nice food that I had at your place. I enjoyed our shopping time very much at the open market in Delft. Angie, during your short visit to TLO, we became very good friends. Thank you for being always there and giving me energy whether you are in Delft or not. Nicolas, thank you for all the great parties and nice French wines. Now I can taste the differences between wines. Shahrzad, thank you for all the office chats. You made my daily office more fun. Rick, thank you for your company at the hEART2017 conference and all the pictures you took of me being "accidentally" with academic celebrities.

Big thanks go to my friends. Xuanzi, thank you for all the talks, care, and support over these years. Whether happy or sad, you are the first person I think of to share. Jian, thank you for all your help. You are one of the most reliable people that I know. Longjian, thank you for all the fun activities. I enjoyed a lot doing Just dance, Karaoke, and mountain biking with you. Niels, you are so open-minded and funny. Thank you for your company during the final phase of my PhD. You brought me a lot of laughter.

Finally, I would like to express a huge thank you to my family. 爸爸妈妈，感谢你们这么多年来对我无私的爱和奉献。妈，我要强不服输的个性一定是随你。爸，我乐观开朗的性格一定是遗传你。感谢你们赋予我这些美好品格和一直以来无条件的支持，才让我博士旅途终落完美句号。宗琛，如果有人问我博士期间最大的收获是什么，我会毫不犹豫地脱口而出: "你!" 我实在是太幸运了，收获了如此美好的爱情。Zongchen, you are my academic role model and life helper, as well as my "financial sponsor" and "chauffeur". Meeting you is the luckiest thing during my PhD journey and marrying you was the best decision I have ever made in my life! Let us create a better future together!

Bing Huang

Zürich, April 2022

Content

1 Introduction

1.1 Background

People make all kinds of choices in their daily lives, such as driving to work rather than taking public transport or buying a particular brand of coffee. Many of these choices have a direct impact on demand for products, services and public infrastructures. Understanding people's choice behaviour can not only infer people's preferences for certain products or services, but more importantly make future demand forecasts.

In the field of transportation, the modelling and analysis of travel choice behaviour have a long history. As travel choice modellers, we are interested in where people go to perform various activities, how they choose destinations, which travel mode they take, which route they choose, etc. These travel-related choices are essential elements of travel demand analysis and transport policymaking. By modelling and analysing travel choices, we can predict the demand for new highways, bus lines, transport terminals, etc., assess the effectiveness of transport services or policies, as well as evaluate the social impact and equity of transport investments.

Discrete choice modelling approaches have been widely used to model and analyse individual and household choice behaviour. In particular, the Random Utility Maximization (RUM) models (McFadden, 1973; Ben-Akiva & Lerman, 1985) emerged to provide a theoretically robust and tractable modelling tool for choice modellers. The RUM models assume that when making choices, decision makers will choose the alternative which brings them the highest utility. The utility of an alternative is often assumed to be a linear-in-parameter function of characteristics of the alternative and associated parameters (decision weights associated with characteristics). Such models are called the linear-additive RUM models.

Due to their tractability and ease of use, the RUM models (especially the linear-additive RUM models) have gained much popularity in travel behaviour analysis. For example, the RUM models are widely used to examine the importance of different travel-related attributes, such as travel time, fare, comfort, and road safety, to name a few (e.g. Ben-Akiva & Lerman, 1985;

McFadden, 2000; Ulfarsson et al., 2006). Moreover, they are also used to analyse various travel behaviour and make forecasts of future demand, such as travel mode choices (e.g. Koppelman & Bhat, 2006; Gan, 2015), route choices (e.g. Erhardt et al., 2003), destinations (e.g. Train, 1998), and car ownerships (Hensher, 2013), etc. The outputs of demand forecasts are routinely used in the cost benefit analyses of new transport infrastructure developments.

Notwithstanding the popularity of the RUM models, recent years have witnessed great interest in incorporating psychological and behavioural elements into the models. The conventional RUM models assume that the utility of an alternative only depends on the performance of this alternative, irrelevant to the status quo or other competing alternatives. This however contrasts with findings in psychology that how people assess an alternative's performance is largely determined by its contrast to the status quo or some reference level (Kahneman & Tversky, 1979; Tversky & Kahneman, 1991; Tversky & Simonson, 1993). This means choice behaviour depends not only on the final state of the choice option but also on the comparison with some reference level. Such behavioural phenomenon that *choice behaviour depends on the reference level* is called *reference dependence*.

The origin of reference dependence can be traced back to Prospect Theory (Kahneman & Tversky, 1979). Prospect Theory was originally proposed to describe risky choices, such as lotteries and gambling. Later on, it has been extended to the context of riskless choices by Tversky & Kahneman (1991), known as the reference-dependent preference theory. The key idea of this theory is that individuals' preferences depend on some reference level; shifts of the reference level may result in the reversal of preferences. To illustrate the reference-dependent effect on people's perceptions and judgement, an example of vision differences is presented in Figure 1.1.

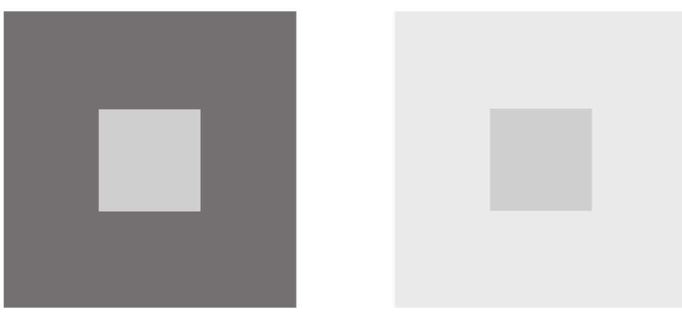


Figure 1.1 Vision differences caused by the reference-dependent effect
Note: This figure is adapted from Kahneman (2002)

In Figure 1.1, the two inner squares are the same colour, but the left one seems to be brighter than the right one. This vision difference is caused by surrounding colours: the inner colour looks brighter when presented in dark surroundings than in bright surroundings. This example well exhibits the reference-dependent nature of people's perceptions. Likewise, people's perceptions or preferences of choice options can be also influenced by some reference or competing choices options.

A noteworthy notion accompanying reference dependence is *loss aversion*. The outcome of a choice is viewed as gains or losses relative to some reference point; gains are desirable outcomes that exceed the reference point, while losses are undesirable outcomes inferior to the reference point. Loss aversion describes a behavioural phenomenon that the disutility caused by losses is larger than the utility caused by equivalent gains. In other words, losses have a greater impact on choices than equivalent gains, and as a result, decision makers tend to exhibit loss aversion behaviour.

Loss aversion can be used to account for many interesting behavioural phenomena, such as

- The endowment effect (Thaler, 1980). It describes a circumstance in which people demand more to give up an object that they own than they would like to pay to acquire it. The explanation for this effect is that the disutility of giving up an object (losses) is greater than the utility of acquiring it (gains).
- The status quo bias (Samuelson & Zeckhauser, 1988). It means that people would rather remain in the status quo than make any changes, because the negative effects of leaving the status quo have a greater impact than its positive effects.
- The compromise effect. This effect has been widely observed and documented in the field of marketing (Simonson, 1989; Tversky & Simonson, 1993; Chernev, 2004; Kivetz et al., 2004). It describes a phenomenon that market share increases when a product has intermediate performance on all attributes rather than having good performance on some attributes but poor performance on others. The existence of this effect can be also attributed to loss aversion. Losses generated by poor performance on one attribute have a greater impact than gains generated by good performance on the other. Therefore, to avoid losses, consumers tend to choose intermediate options rather than extreme options. This effect is also referred to as extremeness aversion.

Another notable aspect of reference dependence is the reference point. Whether an outcome is regarded as gains or losses depends on the reference point; a shift in the reference point may turn gains into losses and vice versa (Tversky & Kahneman, 1991). A question then arises: What is the reference point used by individuals when they are viewing choice options? From an analyst's perspective, it is very hard to answer this question. Because it is difficult to identify individuals' reference points, as the reference point is mostly endogenous. The status quo is widely recognised as the reference point of a person's choice, but decision making can be very context-specific. Expectations (Kőszegi & Rabin, 2006), previous experiences (Brown, 1995), and ideal or acceptable situations are all possible to be individuals' reference points. In addition, the reference point can be exogenously given as well, such as travel information given by mapping services. Furthermore, individuals may have multiple reference points (Zhu & Timmermans, 2010) or have a range for the reference point, for example, the ideal departure time can be a period of time rather than an exact time.

The reference-dependent nature of choice behaviour has been well studied and documented in the literature. Not surprisingly, reference dependence has been incorporated into the choice modelling. The next section will give a brief introduction to some common approaches to modelling reference dependence.

1.2 Modelling reference dependence: A brief introduction

To increase the behavioural realism of choice models, a number of choice modellers have attempted to add behavioural elements to the modelling approaches over the past few decades. Among them, a small group of choice modellers have succeeded in incorporating reference dependence into choice models. One of the most widespread modelling approaches can be traced back to the value function of Prospect Theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). The value function has been widely applied to modelling reference dependence and loss aversion. Such applications can be found in diverse fields, such as marketing (e.g., Hardie et al., 1993), health care (e.g., Rizzo & Zeckhauser, 2003; Sokol-Hessner et al., 2013; Sokol-Hessner et al., 2016), and also transportation (e.g., De Borger & Fosgerau, 2008; Hess et al., 2008; Masiero & Hensher, 2010; Flügel et al., 2015).

More recently, another reference-dependent model has been introduced into travel behaviour research, called the Random Regret Minimization (RRM) model (Chorus et al., 2008). This model postulates that individuals' preferences depend on the relative performance of an alternative against other alternatives in the choice set. Further developments in the RRM model (Chorus, 2010; Van Cranenburgh et al., 2015) have led to the increasing popularity of the application of this model, mostly in the field of transportation, but also in marketing, energy, health care and behavioural economics (Hensher et al., 2013; Boeri & Longo, 2017; Biondi et al., 2019; Masiero et al., 2019; Wong et al., 2020).

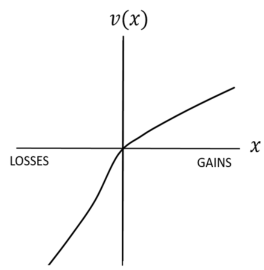
Modelling reference dependence involves dealing with reference points. Most modelling approaches treat reference points as an a priori assumption. For example, the status quo is often assumed as the reference point in many models. Other assumptions about reference points include the least favourite alternative in the contextual concavity model (Kivetz et al., 2004), individuals' expectations in the reference-dependent model proposed in Kőszegi & Rabin (2006), the attribute levels of other competing alternatives in the RRM models, etc. Recently, Bahamonde-Birke (2018) has proposed a reference-dependent model which allows one to estimate the reference point rather than assuming it a priori. The following sections will briefly introduce two reference-dependent models—the value function and the RRM model, as these two models are especially relevant to the studies in this thesis.

1.2.1 The value function

An abundant body of literature has applied the value function to model reference dependence and loss aversion. The value function was first proposed in Kahneman & Tversky (1979). It has three essential features: *reference dependence*—the utility is defined as losses or gains, measured relative to some reference point; *loss aversion*—the value function is steeper for losses than for gains; *diminishing sensitivity*[1]—the function is convex for losses and concave for gains. Later on, they presented the functional form of the value function (Tversky & Kahneman, 1992). It is a piecewise function in which losses and gains are modelled separately.

[1] Diminishing sensitivity means the marginal value of losses and gains decreases as their departure from the reference point.

Kahneman & Tversky (1979) proposed an S-shaped value function which is (1) defined on deviations from the reference point, (2) steeper for losses than gains, and (3) generally convex for losses and concave for gains. Later on, they presented a quantitative description of the value function. It is described by a two-part power function where x denotes the deviations (losses or gains) from the reference point; λ denotes the degree of loss aversion, and loss aversion occurs if $\lambda > 1$; parameters a and b capture the curvature of the value function, for example, the value function is convex for losses and concave for gains if $0 < a, b < 1$.



The S-shaped value function

(Kahneman & Tversky, 1979)

$$v(x) = \begin{cases} x^a & if\ x \geq 0 \\ -\lambda(-x)^b & if\ x < 0 \end{cases}$$

The formulation of the value function

(Tversky & Kahneman, 1992)

The application of the value function can be found in much travel behaviour research (Senbil & Kitamura, 2004; De Borger & Fosgerau, 2008; Hess et al., 2008; Lanz et al., 2010; Masiero & Hensher, 2010; Delle Site & Filippi, 2011; Stathopoulos & Hess, 2012; Jou & Chen, 2013; Flügel et al., 2015). For instance, Senbil & Kitamura (2004) applied the value function to examine reference dependence and loss aversion in departure time decision making. Jou & Chen (2013) applied the value function in the analysis of freeway route choice behaviour. The original form of the value function has three estimable parameters: one parameter captures loss aversion, and the other two govern diminishing sensitivity. Such a functional form is fairly complex. Thus, many studies adopt the variant forms of the value function in practical use. Hess et al. (2008) adopt a simplified modelling approach—each attribute is divided into increase and decrease values by taking differences from its reference level. As a result, increase and decrease values are estimated separately for each attribute. Loss aversion is therefore tested by comparing the estimates of the increase and decrease values. Similar modelling approaches can be found in Masiero & Hensher (2010), Lanz et al. (2010), Stathopoulos & Hess (2012) and Flügel et al. (2015). The applications in these studies involve route and trip choices, freight mode choices, and road safety. Another noteworthy aspect is that, in travel behaviour research, the value function is typically used to test reference preference for time-related or cost-related attributes. As pointed out by Stathopoulos & Hess (2012), reference dependence on situations with complex trade-offs among multiple attributes—a typical feature of real-world choices—has rarely been explored.

1.2.2 The Random Regret Minimization model

Recently, the application of the RRM model to travel behaviour research has received much attention. The RRM model is regarded as a regret-based counterpart of the linear-additive RUM model. Regret is defined as an unpleasant experience when unchosen alternatives perform better than the chosen alternative on one or more attributes. Regret-based models postulate that to avoid regret, decision makers choose the alternative which brings the minimum regret. Seminal work regarding regret-based decision making is done by Loomes & Sugden (1982) and Bell (1982), who mainly focus on risky choices. The RRM model has extended regret-based decision making to the riskless choice situation, and it has evolved into various model specifications, such as the classical RRM model (Chorus, 2010), the μRRM model (van Cranenburgh et al., 2015), and the RRM-Weber model (Jang et al., 2017).

The RRM model exhibits the property of reference dependence. In the RRM model, regret is generated by comparing the chosen alternative with other competing alternatives in the choice set. This means regret is determined by the relative performance of the chosen alternative against its competitors. Thus, in the RRM paradigm, choice behaviour is assumed to be reference-dependent—the reference level is the attribute levels of other competing alternatives in the choice set. Another property of the RRM model is regret aversion. Similar to loss aversion in Prospect Theory, the RRM model postulates that people view regret more importantly than rejoice (the opposite of regret). Therefore, loss aversion and regret aversion pick up the same behavioural pattern that bad experience (losses or regret) has a greater impact on choice behaviour than good experience (gains or rejoice).

In recent years, the RRM model has gained in popularity in transportation for its ease of use. It has been widely used to analyse a variety of travel choice behaviour, such as mode choices (e.g. Jing et al., 2018; Belgiawan et al., 2019), route choices (e.g. Prato, 2014; Mai et al., 2015), destination (e.g. Boeri et al., 2012; Jang et al., 2017; Masiero et al., 2019), departure time (e.g. Chorus & De Jong, 2011) and wildfire evacuation choices (Wong et al., 2020).

1.3 Research goals and research ideas

As shown in the previous section, efforts have been made to incorporate reference dependence into travel choice models. Most of the work is centred on the enhancement of travel choice models to get better model estimates or more accurate behaviour predictions, which is of great importance for travel demand analysis and transport planning. The main aim of this thesis is to extend the frontier of reference dependence choice modelling. Before presenting specific methodological contributions, this thesis starts by reiterating and strengthening the case for considering reference dependence in travel behaviour analysis. Specifically, it demonstrates and highlights the importance of reference dependence in understanding and explaining choice behaviour and in formulating relevant transport polices. The first goal of this thesis is stated as follows:

The first goal of this thesis is to highlight the importance of capturing reference dependence for transport policy analysis and to illustrate how conventional choice models deal with this. In particular, using an empirical case study, we show that failure to accommodate reference dependence may lead to severe bias in the understanding of choice behaviour and the resulting transport policies.

Given the importance of reference dependence, this thesis also aims to make three methodological contributions to the choice modelling regarding reference dependence. The first two contributions are new model specifications. In this thesis, we will introduce a series of reference-dependent model specifications to provide more tools for (travel) choice modellers. More specifically, we incorporate so-called relative thinking into the RRM framework, resulting in several new specifications established on the RRM models, and propose a new model specification to model loss aversion.

The third methodological contribution is concerned with (stated choice) data collection. Introducing alternative models inevitably requires one to conduct model comparison in terms of model fits. This is usually done by estimating all competing models on the same datasets and then comparing their model fits. However, current methods used to collect choice data have not been developed to discriminate between models (in terms of model fits), but to optimize the estimation of a particular model, such as efficient designs (Kanninen, 2002; Rose & Bliemer, 2009). In this thesis, we will put forward an innovative method of collecting choice data (by means of choice experiments) with the aim of discriminating between different choice models. Therefore, the second goal of this thesis can be stated as follows:

This thesis also aims to add three new tools to the "toolbox" for choice modellers when they are to analyse reference-dependent (travel) choice behaviour. The first two tools are new reference-dependent model specifications and the third one is an innovative method of constructing experimental designs that are optimised for discriminating between different choice models.

Tool 1: Incorporating relative thinking into the RRM model

The RRM model, as a reference-dependent model, looks at the relative performance of the chosen alternative against each of the other unchosen alternatives. However, empirical evidence suggests that decision makers exhibit another form of reference dependence, which has so far not been adequately reflected in the RRM model. Take an example of travel time differences. Suppose there are two situations in which both travel time savings are five minutes, but the initial travel time in situation 1 is fifteen minutes, while in situation 2, it is 105 minutes. The travel time differences are the same in these two situations, but how people respond to the differences may be very different: the five-minute travel time saving may be perceived to be larger (or more salient) when the initial travel time is fifteen minutes, compared to the situation where the initial is 105 minutes. Such perception differences can be explained by a famous psychophysical law called *Weber's law* (Weber, 1834; Gescheider, 2013). It relates the *actual* difference between physical stimuli (e.g. heat, weight, brightness) to the size of the difference as *perceived* by people. In the context of choice modelling, Weber's law asserts that how decision makers respond to differences in attribute levels between alternatives is influenced by *the initial attribute levels themselves*.

Besides, how people respond to differences is also influenced by *the range of attribute levels in the choice set*. When decision makers are faced with several choice alternatives, their choices are inevitably restricted to the range of attribute levels in the choice set, and the range of attribute levels provides the frame of reference for decision makers to choose from (Moon & Voss, 2009). A fairly robust empirical finding is called *the range effect*, which means a given attribute-level difference seems larger when presented in a narrow range than in a wide range (Volkmann, 1951; Parducci, 1965; Mellers & Cooke, 1994; Wedell, 1998; Ohler et al., 2000; Rabin & Weizsäcker, 2009; Bushong et al., 2021). Again, we use the example of travel time differences. The reduction of five minutes in travel time from fifteen minutes (10 vs 15 minutes) seems more considerable when choice alternatives are 10, 15, and 20 minutes, compared to the situation where the three alternatives are 10, 15, and 105 minutes.

Decision makers exhibit so-called *relative thinking* (Azar, 2007) when they respond to differences. Relative thinking means that when making choices, decision makers not only care about *absolute* differences between choice alternatives but also *relative* differences. In this thesis, we categorize relative thinking as *level-based*—i.e. relative to the attribute level itself—and *range-based*—i.e. relative to the range of attribute levels in the choice set. Both the attribute level and the range of attribute levels can be regarded as external reference points acquired by decision makers during the decision making process. This thesis incorporates these two categories of relative thinking into the RRM modelling paradigm, respectively, adding another layer of reference dependence to the RRM model.

Tool 2: A new loss aversion model

Currently, the great majority of loss aversion modelling approaches are built on the value function (e.g. Kivetz et al., 2004; Hess et al., 2008; Lanz et al., 2010; Masiero & Hensher, 2010; Flügel et al., 2015). What they have in common is that model specifications are piecewise: the utility functions of losses and gains are modelled separately. Although such models are effective tools of capturing loss aversion, their piecewise specifications come with the drawback of not being twice differentiable around the reference point. To overcome this drawback, this thesis presents a new model with a twice differentiable loss aversion function. The new model specification again makes use of the RRM model specification. As discussed in Section 1.2.2, regret aversion postulated in the RRM model captures the same behavioural pattern as loss aversion: Bad experience (regret) has a greater impact on decision making than good experience (rejoice). Another link between the RRM model and the loss aversion theory is the setting of reference points. The RRM model imposes the reference point in the model specification—the attribute levels of other competing alternatives in the choice set, while the loss aversion theory does not have restrictions about the reference point. Hence, the RRM model can be regarded as a special case of the loss aversion theory where the reference point is the attribute levels of other alternatives. Inspired by the relation between the RRM model and the loss aversion theory, the regret function is adapted in a way that it becomes a smooth loss aversion function (twice differentiable around reference points). Because it is based on the established RRM model, the new loss aversion model is expected to have a tractable and parsimonious structure, which makes it practical for applications in travel behaviour research. In this thesis, we will explore the properties of this model specification and test its empirical performance compared with other loss aversion model specifications.

Tool 3: An innovative method of constructing experimental designs

Recent years have seen increasing interest in introducing new models or model variations into the field of choice modelling; this thesis is one more example of this trend. To test model (specification) performance, new models are usually compared with a series of conventional models or model specifications. For example, the new loss aversion model will be compared with the linear-additive RUM model and an existing loss aversion model. Model comparison is often conducted by estimating all competing models on the same datasets and then comparing their model fits. The main aim is to select the model which can best represent the underlying data generating process from a set of competing models. However, data that are used for model comparison are not always optimally suited to this aim. Most data used for model comparison are stated choice (SC) data from SC experiments. Generating SC data relies on so-called experimental designs, which is the process of assigning attribute levels to the attributes that define every alternative in choice sets. Nowadays, the most common method of experimental designs is what is known as efficient designs (Rose & Bliemer, 2009). Such designs aim to generate data that lead to reliable model parameters for a particular model specification, rather than discriminating between different competing models regarding their model fits. Hence, there is a mismatch between what current experimental designs are optimised for—reliable measurement for model parameters for a particular model—and what they are sometimes used for—identification of the best model from a list of competing models. In this thesis, we will present an innovative method of constructing experimental designs which are optimised for discriminating between different choice models.

Together, this thesis aims to provide a more profound understanding of the importance of reference dependence in travel behaviour analysis and transport policy development, as well as to provide new tools and techniques to effectively model reference dependence and to collect data that are most suitable for comparison between different choice model specifications.

1.4 Thesis outline

To achieve the above-stated research goals, four studies are conducted in this thesis. Each study is a chapter. Before introducing the new formulations of modelling reference dependence, this thesis begins with an empirical case study. It shows how important the reference effect can be in explaining behavioural phenomena and formulating transport policies. The next two chapters present methodological contributions to modelling reference dependence. Specifically, Chapter 2 incorporates relative thinking into the RRM framework, and Chapter 3 presents a new loss aversion model. The last chapter focuses on the method of generating choice experiments. It introduces an innovative way of constructing experimental designs, called discriminatory designs, to discriminate between different choice models in terms of model fit comparison.

Chapter 2: Empirical analysis of reference dependence: Exploring the reference effect on the public opinion about traffic fatalities involving automated vehicles

As a kick-off of this thesis, this study aims to provide empirical insights into the importance of reference dependence in explaining interesting behavioural phenomena in transportation and resulting transport policies. In the meanwhile, it also demonstrates how conventional linear-additive RUM models are traditionally used to model reference dependence.

The case study is conducted on a very cutting-edge topic: future transportation involving automated vehicles (AVs). AVs are expected to be much safer than conventional vehicles (CVs), as they can largely eliminate traffic fatalities that are caused by human operation errors. But the general public does not seem optimistic about AVs concerning safety issues. Recent public debates and academic papers claim that AVs will need to be much safer than CVs before being accepted by the public (Mervis, 2017; Liu et al., 2019; Nees, 2019). This means fatalities caused by AVs are expected to be overweighted by the general public, compared with conventional car fatalities.

In this study, we conduct two stated choice experiments. The first experiment aims to examine *whether and the extent to which AV fatalities carry more weight to the general public, compared to conventional fatalities*. The second experiment is designed in such a way that it allows to answer the following question: *Is the overweighting of AV fatalities caused by the intrinsic, qualitative differences between an AV and a human-operated CV? Or is it caused by the fact that the current number of AV fatalities is so low that each additional AV fatality carries considerable extra weight?* Disentangling these two potential explanations is crucial for policy implications. For example, the latter explanation would suggest that, ironically, the inevitable occurrence of more AV-related road accidents will in time lead to a diminishing degree of the overweighting of AV fatalities.

Chapter 3: Incorporating relative thinking into the Random Regret Minimization model framework

The relative nature of decision making has been well recognized and studied in various domains, most notably in psychology and consumer research (Thaler, 1980; Tversky & Kahneman, 1981; Ranyard & Abdel-Nabi, 1993). Relative thinking refers to a behavioural phenomenon that when making choices, decision makers consider not only absolute differences between choice alternatives but also relative differences. Relative differences can be categorised as differences relative to the attribute level in the choice set, called level-based relative thinking, and differences relative to the range of the attribute, called range-based relative thinking. More specifically, the level-based relative thinking describes a phenomenon that a given absolute attribute-level difference can seem big or small depending on the initial level of this attribute in the choice set (Azar, 2007, 2008; Saini & Thota, 2010; Nicolau, 2012). This phenomenon can also be explained by Weber's law. The range-based relative thinking denotes that the range of attribute levels in the choice set may also influence people's judgement about attribute-level differences: a given absolute difference seems larger when presented in a narrow range than in a wide range (Volkmann, 1951; Parducci, 1965; Mellers & Cooke, 1994; Bushong et al., 2021).

The RRM model is a reference-dependent model, but it fails to capture the reference-dependent nature of how people respond to attribute-level differences (Jang et al., 2017). This chapter aims to *introduce and empirically test a series of new reference-dependent models that incorporate level-based or range-based relative thinking into the RRM framework.* The new models are tested on four empirical data sets in comparison with the RUM-MNL model and the conventional RRM model (i.e., the classical RRM and μRRM).

Chapter 4: A new loss aversion model

Loss aversion denotes that when decision makers make choices, the outcomes of choices are perceived as gains and losses against a reference point, and losses are evaluated more heavily than equivalent gains. In other words, decision makers are more sensitive to losses than gains. This behavioural phenomenon has been well investigated in travel choice behaviour (De Borger & Fosgerau, 2008; Hess et al., 2008; Stathopoulos & Hess, 2012). The modelling approaches used in these studies are mainly built on the value function of prospect theory, which is an unsmooth function. Moreover, loss aversion are typically tested on monetary attributes (e.g. travel cost and fare) and time-related attributes (e.g. travel time), while complex situations or trade-offs involving other attributes have been rarely explored. This may be (partly) due to the complex structures of the existing loss aversion models.

This chapter aims to *propose and empirically test a new loss aversion model that has a smooth function and can be easily applied to capture loss aversion in multiple travel-related attributes.* This new model is systemically compared with several existing loss aversion models in terms of model properties (i.e. reference dependence, loss aversion and diminishing sensitivity) and model performance on empirical data.

Chapter 5: An innovative method of constructing experimental designs optimised for discriminating between different choice models

Data collection is an essential part of travel behaviour research. Analysts observe travellers' choice behaviour by conducting experimental investigations in which usually hundreds or thousands of respondents are requested to fill out a questionnaire designed by analysts. Questionnaires can be used to observe real choices made in the actual market and this type of data is known as revealed preference (RP) data, and they can also be used to observe behavioural intentions in hypothetical choice settings and this type of data is called stated preference (SP) data. Compared to RP data, SP data are more flexible, as they do not have to be constricted by limited variation in the real market. SP data have been widely used to collect choices in travel behaviour research. They are collected by conducting stated choice experiments in which respondents make choices in several hypothetical choice scenarios.

Generating hypothetical choice scenarios relies on experimental designs. There are several methods of constructing experimental designs. At present, the method that is widely used in travel behaviour research is called efficient designs. Efficient designs aim to generate data that lead to reliable model parameters with as small standard errors as possible (Kanninen, 2002; Rose et al., 2008; Rose & Bliemer, 2009). However, in many cases, the aim of studies is to compare model performance between different models in terms of model fits, and consequently, to select the model that is the best representative of choice behaviour (for example, Chapters 3 and 4 in this thesis), as opposed to recovering model parameters for a given model as efficiently as possible. Current experimental designs aimed at efficiently recovering model parameters may not allow obtaining sufficient model discrimination capability.

The last study aims to make a methodological contribution toward designing choice experiments. In particular, it aims to *introduce an innovative method of constructing experimental designs that are optimised for discriminating between competing choice models.* Such experimental designs are named discriminatory designs. In contrast to efficient designs, which aim to maximize collected information about model parameters, discriminatory designs aim to maximize collected information about the underlying data generating process, yielding a result that the most plausible model can be discriminated from a set of competing models. To this end, the method of constructing discriminatory designs is based on the Bayes' theorem and the Kullback-Leibler (KL) divergence (Kullback-Leiber, 1951). In this study, the robustness of discriminatory designs is tested using both synthetic data and empirical data.

Figure 1.2 presents the outline of this thesis. The main chapters of this thesis are composed of one published journal paper (Chapter 2) and three unpublished works (Chapters 3, 4 and 5). The first three main chapters closely surround the main topic of this thesis—reference dependence. In particular, Chapter 2 puts it to the empirical test, and Chapters 3 and 4 present a series of reference-dependent models. Although Chapter 5 is not directly related to reference dependence, the method presented in Chapter 5 focuses on discriminating between different choice models, including different reference-dependent models. In addition, the four main chapters are also classified as empirical and methodological studies. Chapter 6 summarizes and reflects on the four studies, and suggests some directions for future research.

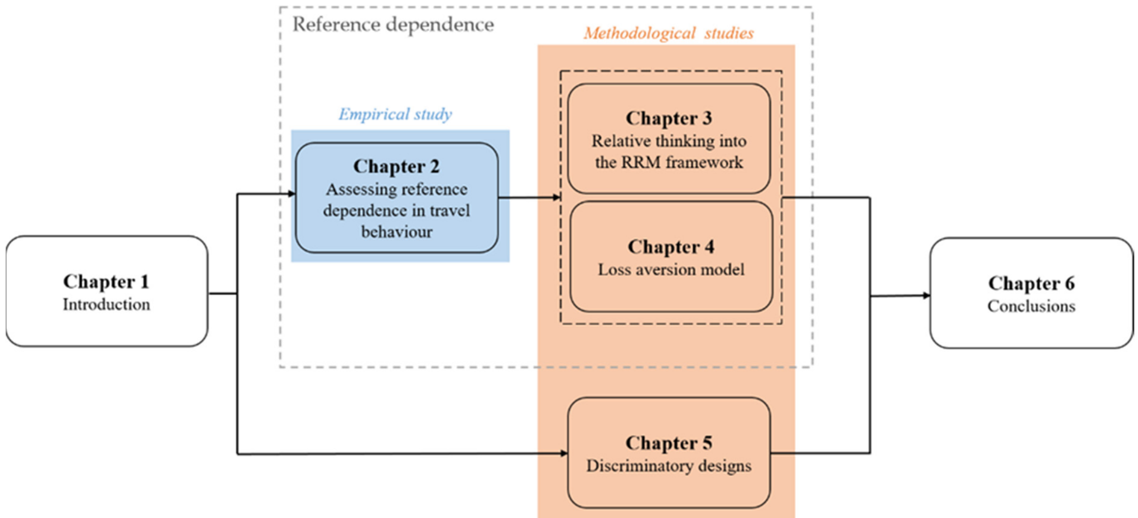


Figure 1.2 The outlines of this thesis

References

Azar, O. H. (2007). Relative thinking theory. *The Journal of Socio-Economics, 36*(1), 1-14.

Azar, O. H. (2008). The effect of relative thinking on firm strategy and market outcomes: A location differentiation model with endogenous transportation costs. *Journal of Economic Psychology, 29*(5), 684-697.

Bahamonde-Birke, F. J. (2018). Estimating the reference frame: A smooth twice-differentiable utility function for non-compensatory loss-averse decision-making. *Journal of choice modelling, 28*, 71-81.

Belgiawan, P. F., Ilahi, A., & Axhausen, K. W. (2019). Influence of pricing on mode choice decision in jakarta: A random regret minimization model. *Case Studies on Transport Policy, 7*(1), 87-95.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations research, 30*(5), 961-981.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. Cambridge: MIT Press.

Biondi, B., Van der Lans, I. A., Mazzocchi, M., Fischer, A. R., Van Trijp, H. C., & Camanzi, L. (2019). Modelling consumer choice through the random regret minimization model: An application in the food domain. *Food Quality and Preference, 73*, 97-109.

Boeri, M., & Longo, A. (2017). The importance of regret minimization in the choice for renewable energy programmes: Evidence from a discrete choice experiment. *Energy Economics, 63*, 253-260.

Boeri, M., Longo, A., Doherty, E., & Hynes, S. (2012). Site choices in recreational demand: a matter of utility maximization or regret minimization? *Journal of Environmental Economics and Policy, 1*(1), 32-47.

Brown, P. M. (1995). Learning from experience, reference points, and decision costs. *Journal of Economic Behavior & Organization, 27*(3), 381-399.

Bushong, B., Rabin, M., & Schwartzstein, J. (2021). A model of relative thinking. *The Review of Economic Studies, 88*(1), 162-191.

Chernev, A. (2004). Extremeness aversion and attribute-balance effects in choice. *Journal of consumer research, 31*(2), 249-263.

Chorus, C. G. (2010). A new model of random regret minimization. *European Journal of Transport Infrastructure Research, 10*(2).

Chorus, C. G., Arentze, T. A., & Timmermans, H. J. (2008). A random regret-minimization model of travel choice. *Transportation Research Part B: Methodological, 42*(1), 1-18.

Chorus, C. G., & De Jong, G. C. (2011). Modeling experienced accessibility for utility-maximizers and regret-minimizers. *Journal of Transport Geography, 19*(6), 1155-1162.

De Borger, B., & 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*(1), 101-115.

Delle Site, P., & Filippi, F. (2011). Stochastic user equilibrium and value-of-time analysis with reference-dependent route choice. *European Journal of Transport Infrastructure Research, 11*(2).

Erhardt, G. D., Koppelman, F. S., Freedman, J., Davidson, W. A., & Mullins, A. (2003). Modeling the choice to use toll and high-occupancy vehicle facilities. *Transportation Research Record, 1854*(1), 135-143.

Flügel, S., Elvik, R., Veisten, K., Rizzi, L. I., Meyer, S. F., Ramjerdi, F., & de Dios Ortúzar, J. (2015). Asymmetric preferences for road safety: Evidence from a stated choice experiment among car drivers. *Transportation research part F: traffic psychology behaviour, 31*, 112-123.

Gan, H. (2015). To switch travel mode or not? Impact of Smartphone delivered high-quality multimodal information. *IET Intelligent Transport Systems, 9*(4), 382-390.

Gescheider, G. A. (2013). *Psychophysics: the fundamentals*: Psychology Press.

Hardie, B. G., Johnson, E. J., & Fader, P. S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing Science, 12*(4), 378-394.

Hensher, D. A. (2013). *Dimensions of automobile demand: a longitudinal study of household automobile ownership and use*: Elsevier.

Hensher, D. A., Greene, W. H., & Chorus, C. G. (2013). Random regret minimization or random utility maximization: an exploratory analysis in the context of automobile fuel choice. *Journal of Advanced Transportation, 47*(7), 667-678.

Hess, S., Rose, J. M., & Hensher, D. A. (2008). Asymmetric preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E: Logistics Transportation Review, 44*(5), 847-863.

Jang, S., Rasouli, S., & Timmermans, H. (2017). Incorporating psycho-physical mapping into random regret choice models: model specifications and empirical performance assessments. *Transportation, 44*(5), 999-1019.

Jing, P., Zhao, M., He, M., & Chen, L. (2018). Travel mode and travel route choice behaviour based on random regret minimization: A systematic review. *Sustainability, 10*(4), 1185.

Jou, R.-C., & Chen, K.-H. (2013). An application of cumulative prospect theory to freeway drivers' route choice behaviours. *Transportation Research Part A: Policy*

Practice, 49, 123-131.

Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel prize lecture, 8*, 351-401.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Kanninen, B. J. (2002). Optimal design for multinomial choice experiments. *Journal of Marketing Research, 39*(2), 214-227.

Kivetz, R., Netzer, O., & Srinivasan, V. (2004). Alternative models for capturing the compromise effect. *Journal of Marketing Research, 41*(3), 237-257.

Koppelman, F. S., & Bhat, C. (2006). A self instructing course in mode choice modeling: multinomial and nested logit models.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The quarterly journal of economics, 121*(4), 1133-1165.

Lanz, B., Provins, A., Bateman, I. J., Scarpa, R., Willis, K., & Ozdemiroglu, E. (2010). Investigating willingness to pay–willingness to accept asymmetry in choice experiments. In *Choice Modelling: The State-of-the-art and The State-of-practice*: Emerald Group Publishing Limited.

Liu, P., Yang, R., & Xu, Z. (2019). How safe is safe enough for self-driving vehicles? *Risk analysis, 39*(2), 315-325.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal, 92*(368), 805-824.

Mai, T., Frejinger, E., & Bastin, F. (2015). *Comparing regret minimization and utility maximization for route choice using the recursive logit model*. Retrieved from

Masiero, L., & Hensher, D. A. (2010). Analyzing loss aversion and diminishing sensitivity in a freight transport stated choice experiment. *Transportation Research Part A: Policy Practice, 44*(5), 349-358.

Masiero, L., Yang, Y., & Qiu, R. T. (2019). Understanding hotel location preference of customers: comparing random utility and random regret decision rules. *Tourism Management, 73*, 83-93.

McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior*. New York: Academic Press.

McFadden, D. (2000). Disaggregate behavioral travel demand's RUM side. *Travel behaviour research*, 17-63.

Mellers, B. A., & Cooke, A. D. (1994). Trade-offs depend on attribute range. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1055.

Mervis, J. (2017). Are we going too fast on driverless cars? *Science*. Retrieved from http://www.sciencemag.org/news/2017/12/are-we-going-too-fast-driverless-cars

Moon, S., & Voss, G. (2009). How do price range shoppers differ from reference price point shoppers? *Journal of Business Research, 62*(1), 31-38.

Nees, M. A. (2019). Safer than the average human driver (who is less safe than me)? Examining a popular safety benchmark for self-driving cars. *Journal of Safety Research, 69*, 61-68.

Nicolau, J. L. (2012). Battle royal: Zero-price effect vs relative vs referent thinking. *Marketing Letters, 23*(3), 661-669.

Ohler, T., Le, A., Louviere, J., & Swait, J. (2000). Attribute range effects in binary response tasks. *Marketing Letters, 11*(3), 249-260.

Parducci, A. (1965). Category judgment: a range-frequency model. *Psychological review, 72*(6), 407.

Prato, C. G. (2014). Expanding the applicability of random regret minimization for route choice analysis. *Transportation, 41*(2), 351-375.

Rabin, M., & Weizsäcker, G. (2009). Narrow bracketing and dominated choices. *American Economic Review, 99*(4), 1508-1543.

Ranyard, R., & Abdel-Nabi, D. (1993). Mental accounting and the process of multiattribute choice. *Acta psychologica, 84*(2), 161-177.

Rizzo, J. A., & Zeckhauser, R. J. (2003). Reference incomes, loss aversion, and physician behavior. *Review of Economics and Statistics, 85*(4), 909-922.

Rose, J. M., & Bliemer, M. C. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews, 29*(5), 587-617.

Rose, J. M., Bliemer, M. C., Hensher, D. A., & Collins, A. T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological, 42*(4), 395-406.

Saini, R., & Thota, S. C. (2010). The psychological underpinnings of relative thinking in price comparisons. *Journal of Consumer Psychology, 20*(2), 185-192.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of risk and uncertainty, 1*(1), 7-59.

Senbil, M., & Kitamura, R. (2004). *Reference points in commuter departure time choice: a prospect theoretic test of alternative decision frames.* Paper presented at the Intelligent Transportation Systems.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of consumer research, 16*(2), 158-174.

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social cognitive and affective neuroscience, 8*(3), 341-350.

Sokol-Hessner, P., Raio, C. M., Gottesman, S. P., Lackovic, S. F., & Phelps, E. A. (2016). Acute stress does not affect risky monetary decision-making. *Neurobiology of stress, 5*, 19-25.

Stathopoulos, A., & Hess, S. (2012). Revisiting reference point formation, gains–losses asymmetry and non-linear sensitivities with an emphasis on attribute specific treatment. *Transportation Research Part A: Policy Practice, 46*(10), 1673-1689.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of economic behavior organization, 1*(1), 39-60.

Train, K. E. (1998). Recreation demand models with taste differences over people. *Land economics*, 230-239.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics, 106*(4), 1039-1061.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of risk uncertainty, 5*(4), 297-323.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management science, 39*(10), 1179-1189.

Ulfarsson, G. F., Kim, S., & Lentz, E. T. (2006). Factors affecting common vehicle-to-vehicle collision types: Road safety priorities in an aging society. *Transportation Research Record, 1980*(1), 70-78.

van Cranenburgh, S., Guevara, C. A., & Chorus, C. G. (2015). New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice, 74*, 91-109.

Volkmann, J. (1951). Scales of judgment and their implications for social psychology.

Weber, E. H. (1834). *De pulsu, resorptione, auditu et tactu: annotationes anatomicae et physiologicae, auctore*: prostat apud CF Koehler.

Wedell, D. H. (1998). Testing models of trade-off contrast in pairwise choice. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 49.

Wong, S. D., Chorus, C. G., Shaheen, S. A., & Walker, J. L. (2020). A Revealed Preference Methodology to Evaluate Regret Minimization with Challenging Choice Sets: A Wildfire Evacuation Case Study. *Travel Behaviour and Society, 20*, 331-347.

Zhu, W., & Timmermans, H. (2010). Cognitive process model of individual choice behaviour incorporating principles of bounded rationality and heterogeneous decision heuristics. *Environment and Planning B: Planning and Design, 37*(1), 59-74.

2 Empirical analysis of reference dependence: Exploring the reference effect on the public opinion about traffic fatalities involving automated vehicles

Huang, B., van Cranenburgh, S., & Chorus, C. G. (2020). Death by automation: Differences in weighting of fatalities caused by automated and conventional vehicles. European journal of transport and infrastructure research, 20(3), 71-86.

Abstract

Although Automated vehicles (AVs) are expected to have a major and positive effect on road safety, recent accidents caused by AVs tend to generate a powerful negative impact on the public opinion regarding safety aspects of AVs. Triggered by such incidents, many experts and policy makers now believe that paradoxically, safety perceptions may well prohibit or delay the rollout of AVs in society, in the sense that AVs will need to become much safer than conventional vehicles (CVs), before being accepted by the public. In this study, we provide empirical insights to investigate and explain this safety paradox. Using stated choice experiments, we show that there is indeed a difference between the weight that individuals implicitly attach to an AV-fatality and a CV-fatality. However, the degree of overweighting of AV-fatalities, compared to CV-fatalities, is considerably smaller than what has been suggested in public opinions and policy reports. We also find that the difference in weighting between AV-fatalities and CV-fatalities is (partly) related to a reference level effect: simply because the current number of fatalities caused by AVs is extremely low, each additional fatality carries extra weight. Our findings suggest that indeed, AVs have to become safer—but not orders of magnitude safer—than CVs, before the general public will develop a positive perception of AVs in terms of road safety. Ironically, our findings also suggest that the inevitable occurrence of more AV-related road accidents will in time lead to a diminishing degree of overweighting of safety issues surrounding AVs.

2.1 Introduction

The advent of Automated—or autonomous, or self-driving—vehicles (AVs) is generally believed to have positive effects on a variety of dimensions, such as road capacity (Shladover et al., 2012), emissions (Greenblatt & Saxena, 2015), travel time (de Almeida Correia et al., 2019), and road safety (Simonite, 2013; Sparrow & Howard, 2017). As for the latter aspect, AVs are expected to have a significant impact on decreasing the number of traffic accidents, as many car crashes are the result of human errors (Singh, 2015), which can be vastly reduced by the assistance of automation technology (Fagnant & Kockelman, 2015; Anderson et al., 2016). However, a considerable degree of societal anxiety exists regarding safety aspects of AVs; colloquially discussed in terms of anxiety and fear of "being killed by a robot". A report from the American automobile association (AAA, 2017) revealed that 78% of American drivers were afraid to ride in AVs. The dominating societal concern about this new technology is software hacking or misuses, according to a large-scale survey that contains 5,000 respondents from 109 countries (Kyriakidis et al., 2015). Other risks that the public is concerned about include hardware or software failure (Piao et al., 2016). Moreover, the public debate has given ample attention to potentially disturbing moral aspects of AVs, such as the necessity for AVs to make life-and-death decisions (Bonnefon et al., 2016; Shariff et al., 2017; Awad et al., 2018); this may lead to severe ethical concerns and hesitates among planners and regulators. All these perceived risks regarding safety aspects can be big barriers standing in the way of AV mass adoption.

As a result of these social concerns, AVs seem to be subject to what may be called a safety paradox: although AVs are expected to substantially contribute to road safety (i.e., eliminating a large share of traffic accidents), concerns among the public (and as a result, policymakers) regarding AV safety issues may well hamper or delay the rollout of AVs. This paradox has found its way to public debates, as illustrated by the following two examples. Gill Prate, CEO of the Toyota Research Institute, stated that "even cutting the number of annual fatalities in half—saving 18,000 lives in the United States for example—would not be enough for AVs to win the public' trust" (Mervis, 2017). Amnon Shashua, a maker of AV technology, even claims that a thousand-fold improvement in safety, compared to conventional vehicles (CVs), is needed for AVs to be accepted by the general public (Economist, 2018). Such discussions can be also found in academic papers. Nees (2019) claimed that although the safety level required for AVs is still unclear, the benchmark of being safer than average human drivers does not seem to be adequate. A very recent study conducted by Liu et al. (2019) attempts to answer this provocative question "how safe is safe enough for AVs?" They examined the relationship between risk frequencies (e.g., one fatality per one million population) and risk-acceptance rates (i.e., the percentage of people accepting presented risk scenarios). The results showed that AVs should be four to five times safer than human drivers in order to be tolerated by the public.

Given the important policy implications of this safety paradox—a delay in the introduction of AVs may in fact inadvertently lead to a failure to avoid thousands of traffic fatalities (Kalra & Groves, 2017), this paper attempts to put this apparent safety paradox to the empirical test, and also to find potential explanations for it. More specifically, we answer the research questions: whether—and if so, why—the fatalities caused by AVs carry more weight to the general public, compared to the fatalities caused by CVs. The starting point for our analysis is a recent stated choice (SC) study (Overakker, 2017) which positively answers the "whether" part of the above research questions; the study finds that AV-fatalities are weighted much more than CV-fatalities. Specifically, AV fatalities caused by software bugs are weighted around four times higher than fatalities caused by CVs; AV fatalities caused by software hacking are weighted even higher—5.5 times. These empirical findings are in line with what was reported in Liu et al. (2019), despite that these two studies used different approaches. In this research, we conduct two SC experiments. Our first experiment, called experiment A, aims to replicate[2] the study by Overakker, to examine whether—and the extent to which—AV-fatalities are weighted more than CV-fatalities. The second experiment, called experiment B, is designed specifically to find explanations for the overweighting of AV-fatalities, compared to CV-fatalities[3]. More specifically, experiment B is designed in such a way that it allows us to answer the following question: is the difference in the weighting between AV-fatalities and CV-fatalities caused by the intrinsic, qualitative differences between an AV and a human-operated CV? Or is it caused by the fact that, quantitatively speaking, the level of AV-related fatalities is currently so low that each additional fatality receives considerable extra weight? Note that disentangling these two potential explanations is crucial, not just from a scientific point of view, but also from a policy point of view. For example, the first explanation would suggest that problematic safety perceptions surrounding AVs may persist at least during the near future; up until the point where humans have truly got used to interacting with AVs. While the second explanation would suggest that, ironically, the inevitable occurrence of more AV-related road accidents will in time lead to a diminishing degree of overweighting of safety issues surrounding AVs.

This paper aims to contribute to the literature on social acceptance of AVs by focusing on a very specific, but highly salient aspect of safety. We use SC experiments to derive and explain the differences in weight attached by citizens to AV- and CV-fatalities, rather than asking them directly about their safety perceptions concerning AVs (relative to CVs); to the best of our knowledge, this approach has not yet been used in the scholarly literature. The remainder of this paper is organized as follows: the next section introduces the experiments and data collection effort. Section 3 presents models and estimation results. Section 4 discusses the obtained findings and related policy implications, as well as limitations.

2.2 Experimental designs and data collections

2.2.1 Experimental designs

Experiments were designed to observe choices that respondents made between hypothetical scenarios in the form of policy packages for the AV era. Specifically, respondents were informed that the government was considering to develop a long-term transport policy to anticipate and facilitate the large-scale introduction of AVs. The background was set at year

---

[2] Note that we do not aim to replicate the exact same experiment as Overakker's. His experiment focused on the social acceptance of AVs which included many AV-related attributes. This research aims to examine the difference in weighting between AV- and CV-fatalities, and to find explanations for the difference; details regarding experimental set-ups are provided in Section 2.2.

[3] Note that the experimental set-up in Overakker (2017) did not allow for studying the possible causes for the overweighting of AV-fatalities.

2045, and participants were informed that about 50% of traffic would consist of full AVs, according to recent studies regarding AV predictions (Litman, 2019; Nieuwenhuijsen et al., 2018). The hypothetical scenarios were described in terms of the consequences of the policy packages, which were presented in choice tasks of three alternatives. Each alternative was described by four following attributes (attribute levels are in brackets):

- The number of fatalities per year caused by CVs (250, 300, 350, 400);
- The number of fatalities per year caused by technical failures (e.g., a software bug) of AVs (50, 100, 150, 200);
- The number of fatalities per year caused by a malicious act (e.g., software hacking) regarding AVs (0, 30, 60, 90);
- The average reduction in car travel time (-30%, -20%, -10%, 0%).

See Figure 2.1 for a detailed wording of these four attributes. The attribute levels of these fatalities were designed according to the current situation in the Netherlands—there are around 600 fatalities each year (SWOV, 2019), and all these fatalities are CV-related. It is uncertain whether the partial introduction of AVs would increase the fatal rate of CV-fatalities or not in the future, thus the attribute level of CV-fatalities was designed to range from 250 to 400, which roughly pivots around 300 (which is half of the current number of CV-related fatalities). Note that while the focus of the experiments is on AV- and CV-fatalities, car travel time was also included in order to help reduce the odds that respondents might add up all fatality numbers and then choose the smallest total. As for other possibly relevant criteria (e.g., cost, feasibility), respondents were informed that they were the same in every policy package.

A D-efficient design was applied to ensure a statistically efficient data collection (Rose & Bliemer, 2009). The priors were obtained by conducting a small pilot study (N=31). Eventually, twelve choice tasks were generated, and they were grouped into two blocks containing six tasks each. Each respondent was asked to complete one block.

2.2.2 Experiment treatment

As mentioned in the Introduction, there are two experiments designed in this study. These two experiments contained the same choice tasks. The only difference was the reference level provided to respondents. More specifically, before performing the choice tasks, respondents were requested to read an introduction page that contained reference levels of CV- and AV-fatalities. The details regarding the reference levels are described as follows.

- Experiment A

Experiment A provided real current levels in the context of the Netherlands: 600 fatalities per year caused by CVs, zero fatalities caused by AVs (either by technical failures or deliberate misuse), and a zero percent reduction in travel time. See Figure 2.1[4] for an example of how these current level-based reference points are visualized.

[4] The original questionnaire was in Dutch.

Q1. Suppose that these are the outcomes, in 2045, of 3 different policy packages. Which policy package would you vote for?

	Policy package 1	Policy package 2	Policy package 3	current situation
Average reduction in car travel time	-10%	-30%	0%	0%
Fatalities caused by **conventional cars** (e.g. the driver not paying attention)	350	250	400	600
Fatalities caused by **technical failure of the AV** (e.g. a software bug)	200	100	100	0
Fatalities caused by **deliberate misuse of the AV by an external party** (e.g. software hack)	0	90	0	0

Figure 2.1 An example of a choice task in experiment A

- Experiment B

Instead of presenting the current reference levels, experiment B provided respondents with projections of future reference levels for each attribute. We varied the reference levels shown to the respondents in experiment B. Such variations were created as follows: respondents were told that four experts had given their predictions concerning AV- and CV-fatality numbers, and travel time reductions in 2045. Respondents were randomly assigned an expert (I, II, III, or IV) to see his/her personal prediction in terms of expected fatality numbers, before revealing their preference for a particular policy package. Table 2.1 shows the levels embedded in the four treatment variations of experiment B, as well as the benchmark levels used in experiment A. Note that in experiment B, reference levels for AV-fatalities were increased, reference levels for CV-fatalities were decreased (both compared to the current situation). Figure 2.2 shows an example choice task accompanied by the reference level given by Expert I's estimates.

Table 2.1 Overview of the reference levels in experiments A and B

	Experiment A	**Experiment B**			
	Current situation	Expert I	Expert II	Expert III	Expert IV
Average reduction in travel time	0%	-10%	0%	-5%	-15%
Fatalities caused by CVs /year	600	400	350	300	260
Fatalities caused by technical failure of the AV/year	0	80	160	110	170
Fatalities caused by a deliberate misuse of the AV/year	0	50	20	80	60

Q1. Suppose that these are the outcomes, in 2045, of 3 different policy packages. Which policy package would you vote for?

	Policy package 1	Policy package 2	Policy package 3	Expert's estimates
Average reduction in car travel time	-10%	-30%	0%	-10%
*Fatalities caused by **conventional cars** (e.g. the driver not paying attention)*	350	250	400	400
*Fatalities caused by **technical failure of the AV** (e.g. a software bug)*	200	100	100	80
*Fatalities caused by **deliberate misuse of the AV by an external party** (e.g. software hack)*	0	90	0	50

Figure 2.2 An example of a choice task in experiment B

The scientific aim behind the treatment is that the creation of different (compared to the current situation) reference points allows for the identification of possible reference point effects as discussed in the Introduction. More specifically, we hypothesize that a (partial) reason behind the overweighting of AV-fatalities relates to the simple fact that current levels of AV-related fatalities are extremely low (zero AV-related fatalities), compared to the current level of AV-fatalities.

- Reflection question

After the choice experiment, respondents were also presented with a reflection question. In it, we asked them to evaluate the extent to which they considered the reference level presented to them. In experiment A, respondents were asked to indicate, on a five-point Likert scale (ranging from strongly disagree to strongly agree, to what extent they agreed with the following proposition: “I considered the current situation when making choices.” , and in experiment B, the proposition read: “I considered the expert’s estimates when making choices.” The distribution of answers will be shown in Section 2.3.

Finally, the whole procedure of our experiments is depicted in Figure 2.3.

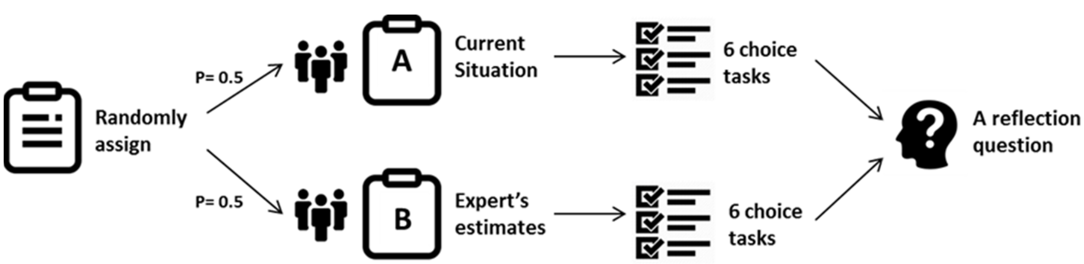


Figure 2.3 The procedure of the choice experiment

2.2.3 Data collection

The data collection was conducted during late April and early May, 2018 in the Zuid-Holland province of the Netherlands; especially within the cities of The Hague and Delft. Respondents were approached at random in the vicinity of public parking facilities within an invitation to fill out, on the spot, a paper-pencil survey, or within a flyer containing the URL and QR-code of the survey. A final sample of 412 completed questionnaires, filled out by individuals who were at least 17 years old, were obtained via either the paper-pencil (N=232) or online (N=180) version. Each participant was randomly assigned to one of the two experiments, leading to a final sample of 197 individuals for experiment A, and of 214 individuals for experiment B. Note that this random assignment to either the survey without or the one with experimental treatment (i.e., the presence of an artificial reference point) provides a stronger mechanism to identify causal treatment-effects than if we were to ask each individual to make choices in the context of both the surveys with and without treatment. The socio-demographic characteristics of samples A and B are shown in Table 2.2.

Table 2.2 Socio-demographic and other relevant characteristics of the sample

Variables		**Experiment A**	**Experiment B**
Gender	Male	62.4%	58.1%
	Female	35.0%	39.5%
	Unknown	2.5%	2.4%
Age	17-25	34.5%	41.4%
	26-50	41.1%	30.7%
	51-70	21.4%	22.8%
	Unknown	3%	5.1%
Completed education	Primary school	1.5%	1.8%
	Middle school	13.2%	16.3%
	High school	20.8%	21.4%
	Bachelor	48.2%	49.8%
	Master or doctor	14.7%	10.7%
	Unknown	1.5%	0.0%
Current usage of Advanced Driver Assistance System (ADAS) (e.g. Adaptive cruise control, lane departure warning, automated park assistant)	Yes	44.7%	42.8%
	No	53.8%	54.0%
	Unknown	1.5%	3.2%
Experience with minor accidents	Yes	40.6%	41.4%
	No	58.4%	57.7%
	Unknown	1.0%	1.0%
Experience with severe accidents	Yes	10.7%	9.2%
	No	89.3%	90.3%
	Unknown	0.0%	0.5%

The Chi-square test shows that the samples of experiments A and B did not differ significantly from one another in terms of gender (P=0.37), age (P=0.13), completed education (P=0.37), current usage of ADAS (P=0.70), experience with minor accidents (P=0.83), and experiences with severe accidents (P=0.61), which indicates that two samples are similar along these lines. This implies that differences between experiments in terms of obtained results are due to the experimental treatment, rather than the differences in the samples. After inspecting the socio-demographic variables, we can notice that both samples are slightly skewed towards males and higher educated people. Specifically, about 60% of the participants were males, and about 60% were highly educated in both samples. However, it should be noted here that no attempt was made to arrive at a representative sample from the Dutch population. The reason for this is that our research aims to search for the first empirical evidence for the overweighting of AV-fatalities, compared to CV-fatalities, which could pave the way for more elaborate explorations and confirmatory studies in different regions and countries. Furthermore, we also checked if the socio-demographic factors have an impact on the degree of overweighting, and found no significant results in that regard.

The distribution of the reflection question is shown in Table 2.3. As expected, respondents to experiment A turned to be more inclined to consider the offered reference point than respondents to experiment B. To respondents, the reference points in experiment B were projections far into the future by an unknown and randomly selected expert, in contrast to reliable information about recent and current levels in experiment A. In next section, we will explore if the level of consideration of reference points has an effect on the overweighting of AV-fatalities.

Table 2.3 To what extent did participants consider the reference points provided to them?

	Experiment A: Reference point is Current situation	**Experiment B**: Reference point is Expert's estimate
Strongly disagree	6.6%	15.3%
Disagree	20.3%	25.6%
Neutral	18.8%	34.0%
Agree	45.7%	23.3%
Strongly agree	8.6%	1.9%

2.3 Modelling methodology and estimation results

2.3.1 Modelling methodology

We develop two models to analyse the choices made in experiments A and B. The first model, which is a baseline model, is used to examine whether, and the extent to which, an AV-fatality is overweighted, compared to a CV-fatality. It is estimated on the data of experiments A and B separately. The second model, called the reference-point model, is applied to explore the effect of reference levels on the (over-)weighting of AV- and CV-fatalities. In order to involve all reference points (i.e., current and artificial reference points), the second model is estimated on the pooled data of experiments A and B.

- Baseline model

The baseline model is a simple Logit-model[5] with the following specification for the systematic utility:

$$V_i = \beta_{CF} CF_i + (1 + \lambda_{AFT}) \beta_{CF} AFT_i + (1 + \lambda_{AFM}) \beta_{CF} AFM_i + \beta_{TR} TR_i. \qquad \text{(Eq. 2.1)}$$

[5] A series of Mixed logit models was estimated as well, to take into account the panel structure of the data and to allow for heterogeneity within the sample in terms of the weight for fatalities and travel time. These models gave qualitatively similar results compared to the Logit model results reported here, and warrant the same conclusions. For clarity of exposition, we limit ourselves to discussing the Logit model outcomes in this paper.

Here, i denotes the alternative (note that alternatives were unlabeled in the experiments). Parameters β_{CF} and β_{TR} refer to fatalities caused by CVs (CF) and average reduction of car travel time (TR) respectively. To facilitate the interpretation of the differences in weighting between (the two types of) AV-fatalities and CV-fatalities, we write coefficients for fatalities caused by AVs as $(1 + \lambda_{AFT})\beta_{CF}$ and $(1 + \lambda_{AFM})\beta_{CF}$ respectively. Here, parameter λ_{AFT} gives the degree to which an AV-fatality caused by technical failure (AFT) is overweighted, compared to a CV-fatality; and parameter λ_{AFM} gives the degree to which an AV-fatality caused by deliberate misuse (AFM) is overweighted, compared to a CV-fatality. If a λ parameter is estimated to be significantly different from zero, this indicates that there is indeed a difference between AV- and CV-fatalities, in terms of the extent to which weight they receive from respondents. By estimating this model on choice data from experiment A, we can examine the overweighting effect of AV-fatalities, compared to CV-fatalities. Furthermore, by comparing estimation results between experiments A and B, we can get a first idea of whether the treatment of artificial reference points embedded in experiment B has had a downward effect on the degree of the overweighting, or not.

Based on the model described in (1) and estimation results of Overakker (2017), we can derive the following hypotheses; we first focus on experiment A:

1. $\lambda_{AFT} > 0$; that is, fatalities caused by technical failure of the AV are overweighted, compared to fatalities caused by CVs.
2. $\lambda_{AFM} > 0$; that is, fatalities caused by deliberate misuse of the AV are overweighted, compared to fatalities caused by CVs.
3. $\lambda_{AFM} > \lambda_{AFT}$; that is, fatalities caused by deliberate misuse of the AV are overweighted, compared to fatalities caused by technical failure of the AV.

If we find support for these three hypotheses, it implies a qualitative (proxy-) replication of results of the results obtained in Overakker (2017). The size of the overweighting found in our experiments could however still be different.

Moving to a comparison between experiments A and B, the following additional hypotheses[6] are considered:

4. $\lambda_{AFT\_A} > \lambda_{AFT\_B}$; that is, the estimate of λ_AFT based on data from experiment A is larger than the corresponding estimate based on data from experiment B, indicating that the overweighting of AV-fatalities caused by technical failure is largest, when the reference point is the current situation (i.e., 0).
5. $\lambda_{AFM\_A} > \lambda_{AFM\_B}$; that is, the estimate of λ_AFM based on data from experiment A is larger than the corresponding estimate based on data from experiment B, indicating that the overweighting of AV-fatalities caused by deliberate misuse is largest, when the reference point is the current situation (i.e., 0).

Together these two hypotheses, if we find support for them, suggest that at least part of the overweighting of AV-fatalities should be attributed not to the intrinsic differences between AVs

[6] Note that we compare the parameters which capture differences in weighting of fatalities (i.e., λ_{AFT} and λ_{AFM}) between two experiments, rather than the coefficients of fatalities themselves.

and CVs, but to the simple fact that AV-fatality levels are currently extremely low, making any additional fatality stand out.

- Reference-point model

To further explore the effect of reference points on the weighting of CV- and AV-fatalities, we propose another Logit model based on the following of specification of the systematic utility, which is estimated on the pooled data of experiments A and B:

$$V_i = (\beta_{CF} + \gamma_{CF}[600 - R_{CF}])CF_i + (\beta_{AFT} + \gamma_{AFT}R_{AFT})AFT_i + (\beta_{AFM} + \gamma_{AFM}R_{AFM})AFM_i + \beta_{TR}TR_i \qquad \text{(Eq. 2.2)}$$

Here, R_{CF} gives the prevailing reference point for CV-fatalities; R_{AFT} and R_{AFM} give the prevailing reference points for AV-fatalities caused by technical failure, and caused by deliberate misuse, respectively[7]. Parameter γ_{CF} represents the effect of a one-unit change in the prevailing reference point for CV-fatalities, which is added to the reference-free weight (β_{CF}) associated with a CV-fatality. Likewise, parameter γ_{AFT} represents the effect of a one-unit change in the prevailing reference point for AV-fatalities caused by technical failure, which is added to the reference-free weight (β_{AFT}) associated with such an AV-fatality; and parameter γ_{AFM} represents the effect of a one-unit change in the prevailing reference point for AV-fatalities caused by deliberate misuse, which is added to the reference-free weight (β_{AFM}) associated with such an AV-fatality. It is important to note here, that in the current situation (i.e., as in experiment A), $R_{CF} = 600$ and $R_{AFT} = R_{AFM} = 0$; so if we substitute the reference levels of the current situation into Equation (2), the equation reduces to the following, simplified utility specification: $V_i = \beta_{CF}CF_i + \beta_{AFT}AFT_i + \beta_{AFM}AFM_i + \beta_{TR}TR_i$.

Based on the model described in (2), we can derive the following additional hypotheses:

6. $\gamma_{CF} < 0$; that is, fatalities caused by CVs receive more (negative) weight, as the reference level becomes lower than the current level of 600 CV-fatalities.
7. $\gamma_{AFT} > 0$ and $\gamma_{AFM} > 0$; that is, as the reference point for AV-fatalities increases, the (negative) weight assigned to such fatalities become smaller.

If we find empirical support for these two hypotheses, it would reinforce the idea that the reference effect plays a role in explaining the higher weight attached to AV-fatalities, compared to CV-fatalities; and more generally, that reference points co-determine the weight attached to additional traffic fatalities of different types.

2.3.2 Model estimation results

We start by estimating the baseline model with systematic utility defined as in Equation (1), on both the data obtained through experiment A and the data obtained through experiment B. Table 2.4 presents model estimation results.

[7] Note that since in this study we are not interested in reference point effects on travel time sensitivity, we include the travel time attribute in our model in a reference-free fashion.

Table 2.4 Estimation results of the base model (Eq. 2.1)

		Experiment A			Experiment B		
		Coeff.	Std. err	t-value	Coeff.	Std. err	t-value
Average reduction in car travel time	β_{TR}	0.032	0.004	7.59	0.016	0.004	4.07
Fatalities caused by conventional cars	β_{CF}	-0.009	0.001	-14.51	-0.008	0.001	-14.09
Fatalities caused by AV technical failure	λ_{AFT}	0.541	0.118	4.60	0.122	0.113	1.08*
Fatalities caused by AV deliberate misuse	λ_{AFM}	1.220	0.170	7.19	0.706	0.164	4.32
Number of observations		1182 (i.e., 197×6)			1290 (i.e., 215×6)		
Null LL		-1299			-1417		
Final LL		-1123			-1274		
Rho-square		0.133			0.101		

* not significantly different from 0 at the 95% confidence level (two-tailed test).

Starting with the column representing experiment A, it is easily seen that support is found for all first three hypotheses. That is, the fatalities caused by AVs are weighted more than fatalities caused by CVs, and the fatalities caused by deliberate misuse of the AV receive higher weight than fatalities caused by technical failure (i.e., $2.2 > 1.5$). The asymptotic t-value for the difference between parameters λ_{AFT} and λ_{AFM} equals 3.28, meaning that these two types of AV-fatalities are weighted differently by participants. If we compare our estimation results with the ones by Overakker (2017), we can find that the degrees of overweighting are smaller: 1.5 as compared to 4 (technical failure) and 2.2 as compared to 5.5 (deliberate misuse). Such differences are expected, given that various changes in experimental designs and also different samples. Our results can, however, still be considered to be in line with those reported by Overakker (2017).

Comparing the estimation results between experiment A and B, we find clear support for hypotheses 4 and 5 (asymptotic t-values for the differences are 2.56 and 2.17 respectively); that is, we find that the degree of overweighting is substantially reduced when reference points for the AV are no longer zero fatalities (but higher), and the reference point for CV is no longer 600 fatalities (but lower). While there is still significant overweighting for deliberate misuse, this is no longer the case for fatalities caused by technical failure of the AV. This provides a strong suggestion that at least part of the overweighting of AV-fatalities compared to CV-fatalities is due to currently very low reference levels for AV-fatalities and a very high level for CV-fatalities.

The reference-point model based on systematic utility specified as Eq. 2.2 is estimated on the pooled data of experiments A and B, the estimation results are presented in Table 5. It can be seen, that no support is found for hypothesis 6; that is, we do not find a lower—than the current level of 600—reference point for CV-fatalities leads to a significantly higher weight carried by per additional CV-fatality. Although the signs of γ_{AFT} and γ_{AFM} are as expected—meaning that as the reference point for AV-fatalities increases from zero, the (negative) weight assigned to such fatalities becomes smaller, these effects are not significant at 95% levels of significance.

Table 2.5 Estimation results of the reference-point model

	Experiment A + Experiment B		
	Coeff.	Std. err	t-value
β_{TR}	0.0231	0.0028	8.17
β_{CF}	-0.0082	0.0006	-14.44
γ_{CF} (per $R_{CF}/100$)	0.0002	0.0003	0.61*
β_{AFT}	-0.0119	0.0008	-14.72
γ_{AFT} (per $R_{AFT}/100$)	0.0010	0.0007	1.59*
β_{AFM}	-0.0172	0.0013	-13.83
γ_{AFM} (per $R_{AFM}/100$)	0.0026	0.0023	1.16*
Number of observations	2472 (i.e.,197×6+215×6)		
Null LL	-2716		
Final LL	-2401		
Rho-square	0.116		

* not significantly different from 0 at the 95% level (two-tailed test).

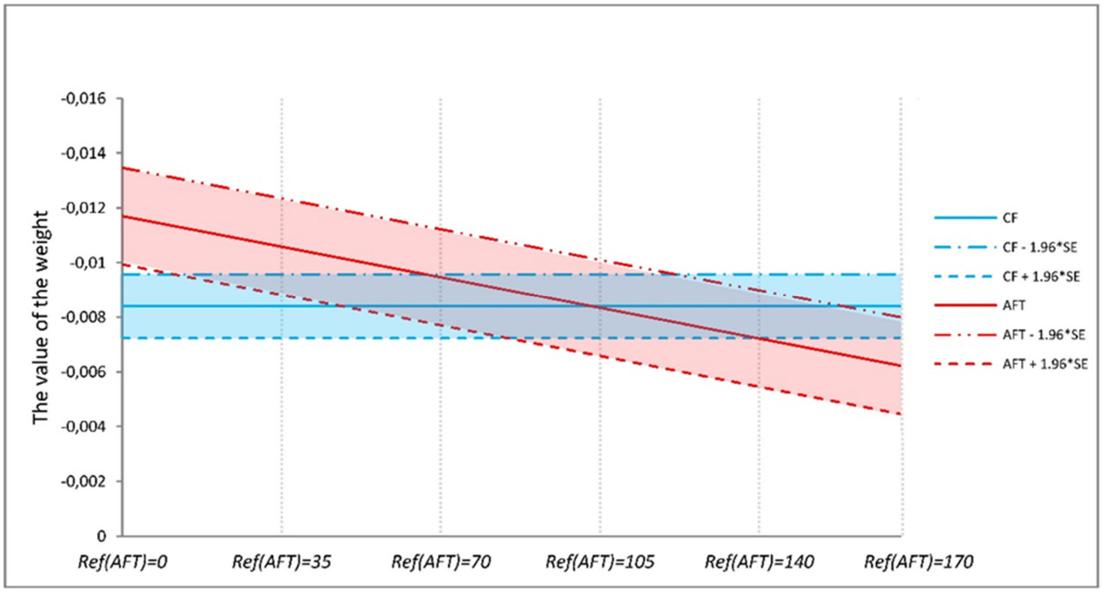
However, it should be noted here that, as highlighted in Section 2.1.2 (Table 2.3), a substantial share of respondents (40.9%) in experiment B indicated that they did not actually consider the artificial reference points when making choice. It is interesting to separately analyze the subsample who indicated that they did consider the reference point. When excluding those participants who indicated that they either disagreed or strongly disagreed with the statement “I considered the expert’s estimates when making choices”, a different picture arises, as can be seen in Table 2.6.

Table 2.6 Estimation results of the reference-point model, subsample of participants who did not express (strong) disagreement with the following proposition: *"I considered the expert's estimates when making choices"*

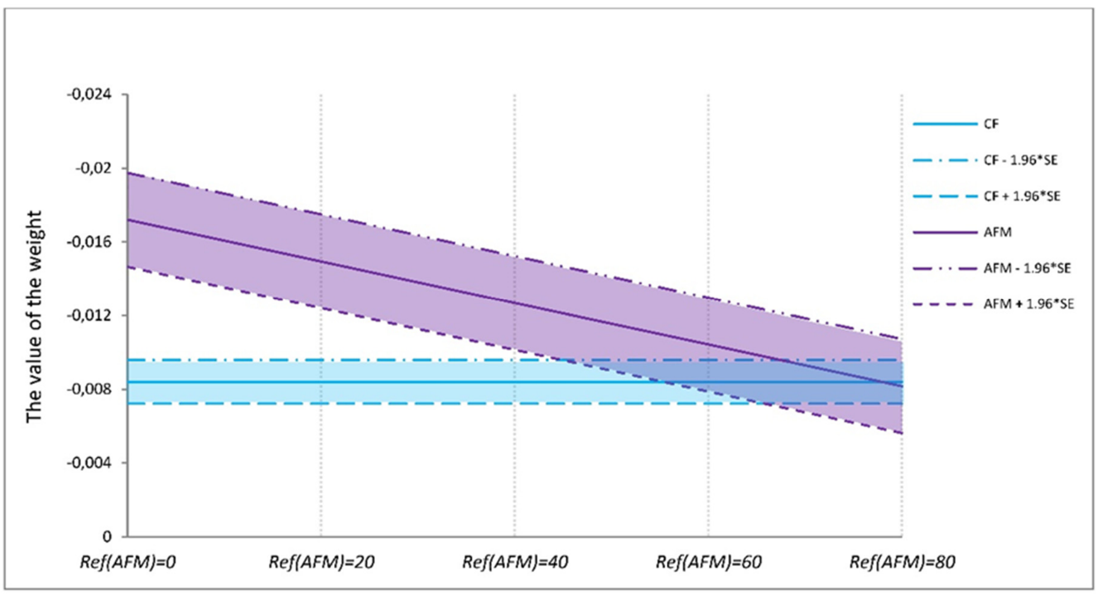
	Experiment A + Experiment B (subsample)		
	Coeff.	Std. err	t-value
β_{TR}	0.0211	0.0032	6.65
β_{CF}	-0.0084	0.0006	-14.22
γ_{CF} (per $R_{CF}/100$)	0.0000	0.0003	0.03 *
β_{AFT}	-0.0117	0.0009	-13.61
γ_{AFT} (per $R_{AFT}/100$)	0.0032	0.0008	4.13
β_{AFM}	-0.0172	0.0013	-12.87
γ_{AFM} (per $R_{AFM}/100$)	0.0113	0.0028	4.10
Number of observations	1944 (i.e.,197×6+127×6)		
Null LL	-2136		
Final LL	-1896		
Rho-square	0.112		

* not significantly different from 0 at the 95% level (two-tailed test)

Table 2.6 shows that, for the subsample of respondents who indicated that they considered the artificial reference points in experiment B, we clearly find support for the hypothesis 7, that is increases in the reference point concerning AV-fatalities lead to a decline in the weight associated with such fatalities. This effect is strongest for fatalities caused by AV-deliberate misuse, but also exists for fatalities caused by AV-technical failure. This finding is further illustrated in Figure 2.4, where we plot the trends in fatality-weights including their 95% confidence intervals. It is clearly seen that as AV-fatalities move away from the current reference point of zero, at some point there is no overweighting left, relative to CV-fatalities. Furthermore, we notice that parameter γ_{CF} is not significant, suggesting that decreasing the reference level does not influence the weighting of CV-fatalities. One possible explanation is that people's perceptions and beliefs regarding CV-fatalities are relatively stable compared to their perceptions and beliefs regarding new types of fatalities, e.g., AV-fatalities; as a result, changing reference levels does not affect how they view and weigh CV-fatalities.



(a) AFT



(b) AFM

Figure 2.4 The trends of the weight associated with AV-fatalities for increasing values of the corresponding reference points (inverted scales)

2.4 Conclusions and discussions

By analysing the choices made by respondents in our first SC experiment, we found that fatalities caused by AVs received more weight than fatalities caused by human drivers in CVs. The difference in weighting equalled 54% for AV-fatalities caused by technical failure of the AV (e.g., software bugs), and 122% for AV-fatalities caused by deliberate misuse of the AV (e.g., software hack). These degrees of overweighting are somewhat smaller than, but are still within the same order of magnitude as, the results in a previous SC experiment (Overakker, 2017). In the second SC experiment, we explored a specific potential reason for this overweighting of AV-fatalities. Specifically, we hypothesized that the current levels of AV-fatalities are so low (i.e., zero) that any additional AV-fatality carries more weight, this having little to do with intrinsic differences between AV- and CV-fatalities in public perception. As hypothesized, we found that by artificially increasing the reference levels of AV-fatalities, we were able to substantially reduce respondents' overweighting of AV-fatalities caused by deliberate misuse, and even eliminate the overweighting of AV-fatalities caused by technical failure.

Delving into the reasons behind the reference level effect, we can find compelling explanations in social science literature. First, *loss aversion* in prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1991) suggests that losses are weighted more heavily than equivalent gains, in this case implying that the increase (loss) in AV-fatalities (relative to the current reference point of zero) will receive more weight than a corresponding decrease (gain) in CV-fatalities. Second, *probability weighting* in prospect theory may also offer a partial explanation: given the currently very low probability of having a fatal accident with an AV, the perceptions of this risk are inflated due to the tendency to overweight small probabilities. Third, and probably most importantly, the so-called Weber effect (Weber, 1834; Gescheider, 2013) suggests that a change relative to a low baseline is perceived as bigger than the same—in an absolute sense—change compared to a higher base-level. In our case, this implies that a change from, for example, zero to ten AV-fatalities weighs more heavily than a change from 600 to 590 or to 610 CV-fatalities.

What implications does this study have for the public debate surrounding the topic of safety issues of AVs, and for policymaking in this domain? The first implication of our results is that the safety paradox affecting AVs—i.e., the notion that while AVs are expected to bring great traffic safety benefits, problematic safety perceptions may delay or even prohibit their introduction—seems to be less salient than many experts have suggested in the public debate. For example, our results suggest that cutting the number of annual fatalities in half by implementing AVs in the Netherlands would already be considered acceptable by Dutch citizens. This is a much smaller decrease than what is often heard in policy and public debates. Furthermore, our findings suggest that as the number of AV-fatalities increases—as will inevitably be the case, once they are more and more allowed to drive on real roads—the extent to which they will be overweighted compared to CV-fatalities will decrease (ceteris paribus). This implies that, somewhat ironically, the occurrence of accidents involving AVs will help redress the above-mentioned safety paradox surrounding AVs. In combination, our findings suggest that once AVs can save at least half of the number of lives lost annually in traffic accidents, there should be no reason to fear that safety perception issues among the general public will backfire to the extent that AV-acceptance becomes highly problematic. In that sense, we concur with Kalra & Groves (2017) that "the perfect should not be the enemy of the good": once AVs have become considerably safer than CVs, they should be allowed on the road as soon as possible, to speed up their learning process (making them even safer) and to save, during this process, lives that would otherwise have been lost in accidents involving CVs.

Clearly, these conclusions and implications are based on a study which has its limitations: first, we used SC-experiments involving participants making hypothetical choices, based on fatality-statistics. Although we believe that this type of experiment is well suited to answer the type of questions posed in this paper—that is, to infer weight from choices, rather than simply asking people for their weight directly—the hypothetical nature of our experiment should be kept in mind when interpreting results. Likewise, although accident statistics (like the ones used in our experiments) tend to play an important role in public debates about road safety, it is likely that other, non-numeric, discussions of AV-safety issues in the media and public debate may affect the safety debate surrounding AVs in ways that go beyond the scope of the conclusions that may be drawn from our study. Furthermore, it should once again be brought to attention that our sample is a relatively small convenience sample recruited in a confined urban area within the Netherlands. Although we did not find evidence for any effect of socio-demographic variables on our main results, it is to be expected that the degree of overweighting of AV-fatalities—compared to CV-fatalities—will be country and (car-) culture specific. Besides, we applied two means of recruiting respondents, immediate responses (i.e., paper-pencil survey) and online responses. Although these two approaches combined are often used in SC experiments, it may bring unknown bias to results as immediate responses usually have less dwell time to think about questions. Last but not least, in this study, we merely look at the reference effect on the overweighting of AV-fatalities, but there can be other possible explanations. An intriguing one from social psychology may also provide an explanation: the general public holds new technologies to higher standards than traditional ones (Fischhoff et al., 1978; Otway & Von Winterfeldt, 1982). Nevertheless, we consider our small-scale study to provide a potential stepping stone for future studies that involve representative samples from different countries where AVs will hit the road in the near future.

Acknowledgement

The first author gratefully acknowledges financial support for her PhD study from the China Scholarship Council (No. 201608310106). The third author wishes to thank the European Research Council for financial support (ERC Consolidator grant BEHAVE—724431).

References

AAA. (2017). Americans feel unsafe sharing the road with fully self-driving cars. Retrieved from https://newsroom.aaa.com/2017/03/americans-feel-unsafe-sharing-road-fully-self-driving-cars/

Anderson, J. M., Nidhi, K., Karlyn, D. S., Paul, S., Constantine, S., & Tobi, A. O. (2016). Autonomous vehicle technology: A guid for policymakers. Retrieved from https://www.rand.org/pubs/research_reports/RR443-2.html

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., . . . Rahwan, I. (2018). The moral machine experiment. *Nature, 563*(7729), 59-64.

Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science, 352*(6293), 1573-1576.

de Almeida Correia, G. H., Looff, E., van Cranenburgh, S., Snelder, M., & van Arem, B. (2019). On the impact of vehicle automation on the value of travel time while performing work and leisure activities in a car: Theoretical insights and results from a stated preference survey. *Transportation Research Part A: Policy Practice, 119*, 359-382.

Economist, T. (2018). How do you define "safe driving" in terms a machine can understand? Writing the robotic rules of the road. Retrieved from https://www.economist.com/science-and-technology/2018/05/10/how-do-you-define-safe-driving-in-terms-a-machine-can-understand

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy Practice, 77*, 167-181.

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy sciences, 9*(2), 127-152.

Gescheider, G. A. (2013). *Psychophysics: the fundamentals*: Psychology Press.

Greenblatt, J. B., & Saxena, S. (2015). Autonomous taxis could greatly reduce greenhouse-gas emissions of US light-duty vehicles. *Nature Climate Changes, 5*(9), 860-863.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Kalra, N., & Groves, D. G. (2017). *The enemy of good: estimating the cost of waiting for nearly perfect automated Vehicles.* Retrieved from https://www.rand.org/pubs/research_reports/RR2150.html

Kyriakidis, M., Happee, R., de Winter, J. C., & behaviour. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation research part F: traffic psychology, 32*, 127-140.

Liu, P., Yang, R., & Xu, Z. (2019). How safe is safe enough for self-driving vehicles? *Risk analysis, 39*(2), 315-325.

Mervis, J. (2017). Are we going too fast on driverless cars? *Science*. Retrieved from http://www.sciencemag.org/news/2017/12/are-we-going-too-fast-driverless-cars

Nees, M. A. (2019). Safer than the average human driver (who is less safe than me)? Examining a popular safety benchmark for self-driving cars. *Journal of Safety Research, 69*, 61-68.

Otway, H. J., & Von Winterfeldt, D. (1982). Beyond acceptable risk: On the social acceptability of technologies. *Policy sciences, 14*(3), 247-256.

Overakker, B. (2017). *The Social Acceptance of Automated Driving Systems: Safety Aspects: A contribution to responsible innovation by using a referendum format, discrete choice model experiment to measure the social acceptance of ADS by Dutch citizens with corresponding heterogeneity.* (Master), Delft University of Technology, Retrieved from https://repository.tudelft.nl/islandora/object/uuid:31159f82-e33d-4124-a5a2-2b605974870e

Piao, J., McDonald, M., Hounsell, N., Graindorge, M., Graindorge, T., & Malhene, N. (2016). Public views towards implementation of automated vehicles in urban areas. *Transportation research procedia, 14*(0), 2168-2177.

Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2017). Psychological roadblocks to the adoption of self-driving vehicles. *Nature Human Behaviour, 1*(10), 694-696.

Shladover, S. E., Su, D., & Lu, X.-Y. (2012). Impacts of cooperative adaptive cruise control on freeway traffic flow. *Transportation Research Record, 2324*(1), 63-70.

Simonite, T. (2013). *Data shows Google's robot cars are smoother, safer drivers than you or I*. Retrieved from https://www.technologyreview.com/s/520746/data-shows-googles-robot-cars-are-smoother-safer-drivers-than-you-or-i/

Singh, S. (2015). *Critical reasons for crashes investigated in the national motor vehicle crash causation survey*. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506

Sparrow, R., & Howard, M. (2017). When human beings are like drunk robots: Driverless vehicles, ethics, and the future of transport. *Transportation Research Part C: Emerging Technologies, 80*, 206-215.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics, 106*(4), 1039-1061.

Weber, E. H. (1834). *De pulsu, resorptione, auditu et tactu: annotationes anatomicae et physiologicae, auctore*: prostat apud CF Koehler.

3 Incorporating relative thinking into the Random Regret Minimization model framework

Abstract

People tend to exhibit what is called “relative thinking”, that is people care about relative differences between choice alternatives, not just absolute differences. This chapter specifically looks at two types of relative thinking: level-based relative thinking—i.e. actual differences between attribute levels may seem big or small depending on the initial level of the attribute—and range-based relative thinking—i.e. actual differences may appear big or small depending on the range size of the choice set. We incorporate these two types into the Random Regret Minimization (RRM) model framework, resulting in several new model specifications. The new model specifications are estimated on four empirical datasets. The estimation results show that the incorporation of level-based or range-based relative thinking may lead to great improvements in model fits of the RRM models, especially when there are various attribute levels in the choice set or various range sizes throughout the whole choice sets. Based on the findings, we make some recommendations and suggest several directions for future research.

3.1 Introduction

The relative nature of decision making has been well recognized and studied in diverse fields, most notably in psychology and consumer behaviour studies. Thaler (1980) presented an example of consumer behaviour in which people may be more willing to save \$5 on a \$25 radio, versus on a \$500 television. Similar experiments were later replicated in many studies (e.g., Tversky & Kahneman, 1981; Mowen & Mowen, 1986; Frisch, 1993; Ranyard & Abdel-Nabi, 1993). The well-known “jacket-calculator” experiment by Tversky & Kahneman (1981) showed that people were more willing to travel 20 minutes to save \$5 on purchasing a \$15 calculator (and a \$125 jacket) than on purchasing a \$125 calculator (and a \$15 jacket).

These experiments are good examples of people's tendency to exhibit so-called *relative thinking*: People care not only about absolute savings but also relative savings (relative to a good's price) when purchasing goods. The term, relative thinking, was first proposed by Azar (2004). It is used to describe the phenomenon of people being less sensitive to a given price change at a higher price level. Relative thinking has received much attention in marketing and consumer behaviour studies, especially in the studies of sale promotions and behavioural pricing (Azar, 2008; Saini et al., 2010; Saini & Thota, 2010; Nicolau, 2012).

Theoretical underpinnings of relative thinking can be traced back to Weber's law (Weber, 1834). It asserts that people's responses to changes in a physical stimulus, such as weight, brightness and loudness, are related to the original intensity of the stimulus. It defines the relationship between *human perceptions* and actual changes in physical stimuli, and now it has been applied to a broader context. For example, travel behaviour research has considered Weber's law in the evaluation of travel time savings (Hensher, 1976; Ben-Akiva & Lerman, 1985; Welch & Williams, 1997; Cantillo et al., 2006). Hensher (1976) pointed out that travel time savings should be evaluated within bounds: for a long-length trip, a small travel time saving may fall into an insensitivity region that people cannot perceive. Therefore, it is debatable whether small travel times should be assigned a value or not. Cantillo et al. (2006) found that the application of the models that overlook relative travel time differences may lead to biases in model estimations and predictions.

More broadly, relative thinking refers to a behavioural phenomenon that people's choice behaviour is influenced by relative differences between choice alternatives, and Weber's law is just one type of it. Weber's law can be regarded as *level-based* relative thinking: *a given absolute attribute difference can seem big or small depending on the original level of the attribute*. Take an example of a 10-minute travel time saving. According to Weber's law, saving 10 minutes on a short trip (e.g., 20 minutes) seems much more considerable than on a long trip (e.g., 110 minutes). Recently, Bushong et al. (2021) have extended this notion to *range-based* relative thinking: *a given absolute attribute difference can seem big or small depending on the range of this attribute in the choice set.* They proposed a model of range-based relative thinking, in which decision makers would attach less weight to a given change when faced with a wider range of the choice set. In the previous example, range-based relative thinking suggests that the range of the choice set may also affect people's perception of the 10-minute saving: the 10-minute travel time savings may seem to be greater when presented in a narrow range (e.g., 10, 20, 30 min) than in a wide range (e.g., 10, 20, 110 min).

Range-based relative thinking also has profound theoretical underpinnings. Seminal work conducted by Volkmann (1951) and Parducci (1965) explored how the size of the range would affect people's judgment of absolute differences. More recently, many empirical findings emerged from the field of marketing showing that a given change in attribute levels has a greater impact when presented in a narrow range than in a wide range (e.g., Mellers & Cooke, 1994; Wedell, 1998; Ohler et al., 2000; Moon & Voss, 2009). We call this finding *the range effect*. Mellers & Cooke (1994) found that the attractiveness difference between a \$200 apartment and a \$ 400 apartment was greater when all alternative apartments ranged from \$200 to \$400 than they ranged from \$100 to \$1000. Wedell (1998) conducted a series of experiments and confirmed the results in Mellers & Cooke (1994) on the range effect. Moreover, Wedell (1998) proposed a so-called value-shifted model in which the weight of the attribute-level difference is assumed inversely proportional to the attribute range in the choice set.

Although relative thinking (level-based or range-based) has attracted much attention in many fields, surprisingly it has not been widely applied to travel choice models. In travel behaviour research, the commonly used models are based on the Random Utility Maximization (RUM) model, for example, a linear-additive utility function of attribute levels and associated weights. Such models mainly look at actual attribute levels of choice alternatives, rather than differences in attribute levels between alternatives. Thus, relative thinking—which is about how people respond to differences—is rarely incorporated into common travel choice models. Recent years have seen a growing popularity in using an alternative model, called the Random Regret Minimization (RRM) model (Chorus et al., 2008; Chorus, 2010; van Cranenburgh et al., 2015). This model looks at the relative performance of the alternatives (attribute-level differences), which provides a natural way of incorporating relative thinking into its model specifications.

The RRM model is regarded as a regret-based counterpart of the RUM model. It postulates that when making choices, a decision maker wishes to minimize regret rather than maximize utility as the RUM model postulates. In the RRM model, the regret function is defined as a function of attribute-level differences, which means the same differences in attribute levels would cause the same amount of regret. Jang et al. (2017) incorporated the idea of level-based relative thinking into one type of the RRM model—the classical RRM model (Chorus, 2010): the regret function is assumed to be a function of relative attribute-level differences (relative to the attribute level of the chosen alternatives), not absolute attribute-level differences assumed in the original model specification. They tested the new model specifications in two empirical datasets and found significant improvements in model fits compared to the original models.

In this study, we will further explore and empirically examine the incorporation of relative thinking into the RRM model framework. Specifically, we first incorporate level-based and range-based relative thinking into the RRM models respectively, resulting in a series of new model specifications—the RRM-Level and RRM-Range models. The former is inspired by Jang et al. (2017)'s work. By extending their method to the most generalised μRRM model (van Cranenburgh et al., 2018), the μRRM-Level model is obtained. The latter is inspired by the value-shifted model (Wedell, 1998) in which the weight attached to an attribute-level difference is assumed to be inversely proportional to the attribute range. Such a transformation is incorporated into two specific RRM models—the μRRM model and the classical RRM model (Chorus, 2010). Second, we empirically test these new models on four datasets in terms of their model fits and discuss the results of one dataset in full detail. The four empirical datasets have been used in the previous travel behaviour studies. One is the revealed preference dataset and the other three are stated preference datasets. The new models will be compared against each other as well as against the existing RRM models and the benchmark RUM model.

The reminding part of this chapter is organized as follows. Section 3.2 first summarizes the existing RRM models, including the classical RRM model, the μRRM model, and Jang et al. (2017)'s RRM-Level model, and then outlines a series of new RRM model specifications. Section 3.3 discusses conceptual differences between the new model specifications. Section 3.4 compares the empirical performance of these models. Conclusions and discussions are given in Section 3.5.

3.2 The RRM, RRM-Level and RRM-Range models

3.2.1 The RRM model

The RRM model is regarded as a regret-based counterpart of the RUM model. Regret occurs when the chosen alternative leads to an undesirable outcome, for example, an unchosen alternative outperforms the chosen one. Regret-based theories or models are built on the premise that to avoid regret (regret aversion), decision makers choose the alternative which brings the minimum regret. In the RRM environment, the regret of an alternative is defined as the sum of all pairwise regrets that are generated from comparisons between this alternative and other competing alternatives in terms of each attribute.

The RRM model has several generations of model specifications. The most generalised version is the μRRM model proposed by van Cranenburgh et al. (2015), see Eq. 3.1. It accommodates different degrees of the regret-averse behaviour and has several special cases when the regret aversion parameter μ is limited to certain values. The model specification of μRRM is given by

$$R_i^{\mu RRM} = \sum_{j \neq i} \sum_m \mu \ln(1 + \exp(\frac{\beta_m}{\mu}[X_{jm} - X_{im}])), \qquad \text{(Eq. 3.1)}$$

where X_{im} and X_{jm} denote the level of attribute m associated with alternative i and competing alternative j respectively, β_m denotes the taste parameter of attribute m, and μ is the regret aversion parameter that can take on any values from the domain (0, +∞). Parameter μ governs the degree of regret aversion in behaviour. In particular, when μ approaches zero (infinitely), the μRRM model becomes the P-RRM model, which implies the strongest regret-averse behaviour; when μ is estimated to be large (in practice, when $\mu > 5$), the μRRM model exhibits the linear-additive RUM behaviour, suggesting no regret aversion in behaviour; and when μ equals one, the μRRM model collapses to the most commonly used RRM model—the classical RRM (CRRM) model; this model implies that there is a moderate regret aversion in behaviour. The model specification of the CRRM model is expressed by

$$R_i^{CRRM} = \sum_{j \neq i} \sum_m \ln(1 + \exp(\beta_m[X_{jm} - X_{im}])), \qquad \text{(Eq. 3.2)}$$

What the RRM model specifications have in common is that regret is defined as a function of attribute-level differences (i.e., $f(X_{jm} - X_{im})$). This means same attribute-level differences will generate the same amount of regret, regardless of the size of its original attribute level or the range of this attribute in the choice set. The next section will introduce the incorporation of relative thinking into the RRM models.

3.2.2 Incorporating relative thinking into the RRM model

- Incorporating level-based relative thinking

In the RRM models, the regret functions are defined as a function of pairwise comparisons between the attribute levels of alternatives, which provides a natural way of incorporating the idea of relative thinking into the model specifications. Jang et al. (2017) were the first who considered Weber's law into the RRM model: the degree of regret is determined by how much people perceive the differences between attribute levels rather than the absolute differences.

The perceived differences are assumed to be inversely proportional to the attribute level of the chosen alternatives. Inspired by their work, we present the μRRM-Level model specification:

$$R_{i\,Level}^{\mu RRM} = \sum_{j \neq i} \sum_{m} \mu \ln(1 + \exp(\frac{\beta_m}{\mu}[\frac{X_{jm} - X_{im}}{X_{im}}])). \qquad \text{(Eq. 3.3)}$$

As mentioned in Section 3.2.1, the μRRM model can capture a wide range of regret aversion in behaviour, and it has the linear-additive RUM model as one special case. This is non-trivial: only by rewriting the linear-additive RUM model as a case of the μRRM model, can we estimate to what extent the original values of attribute-level differences influence the perceived salience of differences between alternatives (in the context of RUM models)[8]. When μ equals one, the model becomes the CRRM-Level model (proposed in Jang et al. (2017)):

$$R_{i\,Level}^{CRRM} = \sum_{j \neq i} \sum_{m} \ln(1 + \exp(\beta_m[\frac{X_{jm} - X_{im}}{X_{im}}])). \qquad \text{(Eq. 3.4)}$$

Compared with conventional RRM models, the new model specifications transform the actual attribute-level differences $(X_{jm} - X_{im})$ into the ratio of the differences to the attribute level of the chosen alternative $(X_{jm} - X_{im})/X_{im}$. Such a transformation leads to a fundamental difference from the conventional RRM models. The RRM-Level models postulate that regret depends on both attribute-level differences and the original attribute level of the chosen alternative. Thus, the same attribute-level differences may generate different amounts of regret, reflecting different regret-averse behaviour. This is however overlooked in the conventional RRM models. The next section will further discuss the differences between RRM and RRM-Level models using specific examples.

- Incorporating range-based relative thinking

Previous studies have shown clear evidence that the size of the attribute range also affects people's perception of attribute-level differences. A given absolute attribute-level difference appears to have a greater impact when presented in a narrow range than in a broad range. To take the attribute range effect into account, we adopt the formulation of the value-shift model proposed by Wedell (1998). The value-shifted model posits that an increase in the range of one dimension will lead to a decrease in the attractiveness difference along with that domain: $U_1 - U_2 = \sum_q w_q (X_{1q} - X_{2q})/\Delta X_q$, where $(U_1 - U_2)$ denotes the utility (attractiveness) difference between alternatives *1* and *2*, $(X_{1q} - X_{2q})$ denotes the attribute-level difference on the dimension q, w_q is the weight attached to the difference, and ΔX_q denotes the range of this dimension: $\Delta X_q = \max\{X_q\} - \min\{X_q\}$. This model assumes that the wider the range, the smaller the corresponding weight attached to the attractiveness difference.

Inspired by the value-shifted model, the regret function can be transformed into a function in which the impact of the attribute-level difference is inversely proportional to the attribute range: $R_i = f\left(\frac{X_{jm} - X_{im}}{\Delta X_{ms}}\right)$. Here ΔX_{ms} denotes the range of attribute m in the choice set *s*: $\Delta X_{ms} =$

[8] This is not achievable in the context of the conventional mathematical formulation of utility as a linear-additive form.

$\max\{X_{ms}\} - \min\{X_{ms}\}$. Substituting the new regret function into the most generalized RRM model derives the model specification of the μRRM-Range model:

$$R_{i_{range}}^{\mu RRM} = \sum_{j \neq i} \sum_{m} \mu ln(1 + exp(\frac{\beta_m}{\mu}[\frac{X_{jm} - X_{im}}{max\{X_{ms}\} - min\{X_{ms}\}}])). \qquad \text{(Eq. 3.5)}$$

Again, the μRRM model can accommodate for different degrees of regret-averse behaviour. When μ equals one, the model collapses to the CRRM-Level model:

$$R_{i_{range}}^{CRRM} = \sum_{j \neq i} \sum_{m} ln(1 + exp(\beta_m[\frac{X_{jm} - X_{im}}{max\{X_{ms}\} - min\{X_{ms}\}}])). \qquad \text{(Eq. 3.6)}$$

From the model specifications, we can see the RRM-Range models postulate that regret depends not only on absolute attribute-level differences, but also on the size of attribute range in the choice set. This leads to a fundamental difference from the conventional RRM models. For example, based on the RRM-Range models, deterioration in the performance of the chosen alternative will cause more regret when it is presented in a narrow range than in a wide range. In contrast, the conventional RRM models fail to consider the range effect; they assume that the same differences cause the same amount of regret, regardless of the size of the attribute range in the choice set.

Apart from the above four model specifications (Eq.3.3 – Eq.3.6), this study also introduces the more generalized versions of the RRM-Level and RRM-Range models respectively. In Jang et al. (2017), they presented a generalized model in which an estimable parameter ϑ is included to govern the degree of the effect of Weber's law (from hereon the Weber effect). Combining their generalization with the μRRM-Level model, we get a more generalized model:

$$R_{i_{level}} = \sum_{j \neq i} \sum_{m} \mu \ln(1 + \exp(\frac{\beta_m}{\mu}[\frac{X_{jm} - X_{im}}{(X_{im})^{\vartheta_m}}])). \qquad \text{(Eq. 3.7)}$$

Similarly, we can also derive a more generalized RRM-Range model:

$$R_{i_{Range}} = \sum_{j \neq i} \sum_{m} \mu \ln(1 + \exp(\frac{\beta_m}{\mu}[\frac{X_{jm} - X_{im}}{(\max\{X_{ms}\} - \min\{X_{ms}\})^{\vartheta_m}}])). \qquad \text{(Eq. 3.8)}$$

In the above model specifications, ϑ_m is an estimable parameter associated with attribute m, governing the size of the Weber effect or the range effect in this attribute. Relevant special cases arise when ϑ_m equals to certain values. For example, when ϑ_m equals one, the model becomes the corresponding μRRM-Level or μRRM-Range model, implying a full Weber effect or range effect in the attribute; when ϑ_m equals zero, the model becomes the μRRM model, implying no Weber effect or range effect in the attribute; when ϑ_m equals a value between zero and one, implying a moderate Weber effect or range effect.

Note that in this study, we do not pay much attention to the generalized versions, nor do we empirically test their model performance in later sections. The main reason is that too many estimable parameters in a model would generate the confounding effect, failing in parameter identification and model convergence[9]. Nevertheless, we wish to present the possible way of creating more generalized models and provide research avenues for interested readers.

3.3 Differences between RRM, RRM-Level and RRM-Range

Section 3.2 has presented the mathematical formulations of the RRM-Level and RRM-Range models. This section highlights differences between the conventional RRM, RRM-Level, and RRM-Range models. To make this more concrete, we use two specific examples in the context of travel time differences. Specifically, we wish to use the examples to demonstrate differences between these three models in measuring regret (caused by the same travel time increases) and the resulting behavioural implications.

Example 1: Choice situation A {10, 20, 30 min} and choice situation B {110, 120, 130 min}

Consider two choice situations A and B, in which a traveller faces three routes characterized by travel time only. In situation A, travel times are 10, 20, and 30 minutes and in situation B, they are 110, 120, and 130 minutes. Therefore, the choice sets that the traveller faces in situations A and B are {10, 20, 30 min} and {110, 120, 130 min}. The two choice sets contain different travel times, but they have similarities: first, the differences between two adjacent travel times are ten minutes; second, the ranges of travel time in the choice sets are the same, 20 minutes.

According to Weber's law, a ten-minute travel time increase seems more significant when two travel times are ten and 20 minutes, compared to the situation in which the travel times are 110 and 120 minutes. Therefore, although travel time differences are 10 minutes in both choice sets, decision makers' perceptions of the 10-minute difference appear to be very different.

This perception difference can be successfully captured by the RRM-Level model. The RRM-level model postulates that regret is determined not only by attribute-level differences but also by the size of the initial attribute level. Therefore, in the RRM-level model, the increase of 10 minutes generates more regret in situation A than in situation B. However, conventional RRM models postulate that regret is merely determined by attribute-level differences. Thus, conventional RRM models derive the same amount of regret for the 10-minute travel time increase in the two sets. As for the RRM-range model, it postulates that regret is also influenced by the attribute range in the choice set. As the ranges are the same in the two choice sets, the RRM-Range model also derives the same regret for the ten-minute travel time increase.

What is the behavioural implication for deriving different regret in the two situations? The RRM-Level model assumes that for the same 10-minute travel time increase, travellers will experience more regret in the first situation than in the second situation. This means the 10-minute route (the best alternative) has a dominant advantage over the other two routes (20 and 30 min) in situation A, as every additional 10-minute increase would cause great regret, while in situation B, the 110-minute route (the best alternative) does not have a significant advantage over the others (120 and 130 min), as a 10-minute increase would not induce much regret in this case.

[9] We tested the generalized versions on our empirical datasets and found that ϑ_m is very likely to be confound with μ.

This can also be seen from the choice probabilities of three routes in the choice sets. Table 3.1 gives choice probabilities when the weight parameter $\beta_{TT} = -1$ and the regret aversion parameter $\mu = 1$. Only the RRM-Level model yields different choice probability distributions of three routes in the two situations. Specifically, in situation A, the best route (10 min) has a choice probability of 63.1%, which is much higher than the choice probabilities of the other two alternatives, while in the situation B, three alternatives have close choice probabilities (37.7% vs. 33.0% vs. 29.3%). Again, this implies that the difference between 110, 120, and 130 minutes may not seem that noticeable to travellers. For the conventional RRM model, it generates an extreme outcome that the best route gains the entire 100% choice probability for both situations A and B. For the RRM-Range model, it postulates that the attribute range also plays a role. According to its model specification, the 10-minute travel time difference is narrowed down (by dividing by the range size), thereby the outcome is not as extreme as RRM: the best routes do not eat away all the 100% probability. But like the conventional RRM model, the RRM-Range model also generates the same choice probability distributions for the two situations.

Table 3.1 Choice probabilities generated by different models (Example 1)

	Situation A			**Situation B**		
	10 min	20 min	30 min	110 min	120 min	130 min
RRM	100%	0%	0%	100%	0%	0%
RRM-Level	63.1%	23.0%	13.9%	37.7%	33.0%	29.3%
RRM-Range	57.5%	29.7%	12.8%	57.5%	29.7%	12.8%

Example 2: Choice set A {10, 20 30 min} and choice set C {10, 20 130 min}

Consider two choice situations A and C. Situation A is the same as in Example 1: the travel times of three routes are 10, 20, and 30 minutes. In situation C, the travel times are 10, 20, and 130 minutes. Therefore, choice sets that the traveller faces in the two situations {10, 20, 30 min} and {10, 20, 130 min} respectively. The main difference between the two choice sets is the travel times of the third route. In choice set C, the travel time of the third route is much longer than that in choice set A, which makes its range much wider, compared to choice set A.

In both situations, travel time increases by 10 minutes when comparing the second alternative with the first. However, what do different RRM models produce for the same 10-minute increase? The conventional RRM model assumes the same attribute-level differences generate the same regret, therefore, in these two situations, increasing 10 minutes leads to the same amount of regret. Similar results can be also obtained by the RRM-level model, as it looks at the original level of the attribute (both are 10 minutes). In contrast, the RRM-Range model generates a very different outcome: the 10-minute increase in travel time will yield a smaller amount of regret in situation C, compared to situation A. This is due to the assumption of the RRM-Range model that the size of the range in the choice set also influences people's perception of differences between alternatives. When one alternative in the choice set has a greater deterioration in travel time (i.e., 130 vs 30 min), a 10-minute travel time increase seems less salient.

What is the behavioural implication for generating different amounts of regret in these two situations? According to the RRM-Range model, when travel time increases from 10 minutes to 20 minutes, traveller would experience more regret in situation A than in situation C. This means the 10-minute route is more attractive than the 20-minute route to travellers when they are presented with a 30-minute route than when they are presented with a 130-minute route.

This is because the 10-minute travel time saving may not be noticeable to travellers when the entire choice set is broad.

The choice probabilities of three routes in two situations are calculated in the case of the weight parameter $\beta_{TT} = -1$ and the regret aversion parameter $\mu = 1$, shown in Table 3.2. We first look at the choice probabilities generated by RRM-Range. In situation A, the choice probability of the best alternative (10 min) is 57.5%, which is around two times higher than it of the second-best alternative (29.7%), while in situation C, the choice probabilities of the two routes become very close (47.6% vs. 42.8%). This again implies that the difference in attractiveness between these two alternatives becomes less pronounced when they are presented in a wider range. For the conventional RRM model, it generates the same outcomes for the two situations; the best alternative (10 min) has a 100% choice probability. As for the RRM-level model, we can see that the 20-minute route gains more choice probabilities in situation C than in situation A, but the 10-minute route still has a dominant advantage over the others.

Table 3.2 Choice probabilities generated by different models (Example 2)

	Situation A			**Situation C**		
	10 min	20 min	30 min	10 min	20 min	130 min
RRM	100%	0%	0%	100%	0%	0%
RRM-Level	63.1%	23.0%	13.9%	61.3%	31.5%	7.2%
RRM-Range	57.5%	29.7%	12.8%	47.6%	42.8%	9.6%

3.4 Empirical applications

This section puts the new RRM models, RRM-Level and RRM-Range, to the empirical test. Specifically, we estimate the new models and several conventional models on four datasets. The estimated models include (1) the linear-additive RUM model, (2) the classical RRM (CRRM) model, (3) the μRRM model, (4) the CRRM-Level model, (5) the μRRM-Level model, (6) the CRRM-Range model, and (7) the μRRM-Range model. The used datasets are existing datasets that have been used in travel behaviour research. For reasons of space limitations, we only discuss the estimation results of one dataset in full detail and discuss the model fit of the four datasets in an overview table.

3.4.1 Description of the datasets

The four datasets were collected for previous travel behaviour studies, and three of them have been recently used for comparing the RRM models against the linear-additive RUM model in terms of empirical model performance. The first three datasets are stated preference (SP) datasets concerning travel route or mode choices, and the last is a revealed preference (RP) dataset about shopping destination choices. Each dataset is briefly introduced below.

Dataset 1: Swiss metro

The first dataset is the one we discuss later in full detail. It is obtained from the BIOGEME website[10]. The data collection effort focused on exploring the impact of introducing a new travel

[10] https://biogeme.epfl.ch/.

mode, Swiss metro, on the mode choice behaviour of commuters in Switzerland (Bierlaire et al., 2001). Three modes were considered in the stated choice experiment: train, car and Swiss metro. The train was described in terms of travel time, travel cost and headway, the car was described in terms of travel time and travel cost, and Swiss metro was described in terms of travel time, travel cost, headway, and seat availability. Considering that estimating RRM models requires at least three pairwise comparisons (three alternatives with the same attributes), and that our aim is to test model performance rather than predicting behaviour, we only include the two generic attributes, *travel time* (TT) and *travel cost* (TC), in the model estimation.

Note that the RRM-Level model cannot be estimated on data that contains zero level attributes, and the RRM-Range model is not applicable for data with zero attribute ranges in the choice set. This means we need to exclude these two types of data when estimating models. The original dataset contains 10,728 choice observations. After cleaning the data, the final data contain 8,288 choice observations.

Dataset 2: Route choices

The second dataset concerns travel route choice behaviour, which was collected in the Netherlands in 2011. It consists of 390 car commuters who were asked to choose among three hypothetical routes described in terms of four following attributes with three levels each: *average door-to-door travel time* (TT) (45, 60, 75 min), *percentage of travel time in traffic jams* (JAM) (10, 25, 40 %), *travel time variability* (VAR) (5, 15, ± 25 min), and *travel cost* (TC) (5.5, 9, €12.5). Each respondent took nine choice sets which were created by a so-called "optimal orthogonal in the differences" design, leading to 3,510 choice observations in total. More details about this dataset can be found in Chorus & Bierlaire (2013).

Dataset 3: Route choices

The third dataset also concerns route choice behaviour. It was originally used for studying the value of time for drivers in Sydney (Rose & Masiero, 2010), and subsequently used for empirical comparisons between the classical RRM model and the linear-additive RUM model in Chorus et al. (2013). The dataset consists of 300 car drivers who were asked to choose between a current route and two alternative routes which differ in terms of five attributes: *free-flow travel time* (FTT), *slowed-down travel time* (STT), *travel time variability* (VAR), *running cost* (RC) and *toll cost* (TC). Sixteen choice sets were created by a pivoted design in which the attribute levels of alternative routes were pivoted around the attribute levels of the current routes. Time-related attributes are in minutes, and cost-related attributes are in Australian dollars. Note that, unlike most SP experiments where attributes are fixed for all respondents, in this experiment, the attribute level is related to a recent trip of each respondent. This means respondents took different choice sets. In addition, the trip length was categorized into three segments: no more than 30 minutes, 30-60 minutes, and more than 61 minutes (capped at 2 hours). This means attribute levels are very different in the whole dataset, but within each category, the magnitude of attribute levels is relatively similar. Note that this is very important for the discussion of model estimation results in section 3.4.3.

As the RRM-Level model cannot be estimated on the data with zero attribute levels, the zero-level data are thus omitted. This leads to 1,572 choices in total.

Dataset 4: Shopping destination choices

This dataset contains 1,503 RP choices concerning grocery shopping destinations in Noord-Brabant province, the Netherlands. It has been used several times in previous studies to compare the empirical performance of the RRM models with the linear-additive RUM model (Chorus, 2010; Chorus et al., 2013; van Cranenburgh et al., 2015). Three attributes describing shopping destinations are *floorspace in* m^2 *concerning groceries in the shopping centre* (FSG), *floorspace in* m^2 *concerning other items* (FSO), and *travel time in* s (TT). There are five shopping destination alternatives in the choice set. More details about this dataset can be found in Arentze et al. (2005).

Excluding data with zero attribute levels and zero attribute ranges results in only 256 choice observations. This means only 256 choices can be used to estimate all seven models. If we only omit data with zero attribute ranges, there are 1,485 choice observations. Thus, for this dataset, the empirical application is first conducted on the data with no zero attribute ranges (1,485 choices), on which only the RUM, RRM and RRM-Range models are estimated. Then all seven models are estimated on the data with neither zero attribute ranges nor zero attribute levels (256 choices).

Descriptive statistics

It is very informative to understand the characteristics of the used data (e.g., variations in attribute levels and in range sizes) when discussing model estimation results, as the new RRM models are used to explore whether the Weber effect or the range effect plays a role in choice behaviour. The descriptive statistics of each dataset are presented in Table 3.3, in which the minimum and maximum variables, mean and standard deviation (SD) of the attribute levels and range sizes are given respectively.

In dataset 1, both two attributes have various levels, for example, the level of attribute TT varies from 8 to 1560. The high SDs also indicate that the values are spread out. Moreover, the range sizes in the choice sets are also very diverse. As mentioned above, dataset 2 contains four attributes, each has three fixed levels. The range size for each attribute is the same throughout all choice sets, which can be seen from the statistics that the min, max, and mean values are the same and SDs are zero. As for dataset 3, it also has different attribute levels and range sizes, but the SDs are very small. This means the values are relatively closely distributed around the mean. In dataset 4, three attributes have very different levels, ranging from zero to very large values, for example, the maximum level of FSG is 20,452. Moreover, the range size of each attribute in choice sets is also very diverse. This means these values are spread out over wide ranges.

Table 3.3 Descriptive statistics of the data

Data set	# of choices		Attribute levels				Range sizes			
			Min	Max	Mean	SD	Min	Max	Mean	SD
1	8288	TT	8	1560	142	80	2	1436	100	66
		TC	8	768	105	64	1	648	57	57
2	3510	TT	45	75	60	12	30	30	30	0
		JAM	10	40	25	12	30	30	30	0
		VAR	5	25	15	8	20	20	20	0
		TC	5.5	12.5	9	3	7	7	7	0
3	1572	FTT	1	108	26	15	1	63	12	8
		STT	2	126	32	21	1	74	15	11
		VAR	1	118	17	17	1	114	30	18
		RC	0.5	27.3	4	2.5	0.3	17.5	2.6	1.7
		TC	1.6	16	4	2.0	0.6	6	1.4	0.8
4-1	1485	FSG	0	20452	1677	2031	195	20385	3623	3219
		FSO	0	95932	7847	15610	558	95932	23046	24911
		TT	0	13796	624	743	67	13796	1224	883
4-2	256	FSG	27	20452	1824	2400	195	20346	3975	4215
		FSO	26	95932	8380	15591	2244	95065	22585	23641
		TT	118	4260	883	6640	67	3492	838	6722

3.4.2 Estimation results of Swiss metro data

Table 3.4 shows the estimation results of all seven models on Swiss metro data. We first look at model fits. In the comparison of RUM, CRRM and μRRM models, the most generalised model μRRM has the best model fit. Its log-likelihood is improved by 99 log-likelihood points and 20 log-likelihood points as compared to RUM and CRRM respectively. When incorporated relative thinking into RRM, the models have very substantial improvements in model fits. Specifically, incorporating Weber's law (level-based relative thinking) into CRRM and μRRM improves the model fits by 279 and 320 log-likelihood points respectively, and incorporating the range effect (range-based relative thinking) improves the model fits by 241 and 221 log-likelihood points respectively. The large differences in model fits indicate that the new models differ significantly from their originals in predicting choice behaviour. Moreover, the significant improvements in model fits imply that relative thinking exists in the data, which can be captured by the new RRM models, but not by the conventional RRM models.

Now we turn to model parameter estimates. First, we see that all taste parameters are statistically significant at the 95% level and their sign is also in the anticipated direction. Then we see that the RUM, CRRM and μRRM models produce the same magnitude of taste parameter estimates, while the RRM-Level and RRM-Range models generate much larger parameter estimates (absolute values). This is in line with expectations. By incorporating Weber's law or the range effect, the values of attribute levels in the denominator are normalized, i.e. divided by the original attribute level or by the size of the range in the choice set, therefore, the parameter estimates are enlarged. Finally, the regret parameter μ is found to be slightly changed when μRRM includes the formula of relative thinking. Specifically, the inclusion of Weber's law increases μ from 0.417 to 0.558, but both imply a strong regret aversion behaviour. The inclusion of the range effect leads to an even larger μ, 1.06, implying a moderate regret aversion behaviour. The estimation result of parameter μ is also consistent with the log-likelihoods of CRRM-Range and μRRM-Range: when μ is close to one, μRRM becomes CRRM, thus the CRRM-Range and μRRM-Range models have the same log-likelihood.

Table 3.4 Estimation results of Swiss metro data

	RUM	**Conventional RRM**		**Level-based relative thinking**		**Range-based relative thinking**	
		CRRM	μRRM	CRRM-Level	μRRM-Level	CRRM-Range	μRRM-Range
	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)
β_{TT}	**-0.019** (*-38.52*)	**-0.014** (*-31.16*)	**-0.015** (*-29.25*)	**-2.540** (*-42.90*)	**-2.880** (*-52.80*)	**-1.610** (*-54.43*)	**-1.060** (*-52.07*)
β_{TC}	**-0.009** (*-13.62)*	**-0.006** (*-14.18*)	**-0.006** (*-14.60*)	**-1.390** (*-24.84*)	**-1.360** (*-23.37*)	**-0.572** (*-24.60*)	**-0.575** (*-22.27*)
μ	-	-	**0.417** (*9.48*)	-	**0.558** (*24.22*)	-	**1.06** (*6.09*)
Number of observations	8288	8288	8288	8288	8288	8288	8288
Null-likelihood	-9105	-9105	-9105	-9105	-9105	-9105	-9105
Final likelihood	-6946	-6867	-6847	-6588	-6527	-6626	-6626

In sum, the above empirical analysis of the RRM-Level and RRM-Range model demonstrates that *i*) by incorporating level-based or range-based relative thinking, RRM models may yield very considerable improvements in model fits; *ii*) the absolute values of taste parameters in RRM-Level and RRM-Range are larger than the corresponding RRM models, as attribute-level differences are divided by the original attribute level or the range size; *iii*) The incorporation of relative thinking into the μRRM model may change the estimate of parameter μ, therefore leading to different inference of regret aversion behaviour. The above observations are only made by estimating one dataset. More empirical tests are needed. The next section gives an overview of empirical results based on all four datasets.

3.4.3 Overview of empirical results

Estimation results of all four datasets are given in Table 3.5 and Table 3.6. For reasons of simplicity, this section only presents the model fits (in terms of final log-likelihoods) of the seven models (Table 3.5) and the estimate results of μ in the μRRM, μRRM-Level, and μRRM-Range models (Table 3.6). Detailed estimation results are given in Appendix 3.1.

We first look at the log-likelihoods of the first three models (i.e., RUM, CRRM, and μRRM). As compared to RUM and CRRM, the μRRM model again has the best model fit on all datasets. This is in line with expectations, as μRRM can accommodate a wide range of regret aversion behaviour, taking RUM and CRRM as special cases. In datasets 2 and 3, the μRRM model has the same log-likelihood as the CRRM model and the RUM model, respectively. This corresponds to the estimates of parameter μ shown in Table3.6, where μ is close to one (i.e., 0.932) in the dataset 2, and μ is estimated to be very large (i.e., >10) in the dataset 3.

We then move to the model fits of the remaining four incorporated models (i.e., CRRM-Level, μRRM-Level, CRRM-Range, μRRM-Range). At first glance, the four incorporated models do perform better than their original models, but not always. Specifically, in datasets 4-2 and 1, the incorporation of level-based relative thinking into CRRM and μRRM results in substantial improvements in model fits, and in datasets 1, 4-1, and 4-2, the incorporation of range-based relative thinking also leads to considerable improvements in model fits. For example, in the dataset 4-1, the CRRM-Range model outperforms the CRRM model by 89 log-likelihood points, and the μRRM-Range model outperforms the μRRM model by 63 log-likelihood points. However, In datasets 2 and 3, we cannot find any improvement in model fits when incorporating level-based or range-based relative thinking into the RRM models. For example, in dataset 2, as compared to CRRM and μRRM, the model fits of their range-based counterparts have no changes, and the model fits of their level-based counterparts are even worse, with 71 and 4 log-likelihood points deterioration. As for dataset 3, the four incorporated models have similar model fits, which are worse compared to the model fits of their original models.

Table 3.5 Summary of the model fits (final log-likelihoods)

	Dataset 1	**Dataset 2**	**Dataset 3**	**Dataset 4-1**	**Dataset 4 -2**
	Swiss metro	Route choices	Route choices	Shopping destination choices	
RUM	-6946	-2613	-1286	-2276	-388
CRRM	-6867	-2605	-1288	-2272	-385
μRRM	-6847	-2605	-1286	-2234	-375
CRRM-Level	-6588	-2676	-1318	-	-370
μRRM-Level	-6527	-2609	-1317	-	-369
CRRM-Range	-6626	-2605	-1317	-2183	-360
μRRM-Range	-6626	-2605	-1317	-2171	-358

Note that for each dataset, different shades of blue represent different model fits. The lighter the blue, the better the model fit.

It seems that incorporating relative thinking into the RRM models (CRRM or μRRM) does not always improve model fit. We first discuss the RRM-Level models. In two out of the four datasets (1 and 4-2), the RRM-Level models outperform their original formulations, while in the remaining datasets (2 and 3), the inclusion of level-based relative thinking deteriorates the model fits. Recall that datasets 1 and 4 have very various attribute levels, as shown in Table 3.3. Therefore, by inspecting the descriptive statistics of the data, we may find commonalities where the RRM-Level models perform better—**the attribute levels of alternatives vary greatly, from small to large values, in the data**. In addition, the cases in which the RRM-Level models perform worse also have a similarity—attribute levels are relatively less various. For example, in dataset 2, attributes only have three fixed levels, and in dataset 3, although attribute levels are various throughout the whole data, attribute levels in one choice set are pivoted around the reference level (current trip) and set within a certain range[11].

Likewise, the RRM-Range models do not seem to always perform better than their original models. In datasets 1, 4-1 and 4-2 where **attribute ranges vary greatly**, the RRM-Range models indeed improve the model fit. In dataset 2 where the range sizes do not change across the whole dataset, the attribute range effect does not play a role, resulting in same log-likelihoods (-2605) as the original models. In dataset 3 in which the range sizes are relatively stable, the RRM-Range models perform even worse than the originals.

Table 3.6 The μ value in μRRM, μRRM-Level, and μRRM-Range models

	ID 1	**ID 2**	**ID 3**	**ID 4-1**	**ID 4-2**
	Swiss metro	Route choices	Route choices	Shopping destination choices	
μRRM	0.417	0.936	>10	0.141	0.175
μRRM-Level	0.558	0.092	>10	-	>10
μRRM-Range	1.06	0.936	>10	0.141	0.089

Table 3.6 presents the estimates of parameter μ in the μRRM, μRRM-Level, and μRRM-Range models. We first look at the μRRM model. As discussed in Section 3.4.2, parameter μ in dataset 1 is estimated to be small (0.417), suggesting a strong regret aversion in behaviour. In dataset 2, parameter μ is close to one (0.936), the μRRM model collapses to its special case—the CRRM model, suggesting a moderate degree of regret aversion. Parameter μ in dataset 3 is larger than ten, implying a linear-additive RUM behaviour. In datasets 4-1 and 4-2, parameters μ are both very small (0.141 and 0.175), suggesting a strong regret aversion behaviour respectively.

When level-based relative thinking is incorporated into RRM, parameter μ changes considerably in two datasets, but remains almost the same in the other two datasets. Specifically, compared to parameter μ in μRRM, parameter μ in μRRM-Level becomes smaller (0.936 vs. 0.092) in dataset 2, and it becomes larger than 10 (0.175 vs. >10) in dataset 4-2. But in datasets 1 and 3, there is no significant change in the estimates of parameter μ.

In terms of the incorporation of range-based relative thinking, we can see there is no significant change in the estimates of parameter μ in four out of five datasets (2, 3, 4-1 and 4-2). Specifically, μ remains the same in datasets 2, 3, and 4-1, and in the dataset 4-2, although μ decreases to 0.089, it is not significantly smaller than 0.175[12]. However, the dataset 1, parameter μ increases to 1.06, which is significantly larger than 0.417 in the μRRM model.

From the model estimation results, we can conclude that *i*) incorporating level based or range-based relative thinking into RRM models does not always improve model fit, however, it is more likely to happen in the data containing various attribute levels in the choice set or various

[11] Recall that the feature of dataset 3 has been discussed in section 3.4.1, that is the data are categorized into three segments: short, medium and long trips. In each segment, attribute levels are relatively close to each other.

[12] The asymptotic t-value for the difference between 0.089 and 0.175 is 1.43, suggesting that these two values are not significantly different from one another.

attribute ranges across the whole dataset; *ii*) the inclusion of relative thinking may cause changes in the estimate of parameter μ, resulting in different predicted behaviour regarding regret aversion. The changes in parameter μ suggest that the Weber effect or the range effect might pick up the effect that is confounded with parameter μ. However, according to the current estimation results, it is hard to conclude how μ would change when the Weber effect or the range effect is taken into account.

3.5 Conclusions and discussions

This chapter mainly focuses on a psychological notion—relative thinking. Relative thinking refers to a behavioural phenomenon that when making choices, people care not only about *absolute* differences between choice alternatives, but also *relative* differences. In this chapter, we mainly distinguish two types of relative thinking: level-based relative thinking and range-based relative thinking. Level-based relative thinking means that actual differences look big or small depending on their initial sizes, while range-based relative thinking means that actual differences look big or small depending on the size of the range in the choice set. Despite the distinction between the two, both level-based and range-based relative thinking reflect the relativity nature of how people perceive differences: the value (difference) looks small when compared with a large value and it looks big when compared with a small value. Range-based relative thinking focuses on the size of ranges in choice sets, but essentially it involves comparing the levels of attributes. For example, a given increase in travel time (e.g. 10 vs. 20 min) seems smaller when another alternative has a much longer travel time (e.g. 130 min) (resulting in a wider range size).

We incorporate level-based and range-based relative thinking into the RRM model framework respectively, resulting in the RRM-Level model and the RRM-Range model. Compared to its original formulations, the RRM-Level model transforms the attribute-level difference between the chosen alternative and a competing alternative into the ratio of the attribute-level difference to the attribute level of the chosen alternative, and the RRM-Range model transforms it into the ratio of the attribute-level difference to the attribute range in the choice set. The former transformation was motivated by Weber's law, which asserts that people's responses to actual changes in stimuli are inversely proportional to the original intensity of stimuli. The latter was inspired by a so-called value-shifted model, which describes the importance of actual differences between alternatives diminishes as the range of the choice set increases. Both transformations attempt to enhance the behavioural realism of the RRM paradigm.

The RRM-Level and RRM-Range models are tested on four empirical datasets. Model estimation results show that there is great potential to improve the model performance (in terms of model fit) by incorporating relative thinking into the RRM models. According to our empirical analyses, the RRM-Level models outperform their originals in two out of four datasets, and the RRM-Range models outperform in three out of five datasets (including a subset). It is worth noting that some of the improvements in model fits are very substantial. But there are also some cases in which the new models perform just equally well as or even worse than their originals.

A question then arises: In which circumstance does the incorporation of relative thinking fit choice data better than the conventional RRM models? Now there seems to be no definitive answer yet, but our empirical evidence so far has indicated some similarities of the datasets on which new models fit better. When the levels of attributes in choice sets are very different (ranging from very small values to large values), it seems very likely to trigger level-based relative thinking. As such, the RRM-Level models are likely to fit the data better than conventional RRM. Likewise, when there are various ranges of attribute levels in the dataset, people's judgments seem likely to be influenced by the range effect. In this case, the RRM-Range models are likely to fit the data better than conventional RRM. Therefore, when datasets have such characteristics, like very different attribute levels or attribute ranges, we recommend considering RRM-Level or RRM-Range models when estimating the data.

The findings of this study provide several avenues for future research. First, in this study, the new RRM models are tested on four empirical datasets. It would be valuable for future research to replicate model comparisons on more empirical datasets, in order to further test the robustness of model performance. Second, more research is needed to gain insight into what types of data will cause the RRM level or RRM range model to perform better or worse, compared to their original counterparts. A possible way is to design customized choice sets which vary in attribute levels or attribute range sizes. It would be interesting to examine to what extent the model performance of the RRM-Level or RRM-Range model depends on the variation in attribute levels or attribute ranges. Third, we find that the incorporation of relative thinking may lead to changes in the estimates of regret aversion parameter μ. But how this parameter changes is still unclear to us. Thus, exploring the changing law of parameter μ is also a research angle. Last but not least, Section 3.2.2 mentions the generalised versions of RRM-Level and RRM-Range models, in which there is a power coefficient governing the degree of the Weber effect or the range effect. This study does not discuss the generalised versions in detail. Interested researchers can future explore and empirically test them in future research.

References

Arentze, T. A., Oppewal, H., & Timmermans, H. J. (2005). A multipurpose shopping trip model to assess retail agglomeration effects. *Journal of Marketing Research, 42*(1), 109-115.

Azar, O. H. (2004). *Psychological motivations and biases in economic behavior and their effects on markets and firm strategy*: Northwestern University.

Azar, O. H. (2008). The effect of relative thinking on firm strategy and market outcomes: A location differentiation model with endogenous transportation costs. *Journal of Economic Psychology, 29*(5), 684-697.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. Cambridge: MIT Press.

Bierlaire, M., Axhausen, K., & Abay, G. (2001). *The acceptance of modal innovation: The case of Swissmetro.* Paper presented at the Swiss Transport Research Conference.

Bushong, B., Rabin, M., & Schwartzstein, J. (2021). A model of relative thinking. *The Review of Economic Studies, 88*(1), 162-191.

Cantillo, V., Heydecker, B., & de Dios Ortúzar, J. (2006). A discrete choice model incorporating thresholds for perception in attribute values. *Transportation Research Part B: Methodological, 40*(9), 807-825.

Chorus, C. G. (2010). A new model of random regret minimization. *European Journal of Transport Infrastructure Research, 10*(2).

Chorus, C. G., Arentze, T. A., & Timmermans, H. J. (2008). A random regret-minimization model of travel choice. *Transportation Research Part B: Methodological, 42*(1), 1-18.

Chorus, C. G., & Bierlaire, M. (2013). An empirical comparison of travel choice models that capture preferences for compromise alternatives. *Transportation, 40*(3), 549-562.

Chorus, C. G., Rose, J. M., & Hensher, D. A. (2013). Regret minimization or utility maximization: it depends on the attribute. *Environment and Planning B: Planning and Design, 40*(1), 154-169.

Frisch, D. (1993). Reasons for framing effects. *Organizational behavior and human decision processes, 54*(3), 399-429.

Hensher, D. A. (1976). The value of commuter travel time savings: empirical estimation using an alternative valuation model. *Journal of Transport Economics and Policy*, 167-176.

Jang, S., Rasouli, S., & Timmermans, H. (2017). Incorporating psycho-physical mapping into random regret choice models: model specifications and empirical performance assessments. *Transportation, 44*(5), 999-1019.

Mellers, B. A., & Cooke, A. D. (1994). Trade-offs depend on attribute range. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1055.

Moon, S., & Voss, G. (2009). How do price range shoppers differ from reference price point shoppers? *Journal of Business Research, 62*(1), 31-38.

Mowen, M. M., & Mowen, J. C. (1986). An empirical examination of the biasing effects of framing on business decisions. *Decision Sciences, 17*(4), 596-602.

Nicolau, J. L. (2012). Battle royal: Zero-price effect vs relative vs referent thinking. *Marketing Letters, 23*(3), 661-669.

Ohler, T., Le, A., Louviere, J., & Swait, J. (2000). Attribute range effects in binary response tasks. *Marketing Letters, 11*(3), 249-260.

Parducci, A. (1965). Category judgment: a range-frequency model. *Psychological review, 72*(6), 407.

Ranyard, R., & Abdel-Nabi, D. (1993). Mental accounting and the process of multiattribute choice. *Acta psychologica, 84*(2), 161-177.

Rose, J. M., & Masiero, L. (2010). A comparison of the impacts of aspects of prospect theory on WTP/WTAEstimated in preference and WTP/WTA space. *European Journal of Transport and Infrastructure Research, 10*(4).

Saini, R., Rao, R. S., & Monga, A. (2010). Is that deal worth my time? The interactive effect of relative and referent thinking on willingness to seek a bargain. *Journal of Marketing, 74*(1), 34-48.

Saini, R., & Thota, S. C. (2010). The psychological underpinnings of relative thinking in price comparisons. *Journal of Consumer Psychology, 20*(2), 185-192.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of economic behavior organization, 1*(1), 39-60.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.

van Cranenburgh, S., Guevara, C. A., & Chorus, C. G. (2015). New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice, 74*, 91-109.

van Cranenburgh, S., Rose, J. M., & Chorus, C. G. (2018). On the robustness of efficient experimental designs towards the underlying decision rule. *Transportation Research Part A: Policy and Practice, 109*, 50-64.

Volkmann, J. (1951). Scales of judgment and their implications for social psychology.

Weber, E. H. (1834). *De pulsu, resorptione, auditu et tactu: annotationes anatomicae et physiologicae, auctore*: prostat apud CF Koehler.

Wedell, D. H. (1998). Testing models of trade-off contrast in pairwise choice. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 49.

Welch, M., & Williams, H. (1997). The sensitivity of transport investment benefits to the evaluation of small travel-time savings. *Journal of Transport Economics and Policy*, 231-254.

Appendix 3.1

The estimation results of dataset 1 have been discussed in detail in Section 3.4.2. This appendix provides estimation results of the remaining three datasets.

Table 3.7 Estimation results of Dataset 2

	RUM	**Conventional RRM**		**Level-based relative thinking**		**Range-based relative thinking**	
		CRRM	μRRM	CRRM-Level	μ RRM-Level	CRRM-Range	μ RRM-Range
	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)
β_{TT}	**-0.067** (*-35.13*)	**-0.047** (*-32.50*)	**-0.047** (*-31.85*)	**-2.920** (*-28.85*)	**-3.540** (*-33.74*)	**-1.410** (*-32.50*)	**-1.410** (*-31.85*)
β_{JAM}	**-0.027** (*-17.39)*	**-0.018** (*-16.66*)	**-0.018** (*-16.63*)	**-0.253** (*-12.03*)	**-0.694** (*-16.29*)	**-0.544** (*-16.66*)	**-0.544** (*-16.63*)
β_{VAR}	**-0.032** (*-11.86)*	**-0.021** (*-11.86*)	**-0.021** (*-11.87*)	**-0.171** (*-9.65*)	**-0.510** (*-11.51*)	**-0.420** (*-11.86*)	**-0.420** (*-11.87*)
β_{TC}	**-0.017** (*-21.52)*	**-0.113** (*-20.28*)	**-0.112** (*-20.78*)	**-1.040** (*-18.17*)	**-1.420** (*-21.02*)	**-0.789** (*-2.28*)	**-0.787** (*-20.78*)
μ	**-**	**-**	**0.936** (*3.38*)	**-**	**0.092** (*22.40*)	**-**	**0.936** (*3.38*)
Number of observations	3510	3510	3510	3510	3510	3510	3510
Null-likelihood	-3856	-3856	-3856	-3856	-3856	-3856	-3856
Final likelihood	-2613	-2605	-2605	-2676	-2609	-2605	-2605

Table 3.8 Estimation results of Dataset 3

	RUM	**Conventional RRM**		**Level-based relative thinking**		**Range-based relative thinking**	
		CRRM	μRRM	CRRM-Level	μRRM-Level	CRRM-Range	μRRM-Range
	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)
β_{FTT}	**-0.061** (*-10.24*)	**-0.043** (*-9.98*)	**-0.041** (*-10.20*)	**-0.876** (*-8.21*)	**-0.699** (*-8.73*)	**-0.685** (*-13.95*)	**-0.663** (*-14.63*)
β_{STT}	**-0.072** (*-13.38*)	**-0.049** (*-12.90*)	**-0.048** (*-13.32*)	**-1.270** (*-10.90*)	**-1.120** (*-13.28*)	**-0.651** (*-13.85*)	**-0.641** (*-14.46*)
β_{VAR}	**-0.008** (*-3.15*)	**-0.006** (*-3.75*)	**-0.006** (*-3.21*)	**-0.021** (*-1.42*)	**-0.023** (*-1.64*)	**-0.296** (*-3.81*)	**-0.290** (*-3.49*)
β_{RC}	**-0.290** (*-8.13*)	**-0.201** (*-8.04*)	**-0.194** (*-8.11*)	**-0.697** (*-7.90*)	**-0.536** (*-8.31*)	**-0.634** (*-12.39*)	**-0.615** (*-12.40*)
β_{TC}	**-0.506** (*-8.22*)	**-0.348** (*-7.73*)	**-0.339** (*-8.15*)	**-1.430** (*-8.85*)	**-1.390** (*-9.66*)	**-0.320** (*-3.77*)	**-0.324** (*-3.75*)
μ	**-**	**-**	**>10**	**-**	**>10**	**-**	**>10**
Number of observations	1572	1572	1572	1572	1572	1572	1572
Null-likelihood	-1727	-1727	-1727	-1727	-1727	-1727	-1727
Final likelihood	-1286	-1288	-1286	-1318	-1317	-1317	-1317

Table 3.9 Estimation results of Dataset 4-1

	RUM	**Conventional RRM**		**Level-based relative thinking**		**Range-based relative thinking**	
		CRRM	μRRM	CRRM-Level	μRRM-Level	CRRM-Range	μRRM-Range
	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)
β_{FSG}	**0.105** (*5.67*)	**0.068** (*4.51*)	**0.131** (*10.57*)	**-**	**-**	**0.472** (*13.53*)	**0.619** (*15.07*)
β_{FSO}	**0.011** (*4.08)*	**0.003** (*1.95*)	**0.001*** (*0.92*)	**-**	**-**	**0.100** (*2.90*)	**0.088** (*2.48*)
β_{TT}	**-0.045** *(-6.48)*	**-0.016** *(-5.35)*	**-0.012** *(-4.23)*	**-**	**-**	**-0.217** *(-7.90)*	**-0.183** *(-6.86)*
μ	**-**	**-**	**0.141** (*8.12*)	**-**	**-**	**-**	**0.141** (*6.42*)
Number of observations	1485	1485	1485	1485	1485	1485	1485
Null-likelihood	-2390	-2390	-2390	-2390	-2390	-2390	-2390
Final likelihood	-2276	-2272	-2234	-	-	-2183	-2171

Table 3.10 Estimation results of Dataset 4-2

	RUM	**Conventional RRM**		**Level-based relative thinking**		**Range-based relative thinking**	
		CRRM	μRRM	CRRM-Level	μRRM-Level	CRRM-Range	μRRM-Range
	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)	Estimate (*t-value*)
β_{FSG}	**0.111** (*2.52*)	**0.089*** (*1.52*)	**0.147** (*4.10*)	**0.042** (*3.57*)	**0.059** (2.90)	**0.529** (*6.04*)	**0.596** (*5.23*)
β_{FSO}	**0.011*** (*1.28)*	**0.001** (*0.15*)	**-0.0003** (*-0.07)*	**0.007** (*2.75*)	**0.08** (*2.22*)	**0.228** (*2.79*)	**0.370** (*3.47*)
β_{TT}	**-0.064** (*-3.13)*	**-0.023** (*-2.91)*	**-0.019** (*-2.62)*	**-0.265** (*-3.32)*	**-0.209** (*2.22*)	**-0.167** (*-2.39)*	**-0.130** (*-2.08*)
μ	**-**	**-**	**0.175** (*3.27*)	**-**	**>10**	**-**	**0.089** (*3.29*)
Number of observations	256	256	256	256	256	256	256
Null-likelihood	-412	-412	-412	-412	-412	-412	-412
Final likelihood	-388	-385	-375	-370	-369	-360	-358

4 A new loss aversion model

Abstract

Loss aversion is a fairly robust phenomenon in choice behaviour. It means losses have a greater impact on choices than equivalent gains. In this chapter, we propose a new loss aversion model. Specifically, the model specification is adapted from the regret function of the Random Regret Minimization (RRM) model, but it is fundamentally different from the RRM model, as the new loss aversion model does not restrict the reference point in the model specification. The new loss aversion model is theoretically contrasted with several existing loss aversion models and empirically compared with the Random Utility Maximization (RUM) and some loss aversion models on three datasets. Empirical results show a promising performance of the new loss aversion model.

4.1 Introduction

Over the past few decades, an increasing number of researchers have questioned the validity of the models built on classical economic theory, and some of them have proposed various alternative models of decision making, involving both risky and riskless contexts. Prospect Theory (Kahneman & Tversky, 1979) is one of the leading theories of decision making under risk. Unlike traditional normative models (e.g., expected utility theory), Prospect Theory is a descriptive model of decision making which is concerned with the choices people *actually* make rather than the choices people *should* make. In Prospect Theory, the process of decision making is described as two phases: "editing" and "evaluation" phases. In the editing phase, the outcome of risky choices (e.g., lotteries) is mapped as gains or losses relative to some reference point; gains are the outcome that exceeds the reference point, and losses refer to the outcome that is inferior to the reference point. The evaluation phase describes how decision makers evaluate the outcome of choices and transform objective probabilities of the outcome into subjective probabilities. Specifically, losses are evaluated more importantly than equivalent gains, and small probabilities are overweighted and large probabilities are underweighted.

Prospect Theory was built on the observations of individual stated choices in risky situations and originally framed as decision making in lottery-based gambling issues. Since its inception, Prospect Theory has rapidly attracted the interest of economists and it has been incorporated into some settings of economic models. Later, it was further developed and extended to the context of riskless choice situations (Tversky & Kahneman, 1991), which accelerates its applications in a wider range of contexts, such as travel behaviour.

Prospect Theory has three key elements: reference dependence, loss aversion and diminishing sensitivity. Reference dependence refers to the phenomenon that individuals' preference is influenced by some reference point. Loss aversion means losses have a greater impact on preference than equivalent gains. Diminishing sensitivity means the marginal utility of losses and gains decreases with the size of losses and gains. An S-shaped value function that covers these three elements was proposed in Kahneman & Tversky (1979), see Figure 4.1. The value function is *i*) defined on deviations from the reference point (reference dependence), *ii*) steeper for losses than gains (loss aversion), and *iii*) generally convex in the loss domain and concave in the gain domain (diminishing sensitivity).

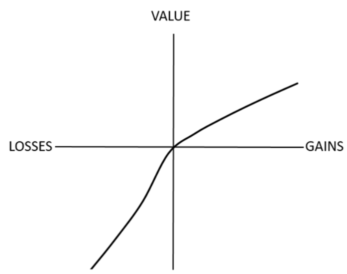


Figure 4.1 The shape of the value function (Kahneman & Tversky, 1979)

These three elements have been extensively applied in the study of (travel) choice behaviour. Below, we briefly introduce their implications, application issues, and specific examples in travel behaviour research.

- *Reference dependence*

Reference dependence refers to a behavioural pattern that preference is influenced by some reference point; shifts of the reference point may give rise to reversals of preference (Tversky & Kahneman, 1991). This notion has been widely incorporated into discrete choice models for analysing various travel-related choice behaviour. In this chapter, we mainly focus on discussing the issues about reference points in travel choice models.

Issues about the setting of reference points have been widely discussed in the literature. Many researchers have recognized the importance of being precise about the reference point (Avineri, 2009; Stathopoulos & Hess, 2012; Avineri & Ben-Elia, 2015). When dealing with monetary outcomes associated with economic decisions, it is generally easy to identify the reference point. For example, €0 is a common value for the reference point. However, when dealing with travel-related choices, from an analyst's perspective, it is very difficult to identify individual reference points, because they are mostly endogenous. The status quo is commonly used as the reference point of individuals' choices, but decision making in travel behaviour can be very context-specific. Expectations, a recent trip, and ideal or acceptable situations are also possible to be the reference points of travellers. Besides, the reference point can be exogenously given as well, for example, travel information given by mapping services. Apart from a single reference point, decision makers may also employ multiple reference points (Mayhew & Winer, 1992; Dholakia & Simonson, 2005; Wang et al., 2019) or have a range for the reference point. For instance, the ideal departure time can be a period of time rather than an exact time.

Attempts to identify actual reference points can be summarized as testing for multiple possible reference points when modelling travel behaviour. Stathopoulos & Hess (2012) tested for three reference points for travel fare in a commuting trip study: the fare of the current trip, the fare of the acceptable trip, and the fare of the ideal trip. Senbil & Kitamura (2004) defined two reference points for modelling departure time choices: the preferred arrival time and the scheduled work start time. Similar efforts can be also found in Jou & Kitamura (2002). Wang et al. (2019) tested multiple reference points in a consumer behaviour study: the most preferred product, the least preferred product, the average product and the status quo.

The majority of existing reference-dependent models deal with the reference point in a way that "it is known as a priori" (Bahamonde-Birke, 2018). The contextual concavity model (Kivetz et al., 2004) defines the least preferred value of the attribute as the reference point. A loss aversion model proposed by Kivetz et al. (2004) employs the mid-point of the attribute-level range as the reference point. Other reference-dependent models implicitly pre-define the reference point in the model specifications. For example, in the Random Regret Minimization (RRM) model (Chorus, 2010; van Cranenburgh et al., 2015), the reference point is implicitly set as the attribute level of unchosen alternatives. Likewise, the Relative Utility model (Zhang et al., 2004) applies other unchosen alternatives as to the reference alternative in the model specification. Apart from assuming the reference point a priori, a handful of studies also incorporate the decision heuristics theory; the reference point is assumed to be updated with choice contexts. For example, Balbontin et al. (2017) have integrated the value learning heuristics in the choice model in which the reference point is assumed to be associated with the best attribute level in the previous choice set. More recently, Bahamonde-Birke (2018) has proposed a loss aversion model in which the reference point can be estimated according to collected data.

- *Loss aversion*

It is widely observed that when people make choices, the outcomes of choices are often mapped as gains or losses against some reference point, and gains and losses receive different responses. Loss aversion means that losses receive a greater response than equivalent gains. In other words, decision makers are more sensitive to losses than gains, which induces a great tendency for decision makers to minimize losses when making choices.

Loss aversion can be used to explain many interesting behavioural phenomena. For example, the endowment effect (Thaler, 1980) refers to a phenomenon that people tend to demand much more to give up an object than they would like to pay to acquire it, often discussed as the discrepancy between the willingness-to-pay (WTP) and the willingness-to-accept (WTA). The explanation for this phenomenon is that the disutility of giving up an object (regarded as losses) is greater than the utility of having it (regarded as gains) (Kahneman et al., 1991). Another famous implication of loss aversion is the status quo bias (Samuelson & Zeckhauser, 1988). It illustrates a phenomenon that remaining at the status quo is strongly preferred by individuals, as the negative impact of leaving the status quo is evaluated greater than the positive impact. In addition, many anomalies in Economics, such as the equity premium puzzle (Benartzi & Thaler, 1995), disposition effects in finance (Weber & Camerer, 1998), framing effect (Tversky & Kahneman, 1981), can also be explained by the notion of loss aversion.

A large and growing body of literature has shown the presence of loss aversion in choice behaviour. In travel behaviour studies, De Borger & Fosgerau (2008) showed evidence of loss aversion in travel time and travel cost in a large-scale route choice study. Hess et al. (2008) found loss aversion in non-commuters' responses to toll cost, and in commuters' responses to slowed-down travel time. Masiero & Hensher (2010) investigated loss aversion in time and cost, as well as punctuality in freight choices, a strong loss aversion was found in response to punctuality. The study conducted by Flügel et al. (2015) revealed loss aversion in road safety; specifically, travellers exhibit loss aversion to the number of road causalities. Empirical evidence for loss aversion can be also found in other fields. For instance, in the field of health care, Neuman & Neuman (2008) found loss aversion in maternity-ward attributes: the number of beds in hospital rooms and travel time from residence to hospital. Ahtiainen et al. (2015) found that residents showed a strong loss aversion in the choices concerning water quality. In the energy field, Bartczak et al. (2017) did research about the preference for implementing renewable energy externalities; they found loss aversion in choices regarding monetary attributes.

Although there is ample evidence to support loss aversion in choice behaviour, some studies have also raised questions of whether loss aversion plays an important role in routine or habit behaviour. Kahneman et al. (1991) pointed out that loss aversion would disappear if individuals experienced financial losses in a market environment with ample learning opportunities. Novemsky & Kahneman (2005) found the absence of loss aversion in money in routine transactions. Coursey et al. (1987) showed the disparity between WTP and WTA decreases with experience in a market. In terms of travel behaviour, as argued by Timmermans (2010), it fundamentally differs from gambling in that "*travellers experience the consequences or outcomes of their decisions, and more importantly adapt their behaviour to influence the experienced outcomes*". Therefore, it is not surprising to find the absence of loss aversion in response to some travel attributes, such as departure time or destination choices.

- *Diminishing sensitivity*

As discussed above, people are more sensitive to losses than gains. However, the sensitivity to both losses and gains is diminishing with the size of gains and losses. This is called diminishing sensitivity. This property is reflected by the curvature of the value function: the value function is less curved with the distance from the original point, see Figure 4.1.

Diminishing sensitivity is often considered in travel demand modelling. It has consistently been observed that travellers' sensitivity to travel time or travel cost appears to decline with the length of trips. Daly (2010) exhaustively discussed a so-called "cost damping" effect, which indicates that the sensitivity to travel cost is diminishing as the trip length increases. Cost damping has been found to provide a better explanation to travel behaviour in many large-scale forecasting studies (e.g., Daly & Carrasco, 2009). In addition to travel cost and travel time, diminishing sensitivity is also found in other travel-related attributes. For example, Masiero & Hensher (2010) showed the presence of diminishing sensitivity in the punctuality attribute in a freight transport study.

Although diminishing sensitivity is a fairly robust property in some travel-related attributes, particularly travel cost, much empirical work also found evidence for non-diminishing sensitivity. For example, Stathopoulos & Hess (2012) found that the sensitivity to fare varies with different reference points: when the current situation was set to the reference point, the sensitivity of both losses and gains in fare was found to decrease; but when the ideal situation was set to the reference point, the sensitivity of losses in fare was found to increase.

Much applied work has extended diminishing sensitivity into non-linear sensitivity (e.g., Stathopoulos & Hess, 2012). In the travel demand analysis, many travel choice models also employ more flexible non-linear functions, which relaxes the assumption of decreasing marginal utility (e.g., Koppelman, 1981; Ben-Akiva & Lerman, 1985).

The above abundant literature has shown the wide application of the three key elements of Prospect theory in (travel) choice behaviour research. The next section introduces the common ways to incorporate loss aversion (and also other two elements) into travel choice models. Section 4.3 presents a new loss aversion model, including the model specification, model properties, and comparisons with existing loss aversion models. Empirical applications of the new model are given in Section 4.4. Finally, this chapter is completed with discussions and future research avenues.

4.2 Loss aversion modelling approaches

Common approaches to modelling loss aversion are often built on the value function of Prospect Theory, which is a piecewise function that models losses and gains separately. This section starts by introducing the functional form of the value function and then discusses two common ways of modelling loss aversion in travel choice models.

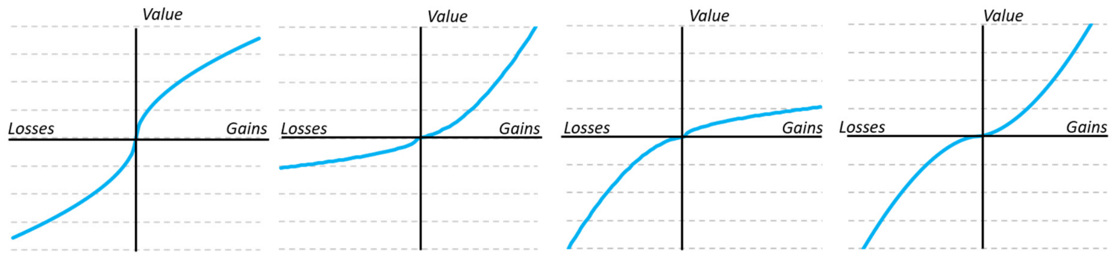
4.2.1 The value function

The formulation form for the value function was first given in Tversky & Kahneman (1992), which is a two-part power function:

$$v(x) = \begin{cases} x^a & if\ x \geq 0 \\ -\lambda(-x)^b & if\ x < 0 \end{cases}, \qquad \text{(Eq. 4.1)}$$

where x denotes the deviations (i.e., losses or gains) from the reference point; λ is a loss aversion parameter and it only takes a positive value; parameters a and b capture the curvature of the value function and they also take positive values. Prospect Theory suggests that the value function is an S-shaped function, which corresponds to the estimation results of $\lambda > 1$, $0 < a < 1$ and $0 < b < 1$. Tversky & Kahneman (1992) gave the estimates of these parameters: $\lambda = 2.25$ and $a = b = 0.88$.

Since these parameters are estimable, it is also possible to obtain a different shape of the value function. For parameter λ, loss aversion occurs if $\lambda > 1$; losses and gains are equivalently weighted if $\lambda = 1$; gains are overweighted than equivalent losses if $0 < \lambda < 1$. For the curvature parameters, if $0 < a < 1$ and $0 < b < 1$, the value function is concave for gains and convex for losses, as shown in Figure 4.2 (a). In this case, there is a diminishing sensitivity in both gain and loss domains (align with Prospect Theory). If $a > 1$ and $0 < b < 1$, the value function is convex for both gains and losses, as shown in Figure 4.2 (b). In this case, there is an increasing sensitivity in the gain domain but a diminishing sensitivity in the loss domain. Figure 4.2 (c) shows a situation where $0 < a < 1$ and $b > 1$, there is a diminishing sensitivity in gains but an increasing sensitivity in losses. Figure 4.2 (d) shows another situation where $a > 1$ and $b > 1$, meaning an increasing sensitivity in both domains.



(a) $0 < a < 1, 0 < b < 1$ (b) $a > 1, 0 < b < 1$ (c) $0 < a < 1, b > 1$ (d) $a > 1, b > 1$

Figure 4.2 The curvature of the value function ($\boldsymbol{\lambda = 1}$)

There is also a circumstance in which the sensitivity to gains or losses does not change with the size of losses and gains from the reference point. Figure 4.3 shows a case where both curvature parameters are equal to one. In this case, the value function is piecewise-linear.

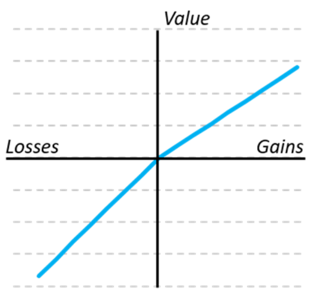


Figure 4.3 A piecewise-linear value function ($\boldsymbol{a = 1, b = 1, \lambda = 1.5}$)

As discussed in Tversky & Kahneman (1992), the curvature of the value function also reflects risk attributes in choice preference. In the loss domain, the increasing marginal value function implies a risk-averse preference for losses, while the diminishing marginal value function implies a risk-seeking preference. In the gain domain, the increasing marginal value function implies risk-seeking, and the diminishing function implies risk-averse. Take Figure 4.2 (a) as an example, its corresponding attitudes are risk-seeking for losses and risk-averse for gains.

4.2.2 Common ways of modelling loss aversion

In travel behaviour research, the common way of modelling loss aversion is to incorporate the value function into the Random Utility Model (RUM) framework (McFadden, 1973). In the RUM, the utility function consists of two parts—a systematic part that is typically assumed to be linear in parameters, and a random part that is assumed to follow an I.I.D extreme value type I distribution:

$$U_i = V_i + \varepsilon_i = \sum_m \beta_m x_{im} + \varepsilon_i \qquad \text{(Eq. 4.2)}$$

where U_i is the total utility of alternative i, V_i denotes the systematic part of the utility, and ε_i denotes the random part. In the systematic part, x_{im} is the attribute level associated with attribute m of alternative i, and β_m is the parameter associated with attribute m. To model loss aversion, the value function can be integrated into the systematic utility, in which the utility for losses and gains are modelled separately:

$$V_i^{loss-gain} = \sum_m \{\beta_{im(loss)} x_{im(loss)}{}^{\delta(loss)} + \beta_{im(gain)} x_{im(gain)}{}^{\delta(gain)}\} \qquad \text{(Eq. 4.3)}$$

This loss-gain utility function is a piecewise-nonlinear form with separate parameters associated with losses and gains relative to some reference level. We call this model *the piecewise-nonlinear model*. In the loss part, $x_{im(loss)}$ denotes the value of losses associated with attribute m, $\beta_{im(loss)}$ is the parameter associated with $x_{im(loss)}$, and $\delta(loss)$ is the curvature parameter for the loss domain. In the gain part, $x_{im(gain)}$ denotes the value of gains associated with attribute m, $\beta_{im(gain)}$ is the parameter associated with the gains, and $\delta(gain)$ is the curvature parameter for the gain domain. The test of loss aversion is conducted by comparing the sizes of $\beta_{im(loss)}$ and $\beta_{im(gain)}$: loss aversion occurs if $\beta_{im(loss)}$ is estimated to be significantly larger than $\beta_{im(gain)}$ (absolute values).

Losses and gains ($x_{im(loss)}$ and $x_{im(gain)}$) are attribute-level differences between alternative *i* and the reference alternative or the reference level. If attribute m is the attribute that any increases in it will increase the utility of the alternative, such as transport service quality, $x_{im(gain)} = x_{im} - x_{rm}$, if $x_{im} \geq x_{rm}$ and $x_{im(loss)} = x_{rm} - x_{im}$, if $x_{im} < x_{rm}$; while if attribute *m* is the attribute that any increases will cause a disutility, such as travel time and fare, $x_{im(gain)} = x_{rm} - x_{im}$ if $x_{im} \leq x_{rm}$, and $x_{im(loss)} = x_{im} - x_{rm}$ if $x_{im} > x_{rm}$. Here x_{rm} denotes the attribute level of the reference.

One challenge that choice modellers may face when estimating this model is the identification of model parameters (Avineri & Bovy, 2008). The estimations of parameter β_{im} and the curvature parameter δ depend on the values of losses or gains. For example, the estimations of parameters $\beta_{im(loss)}$ and $\delta(loss)$ depend on the values of losses $x_{im(loss)}$. These two parameters are very likely to be confounded, which may lead to failure in identification and model convergence. To avoid identification issues, many studies in economics directly use the estimates of the parameters ($\lambda = 2.25$ and $a = b = 0.88$) given in Tversky & Kahneman (1992). This however cannot be applied to travel behaviour, as travel behaviour is very different from and more complicated than gambling-based economic behaviour. As pointed out in Avineri & Bovy (2008): "*Because of the different contexts in which travel journeys are made and because of different modes and trip purposes, it is difficult to estimate a set of parameter values that represent the common decision makers*".

Instead, much travel behaviour research adopts simplified versions of the value function to model loss aversion (Hess et al., 2008; Lanz et al., 2010; Masiero & Hensher, 2010). The most simplified one is *a piecewise-linear model*:

$$V_i^{loss-gain} = \sum_m \{\beta_{im(loss)} x_{im(loss)} + \beta_{im(gain)} x_{im(gain)}\}. \qquad \text{(Eq. 4.4)}$$

Different from Eq. 4.3, this function does not contain the parameters that govern the curvature of the function. This means the model only captures reference dependence and loss aversion, but ignores diminishing sensitivity, or in other words, it assumes that the sensitivity to losses and gains does not change with the size of losses and gains. This model is easy to use, and more importantly it does not have identification issues.

4.3 A new loss aversion model

This section proposes a new loss aversion model. The model specification is adapted from the Random Regret Minimization (RRM) model which is regarded as a regret-based counterpart of the linear-additive RUM model. Regret is defined as an unpleasant experience when unchosen alternatives perform better than the chosen alternative on one or more attributes. The RRM model postulates that when faced with a set of choice alternatives, decision makers will choose the alternative which brings them the minimum regret. An important notion of this model is *regret aversion*: regret caused by the poor performance of an attribute has a greater impact than rejoice caused by the good performance, thus, to avoid regret, decision makers exhibit regret-averse behaviour. It is instructive to note the relation between regret aversion and loss aversion: they both capture the same behavioural pattern—bad experience (regret or losses) has a greater impact on decision making than good experience (rejoice or gains). Moreover, the RRM model is a reference-dependent model. The model specification is composed of pairwise comparisons in attribute levels between the chosen alternative and other competing alternatives. It means that the RRM model implicitly sets the reference point in the model specification: the attribute level of other competing alternatives. Therefore, the RRM model can be regarded as a special case of a loss aversion model in which the reference point is the attribute level of other competing alternatives.

Inspired by the relation between regret aversion and loss aversion, we propose *a loss aversion function* (L_i), which is adapted from the regret function of the μRRM model (van Cranenburgh et al., 2015):

$$L_i = \sum_m \mu_m \ln(1 + \exp(\frac{\beta_m}{\mu_m}[x_{rm} - x_{im}])). \qquad \text{(Eq. 4.5)}$$

Here L_i denotes the systematic (dis)utility of choosing alternative i, β_m denotes the taste parameter (representing attribute importance) associated with attribute m, μ_m is an attribute-specific parameter which governs the degree of loss aversion in attribute m, and x_{im} and x_{rm} are the attribute levels of attribute m associated with alternative i and the reference alternative r respectively. This function postulates that decision makers will choose the choice alternative which brings the minimum loss. To derive choice probabilities, we adopt a similar approach as the RRM model: the minimization of the random disutility L_i is mathematically equivalent to the maximization of the negative of the random disutility. Thus, choice probabilities can be derived using a variant of the linear-additive RUM formulation; the choice probability associated with alternative i is given by

$$P_i = \frac{\exp(-L_i)}{\sum_{j=1,\dots,J}\exp(-L_j)}. \qquad \text{(Eq. 4.6)}$$

It is instructive to note the difference between the loss aversion function and the regret function of the μRRM model, since the former one is adapted from the latter. The regret function is given as follows: $R_i = \sum_{j\neq i}\sum_m \mu_m \ln(1 + \exp\{\frac{\beta_m}{\mu_m}(x_{jm} - x_{im})\})$. The first notable difference between the two functions is the formulation of attribute-level comparison. In the regret function, the attribute-level comparison focuses on bilateral comparisons between the chosen alternative and every other unchosen alternative, while in the loss aversion function, the comparison is only made between the chosen alternative and the reference. The second difference is the number of sum operators in the functions. In the regret function, there are two sum operators: the first sum operator takes the sum of all “binary regret” generated by comparing between the chosen alternative and every unchosen alternative, and the second sum operator is used to calculate the “binary regret” which is specified as the sum of all attribute-level comparisons in terms of each attribute. As for the loss aversion function, it assumes disutility arises from comparing the chosen alternative with the reference alternative rather than every other alternative. Thus, there is only one sum operator in the function and it adds up all attribute-level comparisons for each attribute. Figure 4.4 illustrates the differences in calculation logic between the regret function and the loss aversion function. Specifically, in this example, there are three alternatives (Alt a, Alt b, and Alt c) and a reference alternative (Ref), each alternative is described by two attributes (Att 1 and Att 2).

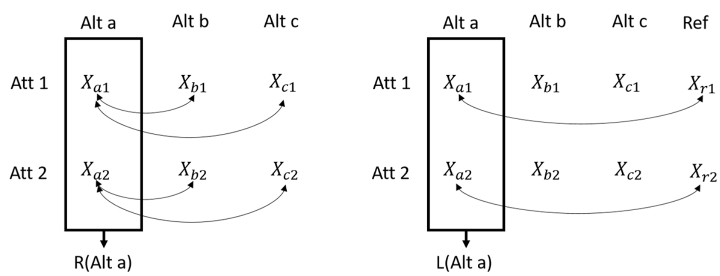


Figure 4.4 Calculation differences between the regret function and the loss aversion function

Figure 4.5 visualizes the shape of the loss aversion function under different values of parameter μ. It specifically focuses on one attribute-level comparison between alternative i and the reference *r*. The X-axis represents the difference between x_r and x_i. Note that the X-axis denotes $y = \ln(2)$ rather than $y = 0$. The right-hand side of the X-axis represents the loss domain in which alternative *i* performs worse than the reference *r*, and the left-hand side is the gain domain in which alternative i outperforms the reference *r*. Parameter μ takes different positive values, varying from a very small value, 0.01, to a large value, 100. As shown in the figure, when parameter $\mu = 100$, the loss aversion function is linear (grey line). In this case, losses and gains are equally important, implying a situation where there is no loss aversion in behaviour. When parameter μ takes a fairly large value, for example, $\mu = 10$, the loss aversion function is almost linear (blue line). When μ takes smaller values, the loss aversion function becomes more curved. For example, when $\mu = 1$, the loss aversion function (yellow line) represents mild loss aversion, that is losses and gains are both important, but losses are weighted more importantly than gains. The red line, i.e., $\mu = 0.01$, represents an extreme situation in which only losses matter the behaviour, in this case, there is the strongest loss aversion in behaviour.

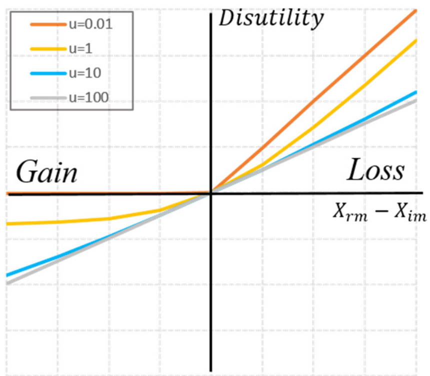


Figure 4.5 The shape of the loss aversion function ($\beta = 1$)

As shown in Figure 4.5, in contrast to the value function and the existing models (piecewise-linear/nonlinear), the loss aversion function is smooth and twice differentiable on the entire X-axis. In addition, the degree of loss aversion associated with each attribute can be more straightforwardly presented by the size of parameter μ: the smaller the size of μ, the stronger loss aversion. This relates to the convenience of applying the new loss aversion model—the degree of loss aversion is disentangled from attribute importance, as they are captured by μ and β respectively, whereas in the piecewise-nonlinear/nonlinear models, loss aversion and attribute importance are entangled (both captured by β).

However, compared to the value function and the existing models, the new loss aversion model is less flexible in terms of capturing an opposite behaviour that gains loom larger than losses. Due to the convexity of the function, the new loss aversion model can only capture the behaviour that losses overweigh gains and no loss aversion, but not the behaviour that gains are evaluated more importantly than losses. Moreover, compared to the piecewise-nonlinear model, this new model is also less flexible in terms of capturing risk attributes (also due to the convexity of the function). It postulates risk-averse attributes in both loss and gain domains (corresponding to Figure 4.2 (c)), which goes partly against the assumptions of Prospect Theory. In the following section, we take a closer look at the properties of the new loss aversion model.

4.3.1 Properties of the new loss aversion model

Property 1: Reference dependence

The new loss aversion model is a reference-dependent model. In contrast to the RRM models, the new loss aversion model does not restrict the reference point in the model specification. Instead, it can easily accommodate any reference points in the loss aversion function. For example, the reference point X_{rm} can be replaced by the value of the status quo, the value of expectations, or ideal or acceptable values. In addition, exogenous reference points and context-dependent reference points can also be easily incorporated into the model specification. The reference level X_{rm} can be replaced by, for example, the least preferred value, the most preferred value, or the mid-point of the attribute range $(max\{X_{im}\} - min\{X_{im}\})/2$, and the like.

Property 2: Loss aversion

The parameter μ governs the shape of the loss aversion function. Similar to the μRRM model, the new loss aversion model has two special cases:

- *Special case 1: μ is arbitrarily small*

When parameter μ is arbitrarily small, the loss aversion function becomes almost piecewise-linear, see the red line in Figure 4.5. It describes a very strong loss aversion behaviour. In the limit case (i.e. $\mu \rightarrow +0$), the loss aversion function is piecewise-linear, which implies the strongest loss aversion behaviour that only losses have a great impact on choices, but gains yield no impact.

- *Special case 2: μ is arbitrarily large*

When parameter μ is arbitrarily large, the loss aversion function hardly yields a difference between the impact of losses and the impact of gains on behaviour. In the limit case (i.e. $\mu \rightarrow +\infty$), the loss aversion function is linear. In this case, the model collapses to the linear-additive RUM model, which exhibits the behaviour with no loss aversion; in other words, losses are weighted equivalently as gains.

- *Loss aversion in behaviour*

As parameter μ governs the shape of the loss aversion function, it can straightforwardly reflect the degree of loss aversion in behaviour. However, in practice, the realization of data (i.e., the observed distribution of attribute-level differences) needs to be jointly taken into consideration. Figure 4.6 presents three plots, each consisting of the same loss aversion function (red line) and the histogram of attribute-level differences (blue bars). Note that the histograms are based on three different synthetic data. In Figure 4.6 (a), most attribute-level differences distributes in the area where the difference between loss-induced (dis)utility and gain-induced utility is small (green area), and a small portion of attribute-level differences locates in the area where the difference between loss-induced (dis)utility and the gain-induced utility is significant (orange area). This refers to a situation where the behaviour imposed by the loss aversion model is only mildly driven by the loss aversion notion. Figure 4.6 (c) illustrates an opposite situation where most attribute-level differences distribute in the orange area, but a small portion is in the green area. This refers to a situation in which the behaviour imposed by the loss aversion model is strongly driven by the loss aversion notion. In Figure 4.6 (b), attribute-level differences are evenly distributed, which presents an intermediate situation.

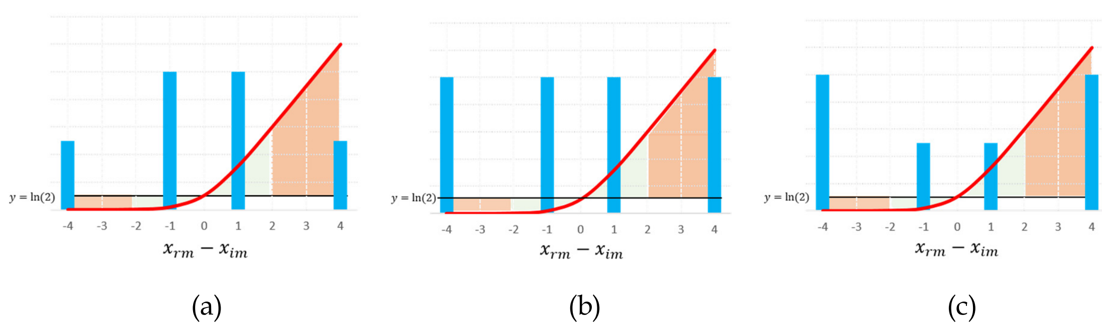


Figure 4.6 Loss aversion in behaviour for three different synthetic datasets

Property 3: non-linearity

The loss aversion function is convex in both loss and gain domains. The convex shape implies that there is a diminishing sensitivity in the gain domain, but an increasing sensitivity in the loss domain. This, however, goes partly against Prospect Theory which postulates that both domains display diminishing sensitivity. In terms of risk attitudes, the convex shape also implies that the new loss aversion model postulates risk-averse attitudes in both loss and gain domains. This may not align with actual behaviour. For example, if the actual behaviour does not align with such risk-averse attributes, it could lead to deterioration in model fit.

4.3.2 How to apply the loss aversion model

This section focuses on some subtleties of the new loss aversion model when applying it in choice modelling. As shown in Figure 4.6 (blue line), the loss aversion function is always above the X-axis ($y = 0$), and it does not approximate zero when the difference in attribute levels between the chosen alternative and the reference alternative ($x_{rm} - x_{im}$) is very small. The loss aversion function implies that both losses and gains lead to disutility. This, at the first glance, seems counterintuitive. However, it should be noted that the absolute level of utility is irrelevant to behaviour (Train, 2003).

The loss aversion function can be modified by subtracting $ln(2)$ from the level associated with one comparison between the chosen alternative and the reference alternative: $L_i{}' = \sum_m \mu_m \ln(1 + \exp(\frac{\beta_m}{\mu_m}[x_{rm} - x_{im}]) - \ln(2))$. As shown in Figure 4.7 (purple line), the modified function merely shifts downwards without changing the shape of the function. Therefore, the loss aversion function and the modified version are completely consistent with each other in terms of every relevant model property, such as model fit, parameter estimates and choice probability predictions.

Therefore, the loss aversion function can be directly applied without modification, except for some special cases. For example, only one or two attributes are tested for loss aversion (commonly time-related or cost-related attributes), but the remaining attributes are modelled using the linear-additive utility function. In this case, the loss aversion function needs to be transferred into the form of a utility function: $V_i^{loss-gain} = -L_i{}' = -\sum_m \mu_m \ln(1 + \exp(\frac{\beta_m}{\mu_m}[x_{rm} - x_{im}]) - \ln(2))$. Thus, the overall utility function including a loss aversion part and a linear-additive utility part can be written as $V_i = V_i^{loss-gain} + \sum_n \beta_{in} x_{in}$.

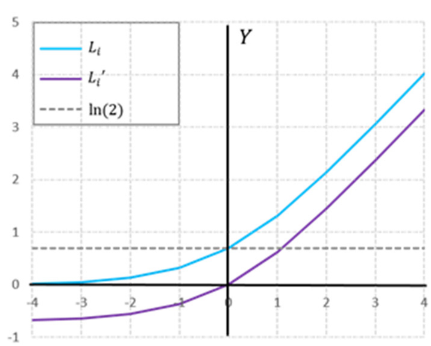


Figure 4.7 The loss function and its modified version ($\boldsymbol{\beta = 1, \mu = 1}$)

4.3.3 Marginal Value-of-Time formulation

Discrete choice analysis is not only used for studying choice behaviour, but also applied for economic appraisals. The notion of Value-of-Time (VoT) is very crucial in economic appraisals of transport projects. It evaluates trade-offs between travel time and travel cost. The VoT is usually interpreted as the WTP for a reduction in travel time by one unit. In the linear-additive RUM models, the derivation of VoT is well known and widely applied, given by the ratio of marginal utilities: $VoT_{RUM} = (\partial U_i/\partial TT_i)/(\partial U_i/\partial TC_i)$, where U_i is the total utility of alternative *i*, and TT_i and TC_i denote travel time and travel cost of alternative *i*. In the case that the utility function is linear in parameters and linear in attributes, the ratio of marginal utilities reduces to the ratio of parameters β_{TT}/β_{TC}.

The new loss aversion model takes on a very different specification from the linear-additive RUM model. As discussed above, its model specification incorporates comparisons between the attribute level of the considered alternative and that of the reference, which introduces a reference dependence not present in the linear-additive RUM model. The VoT measure based on the new loss aversion (LA) model is derived as follows:

$$VoT_{LA} = \frac{\partial L_i/\partial TT_i}{\partial L_i/\partial TC_i}$$

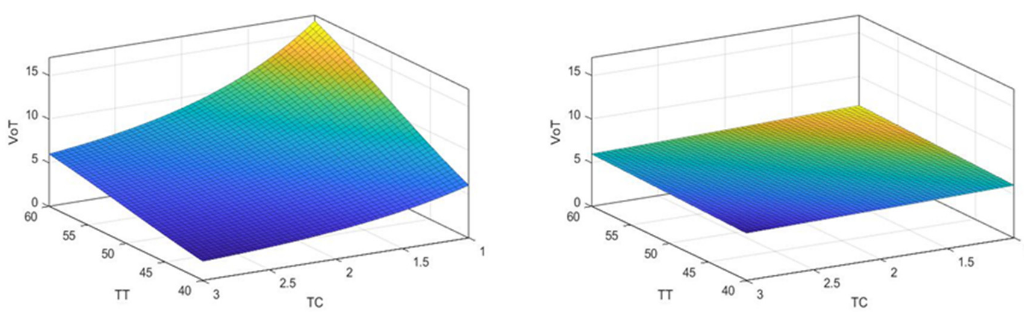
$$= \frac{\beta_{TT}}{\beta_{TC}} \frac{\exp\left(\beta_{TT}/\mu_{TT}\,(TT_r - TT_i)\right)/\left(1+\exp\left(\beta_{TT}/\mu_{TT}\,(TT_r - TT_i)\right)\right)}{\exp\left(\beta_{TC}/\mu_{TC}\,(TC_r - TC_i)\right)/\left(1+\exp\left(\beta_{TC}/\mu_{TC}\,(TC_r - TC_i)\right)\right)}. \qquad \text{(Eq. 4.7)}$$

Similar to the linear-additive RUM model, the ratio of parameters β_{TT}/β_{TC} also plays a role in the LA-based VoT measure. Note that the interpretation of parameters β_{TT} and β_{TC} in the LA model is very different from them in RUM. Whereas in RUM models, parameters represent the additional utility brought about by a unit increase in an attribute, in the LA model they reflect the upper (or lower) bound of the extent to which attribute-level differences of an attribute (compared to the reference) influences the systematic (dis)utility L_i associated with the alternative *i*.

The rest part of Eq. 4.7 shows that the loss aversion parameter μ, and attribute-level differences between alternative *i* and reference *r* also enter the VoT equation, which implies that the LA-based VoT measure is choice set-dependent; it changes when the choice set changes in terms of relative preformation of the alternative compared to the reference. Below, a numerical example is presented to illustrate how the VoT measure changes with attribute levels and the loss aversion parameter.

Consider a situation that the current travel time and cost of a traveller's route is 50 min and €2 respectively: $r_{current} = \{50, 2\}$. Now, this route's travel time and cost will be influenced due to a new transport policy: $r_i = \{TT_i, TC_i\}$ where $40 \le TT_i \le 60$ and $1 \le TC_i \le 3$. Conditional upon $\beta_{TT} = -0.1/min$, and $\beta_{TC} = -1/€$, the VoT measure of r_i based on the linear-additive RUM model is derived, that is €6 per hour. Figure 4.8 (a) plots the VoT values when a moderate loss aversion ($\mu_{TT} = \mu_{TC} = 1$) exists in both travel time and travel cost. We can see that when the route changes to be faster but more expensive than the current situation (e.g. $r_i = \{40, 3\}$), the VoT value based on the LA model is smaller than €6 per hour, dropping to about €2 per hour. When the route changes to be slower but cheaper, $r_i = \{60, 1\}$, the VoT value increases as high as about €16 per hour.

These changes in the VoT values are in line with the disparity between WTP and WTA caused by loss aversion. When the route becomes faster but more expensive than its current situation, the VoT value is interpreted as the WTP for a reduction in travel times. When the route changes to be slower but cheaper, the VoT values are interpreted as the WTA for an increase in travel times. The disparity between WTP and WTA (WTP < WTA) is the result of loss aversion: because the marginal effect of a reduction in travel times is smaller than the marginal effect of an increase in travel costs (which leads to a small WTP) and the marginal effect of an increase in travel times is larger than the marginal effect of a reduction in travel costs (which leads to a large WTA).



(a) $\mu_{TT} = 1;\ \mu_{TC} = 1$ (b) $\mu_{TT} = 10;\ \mu_{TC} = 10$

Figure 4.8 VoT measures based on the new loss aversion model (€/hour)

Figure 4.8 (b) plots the VoT values when a slight (almost no) loss aversion ($\mu_{TT} = \mu_{TC} = 10$) in travel time and travel cost. We can see that the VoT values based on the LA model remain roughly the same—6€ per hour—regardless of the changes in TT and TC, which also approximately equals the VT value based on RUM. This is in line with the property of the LA model: when μ equals a large value, the LA model collapses to the linear-additive RUM model.

Having discussed the LA-based VoT measure, we now turn to the potential of applying it to the aggregated social welfare analysis. As mentioned above, the loss aversion function is adapted from the RRM model, but its differences from RRM enable the possibility of the application in aggregated welfare measures. Specifically, the RRM models are formulated as the sum of all comparisons between the chosen alternative and others across all attributes. As a result, the regret of the chosen alternative is determined by both its own and other alternatives' performance. This however results in aggregation issues when applying the RRM-based VoT measure (Chorus, 2012) in the welfare analysis: the derived VoT only considers the impact of changes in travel times (travel costs) on the regret of the chosen alternative, but neglects the simultaneous change in the regret of other alternatives (caused by the change in the chosen alternative)[13]. In other words, the derived VoT value based on RRM only represents a VoT value of a particular consumer for a particular alternative, which cannot be used for aggregated welfare measures as consumers of all alternatives do not have the same VoT value for the same

[13] Note that Dekker (2014) and Dekker & Chorus (2018) expand RRM-based welfare measures from the perspectives of indifference and consumer surplus, respectively.

changes in travel times. As for the new LA model, the loss aversion function is formulated as a function of comparisons between the chosen alternative and the reference (rather than other alternatives). Changes in travel times (travel costs) of one alternative thereby do not influence the (dis)utility of other alternatives. This leads to the possibility of applying the LA-based VoT measure in transport appraisals.

4.3.4 Model comparisons

Having discussed the model properties of the new loss aversion model, this section recaps the main differences between the new loss aversion model and the piecewise-linear/nonlinear model.

Is the model twice differentiable in the full domain?

As discussed above, the value function is a piecewise function. Thus, it is not twice differentiable around the reference point. Likewise, as the piecewise-linear/nonlinear models are adapted from the value function, they are not twice differentiable around the reference point. As for the new loss aversion model, its loss aversion function is smooth in the full domain. Therefore, unlike the piecewise-linear/nonlinear model, the new model is twice differentiable.

How to apply the model to capture loss aversion?

In the piecewise-linear/nonlinear model, parameters β_{loss} and β_{gain} are associated with losses and gains respectively. The test of loss aversion is conducted by comparing the size of β_{loss} and β_{gain}: loss aversion occurs when β_{loss} is significantly larger than β_{gain} (absolute values). Therefore, to examine whether there is loss aversion, an asymptotic t-test is needed to test the significance of the difference between β_{loss} and β_{gain}. But in the new loss aversion model, no extra calculation is required. The degree of loss aversion is captured by the size of the parameter μ: the smaller the size of μ, the stronger the degree of loss aversion.

How does the model capture nonlinear sensitivity?

The piecewise-linear model fails to capture nonlinear sensitivity, due to its linear form. The new loss aversion relaxes the assumption of linearity. Due to the convexity of the loss aversion function, the model assumes that the sensitivity to losses is increasing and the sensitivity to gains is diminishing. The piecewise-nonlinear model is the most flexible in terms of capturing nonlinear sensitivity. It allows estimating the curvature of the function.

Are there any identification issues in the model parameter estimation?

In the piecewise-linear model, separate parameters (i.e. β_{gain} and β_{loss}) are estimated for losses and gains respectively. Similar to the linear-additive RUM model, parameters in the piecewise-linear model are expected to be easily identified. As for the piecewise-nonlinear model, two extra parameters (i.e. $\delta_{\_gain}$ and $\delta_{\_loss}$) are estimated for the curvatures of the marginal utility of losses and gains respectively. Evidence has shown that these two parameters have a high chance of being correlated with parameters associated with losses and gains (β_{gain} and β_{loss}) (Avineri & Bovy, 2008), leading to a confounding effect between loss aversion and nonlinear sensitivity.

In the new loss aversion model, the loss aversion function is adapted from the regret function of the μRRM model. Ample empirical analyses of the μRRM model have shown that the parameter μ can be identified in most cases (e.g., van Cranenburgh et al., 2015 Table 5.2). Therefore, we expect that parameters of the new loss aversion model are very likely to be identified, but we do not rule out the chance that parameter μ is confounded with parameters associated with attribute importance[14].

4.4 Empirical applications

This section presents the empirical analysis of the new loss aversion model using three datasets. Specifically, we compare the new loss aversion model with the linear-additive RUM model and the piecewise-linear model. The reason for choosing the piecewise-linear model rather than the piecewise-nonlinear model is twofold: *i*) the piecewise-nonlinear model may have parameter identification issues, and *ii*) the comparison with the piecewise-linear model is fair in terms of the number of model parameters. Model estimation is performed on three empirical datasets: two route choice data and one policy package data. The empirical analysis focuses on testing model performance in terms of model fit and the capability of capturing loss aversion.

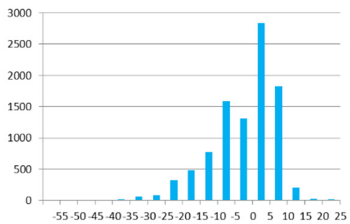
4.4.1 Dataset description

- *Dataset 1: Route choices*

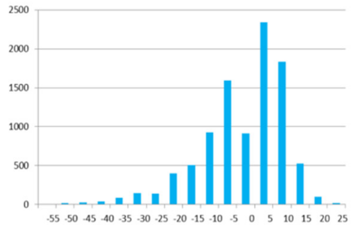
This dataset is part of the study that aimed to estimate the VoT of car drivers in metropolitan regions of Sydney (Rose & Masiero, 2010). It consists of 300 car drivers who made sixteen hypothetical route choices each. These choices were made between *a current route* (the reference) and *two SP routes* which were described by five attributes: free-flow travel time (FF), slowed-down travel time (SLOW), travel time variability (VAR), running cost (RUN) and toll cost (TOLL). Time-related attributes are expressed in minutes, and cost-related attributes are in Australian dollars. The two SP alternatives were generated according to the reference alternative which was provided by the respondents.

As mentioned in Section 4.3, it is very informative to look at the distribution of attribute level differences when examining loss aversion in behaviour. Figure 4.9 shows the histograms of attribute-level differences of the five attributes. We can see that for attributes FF, SLOW and RUN, the attribute-level differences are distributed on both sides around 0, and the majority of the differences are concentrated around 0. For attribute VAR, only a very small number of attribute-level differences are positive, the majority of the differences are negative values; For attribute TOLL, most attribute-level differences are positive values.

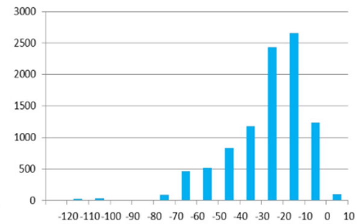
[14] We recommend setting the bounds (0.01 to 5) when estimating parameter μ. When the estimate of μ hits the lower bound, i.e., 0.01, it suggests parameter μ is very small and it approaches 0. When the estimate of μ hits the upper bound, i.e., 5, it means parameter μ is very large.



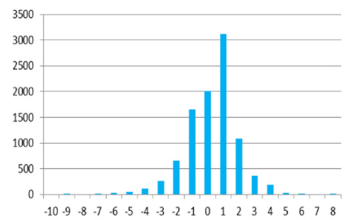
(a) $FF_i - FF_r$



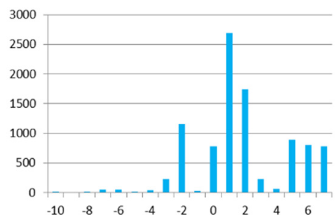
(b) $SLOW_i - SLOW_r$



(c) $VAR_i - VAR_r$



(d) $RUN_i - RUN_r$



(e) $TOLL_i - TOLL_r$

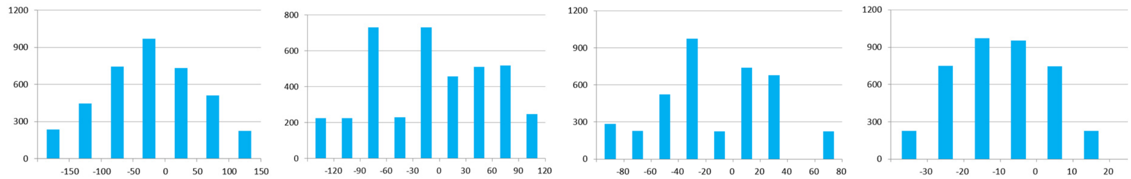
Figure 4.9 Histograms of attribute-level differences in Dataset 1

- *Dataset 2: Policy package choices*

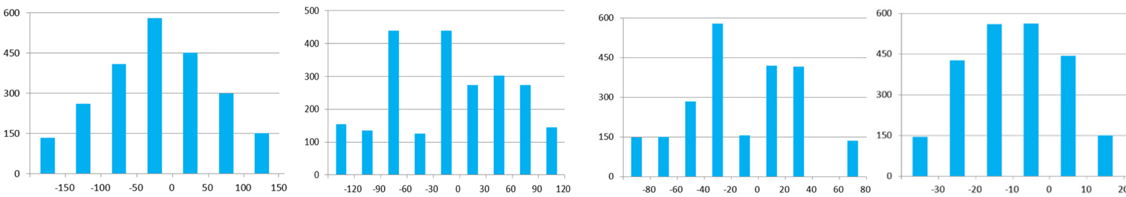
This dataset was collected for the study in Chapter 2. Data collection is aimed at investigating Dutch citizens' opinions on safety issues of automated vehicles (AVs). Choices were made between three hypothetical scenarios which were described as the outcomes of policy packages for anticipating the AV era. The policy packages were described by four attributes, with three levels each: the number of fatalities per year caused by conventional vehicles (CF) (250, 300, 350, 400), the number of fatalities per year caused by technical failure of AVs (AFT) (50, 100, 150, 200), the number of fatalities per year caused by a malicious act regarding AVs (AFM) (0, 30, 60, 90), and the average reduction in car travel time (TR) (30%, 20%, 10%, 0%). In total, twelve choice sets were created. These twelve choice sets were divided into two blocks with six sets each. Each respondent was faced with six choice sets.

In contrast to Dataset 1, in which the reference was included as an alternative in the choice set, the reference in this dataset was provided by the context information. More specifically, before performing six choice sets, respondents were requested to read an introduction page that contained reference levels for conventional fatalities and AV fatalities respectively. There were two types of reference levels: the real current situation and future projections. The detailed reference levels are given in Chapter 2. This chapter only analyses the data which contained the reference levels of future projections. In total, 215 respondents completed the experiment with the reference of future projections, and among them, 127 respondents indicated that they has certain considerations for the provided reference levels. Note that we estimate the models on the whole dataset of 215 respondents and the subset of 127 respondents. With six choice tasks made by each respondent, the two sets contain 1290 and 762 choice observations respectively.

Figure 4.10 shows the histograms of attribute-level differences of the attributes: the whole data set (a), (b), (c), (d) and the subset (e), (f), (g), (h). We can see that first, for each attribute, the frequency distributions of the whole dataset and the subset are roughly the same; second, for attributes CF, AFT, AFM, the positive values and negative values are roughly balanced in number; third, attribute TR has more negative values than positive values.



(a) $CF_i - CF_r$ (b) $AFT_i - AFT_r$ (c) $AFM_i - AFM_r$ (d) $TR_i - TR_r$



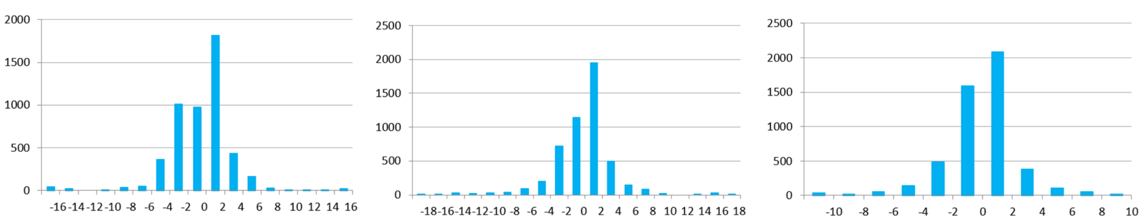
(e) $CF_i - CF_r$ (f) $AFT_i - AFT_r$ (g) $AFM_i - AFM_r$ (h) $TR_i - TR_r$

Figure 4.10 Histograms of attribute-level differences in Dataset 2

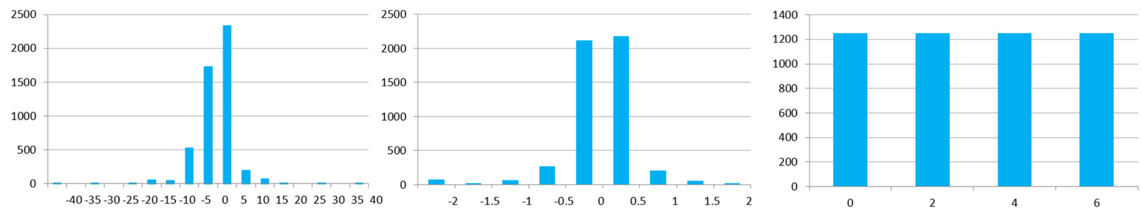
- *Dataset 3: Route choices*

Dataset 3 is also about route choices. It was collected among car drivers in New Zealand. In total, 457 respondents participated in the survey, including 156 local commuters, 153 local non-commuters, and 148 long-distance travellers. Each respondent made sixteen hypothetical choices. In this study, the models are only estimated on two choice segments: commuters and local non-commuters. Choices were made between a reference route and two SP routes which were described by six-time- and cost-related attributes: free-flow travel time (FF), showed-down travel time (SLOW), stopped/crawling time (STOP), contingency time for arrival (CON), running cost (RUN) and toll cost (TOLL). Time-related attributes are expressed in minutes, and cost-related attributes are in New Zealand dollars.

Figure 4.11 presents the frequency of attribute-level differences in the commuter segment. We can see that for the first five attributes, the ranges of attribute-level differences are fairly wide, the most of the differences are concentrated in areas near “zero”. Moreover, they have more positive values, especially the attribute RUN. The last attribute TOLL has non-negative attribute-level differences. This means we cannot test loss aversion on this attribute (as no gains).



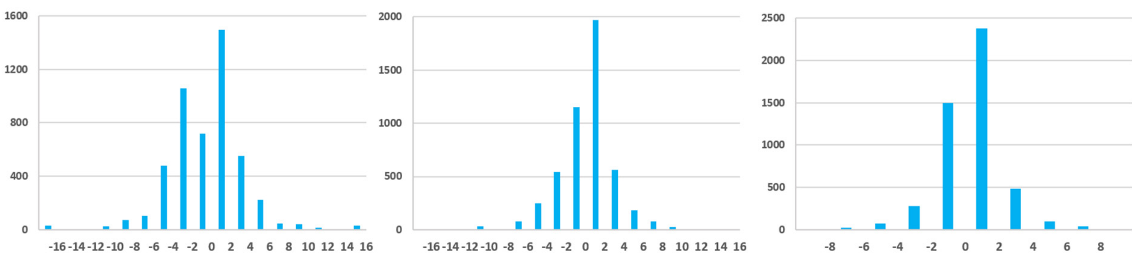
(a) $FF_i - FF_r$ (b) $SLOW_i - SLOW_r$ (c) $STOP_i - STOP_r$



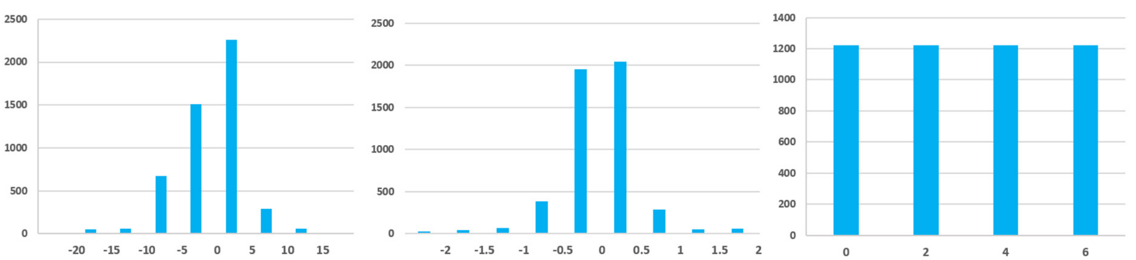
(d) $CON_i - CON_r$ (e) $RUN_i - RUN_r$ (f) $TOLL_i - TOLL_r$

Figure 4.11 Histograms of attribute-level differences in Dataset 3 (Commuters)

Figure 4.12 shows the histograms of attribute-level differences in the non-commuter segment. For each attribute, the histograms for the commuter and non-commuter segments are very similar. Specifically, the attribute TOLL again has no negative values in its attribute-level differences.



(a) $FF_i - FF_r$ (b) $SLOW_i - SLOW_r$ (c) $STOP_i - STOP_r$



(d) $CON_i - CON_r$ (e) $RUN_i - RUN_r$ (f) $TOLL_i - TOLL_r$

Figure 4.12 Histograms of attribute-level differences in Dataset 3 (Non-commuters)

4.4.2 Model estimation results

This section presents model estimation results on the three data sets respectively. The estimated models include the linear-additive RUM model, the piecewise-linear model and the new loss aversion model.

- *Dataset 1: Route choices*

We first briefly look at the estimation result of the linear-additive RUM model. As shown in Table 4.1. Parameters associated with five attributes (β_{FF}, β_{SLOW}, β_{VAR}, β_{RUN} and β_{TOLL}) are all statistically significant and of the expected sign. The constant associated with the reference alternative is positive, but not significantly different from zero. This means people in general have no preference for the reference alternative when all attribute levels are the same across the three alternatives. In other words, no status quo effect is found. The final log-likelihood of the linear-additive RUM model is -4024.

Table 4.1 Estimation results of the linear-additive RUM model (Dataset 1)

	Beta	*t-value (s.e.)*
Constant ASC_{ref}	0.051	*0.80 (0.06)*
Free-flow travel time β_{FF}	-0.057	*-17.48 (0.00)*
Slowed-down travel time β_{SLOW}	-0.074	*-25.64 (0.00)*
Travel time variability β_{VAR}	-0.005	*-2.41 (0.00)*
Running cost β_{RUN}	-0.235	*-12.68 (0.02)*
Toll cost β_{TOLL}	-0.283	*-26.53 (0.01)*
# of observations	4800	
Adjusted Rho-square	0.235	
Null log-likelihood	-5273	
Final log-likelihood	-4024	

We then move to the estimation results of the piecewise-linear model, shown in Table 4.2. We can see that the constant associated with the reference alternative is positive and statistically significant, which is different from the result of the RUM model. For main parameters, parameters associated with increases in FF, SLOW, RUN, and TOLL are negative (as expected) and statistically significant, and parameters associated with decreases in these four attributes are positive and statistically significant. For the attribute VAR, note that there are only a very few increase values in the attribute-level differences (Figure 4.9 c). We can see that the parameter associated with increases in VAR ($\beta_{VAR_{in}}$) is not significantly different from zero, and the parameter associated with decreases ($\beta_{VAR_{de}}$) is positive but only significant at the 88% level. The final log-likelihood of the piecewise-linear model is -4016, outperforming 8 log-likelihood points than the linear-additive RUM model.

The right-hand column of Table 4.2 gives asymptotic t-values, which are used to examine whether there is an asymmetric response to increases and decreases in attributes. For attribute FF, the very low asymptotic t-value means that people's responses to increases and decreases are symmetric. Similar results can be also found in TOLL. For attributes SLOW and RUN, the asymptotic t-values are statistically significant, 2.81 and 3.18 respectively, showing the evidence that an asymmetric response is found in these two attributes. In summary, *i*) no asymmetric responses are found in FF and TOLL, indicating that increases (losses) and decreases (gains) in FF and TOLL are equivalently weighted, and *ii*) an asymmetric response is found in SLOW and RUN, and their decrease parameters are larger than increase parameters (absolute values), indicating that decreases (gains) are weighted more importantly than increases (losses) in attributes SLOW and RUN. Therefore, in this dataset, we do not find loss aversion in any attributes, but we do find a phenomenon that gains are valued more importantly than equivalent losses in attributes SLOW and RUN.

Table 4.2 Estimation results of the piecewise-linear model (Dataset 1)

	Beta	*t-value (s.e.)*	*t-value for diff.*
Constant ASC_{ref}	0.178	*1.99 (0.09)*	-
FF decrease $\beta_{FF_{de}}$	0.057	*10.67 (0.01)*	*0.06*
FF increase $\beta_{FF_{in}}$	-0.058	*-5.00 (0.01)*	
SLOW decrease $\beta_{SLOW_{de}}$	0.082	*16.32 (0.01)*	*2.81*
SLOW increase $\beta_{SLOW_{in}}$	-0.053	*-5.50 (0.01)*	
VAR decrease $\beta_{VAR_{de}}$	0.003	*1.55 (0.00)*	-
VAR increase $\beta_{VAR_{in}}$	0.005	*0.08 (0.07)*	
RUN decrease $\beta_{RUN_{de}}$	0.302	*9.41 (0.03)*	*3.18*
RUN increase $\beta_{RUN_{in}}$	-0.142	*-3.22 (0.04)*	
TOLL decrease $\beta_{TOLL_{de}}$	0.265	*11.72 (0.02)*	*0.88*
TOLL increase $\beta_{TOLL_{in}}$	-0.287	*-19.32 (0.01)*	
# of observations	4800		
Adjusted Rho-square	0.237		
Null log-likelihood	-5273		
Final log-likelihood	-4016		

We now turn our attention to the results of the new loss aversion model in Table 4.3. First, we can see the constant associated with the reference alternative is positive, but not statistically significant. As expected, all main parameters (β) are estimated negative. Besides, these parameters (β_{FF}, β_{SLOW}, β_{RUN} and β_{TOLL}) are statistically significant except β_{VAR}. This is in line with the results of the piecewise-linear model, in which both $\beta_{VAR_{de}}$ and $\beta_{VAR_{in}}$ are not statistically significant.

We now look at the loss aversion parameter μ. Note that the significance test for μ is conducted against the null hypothesis[15]: $\mu = 5$. We can see that parameters μ_{FF}, μ_{SLOW}, μ_{RUN} and μ_{TOLL} are estimated larger than 5, and parameter μ_{VAR} is 1.56, but with a large standard error of 8.83, it is also not significantly smaller than 5. The estimation results of parameters μ indicate that no loss aversion can be found in any attributes, which is in line with the results of the piecewise-linear model. In addition, the phenomenon that gains are valued more importantly than losses, which is found in the piecewise-linear model, cannot be shown in the new loss aversion model. The final log-likelihood of this model is -4022.

[15] Similar to μRRM, when $\mu = 5$, the loss aversion function is nearly linear. We suggest that the significance test is against the null hypothesis $\mu = 5$.

Table 4.3 Estimation results of the new loss aversion model (Dataset 1)

	Beta	***t-value (s.e.)***
Constant ASC_{ref}	0.029	*0.32 (0.09)*
Free flow time β_{FF}	-0.115	*-19.51 (0.01)*
Slowed-down time β_{SLOW}	-0.148	*-28.64 (0.01)*
Travel time variability β_{VAR}	-0.012	*-1.02 (0.01)*
Running cost β_{RUN}	-0.471	*-15.57 (0.03)*
Toll cost β_{TOLL}	-0.551	*26.15 (0.02)*
Loss aversion in FF μ_{FF}	>5	-
Loss aversion in SLOW μ_{SLOW}	>5	-
Loss aversion in VAR μ_{VAR}	1.56	*-0.39 (8.83)*
Loss aversion in RUN μ_{RUN}	>5	-
Loss aversion in TOLL μ_{TOLL}	>5	-
# of observations	4800	
Adjusted Rho-square	0.235	
Null log-likelihood	-5273	
Final log-likelihood	-4022	

- *Dataset 2: Policy package choices (whole sample and subsample)*

Results of the linear-additive RUM model are given in Table 4.4, where the model is estimated on the whole dataset and the subset respectively. In the model of the whole sample, *i*) all parameters are highly significant, and *ii*) as expected, the parameter associated with an average reduction in travel time (β_{TR}) are positive and the parameters associated with different fatalities are negative. In the model of the subset, *i*) parameter β_{TR} is positive but not significantly different from zero, and *ii*) parameters associated with fatalities are all significant and of the expected sign. The final log-likelihoods of these two models are -1274 and -771 respectively.

Table 4.4 Estimation results of the linear-additive RUM model (Dataset 2)

	Whole sample		Subsample	
	Beta	*t-value (s.e.)*	Beta	*t-value (s.e.)*
Average reduction in car travel time β_{TR}	0.016	*4.07 (0.004)*	0.005	*-1.10 (0.005)*
Conventional car fatalities β_{CF}	-0.008	*-14.09 (0.001)*	-0.008	*-10.46 (0.001)*
AV technical failure fatalities β_{AVT}	-0.009	*-9.60 (0.001)*	-0.005	*-4.22 (0.001)*
AV deliberate misuse fatalities β_{AVM}	-0.014	*-9.55 (0.001)*	-0.008	*-4.43 (0.002)*
# of observations	1290		762	
Adjusted Rho-square	0.098		0.079	
Null log-likelihood	-1417		-837	
Final log-likelihood	-1274		-771	

Table 4.5 presents the results of the piecewise-linear models for both the whole dataset and subset. Consistent with intuition, increases in travel time reduction are estimated to be positive and decreases in it are valued negative, and increases in the number of fatalities are estimated to be negative, and decreases in the number of fatalities are estimated to be positive. For TR, both parameters $\beta_{TR_{in}}$ and $\beta_{TR_{de}}$ are statistically significant in the model of the whole sample, while in the model of subsample, only the decrease parameter $\beta_{TR_{de}}$ is significant. Moreover,

the asymptotic t-values are 2.37 and 2.89, indicating that the whole sample and subsample have asymmetrical responses to increases and decreases, specifically, decreases in the average travel time reduction cause a greater response than increases. For CF, parameters $\beta_{CF_{de}}$ and $\beta_{CF_{in}}$ are highly significant in two models. The asymptotic t-value of 2.59 shows the evidence that asymmetrical responses exist in the whole sample: increases in conventional car fatalities incur a greater response than decreases. For AVT, parameters $\beta_{AVT_{de}}$ and $\beta_{AVT_{in}}$ are highly significant in both models, and the low asymptotic t-values of 0.05 and 0.82 show that responses to increases and decreases are almost symmetric. For AVM, both parameters $\beta_{AVM_{de}}$ and $\beta_{AVM_{in}}$ are highly significant for the whole sample, but the difference between these two parameters is not significant (asymptotic t-value = 0.94). In the model of subsample, parameter $\beta_{AVM_{in}}$ is highly significant, but parameter $\beta_{AVM_{de}}$ is only significant at the 90% level. Moreover, with the asymptotic t-value of 3.67, there is clear evidence for an asymmetric response for the subsample; specifically, increases in AVM fatalities cause a greater response than decreases.

In sum, there are asymmetries in response to decreases and increases in this dataset. Specifically, decreases in the travel time reduction (losses) cause a greater response than increases (gains) for both the whole sample and the subsample; increases in the number of conventional car fatalities (losses) cause a greater response than decreases (gains) for the whole sample; and increases in the number of AV deliberate misuse fatalities (loss aversion) cause a greater response than decreases (gains) for the subsample. Therefore, loss aversion is found in the travel time reduction in the whole sample and subsample, in the number of conventional fatalities in the whole sample, in the number of AV deliberate misuse fatalities in the subsample. The log-likelihoods of the two models are -1268 and -760 respectively, with 6 and 10 log-likelihood points improvement respectively.

Table 4.5 Estimation results of the piecewise-linear model (Dataset 2)

	Whole sample			**Subsample**		
	Beta	***t-value (s.e.)***	**t-value for diff.**	**Beta**	***t-value (s.e.)***	**t-value for diff.**
TR increase $\beta_{TR_{in}}$	0.012	*2.45 (0.01)*	2.37	0.003	*0.41 (0.01)*	2.89
TR decrease $\beta_{TR_{de}}$	-0.039	*-3.79 (0.01)*		-0.043	*-3.45 (0.01)*	
CF decrease $\beta_{CF_{de}}$	0.007	*7.58 (0.00)*	2.59	0.007	*6.28 (0.00)*	1.14
CF increase $\beta_{CF_{in}}$	-0.010	*-10.5 (0.00)*		-0.009	*-7.54 (0.00)*	
AVT decrease $\beta_{AVT_{de}}$	0.010	*7.22 (0.00)*	0.05	0.006	*3.51 (0.00)*	0.82
AVT increase $\beta_{AVT_{in}}$	-0.010	*-6.92 (0.00)*		-0.008	*-4.17 (0.00)*	
AVM decrease $\beta_{AVM_{de}}$	0.014	*-7.55 (0.00)*	0.94	0.004	*1.80 (0.00)*	3.67
AVM increase $\beta_{AVM_{in}}$	-0.017	*-6.01 (0.00)*		-0.020	*-5.41 (0.00)*	
# of observations	1290			762		
Adjusted Rho-square	0.100			0.083		
Null log-likelihood	-1417			-837		
Final log-likelihood	-1268			-760		

We now look at the results of the new loss aversion model in Table 4.6. Main parameters β_{TR}, β_{CF}, β_{AVT} and β_{AVM} are statistically significant and of the expected sign in both models. Note that in RUM, β_{TR} is not statistically significant for the subsample, and in the piecewise-linear model only $\beta_{TR_{de}}$ is significant for the subsample. However, in the loss aversion model of the subsample, β_{TR} (attribute importance) is significant.

Again, the significance test for μ is against the null hypothesis of $\mu = 5$. For the whole sample, μ_{TR} is estimated to be small, suggesting a strong loss aversion in TR, while for the subsample, μ_{TR} is even smaller, suggesting that the degree of loss aversion is even greater for the subsample. In the model of the whole sample, parameter μ_{CF} is 2.81, but it is only significant at the 90% level, and in the model of the subsample, parameter μ_{CF} is estimated to be larger than 5. Both suggest that there is no loss aversion. For the attribute AVT, parameter μ_{AVT} is estimated to be larger than 5 for the whole sample, and for the subsample, parameter μ_{AVT} is 3.48, but not significantly different from 5. Again, no loss aversion is found in AVT for either the whole sample or subsample. For the attribute AVM, parameter μ_{AVM} is estimated larger than 5 for the whole sample, indicating there is no loss aversion. For the subsample, it is significantly small (0.561), suggesting a strong loss aversion is found in this attribute.

In brief, a strong loss aversion is found in the attribute TR for both the whole sample and the subsample, and also in the attribute AVM for the subsample. The final log-likelihoods of the new loss aversion models are -1268 and -760 respectively, which are the same as those of the piecewise-linear models.

Table 4.6 Estimation results of the new loss aversion model (Dataset 2)

	Whole sample		**Subsample**	
	Beta	***t-value (s.e.)***	**Beta**	***t-value (s.e.)***
Average reduction in car travel time β_{TR}	0.052	*4.77 (0.01)*	0.047	*3.63 (0.01)*
Conventional car fatalities β_{CF}	-0.017	*-14.33 (0.00)*	-0.016	*-10.93 (0.00)*
AV technical failure fatalities β_{AVT}	-0.020	*-10.21 (0.00)*	-0.014	*-6.17 (0.00)*
AV deliberate misuse fatalities β_{AVM}	-0.030	*-2.58 (0.02)*	-0.024	*-6.31 (0.00)*
Loss aversion in TR μ_{TR}	0.514	*-16.49 (0.27)*	0.121	*-6.46 (0.11)*
Loss aversion in CF μ_{CF}	2.81	*-1.77 (1.24)*	>5	-
Loss aversion in AVT μ_{AVT}	>5	-	3.48	*-0.33 (4.63)*
Loss aversion in AVM μ_{AVM}	>5	-	0.561	*-14.95 (0.30)*
# of observations	1290		762	
Adjusted Rho-square	0.100		0.083	
Null log-likelihood	-1417		-837	
Final log-likelihood	-1268		-760	

Compared with the piecewise-linear model, the new loss aversion model not only shows the existence of loss aversion, but also shows its degree. It is worth noting that there are some inconsistent findings between these two models. Specifically, the piecewise-linear model finds loss aversion in conventional car fatalities for the whole sample, while the new loss aversion model finds no loss aversion in this attribute for the whole sample. The reason is that the two models adopt different criteria for examining loss aversion: one adopts an asymptotic t-test for testing the significance of the difference between increase and decrease parameters, and another adopts a t-test for testing the significance of the loss aversion parameter μ. Inspecting the parameters of the piecewise-linear model, $\beta_{CF_{de}}$ and $\beta_{CF_{in}}$, again, we can find that $\beta_{CF_{in}}$ is only slightly larger than $\beta_{CF_{de}}$ (absolute values, 0.010 vs. 0.007). This corresponds to the estimate of parameter μ_{CF} in the loss aversion model: μ_{CF} is 2.81, suggesting that the loss aversion function is slightly curved. This, in turn, indicates the consistency of the estimation results between the two models.

- Dataset 3: *Route* choices (commuters and non-commuters)

The estimation results of the linear-additive RUM models are shown in Table 4.7, in which models are estimated for commuters and non-commuters respectively. In both models, the constants associated with the reference alternative are positive and significantly different from zero, suggesting there is a general preference for the reference alternative in both groups. Parameter β_{FF} is negative, but not statistically significant at the 95% level in both models. The remaining parameters β_{SLOW}, β_{STOP}, β_{CON}, β_{RUN}, and β_{TOLL} are significant and of the expected sign. The final log-likelihoods for commuter and non-commuter segments are -1602 and -1504, respectively.

Table 4.7 Estimation results of the linear-additive RUM model (Dataset 3)

	Commuters		**Non-commuters**	
	Beta	***t-value (s.e.)***	**Beta**	***t-value (s.e.)***
Constant ASC_{ref}	0.475	*7.77 (0.06)*	0.555	*8.60 (0.06)*
Free-flow travel time β_{FF}	-0.003	*-0.21 (0.02)*	-0.025	*-1.85 (0.01)*
Slowed-down travel time β_{SLOW}	-0.074	*-6.33 (0.01)*	-0.058	*-3.20 (0.02)*
Stopped/crawling time β_{STOP}	-0.170	*-10.12 (0.02)*	-0.143	*-5.95 (0.02)*
Contingency time β_{CON}	-0.026	*-2.64 (0.01)*	-0.044	*-4.08 (0.01)*
Running cost β_{RUN}	-0.294	*-2.98 (0.10)*	-0.849	*-7.34 (0.12)*
Toll cost β_{TOLL}	-0.543	*-18.06 (0.03)*	-0.532	*-17.38 (0.03)*
# of observations	2496		2448	
Adjusted Rho-square	0.413		0.438	
Null log-likelihood	-2742		-2689	
Final log-likelihood	-1601		-1504	

We turn to the results of the piecewise-linear models in Table 4.8, in which separate parameters are estimated for increases and decreases in attribute levels. In both models, the constant associated with the reference is positive and significant, indicating that both commuters and non-commuters have a general preference for the reference route. This is consistent with the results of the RUM models. For main parameters associated with increases and decreases, we see that several parameters' sign is counter-intuitive, e.g., parameters $\beta_{FF_{in}}$, $\beta_{CON_{de}}$, and $\beta_{RUN_{in}}$ in the non-commuter segment and $\beta_{FF_{de}}$ and $\beta_{SLOW_{in}}$ in the commuter segment. This is however not of much concern as these parameters are not significantly different from zero.

For the attribute FF, both parameters $\beta_{FF_{de}}$ and $\beta_{FF_{in}}$ are not significantly different from zero in the commuter segment. In the non-commuter segment, the sign of $\beta_{FF_{de}}$ is counterintuitive, but it is not significant, while $\beta_{FF_{in}}$ is significant and of the expected sign. This means for non-commuters, increases in FF have an impact on route choices, but decreases in this attribute do not play a role. Thus, there is an asymmetric response in free-flow travel time of non-commuters. In the model of commuters, parameters $\beta_{SLOW_{de}}$ and $\beta_{SLOW_{in}}$ are significant and of the expected sign, and its asymptotic t-value shows that the difference between these two parameters is significant at the 90% level. For non-commuters, parameter $\beta_{SLOW_{de}}$ is positive and significant, but parameter $\beta_{SLOW_{in}}$ is not significantly different from zero. This means that only decreases in slowed-down travel time influence non-commuters' route choices. Parameters $\beta_{STOP_{de}}$ and $\beta_{STOP_{in}}$ are significant and of the expected sign in both segments, but the low asymptotical t-values (1.43 and 0.34) show that responses to increases and decreases in stopped or crawling time are not different in both groups. For attribute CON, only parameter $\beta_{CON_{in}}$ is significant in the commuter segment, and only $\beta_{CON_{de}}$ is significant in the non-commuter segment. This implies that increases in contingency time have an influence on commuters' route choices, and decreases in contingency time have an influence on non-commuters' choices. For attribute RUN, only $\beta_{RUN_{de}}$ is significantly different from zero in the model of commuters, suggesting that commuters are only sensitive to the decreases in running cost. In the model of non-commuters, both parameters $\beta_{RUN_{de}}$and $\beta_{RUN_{in}}$are significant and of the expected sign, and the low asymptotic t-value (0.13) suggests a symmetric response to increases and decreases is in running cost for non-commuters. For the last attribute TOLL, as there is no decrease value, only parameter $\beta_{TOLL_{in}}$ is estimated. In both models, parameter $\beta_{TOLL_{in}}$ is negative and highly significant.

In sum, asymmetric responses are found in attributes CON and RUN for commuters and in attributes FF, SLOW, and CON for non-commuters. Specifically, commuters show loss aversion to contingency time for arrival, while they exhibit an opposite behaviour to running cost, that is gains are overweighted than losses. Non-commuters show loss aversion to free-flow travel time, while they show the gain-overweighting behaviour to slowed-down travel time and contingency time.

Table 4.8 Estimation results of the piecewise-linear model (Dataset 3)

	Commuters			**Non-commuters**		
	Beta	***t-value (s.e.)***	**t-value for diff.**	**Beta**	***t-value (s.e.)***	**t-value for diff.**
Constant ASC_{ref}	0.378	*3.97 (0.10)*	-	0.619	*5.46 (0.11)*	-
FF decrease $\beta_{FF_{de}}$	0.044	*1.55 (0.03)*	-	-0.032	*-1.02 (0.03)*	-
FF increase $\beta_{FF_{in}}$	0.050	*1.82 (0.03)*		-0.095	*-2.75 (0.03)*	
SLOW decrease $\beta_{SLOW_{de}}$	0.061	*3.28 (0.02)*	1.88	0.120	*-4.23 (0.03)*	-
SLOW increase $\beta_{SLOW_{in}}$	-0.141	*-3.70 (0.04)*		0.016	*0.42 (0.04)*	
STOP decrease $\beta_{STOP_{de}}$	0.155	*6.24 (0.02)*	1.43	0.136	*-2.93 (0.05)*	0.34
STOP increase $\beta_{STOP_{in}}$	-0.228	*-5.11 (0.04)*		-0.114	*-2.51 (0.05)*	
CON decrease $\beta_{CON_{de}}$	-0.018	*-1.10 (0.02)*	-	0.061	*3.20 (0.02)*	-
CON increase $\beta_{CON_{in}}$	-0.079	*-3.71 (0.02)*		-0.021	*-0.95 (0.02)*	
RUN decrease $\beta_{RUN_{de}}$	0.503	*-3.81 (0.13)*	-	0.884	*4.09 (0.22)*	0.13
RUN increase $\beta_{RUN_{in}}$	0.013	*0.08 (0.16)*		-0.937	*-2.84 (0.33)*	
TOLL decrease $\beta_{TOLL_{de}}$	-		-	-		-
TOLL increase $\beta_{TOLL_{in}}$	-0.549	*-18.02 (0.03)*		-0.535	*-17.46 (0.03)*	
# of observations	2496			2448		
Adjusted Rho-square	0.422			0.441		
Null log-likelihood	-2742			-2689		
Final log-likelihood	-1572			-1497		

We now turn our attention to the results of the new loss aversion model in Table 4.9. The constants associated with the reference alternative are highly significant and of the expected sign in both models. This is consistent with the results of the RUM models and the piecewise-linear models. Parameter β_{FF} is not significantly different from zero in the commuter segment, while in the non-commuter segment, β_{FF} is negative and statistically different from zero. Other main parameters β_{SLOW}, β_{STOP}, β_{CON}, β_{RUN}, and β_{STOP} are all significant and of the expected sign in both models.

We then turn to the estimates of parameter μ. In the model of commuters, first, μ_{con} is very small (0.502) and highly significant, suggesting a strong loss aversion is found in contingency time; parameter β_{SLOW} is 1.93, and it is significantly different from 5 (null hypothesis), suggesting a moderate loss aversion is found in slow-down travel time for commuters; parameter μ_{STOP} is 3.68 but not significantly different from 5; and parameters μ_{FF} and μ_{RUN} are both estimated larger than 5. In the model of non-commuters, μ_{FF} is estimated to be small (0.605), implying there is a strong loss aversion in free-flow travel time; parameter μ_{RUN} is 3.28, but not significantly different from 5; the remaining three parameters μ_{SLOW}, μ_{STOP} and μ_{CON} are all larger than 5, suggesting no loss aversion can be found in these three attributes.

Table 4.9 Estimation results of the new loss aversion model (Dataset 3)

	Commuters		**Non-commuters**	
	Beta	***t-value (s.e.)***	**Beta**	***t-value (s.e.)***
Constant (ASC)	0.400	*6.18 (0.06)*	0.524	*7.91 (0.07)*
Free-flow time (β_{FF})	0.006	*0.25 (0.02)*	-0.074	*-2.75 (0.03)*
Slowed-down time (β_{SLOW})	-0.194	*-5.49 (0.04)*	-0.114	*-3.49 (0.03)*
Stopped/crawling time (β_{STOP})	-0.377	*-8.87 (0.04)*	-0.249	*-5.27 (0.05)*
Contingency time (β_{CON})	-0.061	*-2.58 (0.02)*	-0.088	*-4.09 (0.02)*
Running cost (β_{RUN})	-0.470	*-2.48 (0.19)*	-2.340	*-8.88 (0.26)*
Toll cost (β_{TOLL})	-0.539	*-18.25 (0.03)*	-0.532	*-19.93 (0.03)*
Loss aversion in FF (μ_{FF})	>5	-	0.605	*-6.46 (0.68)*
Loss aversion in SLOW (μ_{SLOW})	1.93	*-4.15 (0.74)*	>5	-
Loss aversion in STOP (μ_{STOP})	3.68	*-0.70 (1.89)*	>5	-
Loss aversion in CON (μ_{CON})	0.502	*-10.00 (0.45)*	>5	-
Loss aversion in RUN (μ_{RUN})	>5	-	3.28	*-1.64 (1.05)*
# of observations	2496		2448	
Adjusted Rho-square	0.413		0.441	
Null log-likelihood	-2742		-2689	
Final log-likelihood	-1598		-1495	

The new loss aversion model is basically consistent with the piecewise-linear model in capturing loss aversion. Specifically, according to the results in Table 4.9, a strong loss aversion can be found in commuters' contingency time for arrival and in non-commuters' free-flow travel time. In addition, a moderate loss aversion is found in slowed-down travel time of commuters. This seems to be against the finding of the piecewise-linear model. However, inspecting parameters $\beta_{SLOW_{in}}$ and $\beta_{SLOW_{de}}$, we can see the difference between these two parameters is significant at the 90% level (t-value=1.88). As discussed above, the subtle differences between the two models are simply due to the different criteria that the two models adopt.

Table 4.10 summarizes the comparison between models in terms of model fits. Several observations can be made. At the first glance, the new loss aversion model and the piecewise-linear model seem to perform better than the RUM model on the five datasets (including two subsets). After closer inspecting final log-likelihoods, we see that there is no substantial difference in model fits between the loss aversion model and the RUM model, when no loss aversion occurs in behaviour (Dataset 1). Moreover, when there is only loss-averse behaviour, the loss aversion model performs as well as the piecewise-linear model (Dataset 2). But when the opposite behaviour pattern—gains are valued more importantly than losses—also presents in behaviour, the new loss aversion model is outperformed by the piecewise-linear model (Datasets 1 and 3).

Table 4.10 Model fit comparison

	# of choice observations	Final LL			
		Null-LL	RUM	Piecewise-linear	New LA
Dataset 1	4800	-5237	-4024	-4016	-4022
Dataset 2: whole set	1290	-1417	-1274	-1268	-1268
Dataset 2: subset	762	-837	-771	-760	-760
Dataset 3: commuters	2496	-2742	-1601	-1572	-1598
Dataset 3: non-commuters	2448	-2689	-1504	-1497	-1495

4.5 Conclusions

Loss aversion, as an important element of Prospective Theory, has been widely applied in travel behaviour research. To model loss aversion, most travel behaviour research either directly applies the value function of Prospect Theory or adopts its variants in the way that losses and gains are separately modelled in a piecewise-linear/nonlinear manner. This chapter puts forward a new way of modelling loss aversion. Specifically, the new loss aversion function is adapted from the regret function of the μRRM model. It thus inherits several properties of the μRRM model. First, the new loss aversion function is smooth and twice-differentiable in the full domain, whereas the value function and its variants are not smooth at the reference point. Second, the new loss aversion model is mathematically tractable; it can be easily estimated and adopted without the need for a computer-intensive estimation process. Third, in the loss aversion model, there are separate parameters that govern the degree of loss aversion and attribute importance respectively, as a result, the two effects can be disentangled from each other.

Compared to the existing loss aversion models, the new loss aversion model lacks flexibility in two aspects. First, due to the convexity of the loss aversion function, the new model fails to capture the situation in which gains are evaluated more importantly than equivalent losses. In contrast, the value function and its variants are more flexible in capturing a wider range of behaviour. Second, the convexity of the loss aversion function implicitly constrains the assumption of risk attitude towards losses and gains. The loss aversion model assumes that people have a risk-averse attitude to both losses and gains, which may violate people's true risk attitudes.

Our empirical analyses illustrate these methodological properties of the new loss aversion model. The new loss aversion model successfully captures loss aversion in behaviour. Moreover, the degree of loss aversion in certain attributes is straightforwardly reflected by the magnitude of loss aversion. In terms of model fit, the new loss aversion model performs as well as the

RUM model when there is no loss aversion in behaviour; it fits the data just as well as the piecewise-linear model when only loss aversion presents in behaviour; but it is outperformed by the piecewise-linear model when an opposite behavioural pattern—gains are valued more importantly than losses—exists in behaviour.

The findings of this chapter provide several avenues for future research. First, the estimation results of Dataset 2 show that loss aversion (for the number of fatalities caused by AV deliberate misuse) is observed only in the subsample, but not in the whole sample. This means there is heterogeneity in their choice behaviour in terms of loss aversion. An interesting research direction is to explore different latent classes by assigning decision makers to the RUM and new loss aversion model classes, depending on their characteristics. Second, although ample evidence has supported loss aversion, loss aversion may not exist in certain sorts of choices. Thus, another interesting research direction is to explore under what circumstances will trigger people to show loss aversion and under what circumstances will trigger people to show the opposite behaviour, especially in the travel choice context. The last but not least, in this study, the new loss aversion model is only compared empirically with the piecewise-linear model, other loss aversion models, like the piecewise-nonlinear model, need to be involved in model comparison.

References

Ahtiainen, H., Pouta, E., & Artell, J. (2015). Modelling asymmetric preferences for water quality in choice experiments with individual-specific status quo alternatives. *Water Resources Economics, 12*, 1-13.

Avineri, E. (2009). Incorporating fuzzy reference points into applications of travel choice modelling. In *Applications of Soft Computing* (pp. 221-229): Springer.

Avineri, E., & Ben-Elia, E. (2015). Prospect theory and its applications to the modelling of travel choice. In *Bounded Rational Choice Behaviour: Applications in Transport*: Emerald Group Publishing Limited.

Avineri, E., & Bovy, P. H. (2008). Identification of parameters for a prospect theory model for travel choice analysis. *Transportation Research Record, 2082*(1), 141-147.

Bahamonde-Birke, F. J. (2018). Estimating the reference frame: A smooth twice-differentiable utility function for non-compensatory loss-averse decision-making. *Journal of choice modelling, 28*, 71-81.

Balbontin, C., Hensher, D. A., & Collins, A. T. (2017). Integrating attribute non-attendance and value learning with risk attitudes and perceptual conditioning. *Journal of choice modelling, 97*, 172-191.

Bartczak, A., Chilton, S., Czajkowski, M., & Meyerhoff, J. (2017). Gain and loss of money in a choice experiment. The impact of financial loss aversion and risk preferences on willingness to pay to avoid renewable energy externalities. *Energy Economics, 65*, 326-334.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. Cambridge: MIT Press.

Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The quarterly journal of economics, 110*(1), 73-92.

Chorus, C. (2012). Random regret minimization: an overview of model properties and empirical evidence. *Transport Reviews, 32*(1), 75-92.

Chorus, C. G. (2010). A new model of random regret minimization. *European Journal of Transport Infrastructure Research, 10*(2).

Coursey, D. L., Hovis, J. L., & Schulze, W. D. (1987). The disparity between willingness to accept and willingness to pay measures of value. *The quarterly journal of economics, 102*(3), 679-690.

Daly, A. (2010). Cost damping in travel demand models: Report of a study for the department for transport.

Daly, A., & Carrasco, J. (2009). The influence of trip length on marginal time and money values. *The Expanding Sphere of Travel Behaviour Research*.

De Borger, B., & 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*(1), 101-115.

Dekker, T. (2014). Indifference based value of time measures for Random Regret Minimisation models. *Journal of choice modelling, 12*, 10-20.

Dekker, T., & Chorus, C. G. (2018). Consumer surplus for random regret minimisation models. *Journal of Environmental Economics, 7*(3), 269-286.

Dholakia, U. M., & Simonson, I. (2005). The effect of explicit reference points on consumer choice and online bidding behavior. *Marketing Science, 24*(2), 206-217.

Flügel, S., Elvik, R., Veisten, K., Rizzi, L. I., Meyer, S. F., Ramjerdi, F., & de Dios Ortúzar, J. (2015). Asymmetric preferences for road safety: Evidence from a stated choice experiment among car drivers. *Transportation research part F: traffic psychology behaviour, 31*, 112-123.

Hess, S., Rose, J. M., & Hensher, D. A. (2008). Asymmetric preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E: Logistics Transportation Review, 44*(5), 847-863.

Jou, R.-C., & Kitamura, R. (2002). *Commuter departure time choice: a reference point approach.* Paper presented at the 9th Meeting of the Euro Working Group on Transportation.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives, 5*(1), 193-206.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Kivetz, R., Netzer, O., & Srinivasan, V. (2004). Alternative models for capturing the compromise effect. *Journal of Marketing Research, 41*(3), 237-257.

Koppelman, F. S. (1981). Non-linear utility functions in models of travel choice behavior. *Transportation, 10*(2), 127-146.

Lanz, B., Provins, A., Bateman, I. J., Scarpa, R., Willis, K., & Ozdemiroglu, E. (2010). Investigating willingness to pay–willingness to accept asymmetry in choice experiments. In *Choice Modelling: The State-of-the-art and The State-of-practice*: Emerald Group Publishing Limited.

Masiero, L., & Hensher, D. A. (2010). Analyzing loss aversion and diminishing sensitivity in a freight transport stated choice experiment. *Transportation Research Part A: Policy Practice, 44*(5), 349-358.

Mayhew, G. E., & Winer, R. S. (1992). An empirical analysis of internal and external reference prices using scanner data. *Journal of consumer research, 19*(1), 62-70.

McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior*. New York: Academic Press.

Neuman, E., & Neuman, S. (2008). Reference-dependent preferences and loss aversion: A discrete choice experiment in the health-care sector. *Judgment Decision Making, 3*(2), 162-173.

Novemsky, N., & Kahneman, D. (2005). The boundaries of loss aversion. *Journal of Marketing research, 42*(2), 119-128.

Rose, J. M., & Masiero, L. (2010). A comparison of the impacts of aspects of prospect theory on WTP/WTAEstimated in preference and WTP/WTA space. *European Journal of Transport and Infrastructure Research, 10*(4).

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of risk uncertainty, 1*(1), 7-59.

Senbil, M., & Kitamura, R. (2004). *Reference points in commuter departure time choice: a prospect theoretic test of alternative decision frames.* Paper presented at the Intelligent Transportation Systems.

Stathopoulos, A., & Hess, S. (2012). Revisiting reference point formation, gains–losses asymmetry and non-linear sensitivities with an emphasis on attribute specific treatment. *Transportation Research Part A: Policy Practice, 46*(10), 1673-1689.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of economic behavior organization, 1*(1), 39-60.

Timmermans, H. (2010). On the (ir) relevance of prospect theory in modelling uncertainty in travel decisions. *European Journal of Transport and Infrastructure Research, 10*(4).

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics, 106*(4), 1039-1061.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of risk uncertainty, 5*(4), 297-323.

van Cranenburgh, S., Guevara, C. A., & Chorus, C. G. (2015). New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice, 74*, 91-109.

Wang, P., Sun, L., Niraj, R., Chung, J., & Su, M. (2019). Incorporating reference products into modeling consumer choice decision: A mixtures-of-experts model. *Decision Support Systems, 119*, 85-95.

Weber, M., & Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis. *Journal of economic behavior organization, 33*(2), 167-184.

Zhang, J., Timmermans, H., Borgers, A., & Wang, D. (2004). Modeling traveler choice behavior using the concepts of relative utility and relative interest. *Transportation Research Part B: Methodological, 38*(3), 215-234.

5 Optimal experimental designs for discriminating between choice models

Abstract

Experimental designs optimized for model discrimination have been widely used in many fields, such as geoscience, hydrology, and epidemiology. When compared with the wealth of literature in the above-mentioned fields, experimental designs for model discrimination have not attracted much attention in the choice modelling community. This chapter introduces an innovative method of constructing model discrimination designs optimised for model discrimination, called discriminatory designs. In contrast to efficient designs, which aim to identify model parameters in a statistically efficient way, discriminatory designs aim to discriminate between different choice models, yielding the result that the true data generating process (DGP) (given collected data) can be distinguished from a set of competing models. To construct discriminatory designs, we adopt Bayes' theorem and information theory. The robustness of discriminatory designs is tested on synthetic data and empirical data. The tests show that an elaborate discriminatory design can discriminate between models. Specifically, for synthetic data, our method can generate larger model fit differences, yielding a result that clearly favours the true DGP, and for empirical data, if there is the true DGP given data (among all competing models), using discriminatory designs can maximize the likelihood of discriminating it from other competing models.

5.1 Introduction

Stated choice experiments are wildly used to collect data on choice behaviour. Generating such experiments relies on the so-called experimental designs, which are the process of assigning attribute levels to choice sets in stated choice experiments. Currently, in the field of choice modelling, the mainstream experimental designs are what are known as efficient designs (Kanninen, 2002; Rose & Bliemer, 2009; Bliemer & Rose, 2011). These designs aim to generate experiments that can maximize the collected information (data) about model parameters, thereby yielding reliable parameter estimates for a given model (Rose et al., 2008; Bliemer & Rose, 2010; van Cranenburgh et al., 2018).

Recent years have seen increasing interest in introducing new model specifications into the choice modelling field, such as Random Regret Minimization models (Chorus, 2010), Decision Field Theory (Busemeyer & Townsend, 1993; Hancock et al., 2018) and Quantum choice models (Yu & Jayakrishnan, 2018; Hancock et al., 2020), etc. To test model performance, new models are often compared with a series of conventional models. Typically, such model comparison is conducted by estimating all competing models on the same data and then comparing their model fits. However, the data used for model comparison are usually generated from efficient designs that are optimized to efficiently recover model parameters, rather than to select the model that best describes the underlying choice behaviour from a set of competing models. Therefore, there is a mismatch between what stated choice experiments are typically optimized for—i.e., the reliable measurement for model parameters for a particular model—and what they are used for—i.e., the identification of the most plausible model from a set of competing models. This indicates the need for introducing experimental designs tailored for model discrimination into the field of choice modelling.

The idea of model discrimination can be traced back to Hunter & Reiner (1965). They defined model discrimination as the likelihood ratio in the case of two competing models. Another seminal work was done by Box & Hill (1967). They took a different approach: model discrimination was achieved by maximizing the expected information (entropy) change before and after data collection. In recent years, an increasing number of studies have taken model discrimination as the objective of experimental designs, especially in the fields of hydrology (Schöniger et al., 2014; Kikuchi et al., 2015; Nowak & Guthke, 2016; Pham & Tsai, 2016), psychology and cognitive science (Myung & Pitt, 2009; Cavagnaro et al., 2010), chemical engineering (Alberton et al., 2011), as well as epidemiology (Dehideniya et al., 2018). Most of these studies applied or extended the approach of Box & Hill (1967): the worth of experiments to achieve model discrimination is quantified using information theory.

Inspired by the above literature on model discrimination, this study puts forward a method of constructing experimental designs that are optimized for discriminating between choice models. We call such designs *discriminatory designs*. The aim of discriminatory designs is to generate experiments that lead to the data with the maximum model discrimination capability, thereby yielding the result that clearly favours the most plausible model over other competing models based on a given, limited number of choice observations. Specifically, we adopt Bayes' theorem and information theory. To evaluate the model discrimination capability of an experimental design, we first use Bayes' rule to calculate the expected probability of each competing model and then adopt the Kullback-Leiber (KL) divergence (Kullback & Leibler, 1951) to quantify expected changes in model probabilities before and after data collection. Finally, we define optimal discriminatory designs as the design consisting of the choice sets with the maximum KL divergence.

The remainder of this chapter is organised as follows. The next section presents a conceptual framework of how to construct discriminatory designs. Section 5.3 outlines all mathematical equations and derivations used for design construction. Sections 5.4 and 5.5 test the robustness of discriminatory designs using synthetic data and empirical data, respectively. The final section summarizes findings and discusses a possible avenue for future research.

5.2 Constructing discriminatory designs: Framework

This section presents a conceptual framework of how to construct discriminatory designs. Before diving into the framework, we first outline basic ideas behind the construction of these designs. Models are seen as the hypotheses of the true underlying data generating process (DGP). In general, several candidate models are all plausible to be the true DGP. To select the most plausible model, each candidate model is tested against the observed data and their goodness of fit is then compared. However, analysts are often confronted with a situation in which several competing models have very similar model fits, which makes it difficult to identify the true underlying DGP from the set of models. To tackle this issue, discriminatory designs come into play. The aim of discriminatory designs is to generate experiments (i.e., stated choice sets) that lead to the data with maximum model discrimination capability, yielding a result that the true DGP can be greatly discriminated from a set of competing models. Or in other words, the aim of discriminatory designs is to generate experiments that can maximize the collected information (data) about the true underlying DGP.

This study makes use of *model probabilities* (likelihoods) to construct discriminatory designs. The model probability is defined as the likelihood of each competing model to be the true DGP (given the data). Before collecting data, analysts usually have several model hypotheses, and each of these model hypotheses has the same model probability. Once data are collected, model probabilities will update to different ones based on the collected information. Figure 5.1(a) depicts an example of three competing models. It shows that model probabilities update from the same to different ones after data collection. The collected data contain information about the true DGP, therefore they cause changes in model probabilities. The extent to which data cause changes in model probabilities is determined by the design used during data collection. To illustrate this, Figure 5.1(b) presents three possible outcomes of model probability changes brought about by three designs. Suppose model M_1 is the true DGP given the data. We can see that, in this example, Design 3 is the best compared with the other two. This design leads to a result that the probability of model M_1 to be the true DGP is very large, 0.9, and the probabilities of the other two models are very small. Such a result makes it easier to distinguish the true DGP (model M_1) from the set of models, compared with the results caused by the other designs.

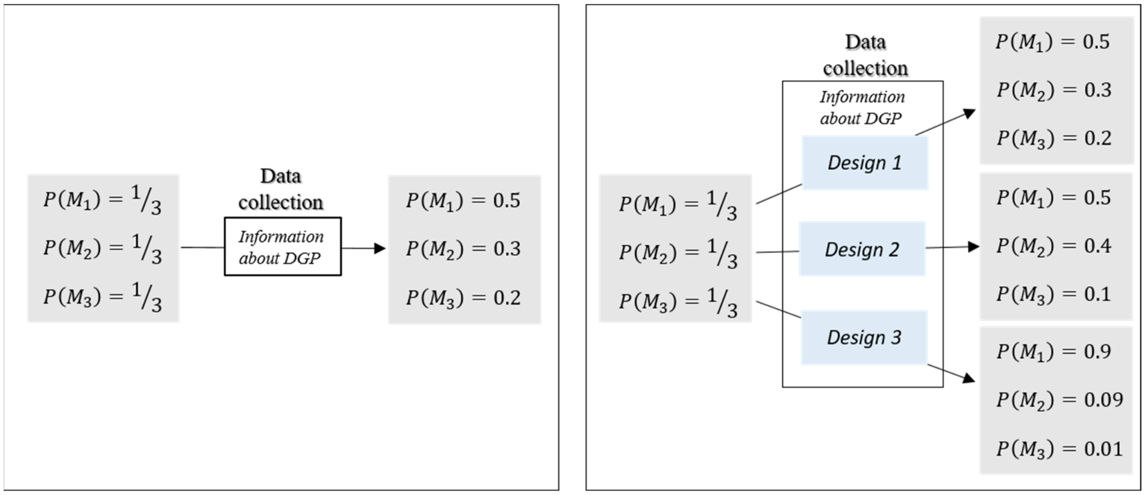


Figure 5.1(a) An example of three competing models. Model M_1 has the largest model probability, meaning that this model is more likely to be the true DGP than the other two according to the collected data. **(b) Three possible outcomes of the model probabilities caused by three designs.** When comparing Design 1 and Design 2, although they result in the same maximum model probability of 0.5, the model probabilities caused by Design 1 are more separate from each other than those caused by Design 2. Design 3 results in a maximum probability of 0.9 and three probabilities that significantly differ from each other. In this case, Design 3 is the best design, followed by Design 1, and Design 2 probably leads to a failure in model discrimination because two of the three probabilities are too close (0.5 vs 0.4).

Model probabilities after data collection represent the likelihood of the model in light of the collected data. To evaluate the likelihood of each competing model, we need data. However, data are not yet available during the phase of designing experiments, as designs need to be created before the formal data collection. The essential idea of constructing discriminatory designs is to optimize *future* collected data. *Although data are not yet available at the stage of designing experiments, we wish to estimate the information about the true DGP that future data may bring, and attempt to find what kind of data (designs) can bring the maximum information*. This idea is based on the Bayesian viewpoint of inference: use the prior knowledge to predict the distribution of possible future model probabilities. Model probabilities before data collection correspond to *prior model probabilities*, and possible future model probabilities are referred to as *posterior model probabilities*.

Figure 5.2 presents the conceptual framework used in this study for constructing discriminatory designs. For the computation of posterior model probabilities, this study adopts techniques used in Bayesian model selection (Raftery, 1995), namely *Bayes' theorem* and *Bayesian Information Criterion (BIC) approximation*. Bayesian model selection is a method of using probabilistic statistical measures to quantify model performance and model complexity. It ranks competing models based on Bayes' theorem. Bayes' theorem provides a straightforward means of connecting posterior probabilities with prior probabilities. By Bayes' theorem, posterior model probabilities are derived by prior model probabilities and the marginal likelihood of the observed data integrated over model parameters. The calculation of this integral is highly challenging as exact analytical solutions are often not tractable due to its high dimensionality[16]. Common methods for dealing with this integral include numerical evaluation, e.g., Monte Carlo

[16] This integral is as high-dimensional as the number of model parameters.

simulation, and mathematical approximation, e.g., BIC approximation. In contrast to Monte Carlo simulation and importance sampling, mathematical approximation does not use any types of sampling and it is easy to apply, and most importantly it has an explicit form. But as discussed in Zhao & Severini (2017), its accuracy cannot be guaranteed if the sample size is too small. In a seminal paper on Bayesian model selection, Raftery (1995) introduced the BIC approximation and gave explicit expressions for the integral. In this study, we use Raftery's approach to derive the approximation of the posterior model probability.

Using Bayes' theorem and the BIC approximation, we are able to predict each realization of possible future data (i.e., the posterior probability distribution). When prior probabilities update to posterior probabilities, the gained information is evaluated using the *Kullback-Leiber (KL) divergence*, which is commonly applied in previous model discrimination studies, such as Kikuchi et al. (2015) and Huan & Marzouk (2013). The KL divergence is a measure of statistical "distance" between two probability distributions of a random variable. It quantifies the difference between two probability distributions[17]. In information theory, the KL divergence is also known as relative entropy, which measures the information gain when one probability distribution changes to another. As discussed above, discriminatory designs aim to generate experiments that lead to the maximum information about the true DGP. Therefore, discriminatory designs entail finding a design that maximizes the information gain when prior model probabilities update to posterior model probabilities. As the information gain is measured by the KL divergence, discriminatory designs are about finding a design that results in the largest value of the KL divergence between prior and posterior model probabilities.

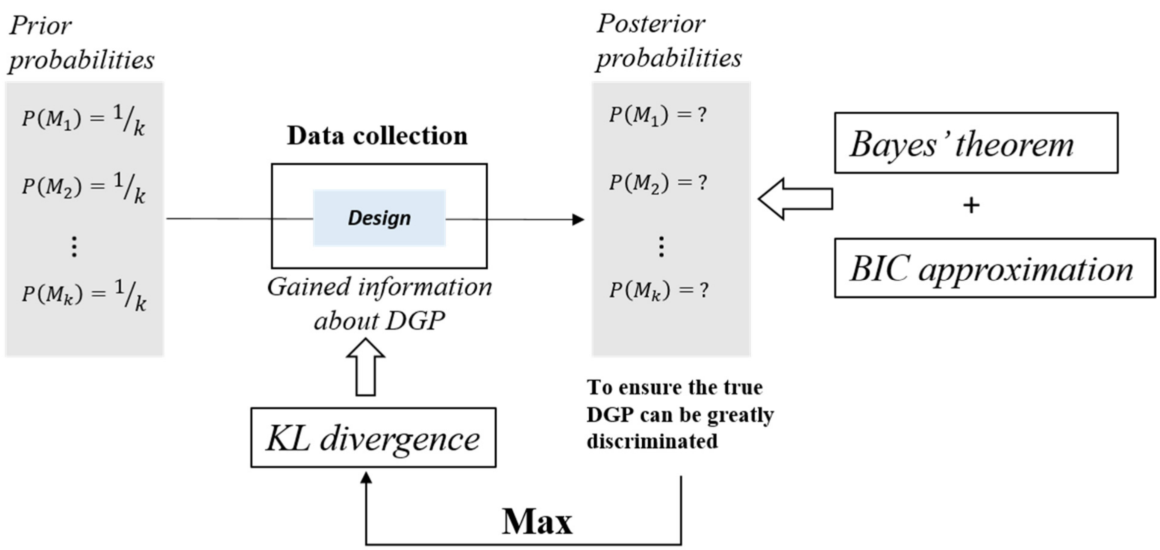


Figure 5.2 Conceptual framework of constructing discriminatory designs

[17] Note that the KL divergence is not symmetric, that is $KL(p||q) \neq KL(q||p)$.

5.3 Constructing discriminatory designs: Methods

After introducing the conceptual framework, we look into the details of mathematical equations and derivations used in the construction of discriminatory designs. Specifically, we first present the model discrimination criterion, give the mathematical derivations of Bayes' theorem and BIC approximation, and finally propose a cookbook for constructing discriminatory designs step by step.

5.3.1 Model discrimination criterion

Constructing experimental designs relies on design criteria, which reflect the goal of the experiments. Take efficient designs as an example. Efficient designs aim to generate experiments with the maximum collected information about model parameters for a given model, thereby yielding more reliable model parameter estimates. Such aim is achieved by the so-called D-optimum criterion[18], under which the design that minimizes the D-error—i.e. the determinant of the asymptotic variance-covariance matrix of models estimated on collected data—is to be chosen (Bliemer & Rose, 2011).

In contrast to efficient designs, discriminatory designs aim to generate experiments that maximize the collected information about the true underlying DGP, rather than parameter estimates of a given model. To achieve this aim, we put forward a model discrimination criterion that is based on information theory. Specifically, the criterion is derived by maximizing the KL divergence between prior model probabilities and posterior model probabilities. Suppose there are a number of k plausible, competing models M_k with $k = 1, 2, \dots, K$. The KL divergence between model probability distributions before and after data collection can be expressed by

$$D_{KL} = \sum_{k=1}^{K} P(M_k|y,d) ln\left(\frac{P(M_k|y,d)}{P(M_k)}\right), \qquad \text{(Eq. 5.1)}$$

Where $P(M_k)$ denotes the prior model probability, and $P(M_k|y,d)$ is the posterior probability of model M_k conditional on the observed data y and the design d. The KL divergence D_{KL} takes values from the domain $[0, \ln(k)]$, where k is the number of competing models. The value of zero is taken when there is no change in model probabilities, and the maximum value $ln(k)$ is taken when one of the posterior model probabilities reaches the maximum, $P = 1$, and the rest are zero. The bigger the changes in model probabilities, the larger the KL divergence. We define the discriminatory power of a design as the extent to which future data cause the expected change in model probabilities. An optimal design with the greatest discriminatory power is the one that leads to the largest value of KL divergence. Formally, experimental designs for model discrimination entail finding an optimal design d^* that maximizes the value of the KL divergence D_{KL} between the prior and the posterior model probability distributions:

$$d^* = \arg\max_d \{D_{KL}\}$$

[18] Efficient designs based on the D-optimal criterion are called D-efficient designs. Although there are other efficient designs which are built on different criterion, such as A-efficient designs, D-efficient designs are most-widely used in the stated choice experiments. In general, efficient designs refer to D-efficient designs.

$$= \arg\max_{d} \left\{ \sum_{k=1}^{K} P(M_k|y,d) ln \left(\frac{P(M_k|y,d)}{P(M_k)} \right) \right\}. \qquad \text{(Eq. 5.2)}$$

The computation of the KL divergence requires first determining prior and posterior model probabilities, $P(M_k)$ and $P(M_k|y,d)$. The prior model probability $P(M_k)$ is assumed to be the same across all competing models before data collection: $P(M_k) = {}^1/_k$ and $\sum_{k=1}^{K} P(M_k) = 1$. For the computation of posterior model probabilities, we adopt Bayes' theorem and BIC approximation. The mathematical equations and derivations are provided in the next section.

5.3.2 Posterior model probability

Bayes' theorem provides a straightforward means of connecting the prior probability and the posterior probability. By Bayes' theorem, the probability of model M_k conditional on the observed data *y* and the design *d* is given by

$$\mathrm{P}(M_k|y,d) = \frac{P(y|M_k,d)P(M_k)}{\sum_{k=1}^{K} P(y|M_k,d)P(M_k)}, \qquad \text{(Eq. 5.3)}$$

where $P(y|M_k,d)$ is the marginal likelihood of the observed *y* given model M_k and design *d*. It is also referred to as *prior predictive* as it represents the likelihood of the observed data based on the prior distribution of model parameter β_k(Schöniger et al., 2014):

$$P(y|M_k,d) = \int P(y|M_k,\beta_k,d) f(\beta_k|M_k) d\beta_k. \qquad \text{(Eq. 5.4)}$$

As mentioned in the previous section, the computation of this integral is fairly challenging. This study adopts the BIC approximation.

BIC approximation

The integrated likelihood $P(y|M_k,d)$ is approximated by taking a Taylor series expansion along with the Laplace approximation. Detailed derivations are given in Appendix 5.1. Eq. 5.5 is the Laplace estimator for the integrated likelihood:

$$P(y|M_k,d) \approx P\big(Y\big|M_k,\hat{\beta}_k,d\big) P\big(\hat{\beta}_k\big|M_k\big) (2\pi)^{N_p/2} \big|\hat{\Sigma}\big|^{\frac{1}{2}}, \qquad \text{(Eq. 5.5)}$$

where $\hat{\beta}_k$ are the maximum likelihood estimates (MLE) of parameters β_k, $|\hat{\Sigma}|$ denotes the determinant of the variance-covariance matrix about $\hat{\beta}_k$, and N_p denotes the number of parameters. The idea behind the Laplace approximation means for a large sample, parameters β_k which are near the MLE $\hat{\beta}_k$ make the most contribution to the integral. This also implies that for a small sample, the accuracy cannot be guaranteed. Details about the Laplace method for integrals can be found in Tierney & Kadane (1986).

As the Fisher information matrix is the inverse of the variance-covariance matrix, the variance-covariance matrix Σ in Eq. 5.5 can be substituted by the Fisher information matrix *F*. Given that the Fisher information matrix is often normalized by the sample size N_s, Eq. 5.5 is adapted as follows:

$$P(y|M_k,d) \approx P\big(Y\big|M_k,\hat{\beta}_k,d\big) P\big(\hat{\beta}_k\big|M_k\big) \left(\frac{2\pi}{N_s}\right)^{N_p/2} \big|\hat{F}_1\big|^{-\frac{1}{2}}. \qquad \text{(Eq. 5.6)}$$

Here $F_1 = F/N_s$ and $|F| \approx {N_s}^{N_p}|F_1|$. If both sides of Eq. 5.6 apply ln, we get

$$ln\, P(y|M_k, d) \approx lnP\left(y\middle|M_k, \hat{\beta}_k, d\right) + lnP\left(\hat{\beta}_k|M_k\right) - \frac{N_p}{2} ln \frac{N_s}{2\pi} - \frac{1}{2}\ln\left|\hat{F}_1\right|. \quad \text{(Eq. 5.7)}$$

If the sample is large enough, the terms without N_s can be omitted. The further approximation of Eq. 5.7 is

$$ln\, P(y|M_k, d) \approx lnP\left(Y\middle|M_k, \hat{\beta}_k, d\right) - \tfrac{N_p}{2} lnN_s. \quad \text{(Eq. 5.8)}$$

It is seen that the right side of Eq. 5.8 is the negative half of the BIC. Thus, the integrated likelihood can be expressed as a function of the BIC:

$$P(y|M_k, d) \approx \exp\left(-{}^1\!/_2\, BIC_{M_k}\right). \quad \text{(Eq. 5.9)}$$

Substituting Eq. 5.9 into Eq. 5.3, we obtain the approximation of the posterior model probability (Raftery, 1995):

$$P(M_k|y,\ d) \approx \frac{\exp\left(-{}^1\!/_2\, BIC_{M_k}\right)}{\sum_{k=1}^{K} \exp\left(-{}^1\!/_2\, BIC_{M_k}\right)}. \quad \text{(Eq. 5.10)}$$

Calculation of the BIC

The BIC value of model M_k consists of two components: $BIC_{M_k} = -2lnP\left(y\middle|M_k, \hat{\beta}_k, d\right) + N_p lnN_s$. The first component is the maximized log-likelihood of the observed data given the model, and the second component is a penalty term for the number of model parameters.

The calculation of the maximized log-likelihood $lnP\left(y\middle|M_k, \hat{\beta}_k, d\right)$ depends on the model M_k, the choice sets of design *d*, the maximum likelihood estimates $\hat{\beta}$, and also the outcomes of data collection $y = [y_{nsj}]$, where $y_{nsj} = 1$ if a respondent *n* chooses alternative *j* in the choice set *s*, and $y_{nsj} = 0$ otherwise. However, the maximum likelihood estimates $\hat{\beta}$ and data collection outcomes y_{nsj} are not yet known before data collection. For $\hat{\beta}$, we use the same techniques used to create efficient designs: prior parameters $\tilde{\beta}_k$ are used as the best guesses. As for y_{nsj}, we adopt the expected outcome $E(y_{sj})$ instead. The expected outcomes $E(y_{sj})$ are the probabilities that alternative *j* is chosen in the choice set *s* in the case of infinite observations. This is a probability event, so the expected outcomes $E(y_{sj})$ is irrelevant to a certain respondent *n*. Note that $lnP\left(y\middle|M_k, \hat{\beta}_k, d\right)$ is calculated for one choice set of the design. In efficient designs, D-errors are determined by the combination of all choice sets in the design. However, in discriminatory designs, all calculations are performed on one choice set, not on the combination of all choice sets. Therefore, $lnP\left(y\middle|M_k, \hat{\beta}_k, d\right)$ should be expressed by $lnP\left(y\middle|M_k, \hat{\beta}_k, s_d\right)$, where s_d denotes the choice set *s* in the design *d*.

The calculation of the penalty term depends on the number of model parameters N_p and the sample size N_s. The sample size is often decided before data collection (i.e., the target sample size of the survey). If competing models have the same number of parameters, the penalty term becomes irrelevant.

We use the following equation to calculate the (expected) BIC value of each competing model:

$$BIC_{M_k} = -2lnN_s \sum_s \sum_j E(y_{sj}) lnP_{sj}(M_k, \tilde{\beta}_k, s_d) + N_p lnN_s, \qquad \text{(Eq. 5.11)}$$

where $P_{sj}(M_k, \tilde{\beta}_k, s_d)$ denotes the choice probabilities of alternative j given model M_k, prior estimates $\tilde{\beta}_k$ and the choice set *s* of the design *d*. It is first computed for one respondent, then multiplied by the expected sample size N_s.

5.3.3 Cookbook for constructing discriminatory designs

Based on the methods given above, we provide general steps for constructing discriminatory designs. Major steps are summarized in the workflow of the construction framework, as shown in Figure 5.3.

Step 1. Specify the plausible, competing models M_k.

Step 2. Specify experimental design set-ups.

This includes specifying (*i*) attributes (generic or alternative specific) and attribute levels; (*ii*) the number of alternatives; (*iii*) the number of choice sets *s* that will be presented to each respondent; (*iv*) the targeted sample size N_s, and so on.

Step 3. Generate all possible choice situations (choice sets).

Given the design set-ups, we can generate all possible choice situations, which is called full factorial designs. Suppose there are a number of *J* alternatives, each with a number of K_j attributes, where attribute $k \in K_j$ has L_{jk} levels, the total number of all possible choice situations is $S = \prod_{j=1}^{J} \prod_{k=1}^{K_j} l_{jk}$. Full factorial designs contain duplicate alternatives and dominant alternatives. In the case of unlabelled alternatives, choice sets with duplicate alternatives or dominant alternatives cannot gain more information during data collection. Such choice sets are often omitted in the final experiment.

Step 4. Obtain parameter priors $\tilde{\beta}_k$ for each competing model.

Similar to efficient designs, prior information for model parameters can be obtained through the following ways: *i*) according to the literature, *ii*) by conducting a pilot study. For the studies aiming to test a new proposed model, priors are not available in the literature. In this case, a small scale pilot study ($\text{e.g.}\, N = 30$) with an experiment generated by orthogonal designs is conducted first to obtain the priors.

Step 5. For each choice set, calculate the choice probability of each alternative for every competing model.

With the prior parameters $\tilde{\beta}_k$, we can calculate the choice probability $P_{sj}(M_k, \tilde{\beta}_k, s_d)$. Take the linear-additive RUM model as an example. The choice probability of alternative j in one choice set *s* is $\frac{\exp(V_{js})}{\sum_{j=1,2,..J} \exp(V_{js})}$, where V_{js} is the systematic utility of alternative j in the choice set *s*.

Step 6. For each choice set, calculate the maximized log-likelihood of every competing model.

The equation of computing the maximized log-likelihood is given in Eq. 5.11. The computation depends on the sample size N_s, the expected outcome $E(y_{sj})$, and the choice probability

$P_{sj}(M_k, \tilde{\beta}_k, s_d)$. The sample size is a pre-set variable. The choice probability is obtained from Step 5. Now we discuss how to calculate the expected outcome $E(y_{sj})$. The expected outcome is calculated as the choice probability, but this choice probability is different from the choice probability obtained from Step 5. This choice probability represents how likely alternative j is chosen given the true DGP. For example, if the true DGP follows the RUM model, the expected choice probability of alternative j for choice set s is given by $E(y_{js}) = \frac{exp(V_{js})}{\sum_{j=1,2,..J} exp(V_{js})}$, where V_{js} is the systematic utility. The choice probability $P_{sj}(M_k, \tilde{\beta}_k, s_d)$ refers to the probability of choosing alternative j given the estimated model M_k. This means if the estimated model is identical with the true DGP, then the calculations of $E(y_{sj})$ and $P_{sj}(M_k, \tilde{\beta}_k, s_d)$ are the same, otherwise, they are different. However, the true DGP is unknown. We assume that every competing model in turn becomes the true DGP. Therefore, the log-likelihood of each model needs to be calculated once under each assumption about the true DGP. For example, in the case of two competing models M_1 and M_2, the log-likelihood of the model M_1 has two values: one is under the assumption that model M_1 is the true DGP; the other is under the assumption that model M_2 is the true DGP. Further details will be illustrated by an example of two competing models in the next section.

Step 7. For each choice set, calculate the (expected) BIC for each competing model.

A model's BIC consists of the log-likelihood and a penalty term: the log-likelihood is obtained from the last step, and the penalty term $N_p ln N_s$ is easy to compute. Note that, as described in Step 6, each competing model is in turn assumed to be the true DGP. Therefore, under every assumption about the true DGP, each competing model has a BIC value.

Step 8. For each choice set, calculate the posterior probability of each competing model.

Once the BICs are obtained, we can calculate the posterior probability according to Eq. 5.10. Again, each competing model has a posterior probability under every assumption about the true DGP.

Step 9. For each choice set, calculate the value of KL divergence.

Substituting both prior and posterior probabilities into Eq. 5.1, the value of KL divergence for each choice set can be obtained. Note that each choice set has more than one KL divergence value, as it is calculated under every assumption.

Step 10. To construct a discriminatory design, select the choice sets with the largest KL divergence values.

Since the KL divergence values for each choice set are calculated, the final step is to sort them from the largest value to the smallest value, and then select the choice sets with the largest values and make them up of a discriminatory design. As one choice set has more than one KL divergence value, we need to select the choice sets which have all large values of KL divergence under different assumptions.

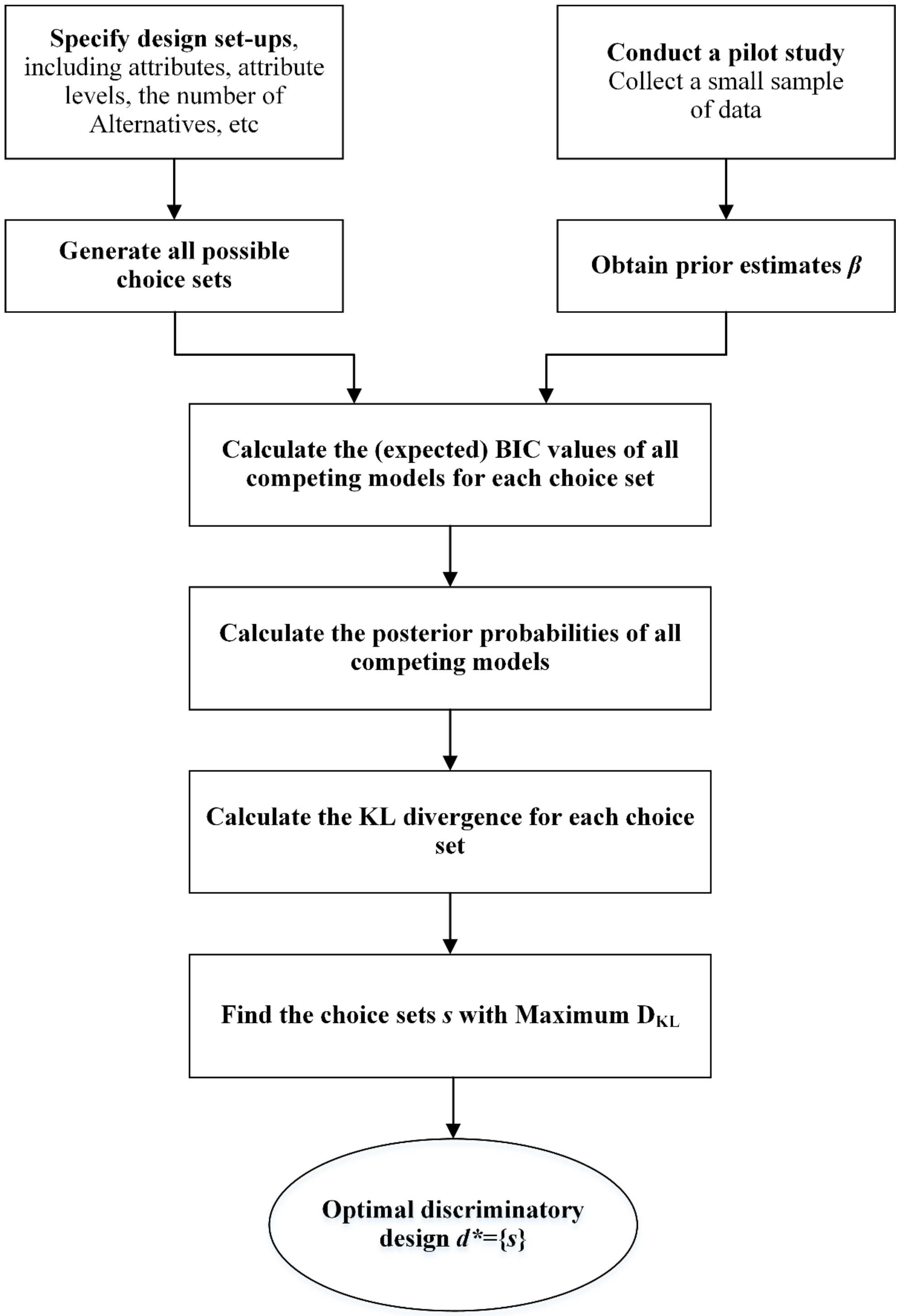


Figure 5.3 Schematic diagram showing the workflow of constructing disicmrinatory designs

5.4 Testing the robustness of discriminatory designs

This section examines the robustness of discriminatory designs using synthetic data. First, we present an example of constructing a discriminatory design for two competing models by following the steps given in section 5.3.3. In addition, a randomly-generated design is also constructed to serve as a benchmark. The two designs are then compared in terms of the capability of model discrimination. Finally, this section explores the relationship between discriminatory power and efficiency.

5.4.1 Constructing a discriminatory design for two competing models

This section presents an example of constructing an experimental design for discriminating between two models[19]. In this example, model discrimination is conducted between the linear-additive RUM model (fromhereon, the RUM model) and the P-RRM model. Recent years have seen extensive comparisons between the RUM model and the RRM model in various choice contexts (e.g., C. Chorus, 2012; Boeri & Longo, 2017; Masiero et al., 2019; Sharma et al., 2019; Wong et al., 2020; Iraganaboina et al., 2021). The RRM model is one of the most widely used non-RUM models, and the P-RRM model is an extreme case of the RRM model, which yields a very different behaviour (the stongest regret aversion) from the RUM model (van Cranenburgh et al., 2015).

The next step is to specify design settings. In this example, we create an experiment that contains four generic attributes with three levels each: travel time (TT) (45, 60, 75 minutes), percentage of travel time in congestion (CONG) (10, 25, 40%), number of traffic fatalities per year (TF) (1, 2, 3), and travel cost (TC) (5.5, 9, 12 euro). There are three unlabeled alternatives in each choice set and five choice sets make up the experiment. In total, there are 24,025 unique choice sets without duplicate or dominant alternatives. Table 5.1 presents the design set-ups used in this example.

Table 5.1 Design set-ups used in the example

Design set-ups	
Number of alternatives	3
Number of generic attributes	4
Number of attribute levels	3
Number of choice sets per respondent	5
Number of unique choice sets without duplicate or dominant alternatives	24,025

Prior parameters are obtained from a previous empirical study (C. G. Chorus, 2012), which has the same design set-ups. This empirical study contains 3,897 choice observations. Instead of using perfect priors, we use the estimates of the models on a small sub-dataset. Specifically, the two models are estimated on a subset of 150 choice observations randomly drawn from all choice observations. This is to mimic the general way of obtaining priors, in which usually a small number of target respondents are surveyed in a pilot study. Table 5.2 presents the priors used in this example.

[19] Note that the general steps given in the previous section can be also applied for the situation where more than two models are to be discriminated, but in this study we focus on a simple situation of two competing models.

Table 5.2 Parameter priors for the two competing models in the example

	Prior (TT)	Prior (CONG)	Prior (TF)	Prior (TC)
The RUM model	-0.06	-0.02	-0.21	-0.17
The P-RRM model	-0.05	-0.02	-0.13	-0.10

The next steps involve first obtaining the systematic utility of the RUM model and the systematic regret of the P-RRM model, then calculating the corresponding choice probabilities of the two models for every unique choice set, and calculating the (expected) maximized log-likelihood for each model. In these steps, the RUM model and the P-RRM model are assumed to be the true DGP respectively. Specifically, the equations for calculating the (expected) maximized log-likelihood are given in Table 5.3, in which V_{js} denotes the systematic utility of alternative j in choice set s, R_{js} denotes the systematic regret of alternative j, and N_s is the sample size. In this example, we set $N_s = 100$.

Table 5.3 Equations of (expected) maximized log-likelihoods

(a) Assume that the RUM model is true DGP

Model	RUM	P-RRM
Expected LL	$N_s \sum_j \frac{\exp(V_{js})}{\sum_j \exp(V_{js})} ln \frac{\exp(V_{js})}{\sum_j \exp(V_{js})}$	$N_s \sum_j \frac{\exp(V_{js})}{\sum_j \exp(V_{js})} ln \frac{\exp(-R_{js})}{\sum_j \exp(-R_{js})}$

(b) Assume that the P-RRM model is true DGP

Model	RUM	P-RRM
Expected LL	$N_s \sum_j \frac{\exp(-R_{js})}{\sum_j \exp(-R_{js})} ln \frac{\exp(V_{js})}{\sum_j \exp(V_{js})}$	$N_s \sum_j \frac{\exp(-R_{js})}{\sum_j \exp(-R_{js})} ln \frac{\exp(-R_{js})}{\sum_j \exp(-R_{js})}$

Note that the equations in blue denote the expected probabilities given true DGP.

With the (expected) maximized log-likelihood and the penalty term, we can obtain the (expected) BIC value for each model. With the BIC values, posterior probabilities are computed according to Eq. 5.10. As there are two competing models in this example, the prior probability of each model is $1/2$. Substituting both prior probabilities and posterior probabilities into Eq. 5.1, we can get the KL divergence for each choice set. In order to clearly demonstrate the calculation process, a specific choice set is taken as an example. Table 5.4 specifies the attribute levels of the three alternatives in the choice set and gives the corresponding systematic utilities and systematic regrets respectively for each alternative. Table 5.5 shows the calculation results of the KL divergence under different assumptions about the true DGP.

Table 5.4 Attribute levels and corresponding systematic utilities/regrets for one choice set

Choice set	Attribute levels (TT, CONG, TF, TC)			Systematic utility			Systematic Regret		
	Alt 1	Alt 2	Alt 3	Alt 1	Alt 2	Alt 3	Alt 1	Alt 2	Alt 3
1	45, 10, 1, 12.5	45, 10, 2, 9	45, 10, 3, 5.5	-5.24	-4.85	-4.47	-1.08	-0.51	-0.42

Table 5.5 Values of the KL divergence for one choice set

(a) Assume that the RUM model is true DGP

Choice set	Choice probabilities						Expected LL		BIC (N_s=100)		Posterior probability		D_{KL}
	RUM			P-RRM									
	Alt1	Alt2	Alt3	Alt1	Alt2	Alt3	RUM	P-RRM	RUM	P-RRM	RUM	P-RRM	
1	0.2	0.3	0.5	0.2	0.4	0.4	-1.05	-1.06	210	212	0.70	0.30	**0.08**

(b) Assume that the RUM model is true DGP

Choice set	Choice probabilities						Expected LL		BIC (N_s=100)		Posterior probability		D_{KL}
	RUM			P-RRM									
	Alt1	Alt2	Alt3	Alt1	Alt2	Alt3	RUM	P-RRM	RUM	P-RRM	RUM	P-RRM	
1	0.2	0.3	0.5	0.2	0.4	0.4	-1.07	-1.06	214	212	0.30	0.70	**0.08**

The discriminatory power of designs is measured by the KL divergence. As shown in Table 5.5, each choice set corresponds to a KL divergence value (under one assumption). A discriminatory design consists of the choice sets which have large values of the KL divergence. After obtaining the KL divergence for each choice set, we select the choice sets with the largest KL divergence values to form a discriminatory design. Since one choice set has multiple KL divergence values under different assumptions, it is required that the selected choice sets have a large KL divergence under each assumption. In general, a choice set with a large (or small) KL divergence under one assumption usually has a large (or small) KL divergence under another assumption, see the example given in Table 5.5. In this example, we select five choice sets with the largest values of KL divergence to form a discriminatory design. The design is presented in Table 5.6.

Table 5.6 The discriminatory design generated in this example

Choice set	Alternative levels (TT, CONG, TF, TC)			D_{KL}	
	Alt1	*Alt2*	*Alt3*	DGP:RUM	DGP: P-RRM
1	45, 10, 2, 12.5	45, 40, 1, 12.5	75, 10, 1, 5.5	0.693	0.693
2	45, 25, 1, 12.5	45, 40, 3, 5.5	75, 10, 1, 12.5	0.693	0.693
3	45, 40, 1, 12.5	45, 40, 2, 9	60, 10, 3, 5.5	0.688	0.685
4	45, 40, 1, 5.5	75, 10, 1, 12.5	75, 10, 3, 9	0.692	0.693
5	60, 25, 3, 12.5	60, 40, 2, 12.5	75, 25, 1, 5.5	0.693	0.693

To evaluate how well the method works, we create a random design that serves as a benchmark. Specifically, the design is created by randomly selecting five choice sets from all possible choice sets. The design and corresponding KL divergence values are given in Table 5.7. This design is non-discriminatory, as all choice sets have small values of the KL divergence. The next section will compare these two designs in terms of their discrimination capability.

Table 5.7 The non-discriminatory design used in this example

Choice set	Alternative levels (TT, CONG, TF, TC)			D_{KL}	
	Alt1	*Alt2*	*Alt3*	DGP:RUM	DGP: P-RRM
1	45, 10, 2, 12.5	60, 40, 3, 5.5	75, 40, 1, 12.5	0.339	0.321
2	45, 10, 3, 9	60, 25, 1, 9	75, 10, 2, 12.5	0.088	0.075
3	45, 25, 3, 9	60, 40, 1, 9	75, 25, 2, 9	0.088	0.079
4	60, 40, 2, 5.5	75, 25, 3, 12.5	75, 40, 2, 9	0.009	0.009
5	45, 10, 1, 9	60, 40, 3, 5.5	75, 25, 1, 5.5	0.067	0.060

5.4.2 Robustness of discriminatory designs (synthetic data test)

In this section, a synthetic data test is conducted to assess how well the discriminatory design (Table 5.6) can discriminate between the RUM model and the P-RRM model. For comparison, the test is also conducted on the non-discriminatory design given in Table 5.7. These two designs generate two different datasets respectively: dataset A and B. The generation of synthetic data relies on a certain decision rule. In this example, the true DGP of synthetic data in these two datasets follows the RUM model. Parameters used for the true DGP are obtained by estimating the RUM model on the actual observed choices of an empirical study (C. G. Chorus, 2012): $\beta_{TT} = -0.07$, $\beta_{CONG} = -0.03$, $\beta_{TF} = -0.29$, and $\beta_{TC} = -0.18$. Note that these parameters are different from the priors used for constructing the designs in the example. We assume that one hundred respondents take the survey, each facing five choice sets. Therefore, there are 500 synthetic choices in each dataset.

After generating the synthetic data, the RUM model and the P-RRM model are estimated on the two datasets respectively, the model estimation results are given in Table 5.8. We first look at model fits. For the estimation results of dataset A, there is a distinct difference in model fits between the two models. Specifically, the RUM model (true DGP) performs better than the P-RRM model with a 12-point log-likelihood improvement. Such a difference in model fits is significant for 500 choice observations, which gives a clear signal that the RUM model is the true DGP compared to the P-RRM model. For dataset B, the two competing models have the same model fit, which makes it difficult to distinguish which of the two models is the true DGP. Therefore, we can see with an elaborate design that has high discriminatory power, the true DGP can be clearly identified; while with a randomly generated design or a non-discriminatory design, the true DGP is very likely to be incorrectly identified.

Table 5.8 Estimation results of the synthetic data (True DGP is RUM)

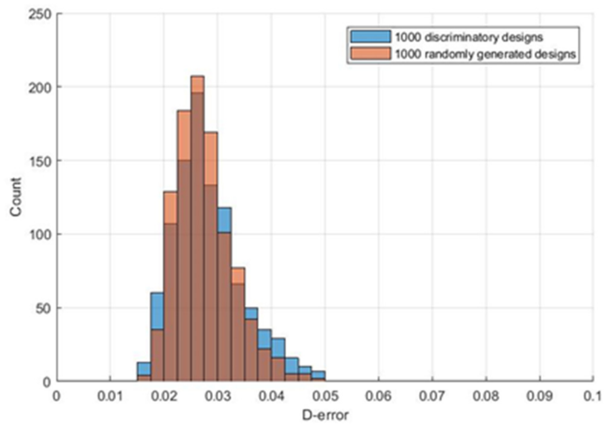
	Dataset A Discriminatory design		**Dataset B** non-discriminatory design	
	RUM	**P-RRM**	**RUM**	**P-RRM**
β_{TT}	-0.07	-0.03	-0.06	-0.04
(*s.e. t-value*)	(*0.01 -9.6*)	(*0.03 -7.5*)	(*0.01 -10.6*)	(*0.00 -10.6*)
β_{CONG}	-0.03	-0.01	-0.02	-0.01
(*s.e. t-value*)	(*0.01 -5.2*)	(*0.00 -3.5*)	(*0.01 -2.2*)	(*0.01 -2.9*)
β_{TF}	-0.39	-0.25	-0.24	-0.14
(*s.e. t-value*)	(*0.07 -6.0*)	(*0.04 -6.9*)	(*0.08 -3.0*)	(*0.05 -2.9*)
β_{TC}	-0.17	-0.17	-0.08	-0.03
(*s.e. t-value*)	(*0.02 -9.6*)	(*0.02 -9.5*)	(*0.04 -1.8*)	(*0.03 -0.9*)
Number of choices	500	500	500	500
Null LL	-549	-549	-549	-549
Final LL	**-457**	**-469**	**-437**	**-437**

In terms of parameter estimates, it is seen that all parameters are highly significant and of the expected sign in the results of dataset A. Moreover, all parameters are successfully recovered. As for the results of dataset B, we see that parameter β_{TC} is not significant in both models, which means this parameter cannot be recovered, and the rest parameters are significant and of the expected sign. In this example, we see that the discriminatory design is more efficient than the randomly-generated design in terms of recovering model parameters (as efficient designs). Several questions then arise: Are discriminatory designs more efficient than non-discriminatory designs in terms of recovering model parameters? Is there a close correlation between the discriminatory power of designs and their efficiency in terms of recovering parameters? Compared to a state-of-practice D-efficient design, how much less efficient is a discriminatory design for recovering parameters? Or how much less discriminatory is a D-efficient design than the best discriminatory design? To answer these questions, the next section will explore the relationship between discriminatory power and efficiency.

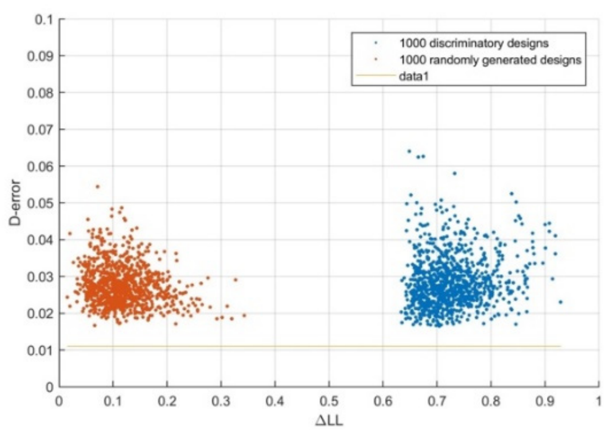
5.4.3 Relationship between discriminatory power and efficiency

To demonstrate the relationship between discriminatory power and efficiency, we create two groups of designs. The first group consists of 1,000 discriminatory designs: five choice sets are randomly drawn from the top 100 choice sets which have large values of the KL divergence to form a discriminatory design, and this procedure repeats 1,000 times. The second group is made of 1,000 random designs, each consisting of five randomly selected choice sets. After generating two groups of designs, we calculate the D-error for each design.

Figure 5.4 (a) presents the histograms of D-errors of the designs in the two groups. Specifically, the blue histogram shows the distribution of D-errors of the discriminatory designs, and the orange one is the distribution of D-errors of the random designs. We can see that these two groups have very similar distributions of D-errors: two histograms have similar shapes and they have a shared range on the X-axis (the size of D-errors). Figure 5.3 (b) further illustrates the relationship between discriminatory power and efficiency. From the scatter plots, we can see there is no certain relationship between discriminatory power and efficiency, or in other words, they are not much correlated to each other. In this example, the smallest D-error is 0.011, shown as the yellow line in Figure 5.4 (b). it is seen that all blue dots (discriminatory designs) are above the yellow line, but some of them are not far from it, meaning some discriminatory designs have relatively small D-errors. This indicates that constructing discriminatory designs does not necessarily jeopardize the reliable recovery of model parameters. Therefore, it is possible to construct designs that are both discriminatory and efficient.



(a) Histograms of D-errors of two groups of designs



(b) Scatter plots of discriminatory power of designs and their efficiency

Figure 5.4 The relationship of discriminatory power and efficiency

5.5 Empirical applications

In this section, we investigate how well the present method of constructing discriminatory designs performs on empirical data. Specifically, we aim to compare the empirical performance of discriminatory designs with non-discriminatory designs in terms of discriminating between the RUM model and the P-RRM model. First, this section starts by an empirical study, in which the choice experiment consists of choice sets generated by two different designs—a discriminatory design and an efficient but non-discriminatory design. The aim is to see whether the discriminatory design can lead to a greater difference in model fits, compared to the non-discriminatory design. Second, we also test the present method on five existing datasets. As the choice sets of these datasets were generated before by different designs, the technique used here is to separate the choice sets of each dataset into two groups—one with high discriminatory power and the other with low discriminatory power and then estimate the RUM and P-RRM models on two groups respectively. Using this technique, we wish to test whether the one with high discriminatory power performs better than the one with low discriminatory in terms of model discrimination.

5.5.1 Case study: Value of train travel time

A case study is conducted to assess the empirical performance of discriminatory designs. Specifically, two different designs are used to generate the choice sets of the stated choice experiment in this study: one is a discriminatory design constructed by following our method, the other is an efficient design generated by the software Ngene. The context of this stated choice experiment concerns the valuation of train travel time. Respondents were first provided with the travel information about the current trip by train: the total travel time is 60 minutes, the total cost is ten euros, and there is a transfer during the trip. They were then asked to choose between the current trip and two additional trip options. Table 5.9 presents the specific design setting used in this study. Each choice set consists of three unlabelled alternatives described by travel time, travel cost and the number of transfers. The settings of attribute levels are pivoted around the levels of the current trip. Specifically, the level of travel time ranges from 42 to 78 minutes with equal intervals of 9 minutes, the level of travel cost ranges from 7 to 13 euro with equal 1.5 euro intervals, and the level of transfers is 0, 1 or 2. The experiment consists of twelve choice sets: six are generated by a discriminatory design, and other six are generated by an efficient design.

Table 5.9 Design settings in the case study

Number of alternatives	3
Number of generic attributes	3
Attribute levels: travel time (minute)	45, 51, 60, 69, 78
Attribute levels: travel cost (euro)	7, 8.5, 10, 11.5, 13
Attribute levels: transfers	0, 1, 2
Number of choice sets per respondent	12 (i.e. 6 + 6)

The construction of discriminatory designs or efficient designs requires priors for model parameters. Priors used in this study are based on the pricing of Dutch train tickets. In the Netherlands, one hour trip by train roughly costs ten euros, thus the value of train travel time is approximately €10 per hour (€0.17/min). Assuming that the average transfer time is ten minutes, then the value of one transfer equals €1.7. Therefore, for the RUM model, the setting of priors needs to satisfy $\beta_{Time}/\beta_{Cost} \approx 0.17$ and $\beta_{Trans}/\beta_{Cost} \approx 1.7$. In this experiment, the priors for travel time, travel cost, and transfer are set to -0.09, -0.53, and -1 respectively in the RUM model, and the priors in the P-RRM model are set to two-thirds of those in the RUM model (see van Cranenburgh et al., 2015 for the relation between RRM and RUM).

The discriminatory design constructed in this study needs to be efficient as well, in order to ensure the reliable recovery of model parameters with a limited number of respondents. In this study, the experiment design generated by Ngene has a D-error of 0.047 and the discriminatory design used here also has a small D-error, 0.048. The choice sets generated by the two designs and corresponding KL divergence values and D-errors are presented in Table 5.10. From the table, we can see that both designs are efficient, but they have different degrees of discriminatory power. The efficient design has several choice sets which have very small KL divergence values, this design therefore can be seen as a non-discriminatory design.

Table 5.10 Choice sets of the case study

Discriminatory design

Choice	*Current trip*			*Trip 1*			*Trip 2*				
set	Time	Cost	Trans	Time	Cost	Trans	Time	Cost	Trans	D_{KL}	*D-error*
1	60	10	1	42	11.5	2	78	7	0	*0.69*	*0.048*
2	60	10	1	78	7	2	51	13	0	*0.69*	
3	60	10	1	51	7	2	78	10	0	*0.67*	
4	60	10	1	78	8.5	0	42	13	2	*0.69*	
5	60	10	1	42	13	2	78	7	0	*0.69*	
6	60	10	1	42	13	0	78	7	2	*0.69*	

Efficient design

Choice	*Current trip*			*Trip 1*			*Trip 2*				
set	Time	Cost	Trans	Time	Cost	Trans	Time	Cost	Trans	D_{KL}	*D-error*
1	60	10	1	42	11.5	0	60	7	2	*0.59*	*0.047*
2	60	10	1	42	10	2	69	7	0	*0.60*	
3	60	10	1	60	7	1	51	11.5	1	*0.11*	
4	60	10	1	69	13	1	78	10	1	*0.00*	
5	60	10	1	51	7	2	42	13	0	*0.32*	
6	60	10	1	78	8.5	0	42	8.5	2	*0.21*	

Data collection was carried out during March and April, 2020 in the Netherlands. Respondents were recruited from Bachelor or Master students at the Delft University of Technology. In total, 56 respondents completed the survey. Each respondent was faced with twelve choice sets which were generated by either the discriminatory design or the efficient design, leading to 672 choice observations. To avoid ordering effects, choice sets generated by the two designs were presented in a random order.

Table 5.11 Model estimation results of the case study

	Whole dataset		**Subset 1** Discriminatory design		**Subset 2** Efficient design	
	RUM	***P-RRM***	***RUM***	***P-RRM***	***RUM***	***P-RRM***
β_{Time} (*s.e.* *t-value*)	-0.08 (*0.01 -14.2*)	-0.05 (*0.00 -12.4*)	-0.07 (*0.01 -8.0*)	-0.04 (*0.01 -8.5*)	-0.09 (*0.01 -11.5*)	-0.07 (*0.01 -10.3*)
β_{Cost} (*s.e.* *t-value*)	-0.36 (*0.03 -11.3*)	-0.23 (*0.02 -10.4*)	-0.38 (*0.05 -7.0*)	-0.22 (*0.03 -7.2*)	-0.31 (*0.04 -7.1*)	-0.22 (*0.03 -6.2*)
β_{Trans} (*s.e.* *t-value*)	-0.70 (*0.06 -10.9*)	-0.48 (*0.05 -9.0*)	-0.69 (*0.08 -8.6*)	-0.47 (*0.07 -7.0*)	-0.63 (*0.11 -5.7*)	-0.48 (*0.09 -5.4*)
Choice observations	672	672	336	336	336	336
Null-LL	-738	-738	-369	-369	-369	-369
Final LL	**-638**	**-644**	**-332**	**-322**	**-301**	**-312**

Table 5.11 presents the model estimation results. For the whole dataset, the RUM model outperforms the P-RRM model by 6 log-likelihood points. We expect that a larger difference in model fits can be observed in the estimation result of Subset 1, and a smaller difference can be observed in the estimation result of Subset 2. However, the results do not seem to be as expected. Specifically, the model estimation of two subsets gives opposite results about the better fitting model: for Subset 1, the P-RRM model performs better than the RUM model with 10 log-likelihood points, while for Subset 2, the RUM model performs better with 11 log-likelihood points. Therefore, the discriminatory designs do not appear to amplify the difference in model fits between the two competing models.

Possible reasons for such results might lie in the implicit assumption in the design construction. The construction of discriminatory designs is based on an implicit assumption that one of the competing models is the true underlying DGP. In this example, we assume that either RUM or P-RRM would be the true DGP. However, in the actual data, it is possible that i) the respondents followed neither of these two decision rules when making choices, or ii) the respondents followed multiple decision rules, for example, some followed the RUM decision rule and others used the P-RRM decision rule.

To further explore the reasons, the whole dataset is estimated on the μRRM model, which is the most generalized model that can accommodate the RUM and P-RRM choice behaviour. The estimated parameter μ governs the degree of regret aversion in behaviour: when μ approaches zero, it implies a strong regret aversion in behaviour (the P-RRM choice behaviour); when μ is large, i.e. $\mu > 10$, it implies no regret aversion (the RUM choice behaviour). The estimation results (given in Appendix 5.2) show that parameter μ equals 1.05, implying a moderate degree of regret aversion, neither strong regret aversion nor no regret aversion at all. This indicates that neither of the two models is the true DGP of the data.

5.5.2 Overview of empirical results

We also test the empirical performance of discriminatory designs on five existing datasets. These datasets have been used to compare the RUM model and the RRM model. The choice sets of these datasets were generated following different designs, which were not optimized for discriminating between models. Thus, we apply a special way of conducting the empirical test. For each dataset, we first split the choice sets into **Subset 1**—consisting of choice sets with larger KL divergence values—and **Subset 2**—consisting of choice sets with smaller KL divergence values. Then, the RUM model and P-RRM model are estimated on the two subsets respectively. The final step is to see whether subset 1 can lead to greater differences in model fits, compared with subset 2.

Details of each dataset are given in Table 5.12. Four of the five datasets concern route choices, and the rest is about dating choices. Specifically, these choice sets were generated by different designs: orthogonal designs, pivoted designs, and efficient designs. To split choice sets, we first calculate the value of KL divergence for each choice set, and select the choice sets with higher values to form Subset 1 and the remaining sets enter Subset 2. As seen in the table, both datasets 1 and 4 contain nine choice sets. In this case, Subset 1 consists of the top four choice sets with larger values and Subset 2 is composed of the last four choice sets with smaller values, the "middle" choice set is discarded. Dataset 2 are made up of 2,561 different choice sets. In this case, we calculate the value of KL divergence for each choice set, and the whole dataset is evenly divided into two sets based on their KL divergence values. Datasets 3 and 5 have an even number of choice sets, which are equally split into two sets according to the KL divergence values.

Table 5.12 Details about five existing datasets

ID	Reference	Choice context	Design	Alternatives	# of choice set	Choice observations		
						Whole	Subset 1	Subset 2
1	Chorus, 2012	Route choice	Optimal orthogonal	3 unlabelled routes	9	3897	1725	1732
2	Chorus and Rose, 2012	Dating	Unknown	3 unlabelled alternatives	2,561 different sets	4299	2149	2150
3	Chorus et al., 2013	Route choice	Pivoted	a current trip and 2 alternative routes	16	4800	2400	2400
4	Chorus and Bierlaire, 2013	Route choice	Optimal orthogonal	3 unlabelled routes	9	3856	1560	1560
5	Van Cranenburgh et al., 2018	Route choice	Efficient	3 unlabelled routes	10	1060	530	530

Table 5.13 presents the final log-likelihoods of the RUM model and the P-RRM model in the five datasets. For all five datasets, we first estimate these two models on the whole dataset, and then estimate them on Subset 1 and Subset 2 respectively. From the estimation results, several observations can be made. First, the results of the first three datasets are in line with the expectation: the difference in model fits between RUM and P-RRM is greater in Subset 1, compared with Subset 2. Especially for Dataset 3, the model fit difference of the whole dataset is 92-point log-likelihood; the difference of Subset 1 (containing half of the choice observations) reaches 82-point log-likelihood, while the difference of Subset 2 (containing another half of choice observations) is very small, only 7-point log-likelihood. These two subsets contain the same number of choice observations (i.e., 2400), but the choice sets of Subset 1 lead to a much larger discrepancy in model fits, compared with the choice sets of Subset 2. However, for Datasets 4 and 5, we cannot observe a larger model fit difference in Subset 1 than in Subset 2. After inspecting the final log-likelihood of each model, we notice that Datasets 4 and 5 have a similar situation to the case study in Section 5.5.1. The RUM model offers better performance than the P-RRM model for the whole dataset, while the estimation results of the subsets give inconsistent outcomes regarding the better-performing model. As discussed above, possible explanations include that 1) different respondents employed different decision rules, e.g., some employed RUM and others employed P-RRM, 2) certain choice sets may trigger respondents to employ different decision rules. These reasons can lead to an outcome that neither RUM nor P-RRM is the true DGP of the entire data.

Table 5.13 Model fits of RUM and P-RRM in five datasets

ID	Whole dataset			Subset 1			Subset 2		
	LL_{RUM}	$LL_{P\text{-}RRM}$	ΔLL	LL_{RUM}	$LL_{P\text{-}RRM}$	ΔLL	LL_{RUM}	$LL_{P\text{-}RRM}$	ΔLL
1	-2847	-2893	***46***	-1449	-1477	***28***	-1363	-1380	***17***
2	-3689	-3607	***82***	-2014	-1956	***58***	-1665	-1644	***21***
3	-4024	-4116	***92***	-1954	-2036	***82***	-2063	-2070	***7***
4	-2613	-2617	***4***	-1293	-1287	***6***	-1235	-1240	***5***
5	-1123	-1128	***5***	-567	-579	***5***	-554	-549	***12***

Better log-likelihoods are highlighted in blue.

5.6 Conclusions and discussions

Despite wide applications in many fields such as geoscience and hydrological science, model discrimination has yet to become an idea of experimental designs in the field of choice modelling for stated choice studies. In this chapter, we put forward an innovative method of constructing experimental designs optimized for discriminating between competing choice models. Such designs are called discriminatory designs. In contrast to efficient designs, which maximize collected information about model parameters, discriminatory designs aim to generate experiments that maximize collected information about the true underlying DGP, rather than parameter estimates of a given model.

We tested the robustness of discriminatory designs on synthetic data. The synthetic data test showed that discriminatory designs are a robust approach to experimental designs for model discrimination. More specifically, an elaborate discriminatory design can lead to a greater difference in model fits, yielding a result that clearly favours the true DGP over the competing model; while a non-discriminatory design can lead to a misleading result that the true DGP and the competing model have the same model fit. Furthermore, the synthetic data test also examined the relationship between a design's discriminatory power and its efficiency in terms of recovering model parameters. It showed that the construction of discriminatory designs does not necessarily jeopardize the reliable recovery of model parameters, implying that we can create designs that are both discriminatory and efficient.

We also tested discriminatory designs on empirical data. Specifically, we conducted a case study in which the experiment consisted of both discriminatory and non-discriminatory designs, and also did empirical tests on five existing datasets, each of which was split into two subsets with high and low discriminatory power, respectively. The empirical findings showed that our approach can enlarge the model fit differences between competing models, leading to a greater likelihood of differentiating the competing models. However, we also found that for empirical data, discriminatory designs do not always produce ideal results regarding expanding the difference in model fits. For example, in the case study, we did not find a larger difference in model fits for the discriminatory design and a smaller difference for the non-discriminatory design. Similar results were also found in the empirical analyses of two existing datasets.

Discriminatory designs are constructed based on an implicit assumption that one of the competing models is the true underlying DGP. Such an assumption also exists in the approaches to model selection, for example (Bayesian) hypothesis testing. The literature of model selection often assumes that there is a "true" model and this "true" model belongs to a specified class of models considered. As discussed in Bernardo & Rueda (2002), assuming the existence of a "true" model is appropriate in the situation where "one knew for sure that the real world mechanism which has generated the available data was one of a specified class." This is the case when data are generated by computer simulation, such as synthetic data. However, empirical data are much more complicated than synthetic data. In actual data collection, decision makers may employ various decision rules when facing the same choice situations, or a single decision maker may use more than one decision rule in all choice tasks. Thereby, for empirical data, discriminatory designs may fail to yield a result that clearly favours one model over others. This is however not as bad as it sounds. The purpose of creating discriminatory designs is to increase the likelihood of differentiating between models which will be compared in the later model selection, and as mentioned above, model selection is also based on the same assumption. Therefore, in the case that there is one favoured model ("true" model) given the data, then using discriminatory designs can maximize the likelihood of discriminating this model from other competing models.

Reliable priors are very crucial to the construction of discriminatory designs, because they are the key values in determining how much the discriminatory power a design is expected to have. Similar to efficient designs, ill-chosen priors may undermine the goal of discriminatory designs, leading to designs with low discriminatory power. In our synthetic data test and empirical data test, we did not use perfect priors to construct designs; instead, we used a small portion of data to generate priors, so as to be as close as possible to the actual situation of data collection (i.e., priors obtained from a small pilot study). Our results showed that designs generated with reliable (but not perfect) priors can yield the experiments with sufficient discriminatory power. The way to obtain reliable priors can refer to the literature on efficient designs, which has been extensively discussed. Common ways include conducting a pilot study, obtaining priors based on the previous literature, and using so-called Bayesian priors (Bliemer & Collins, 2016). Note that zero priors with the sign are also informative for creating efficient designs, but this is not the case for discriminatory designs which rely on the expectation regarding the size of parameters.

Finally, this chapter concludes with a possible avenue for future research. Inspired by the Bayesian model selection, we adopted the BIC approximation to compute the integrated likelihood, which is an important component for calculating the posterior probability distribution of the models. Although the BIC approximation is easy to apply and has an explicit form, its accuracy has been questioned especially for the small sample size. In the literature on model discrimination, other popular ways for computing the integrated likelihood include Monte Carlo simulation and importance sampling. Zhao & Severini (2017) reviews twelve different approaches to computing the integrated likelihood. Future research may consider different approaches to deriving the posterior probability distributions.

References

Alberton, A. L., Schwaab, M., Lobão, M. W. N., & Pinto, J. C. (2011). Experimental design for the joint model discrimination and precise parameter estimation through information measures. *Chemical Engineering Science, 66*(9), 1940-1952.

Bernardo, J. M., & Rueda, R. (2002). Bayesian hypothesis testing: A reference approach. *International Statistical Review, 70*(3), 351-372.

Bliemer, M. C., & Collins, A. T. (2016). On determining priors for the generation of efficient stated choice experimental designs. *Journal of Choice Modelling, 21*, 10-14.

Bliemer, M. C., & Rose, J. M. (2010). Construction of experimental designs for mixed logit models allowing for correlation across choice observations. *Transportation Research Part B: Methodological, 44*(6), 720-734.

Bliemer, M. C., & Rose, J. M. (2011). Experimental design influences on stated choice outputs: an empirical study in air travel choice. *Transportation Research Part A: Policy and Practice, 45*(1), 63-79.

Boeri, M., & Longo, A. (2017). The importance of regret minimization in the choice for renewable energy programmes: Evidence from a discrete choice experiment. *Energy Economics, 63*, 253-260.

Box, G. E., & Hill, W. J. (1967). Discrimination among mechanistic models. *Technometrics, 9*(1), 57-71.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432.

Cavagnaro, D. R., Myung, J. I., Pitt, M. A., & Kujala, J. V. (2010). Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural computation, 22*(4), 887-905.

Chorus, C. (2012). Random regret minimization: an overview of model properties and empirical evidence. *Transport reviews, 32*(1), 75-92.

Chorus, C. G. (2010). A new model of random regret minimization. *European Journal of Transport and Infrastructure Research, 10*(2), 181-196.

Chorus, C. G. (2012). Logsums for utility-maximizers and regret-minimizers, and their relation with desirability and satisfaction. *Transportation Research Part A: Policy and Practice, 46*(7), 1003-1012.

Dehideniya, M. B., Drovandi, C. C., & McGree, J. M. (2018). Optimal Bayesian design for discriminating between models with intractable likelihoods in epidemiology. *Computational Statistics and Data Analysis, 124*, 277-297.

Hancock, T. O., Broekaert, J., Hess, S., & Choudhury, C. F. (2020). Quantum choice models: A flexible new approach for understanding moral decision-making. *Journal of choice modelling, 37*, 100235.

Hancock, T. O., Hess, S., & Choudhury, C. F. (2018). Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. *Transportation Research Part B: Methodological, 107*, 18-40.

Huan, X., & Marzouk, Y. M. (2013). Simulation-based optimal Bayesian experimental design for nonlinear systems. *Journal of Computational Physics, 232*(1), 288-317.

Hunter, W. G., & Reiner, A. M. (1965). Designs for discriminating between two rival models. *Technometrics, 7*(3), 307-323.

Iraganaboina, N. C., Bhowmik, T., Yasmin, S., Eluru, N., & Abdel-Aty, M. A. (2021). Evaluating the influence of information provision (when and how) on route choice preferences of road users in Greater Orlando: Application of a regret minimization approach. *Transportation Research Part C: Emerging Technologies, 122*, 102923.

Kanninen, B. J. (2002). Optimal design for multinomial choice experiments. *Journal of Marketing Research, 39*(2), 214-227.

Kikuchi, C., Ferré, T., & Vrugt, J. (2015). On the optimal design of experiments for conceptual and predictive discrimination of hydrologic system models. *Water Resources Research, 51*(6), 4454-4481.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics, 22*(1), 79-86.

Masiero, L., Yang, Y., & Qiu, R. T. (2019). Understanding hotel location preference of customers: comparing random utility and random regret decision rules. *Tourism Management, 73*, 83-93.

Myung, J. I., & Pitt, M. A. (2009). Optimal experimental design for model discrimination. *Psychological review, 116*(3), 499.

Nowak, W., & Guthke, A. (2016). Entropy-based experimental design for optimal model discrimination in the geosciences. *Entropy, 18*(11), 409.

Pham, H. V., & Tsai, F. T. C. (2016). Optimal observation network design for conceptual model discrimination and uncertainty reduction. *Water Resources Research, 52*(2), 1245-1264.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological methodology*, 111-163.

Rose, J. M., & Bliemer, M. C. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews, 29*(5), 587-617.

Rose, J. M., Bliemer, M. C., Hensher, D. A., & Collins, A. T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological, 42*(4), 395-406.

Schöniger, A., Wöhling, T., Samaniego, L., & Nowak, W. (2014). Model selection on solid ground: Rigorous comparison of nine ways to evaluate B ayesian model evidence. *Water resources research, 50*(12), 9484-9513.

Sharma, B., Hickman, M., & Nassir, N. (2019). Park-and-ride lot choice model using random utility maximization and random regret minimization. *Transportation, 46*(1), 217-232.

Tierney, L., & Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the american statistical association, 81*(393), 82-86.

van Cranenburgh, S., Guevara, C. A., & Chorus, C. G. (2015). New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice, 74*, 91-109.

van Cranenburgh, S., Rose, J. M., & Chorus, C. G. (2018). On the robustness of efficient experimental designs towards the underlying decision rule. *Transportation Research Part A: Policy and Practice, 109*, 50-64.

Wong, S. D., Chorus, C. G., Shaheen, S. A., & Walker, J. L. (2020). A Revealed Preference Methodology to Evaluate Regret Minimization with Challenging Choice Sets: A Wildfire Evacuation Case Study. *Travel Behaviour and Society, 20*, 331-347.

Yu, J. G., & Jayakrishnan, R. (2018). A quantum cognition model for bridging stated and revealed preference. *Transportation Research Part B: Methodological, 118*, 263-280.

Zhao, Z., & Severini, T. (2017). Integrated likelihood computation methods. *Computational Statistics, 32*(1), 281-313.

Appendix 5.1

As shown in Section 5.3.2, the marginal likelihood $P(y|M_k, d)$ is defined as an integral over the distribution $f(\beta_k|M_k)$ of model parameters β_k. For simplification, we omit the general notions M_k and d in the Eq.5.4, the marginal likelihood $P(y)$ is expressed by

$$P(y) = \int P(y|\beta_k)f(\beta_k)d\beta_k.$$

Let $g(\beta_k) = ln(P(y|\beta_k)f(\beta_k))$. Consider a Taylor series expansion of $g(\beta_k)$ about $\bar{\beta}_k$, and $\bar{\beta}_k$ are the values of β_k that maximize the function $g(\beta_k)$. We only consider the first- and second-order terms of the Taylor series expansion, and neglect the third- and higher-order terms. Thus, the Taylor series expansion of $g(\beta_k)$ about $\bar{\beta}_k$ is given by

$$g(\beta_k) = g(\bar{\beta}_k) + (\beta_k - \bar{\beta}_k)^T g'(\bar{\beta}_k) + \frac{1}{2}(\beta_k - \bar{\beta}_k)^T g''(\bar{\beta}_k)(\beta_k - \bar{\beta}_k) + o(\| \beta_k - \bar{\beta}_k \|^2),$$

where the superscript ^{T} denotes the matrix transpose, g' is the vector of the first partial derivative of $g(\beta_k)$, and g'' is the Hessian matrix of second partial derivatives of $g(\beta_k)$. Since $g'(\bar{\beta}_k) = 0$, the first-order term drops out:

$$g(\beta_k) = g(\bar{\beta}_k) + \frac{1}{2}(\beta_k - \bar{\beta}_k)^T g''(\bar{\beta}_k)(\beta_k - \bar{\beta}_k) + o(\| \beta_k - \bar{\beta}_k \|^2).$$

The approximation is close to the true value only if the sample size is large enough. Now take the exponent of $g(\beta_k)$, and substitute it in the first equation, we get

$$P(y) = \int exp[g(\beta_k)]\, d\beta_k$$

$$\approx exp[g(\bar{\beta}_k)] \int exp[\tfrac{1}{2}(\beta_k - \bar{\beta}_k)^T g''(\bar{\beta}_k)(\beta_k - \bar{\beta}_k)]d\beta_k.$$

The Laplace method for integrals changes the integral of the above question into

$$P(y) \approx exp[g(\bar{\beta}_k)](2\pi)^{N_p/2}|\bar{\Sigma}|^{1/2},$$

where N_p is the number of parameters, and $\bar{\Sigma}$ is the variance-covariance matrix about $\bar{\beta}_k$.
Thus, we get

$$P(y) \approx P(y|\bar{\beta}_k)P(\bar{\beta}_k)(2\pi)^{N_p/2}|\bar{\Sigma}|^{1/2}$$

For the large sample, the approximation can take place at $\hat{\beta}_k$, which maximizes the likelihood $P(y|\,\beta_k)$. Adding the omitted $\mathrm{M_k}$ and d, the final equation of the mathematical approximation is given by

$$P(y|M_k, d) \approx P(Y|M_k, \hat{\beta}_k, d)P(\hat{\beta}_k|M_k)(2\pi)^{N_p/2}|\hat{\Sigma}|^{\frac{1}{2}}.$$

Here $\hat{\Sigma}$ denotes the variance-covariance matrix about $\hat{\beta}_k$, its inverse is the Fisher Information Matrix *F*.

Appendix 5.2

In the case study of Section 5.5.1, the whole dataset is also estimated on the µRRM model. The estimation result is presented in the below table.

Table 5.14 Model estimation results of the µRRM model

	Whole dataset
β_{Time}	-0.06
(*s.e. t-value*)	(*0.00 -14.2*)
β_{Cost}	-0.28
(*s.e. t-value*)	(*0.02 -11.9*)
β_{Trans}	-0.55
(*s.e. t-value*)	(*0.05 -10.7*)
μ	1.05
(*s.e. t-value*)	(*0.25 4.12*)
Choice observations	672
Null-LL	-738
Final LL	-626

6 Conclusions, reflections and future research

The past few decades have witnessed many advances in the development of choice behaviour models, one of which is the enhancement of the models' behavioural realism. Such a research line more or less coincides with developments in psychology, as psychology provides many valuable insights on behavioural aspects of decision making. Not surprisingly, many interesting behavioural phenomena revealed in psychology have been considered in choice modelling. In this thesis, we have particularly looked at a behavioural phenomenon, namely reference dependence. Reference dependence refers to a phenomenon that how people assess the outcome of a choice is largely determined by its comparison with some reference point; shifts of the reference point may give rise to reversals of preferences.

This thesis has investigated reference dependence empirically and provided new tools and techniques to effectively model it. Moreover, the last study of this thesis involves experimental design methods; it provides an innovative approach to collecting data that are most suitable for comparisons between different choice models. In this chapter, I would like to summarize the main findings of this work and, in light of these findings, discuss some broader implications and directions for future research.

This chapter is organised as follows. Section 6.1 gives an overview of the four studies conducted in this thesis and summarises the main findings of each study. Sections 6.2 and 6.3 reflect on this thesis from the behavioural and modelling perspectives respectively. Section 6.4 mainly focuses on some practical ways to deal with reference points. Section 6.5 discusses several interesting implications and strategies. The last section proposes several future research avenues.

6.1 Overview of studies in this thesis

- ***Empirical evidence for reference dependence***

This thesis starts with an empirical study (Chapter 2) to illustrate the importance of considering reference dependence in choice behaviour research. The study was conducted in the context of future transportation involving automated vehicles (AVs). In particular, it looks at the phenomenon of people having different opinions about death caused by AVs and conventional vehicles. Recently, as more and more AVs come into view, there is growing concern about safety issues surrounding this new technology. In some public debates and academic work, a view has been expressed that compared to fatalities caused by conventional vehicles, fatalities caused by AVs carry more weight to the general public, or in other words, the general public overweight the fatalities caused by AVs.

In the first study, we have put this view to the empirical test and explained it using reference dependence. In the choice experiments, *i*) we tested whether fatalities caused by AVs carry more weight than fatalities caused by conventional vehicles, and *ii*) we varied the reference levels of the number of fatalities caused by AVs and conventional vehicles respectively, to see whether the weight attached to different fatalities would change accordingly. Our results showed that AV fatalities did carry more weight than conventional fatalities; specifically, AV fatalities caused by technical failure were weighted around 1.5 times higher than conventional fatalities, and AV fatalities caused by deliberate misuse were weighted more than 2.2 times higher than conventional fatalities. More importantly, we found that the overweighting of AV fatalities was diminishing as the reference level of AV fatalities increased. This indicates that how people perceive the death caused by AVs is largely influenced by its reference level. Simply because the current number of AV fatalities is so low (currently there are no accidents involving AVs in many cities) that each additional AV fatality carries extra weight.

These findings are very important for transport policy implications. They suggest that *i*) cutting the number of annual traffic fatalities in half by implementing AVs would be already considered acceptable by the general public (in the Netherlands), *ii*) as the number of AV fatalities increases—as will inevitably be the case once they are allowed on roads—the extent to which they are overweighted will decrease. Overall, our findings suggest that the occurrence of traffic accidents involving AVs will help ameliorate people's excessive fear and concern about AV safety issues; once AVs become considerably safer than conventional vehicles, they should be allowed on the roads as soon as possible, to save lives that would otherwise have been lost in accidents involving conventional vehicles.

This empirical study highlights the importance of reference dependence in explaining choice behaviour—failure to accommodate reference dependence may lead to severe bias in understanding choice behaviour and resulting transport policies. In addition, this study also shows how conventional choice models deal with reference-dependent choices, which indicates the need for (new) models that can easily handle reference dependence.

- ***New tools for modelling reference dependence***

This thesis introduces two new tools for modelling reference dependence. The first tool is a series of new model specifications (Chapter 3). Built on the Random Regret Minimization (RRM) models (Chorus, 2010; van Cranenburgh et al., 2015), the new model specifications can not only capture reference dependence but also accommodate the relativity nature of people's response to differences. Specifically, so-called relative thinking was incorporated into the RRM model framework. Relative thinking means how people respond to differences is influenced by some reference level. In other words, a given difference seems big or small depending on the reference level. Based on the types of reference levels, we categorize relative thinking into "level-based relative thinking" and "range-based relative thinking": the former means that a given difference seems big or small depending on its original level, and the latter means a given difference seems big or small depending on its range size of the choice set. The two types of relative thinking were incorporated into the RRM models respectively, resulting in RRM-Level and RRM-Range models. The new models were compared with the linear-additive RUM model and conventional RRM models on four empirical datasets. The results showed that incorporating relative thinking into the RRM model can improve model fit and some of the improvements were very considerable, but this did not hold for all estimation results. There were also cases in which the new models perform worse than conventional RRM models. Moreover, when the RRM-Level and RRM-Range models were compared, there was no consistent result of which one was better than the other. Therefore, which model performs better than others is actually very dataset-specific.

The second tool is a new loss aversion model (Chapter 4). Loss aversion always comes with reference dependence. When people make choices, the outcomes of choices are often mapped as gains or losses against some reference point, and gains and losses receive a different response. Loss aversion means that losses receive a greater response than equivalent gains, or in other words, losses have a greater impact on choices than gains. The new loss aversion model is adapted from the specification of the μRRM model (van Cranenburgh et al., 2015). Compared with many other existing loss aversion models, the new model has a smooth function and its function is twice-differentiable in the full domain. Moreover, the new loss aversion model has separate parameters which govern the degree of loss aversion and the attribute importance respectively, as a result, the two effects can be disentangled from each other. We tested the robustness of the new loss aversion model on three empirical data. The model was compared with the linear-additive model and a piecewise-linear model (commonly used to model loss aversion). Our results showed that the new loss aversion model successfully captured loss aversion in behaviour; moreover, the degree of loss aversion in certain attributes was straightforwardly reflected by the magnitude of loss aversion. In terms of model fit, the new loss aversion model performed as well as the RUM model when there was no loss aversion in behaviour; it fit the data just as well as the piecewise-linear model when only loss aversion is presented in behaviour; but it was outperformed by the piecewise-linear model when a reverse behavioural pattern—gains are valued more importantly than losses—existed in behaviour.

- ***A new tool for collecting data***

This thesis also provides a new tool for collecting stated preference choice data. Specifically, the last study introduces an innovative method of constructing experimental designs into the choice modelling field. Currently, the mainstream design method is the so-called efficient designs, which focus on the reliable recovery of model parameters for a given model in a statistically efficient way. However, our method is from a different perspective: it focuses on how to generate data with the maximum model discrimination capability. Like many other recent choice modelling studies, this thesis put forward several new model specifications. The main business of such research is not to find the most reliable parameters for a specific model, but to select the model that best describes the underlying choice behaviour from a set of competing models. Thus, there is a need to propose an experimental design method suitable for this aim. In this thesis, we propose an innovative method of constructing experimental designs called *discriminatory designs*, which aim to generate experiments that lead to the data with the maximum model discrimination capability, thereby yielding the result that clearly favours the most plausible model over other competing models based on a given, limited number of choice observations. To construct such designs, we applied Bayes' theorem and the Bayesian Information Criterion approximation. The robustness of discriminatory designs was tested on both synthetic data and empirical data. The test results show that an elaborate discriminatory design can discriminate between different competing models. Specifically, for synthetic data, our method can generate larger model fit differences, yielding the result that clearly favours the "true" model (i.e., true DGP given the data), and for empirical data, if there is the "true" model given data (among all competing models), using discriminatory designs can maximize the likelihood of discriminating it from other competing models.

6.2 Behavioural reflections

6.2.1 Reference dependence in travel behaviour

This thesis has provided empirical evidence for reference dependence in travel choice behaviour. In Chapter 2, we found that how people assessed death caused by AVs was (largely) influenced by the current low number of AV fatalities (i.e. the current reference level). Chapter 3 has shown that travel behaviour can be also influenced by some external reference points provided by the choice context. For example, the attractiveness of a travel mode depends not only on how much faster (or cheaper) it is than other travel modes in the choice set but also on the range of travel times (or costs) of all travel mode options. In Chapter 4, the empirical tests of the loss aversion model have also shown that travellers' route choices are dependent on their current routes or reference routes.

When it comes to "reference-dependent preferences", a well-known statement is often jointly mentioned: "shifts in reference points may cause reversals of preferences". This can be explained by another notion—loss aversion. Loss aversion is closely related to the setting of the reference point: shifting the reference point is likely to turn a loss into a gain, and vice versa. Evidence for this statement can be found in this thesis. Chapter 2 has shown that people overweight the fatalities caused by AVs, compared to the fatalities caused by conventional vehicles. However, by artificially increasing the reference levels of AV fatalities, the extent to which they are overweighted (compared to conventional fatalities) substantially decreased or even eliminated. This indicates that shifts in the reference levels could change people's preferences. Interestingly, we also attempted to change the reference level of fatalities caused by conventional vehicles, but could not find the same effect as for AV fatalities. This means that how people assess fatalities caused by conventional vehicles could not be influenced by artificially changing their reference levels. A possible reason is that compared to new or emerging things (e.g., fatalities caused by AVs), people's judgment of conventional things (e.g., fatalities caused by conventional cars) is much less susceptible to some "external force". In other words, their reference levels of conventional things are relatively stable and not easily affected.

In this thesis, we have mainly looked at reference-dependent choice behaviour, but for analysts, it is very difficult to observe or identify people's reference points when they are making travel-related decisions. In Chapter 2, we attempted to "manipulate" people's reference points for AV fatalities and conventional fatalities by offering them information about reference points. We found that not every respondent complied with the reference points that we provided. In Chapter 3, it seems that any information gained from the choice context can become a reference point for decision makers. Specifically, we found that decision makers may use "the original level of attributes" or "the range of attribute levels in the choice set" as a reference when assessing attribute-level differences between choice alternatives.

Timmermans (2010) argues that an important difference between travellers and gamblers in Prospect Theory is that "travellers experience the consequences or outcomes of their decisions, and more importantly, they can adapt their behaviour to influence the experienced outcome". Therefore, for travellers, they may have a very stable reference point for routine or habitual travel choices, while for non-routine choices (e.g. new transport services), they may continue to update their reference points until some stable point is reached. As discussed above, in Chapter 2, we found people's reference points for fatalities caused by conventional vehicles were relative stable, while their reference points for AV fatalities could be influenced by the provided information. We concluded that people overweight the fatalities caused by AVs, because of the current low number of traffic accidents involving AVs. Once more and more AV-related accidents inevitably happen, people will also learn and update their reference points for AV fatalities, and as a result, the extent to which they are weighted more than conventional fatalities will eventually disappear.

6.2.2 Loss aversion in travel behaviour

Ample experimental and empirical evidence has shown that there is loss aversion in decision making, especially when it involves the risk of monetary loss. This thesis also provides empirical evidence for loss aversion in travel behaviour. However, some studies have raised questions about the relevance of loss aversion in the context of travel choice behaviour. A frequently debated question is *whether the loss (or gain) of travel behaviour is truly perceived as a loss (or gain) by travellers*? Take departure time choices as an example. Usually, we consider late arrivals as losses, but how about early arrivals? Timmermans (2010) argued that early arrivals can be also seen as losses for travellers. Similarly, Senbil & Kitamura (2004) and Jou & Kitamura (2002) treated early arrivals (earlier than the acceptable earliest time) as losses when modelling departure time choices. As for late arrivals, they may be viewed as losses, the resulting outcome can be easily compensated by, for example, working late or working more efficiently. This implies that, unlike gambling problems, in travel behaviour, there seems to be no clear-cut distinction between losses or gains; a loss (or gain) can be easily transferred, changed, or even compensated.

The second question related to loss aversion in travel behaviour is *whether loss aversion plays an important role in routine behaviour such as commuting route choices*? In Chapter 4, we assessed loss aversion in route choice behaviour and found that loss aversion did not occur in some conventional travel attributes, like free-flow travel time and running cost, but in the attribute involving uncertainty. For example, in our first route choice data, no loss aversion was found in any time or cost attributes. In our second route choice data, for commuters, we found that loss aversion only occurred in attributes such as *stopped/crawling time* and *contingency* *time for arrival,* but not in attributes that commuters are accustomed to, such as free-flow time and slowed-down time. But for non-commuters, loss aversion was found in free-flow travel time. Interestingly, in these two route choice data, we found an opposite behavioural pattern, that is gains are weighted more importantly than losses. For example, such behavioural pattern was found in slow-down time and running cost in the first route choice data, and it was found in running cost for commuters in the second route choice data. Moreover, we also examined whether loss aversion played a role in how people assess traffic fatalities caused by AVs and by conventional vehicles. Loss aversion was only found in AV fatalities (deliberate misuse of AVs), but not in the fatalities caused by conventional vehicles. Overall, our empirical findings seem to provide evidence that loss aversion is more likely to occur in the response to choices involving uncertainty, but may disappear in certain routine or conventional choices.

6.3 Methodological reflections

6.3.1 Modelling challenges of identifying the reference point

The reference point is the threshold that distinguishes losses from gains, so determining the reference point is undoubtedly the key to model reference dependence. In the context of travel behaviour, there seems no consensus value on the reference point when dealing with travel-related decision-making (Avineri & Bovy, 2008). In gambling problems, a natural reference point is whether the gambler wins money (i.e., €0). But there is no such natural reference point for travel behaviour, thus, determining the reference point in travel behaviour is much more complicated. For example, the reference point for travel behaviour choices is often endogenous, and it may vary from traveller to traveller, as well as over time and circumstances. In addition, a traveller's reference point can be also influenced by some external factors, such as implicit or explicit information, norms, and social comparisons (Kahneman & Tversky, 1979).

Here, I summarize two approaches that are mainly used to deal with reference points in the current reference-dependent models. The first approach is to explicitly or implicitly define the reference point in the model specification. Two well-known examples are the contextual concavity model (Kivetz et al., 2004) and the RRM model. Specifically, in the contextual concavity model, the reference point is set to *the least preferred attribute level*, and in the RRM models, the reference point is implicitly defined as *the attribute level of every other unchosen alternative*. Chapter 3 put forward a series of new model specifications, RRM-Level and RRM-Range models. Like conventional RRM models, the new models assume that people use unchosen alternatives as to the reference when making comparisons between choice alternatives. In addition, the models also incorporate another layer of reference dependence. They explicitly assume that how people respond to differences is also dependent on some reference level, that is the attribute level of the chosen alternative or the range size of the attribute in the choice set.

The second approach is to leave the reference point undefined in the model specification, allowing analysts to test possible reference points. The value function and the new loss aversion model (Chapter 4) both fall into this category. In the model specification, there is a variable (e.g. X_r) that denotes the level of the reference, but the reference is not specified. Compared to the first approach, this approach relaxes the restriction on the setting of reference points in the model, allowing multiple potential reference points to be tested when examining reference dependence.

What both approaches have in common is that the reference point is assumed a priori (by model specifications or by analysts), which greatly simplifies, or even neglects, the complexity and variability of actual reference points of decision makers. Recently, there have been several models which allow estimating reference points rather than assuming it a priori, such as Bahamonde-Birke (2018). However, these models are often limited to small-scale applications (e.g., reference dependence is only tested on one attribute), due to the dependence on other parameters of the model that need to be estimated as well and the lack of systematic search algorithms that guarantee the model convergence (Avineri & Bovy, 2008). Another noteworthy approach of framing reference points is to use a fuzzy representation of reference points. Avineri (2009) argued that '*a traveller does not necessarily have a crisp and sharp definition of a reference point in mind, therefore a sound assumption may be that the perception of reference point in the mind of a traveller is vague or fuzzy rather than crisp*.' This study assumes that the actual reference point of a traveller is not a specific value, but an interval of values. Similarly, Kőszegi & Rabin (2006) assume that decision makers' reference points are not a specific value but stochastic, and that the mean value is based on rational expectations they held in the recent past.

Overall, it is still very challenging to identify and frame travellers' actual reference points when modelling reference dependence. This thesis has put forward several new reference-dependent model specifications. But our modelling approaches do not address the question of how to determine travellers' reference points, as they deal with it in the same way as the mainstream models, assuming that the reference point is a priori. The next section will discuss some practical ways of dealing with reference points in travel behaviour research.

6.3.2 Model performance dependent on data (sets)

In this thesis, we empirically compared the new model specifications with several benchmark or conventional models. We found that none of these models could outperform others on all data sets in terms of model fits. This means that the model's goodness of fit was highly *dataset-specific*.

The RRM-Level and RRM-Range models are used to capture relative thinking (Weber effect or the range effect) in behaviour respectively. The models are expected to perform better than their original counterparts when relative thinking plays a role in the decision-making process. This raises a question as to *what kind of choice situation or context is more likely to trigger people to show relative thinking when making choices*. Chapter 3 has given one possible answer: when choice situations contain various attribute levels or various range sizes in the choice set, relative thinking is more likely to be triggered, and therefore, the RRM-Level and RRM-Range models would perform better on such data. Another question then arises as to *what kind of data is likely to have the characteristics of various attribute levels or range sizes*. The answer that comes to my mind is revealed preference (RP) data. RP data are choices that are actually made during real travel or activities. In real choice situations, choice options may vary considerably in terms of their certain attributes. For example, in the case of real shopping choices, you may go shopping at a nearby convenience store, or go to a large shopping centre that is an hour's drive away. In this case, travel time may vary greatly. But for stated preference

(SP) data, which are collected via hypothetical choice scenarios, the levels of attributes between choice alternatives usually differ little. For example, in stated choice experiments, attributes are often assigned levels of the same order of magnitude, resulting in little variation in attribute levels. We can expect that relative thinking is more likely to occur with RP data than with SP data. Therefore, when estimating RP data, we recommend examining relative thinking in choice behaviour.

In Chapter 4, the new loss aversion model was compared with the linear-additive RUM model and the piecewise-linear model. The new loss aversion model is a generalized model that can accommodate different degrees of loss aversion choice behaviour. Our model estimation results showed that the new loss aversion model performs as well as the linear-additive RUM model on the data where there is no loss aversion, and performs as well as the piecewise-linear model on the data where there is only loss aversion in choice behaviour. But when the opposite behaviour pattern—gains are valued more importantly than losses—presents in behaviour, the new loss aversion model is outperformed by the piecewise-linear model. These results mean that the new loss aversion model is a promising model for capturing loss aversion, but its performance (in terms of model fit) is dependent on specific data.

6.4 How to deal with the reference point?

The previous section has discussed the challenges of identifying the reference point in travel behaviour. Based on the literature and our findings, I would like to provide several methods on how to deal with the reference point in practice. These methods are not limited to the models in the thesis, but rather can be used for the reference dependence research in general.

- *Frame attributes as changes to the reference point*

The first method is to frame attributes as changes to the reference point. For example, travel time is set to “30 minutes more than your current trip”. By framing attributes as changes (losses or gains), we do not have to bother to identify traveller’s reference points. The modelling of reference dependence is only concerned with losses or gains.

This method is easy to apply, but we need to pay attention to two aspects when using it. First, although we do not have to identify the specific reference point of each respondent, we need to ensure that respondents facing the same choice sets have a similar magnitude of reference points. In the previous example, a 30-minute increase in travel time may seem like a lot (unacceptable) for travellers who currently travel 10 minutes, while it might be acceptable for travellers who currently travel 3 hours. Therefore, SP choice experiments should be customised for different categories of respondents (or data analysis is performed separately for different categories of respondents). For example, travellers are categorized into commuters and non-commuters, or short-distance travellers and long-distance travellers, etc. Second, it should be noted that the framing effect may affect the degree of loss aversion. The framing effect means that individuals’ preferences may be influenced by the way that the decision problem is presented (Tversky & Kahneman, 1981). When the loss is explicitly stated (or emphasized) in the choice set, it is very likely to increase the degree of loss aversion. Therefore, when attributes are framed as only losses or gains, it might increase or decrease the degree of loss aversion.

- *Test for multiple possible reference points*

The second method is to test for multiple possible reference points when modelling choice behaviour. Similar efforts can be seen in many studies, such as Jou & Kitamura (2002), Senbil & Kitamura (2004), Stathopoulos & Hess (2012), and Wang et al. (2019). For example, Stathopoulos & Hess (2012) tested for three plausible reference points of travel fare in modelling commuting choices, namely the fare of the *current*, *acceptable*, and *ideal* trip. They found that compared with the commonly used current fare, using the acceptable and ideal fares as the reference point could lead to better model performance and higher degrees of asymmetry in response to losses and gains. Therefore, it would be interesting and valuable to test for multiple possible reference points and then compare their results, and explore, for example, which reference point can lead to the best model fit. In addition to the current, acceptable and ideal conditions, possible reference points can be the latest trip, expectations, the preferable condition, or trip information provided by Google maps. It should be noted that using the reference point reported by respondents, such as "your ideal travel time", may lead to some endogeneity bias.

- "*Manipulate" respondents' reference points*

As discussed above, individuals' reference points can be influenced by some external factors (Kahneman & Tversky, 1979). So the third method that I come up with is to "manipulate" people's reference points by proving external information. In this thesis, we have applied this method in our first study. By providing the current situation and experts' projections of the numbers for AV fatalities and conventional fatalities, we attempted to shape respondents' reference levels for these fatalities. However, it is important to note that not every respondent would comply with the provided reference points, just like what has been shown in our results. Moreover, our results suggest that people's beliefs and judgement about old or conventional things are relatively stable and less likely to change than about new or emerging things. Therefore, this method would not be very effective if we try to manipulate reference points for conventional things.

- *Reference points for travel time: based on average commuting time*

The fourth one is concerned with the reference point for travel time. Avineri & Bovy (2008) suggest that in the absence of the reference point for travel time, one can assume that the reference level of travel time is related to the average travel time experienced by the target traveller group. They mentioned that, for example, the average one-way commuting time is about 30 minutes in many countries, which can be used as the reference point of travel time in empirical studies. Of course, this value varies depending on different counties and areas. In addition to average values, Avineri & Bovy (2008) also suggested median values or mode values. These values, average/median/mode travel times, can be obtained from large-scale national or regional travel surveys. Importantly, when using average/median/mode values of commuting travel time as the reference point, attention should be paid to spatial and regional characteristics, travel purposes, target respondents, etc.

6.5 Implications and strategies

This thesis has looked at three behavioural phenomena: reference dependence, relative thinking and loss aversion. They have important policy implications for areas related to economics and management. In this section, I would like to list several interesting implications, primarily in but not limited to the field of transportation.

- Because of (level-based) relative thinking, people are more willing to travel 20 minutes to save $5 on a $15 calculator than on a $125 jacket (Tversky & Kahneman, 1981). This indicates that when the same amount of money is saved, people are more willing to travel a long distance to the stores that sell lower-priced goods. For example, we can expect that people are willing to drive one hour to a supermarket to save €50 on grocery purchases[20], but not willing to drive one hour to a computer shop to save €50 on laptops. This has important implications for urban planning and transport planning. For instance, if the government wants to boost the economy of suburban areas, building a discount supermarket would attract more people to come than building a discount electronics store. Moreover, if there is a large discount shopping centre in the suburbs, it would be cost-effective to have a direct bus line (as people are willing to come to this place for shopping).
- Relative thinking also has many implications for product marketing. For example, if a car dealer wants to sell car maintenance services (e.g. €300) to the client, it would be better to increase the car price from, e.g., €3000 to €3300 (and free maintenance) than to sell them separately. Another example is about applying relative thinking into the sales strategy of large stores, such as department stores. Consumers are less sensitive to small price increases of high-priced products, compared to low-priced products. So the strategy is to increase the prices of high-priced products and decrease the prices of low-priced products to gain more profits.
- Furthermore, range-based relative thinking has important implications for marketing. For example, for the same type of products, it is better to provide consumers with more choice options at different prices (enlarge the range size). For example, consumers would find a €500 TV more attractive when it is sold alongside other €400, €600, and €700 TV sets than when it is sold only with €400 TVs. Because when the price range is expanded (from €400 to €700), the price increase of €100 seems to be less, compared to the situation in which the range is from €400 to €500. Similarly, when the company launches new products, it is better to launch multiple product lines at different prices. This would prompt commuters to choose higher-priced products.
- Reference dependence can be applied to some government incentive policies, for example incentives for electric vehicles (EVs). EVs are seen as a promising solution to transportation decarbonisation, therefore governments around the world are offering all kinds of monetary and non-monetary incentives to accelerate EV adoption. In terms of monetary incentives, cash-back subsidies would be more attractive to consumers than

[20] Evidence can be found in reality. Every Saturday, many Swiss citizens (e.g., those from Zurich) spend about one hour travelling to a German supermarket (i.e., Aldi in Konstanz) on the border for their grocery shopping, because prices are lower in Germany.

equivalent discounts. For example, a cashback of €4,000[21] on a €40,000 new EV sounds more beneficial to consumers, compared to the discount of 10% off. Because the €4000 cashback is seen as an increase from €0 to €4000, while the 10% off is seen as a discount from 0% to 10%. The total amount of subsidies is the same, but different frames of reference may elicit different responses.

- Reference dependence and loss aversion have important implications for travel demand forecasts. According to loss aversion, people would evaluate losses more importantly than equivalent gains. Therefore, the loss brought about by a new transport project would have a greater impact on travellers' choices than the gain. In travel demand analysis, without accounting for loss aversion, benefits brought about by a new travel project would be overestimated and its disadvantages would be underestimated. As a consequence, the travel demand of the new transport project would be over-optimistically estimated. In fact, the over-optimistic forecasting of travel demand is a common problem in actual demand forecast in many countries (Næss et al., 2006; Schmitt, 2016; Voulgaris, 2019). One possible reason might lie in the used models which fail to account for loss aversion. However, loss aversion may be absent in routine or habitual choices. Thus, as time goes by, loss aversion may diminish or even vanish. Therefore, in my point of view, loss aversion needs to be considered (at least) in short-term demand forecasts, for example, the initial attractiveness of a new light rail.
- Loss aversion also has implications for transport appraisals. Transport appraisals are mainly about quantifying economic benefits of (new) transport projects. A considerable part of the quantified benefits of transport projects (road schemes) come from travel time savings, especially *small* travel time savings (Daly et al., 2014). The treatment of small travel time savings has been much debated. The debate revolves around the question of whether the value attached to small travel time savings should be the same as the value attached to large travel time savings. According to the answers to this question ("should" or "should not"), these are two corresponding approaches to the evaluation of travel time savings: the constant unit value (CUV) approach—a constant unit value is attached to all travel time savings irrespective of the size—and the discounted unite value (DUV) approach—below a certain threshold, the unit value to travel time savings is discounted or even zero. In this thesis, some findings seem to support the DUV approach: according to Weber's law (Chapter 3), saving a small travel time (e.g., five minutes) seems less important if the trip length is long (e.g., one hour), and as gains carry fewer weights than losses (Chapter 4), a small travel time saving (gains) might be negligible. Therefore, it seems that travellers cannot benefit from small travel time savings. However, in actual transport appraisals, small travel time savings (less than five minutes) account for a significant portion of travel time savings. Adding up small travel times can lead to significant travel time savings that would be valued by travellers. In my point of view, although some stated preference results may suggest that small travel time savings (below 5 minutes) are not valued by travellers, small travel time savings should be valued (regardless of the trip length) in transport appraisals.

[21] According to the Netherlands' new subsidy scheme (June, 2020), the purchase or lease of a new EV with an original value of €12,000 to €45,000 will receive a €4,000 subsidy.

However, It seems that more empirical work is needed to answer the question of whether the values should be constant or discounted.

- Another controversial issue is whether loss aversion should be considered in the appraisal of long-term transport projects. Chapter 3 has introduced a new loss aversion model and also discussed its potential for application in transport appraisals. But the use of loss aversion remains questionable. As discussed in the behavioural findings, our empirical results have shown that loss aversion is more likely to occur in people's response to new or emerging things, compared to routine or conventional things. As time goes by, a new transport project (e.g., transport infrastructure) may no longer be a fresh stimulus for travellers, then loss aversion may disappear. Moreover, as travellers become more accustomed to the new transport infrastructure, their reference points will also change over time, and so will the losses and gains. Thus, the loss aversion penalty may diminish over time. Therefore, accounting for loss aversion in long-run transport appraisals may lead to some bias.

6.6 Future research

Finally, I would like to end this thesis by presenting several future research directions. In this thesis, each chapter has concluded several avenues for future research and these will be not reiterated here. I will mainly discuss some other recommendations from the modelling aspects.

First, given that new modelling tools are introduced into the field, an obvious direction for future research is further empirical comparisons between the new model specifications and conventional models. Model comparisons are better based on data that allow for clear interpretation of estimation results in terms of conceptual differences between the models. For example, in Chapter 3, we concluded that the RRM-Level or RRM-Range model is expected to perform better than conventional RRM models on data containing various attribute levels or various ranges across choice sets. More empirical comparisons should be conducted on such data. For loss aversion, we found that it may not occur in response to habitual or routine choices. Thus, when comparing the new loss aversion model with the RUM-MNL model, it is better not to estimate on habitual or routine choice data.

When it comes to model comparison, discriminatory designs proposed in Chapter 5 can come to play a role. In addition to using above mentioned certain sorts of data (like non-habitual choices), we can also use discriminatory designs to generate data that are optimised for discriminating between models. In this thesis, we only test the robustness of discriminatory designs on model comparison between the RUM model and the RRM model. Thus, the second research direction is to test the robustness of discriminatory designs on discrimination between other new recently-introduced choice models, including the new model specifications proposed in this thesis.

Another research direction related to discriminatory designs is to explore whether discriminatory designs can be used for model comparison between any different choice models. In this thesis, we focused on model comparison between non-nested models, for example, the RUM model and the RRM model are based on two different decision rules. But for the RRM-Level and RRM-Range models, they are built based on the RRM models. In this case, can we still use discriminatory designs to generate optimal data for discriminating between the RRM-level/RRM-Range model and the RRM model? What is the discrimination capability of discriminatory designs if we use them for model comparison between nested choice models? Since this thesis is the first to explore such a design method in the choice modelling field, there is certainly much more to be explored in the future.

References

Avineri, E. (2009). Incorporating fuzzy reference points into applications of travel choice modeling. In *Applications of Soft Computing* (pp. 221-229): Springer.

Avineri, E., & Bovy, P. H. (2008). Identification of parameters for a prospect theory model for travel choice analysis. *Transportation Research Record, 2082*(1), 141-147.

Bahamonde-Birke, F. J. (2018). Estimating the reference frame: A smooth twice-differentiable utility function for non-compensatory loss-averse decision-making. *Journal of choice modelling, 28*, 71-81.

Chorus, C. G. (2010). A new model of random regret minimization. *European Journal of Transport Infrastructure Research, 10*(2).

Daly, A., Tsang, F., & Rohr, C. (2014). The value of small time savings for non-business travel. *Journal of Transport Economics and Policy (JTEP), 48*(2), 205-218.

Jou, R.-C., & Kitamura, R. (2002). *Commuter departure time choice: a reference point approach.* Paper presented at the 9th Meeting of the Euro Working Group on Transportation.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Kivetz, R., Netzer, O., & Srinivasan, V. (2004). Alternative models for capturing the compromise effect. *Journal of Marketing Research, 41*(3), 237-257.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics, 121*(4), 1133-1165.

Næss, P., Flyvbjerg, B., & Buhl, S. r. (2006). Do road planners produce more 'honest numbers' than rail planners? An analysis of accuracy in road-traffic forecasts in cities versus peripheral regions. *Transport Reviews, 26*(5), 537-555.

Schmitt, D. (2016). *A Transit Forecasting Accuracy Database: Beginning to Enjoy the 'Outside View'*. Retrieved from

Senbil, M., & Kitamura, R. (2004). *Reference points in commuter departure time choice: a prospect theoretic test of alternative decision frames.* Paper presented at the Intelligent Transportation Systems.

Stathopoulos, A., & Hess, S. (2012). Revisiting reference point formation, gains–losses asymmetry and non-linear sensitivities with an emphasis on attribute specific treatment. *Transportation Research Part A: Policy Practice, 46*(10), 1673-1689.

Timmermans, H. (2010). On the (ir) relevance of prospect theory in modelling uncertainty in travel decisions. *European Journal of Transport and Infrastructure Research, 10*(4).

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.

van Cranenburgh, S., Guevara, C. A., & Chorus, C. G. (2015). New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice, 74*, 91-109.

Voulgaris, C. T. (2019). Crystal Balls and Black Boxes: What Makes a Good Forecast? *Journal of Planning Literature, 34*(3), 286-299.

Wang, P., Sun, L., Niraj, R., Chung, J., & Su, M. (2019). Incorporating reference products into modeling consumer choice decision: A mixtures-of-experts model. *Decision Support Systems, 119*, 85-95.

Summary

People make all kinds of choices every day, such as driving to work rather than taking public transport. Many of these choices have a direct impact on demand for products, services or public infrastructures. Understanding people's choice behaviour can not only infer people's preferences for certain products or services but more importantly make future demand forecasts. Over the last fifty years, there has been a steadily growing interest in applying a quantitative statistical method, discrete choice modelling, to study individual and household choice behaviour. Discrete choice models provide a theoretically robust and tractable tool for modelling and analysing various choices across many fields such as transport, health, and marketing.

The past few decades have witnessed many advances in the development of discrete choice models, one of which is the enhancement of the models' behavioural realism. Such a research line more or less coincides with developments in psychology, as psychology provides many valuable insights on behavioural aspects of decision making. Not surprisingly, many interesting behavioural phenomena revealed in psychology have been considered in choice modelling. In this thesis, we particularly look at a behavioural phenomenon, namely reference dependence and incorporate it into choice modelling techniques. Reference dependence means that how people assess the outcome of a choice is largely determined by its comparison with some reference point; shifts of the reference point may give rise to reversals of preferences. Another behavioural phenomenon often accompanying reference dependence is loss aversion. It means that the outcome of a choice is viewed as gains or losses relative to the reference point, and losses have a greater impact on the choice than equivalent gains.

Currently, there are several choice modelling methods that account for reference dependence. However, we still see many research opportunities to (further) develop reference dependence modelling techniques specifically for travel behaviour research. This thesis aims to enrich the understanding of reference dependence and extend the frontier of reference dependence modelling methods. More specifically, there are three main goals of this thesis:

- to provide a more profound understanding of the importance of reference dependence in travel behaviour analysis and transport policy development;
- to provide new tools and techniques to effectively model reference dependence;
- to provide an innovative method of collecting data that are most suitable for comparison between different choice model specifications.

To this end, four studies have been conducted in this thesis. Each study is a chapter. Before introducing the new formulations of modelling reference dependence, this thesis begins with an empirical case study. It shows how important the reference effect can be in explaining behavioural phenomena and formulating transport policies. The next two chapters present methodological contributions to modelling reference dependence. Specifically, Chapter 2 incorporates relative thinking into the RRM framework, and Chapter 3 presents a new loss aversion model. The last chapter focuses on the method of generating choice experiments. It introduces an innovative way of constructing experimental designs, called discriminatory designs, to discriminate between different choice models in terms of model fit comparison. The main findings of each chapter are summarised below.

Chapter 2 provides empirical insights into the importance of considering reference dependence in explaining behavioural phenomena in transportation and resulting transport policies. The study is conducted in the context of future transportation involving automated vehicles (AVs). Specifically, it looked at whether—and if so, why—fatalities caused by AVs are weighted more heavily by the general public than fatalities caused by human drivers in conventional vehicles. The main empirical findings are as follows. First, fatalities caused by AVs indeed carry more weight than fatalities caused by conventional vehicles. Specifically, AV fatalities caused by technical failure are weighted around 1.5 times higher than conventional fatalities, and AV fatalities caused by deliberate misuses are weighted more than 2 times higher than conventional fatalities. Second, the overweighting of fatalities caused by AVs, compared to conventional fatalities, can be (partly) explained by the reference level effect: because the current levels of AV fatalities are so low that any additional AV fatality carries more weight. By artificially increasing the reference levels of AV fatalities, we are able to substantially reduce and even eliminate the overweighting of AV fatalities. The findings have important transport policy implications. First, this study suggests that cutting the number of annual fatalities in half by implementing AVs would already be considered acceptable by citizens in the Netherlands. Second, it also implies that as the number of AV fatalities goes up, the extent to which they will be overweighted will decrease. This means somewhat ironically, the occurrence of accidents involving AVs will help alleviate concerns and anxieties about AV safety issues. Third, as for transport policymaking, once AVs have become considerably safer than conventional vehicles, they should be allowed on the road as soon as possible, to speed up their learning process (making them even safer) and to save, during this process, lives that would otherwise have been lost in accidents involving conventional vehicles.

Chapter 3 is a methodological contribution to enrich the literature on reference-dependent modelling by introducing new model formulations. The new model specifications were based on the well-established RRM model framework; these new models can not only accommodate reference-dependent behaviour but also take into account the relative nature of how people respond to differences—referred to as relative thinking. Specifically, we consider two types of relative thinking: *i*) level based—people's response to an attribute difference is influenced by the original level of that attribute—and *ii*) range-based—people's response to an attribute difference is influenced by the range of attribute levels in the choice set. These two types were incorporated into the RRM model framework respectively, resulting in several new model specifications. These new model specifications were comprehensively compared with a series of benchmark or conventional models, the linear-additive RUM model and the RRM models, using four empirical data sets. The used data involved both stated and revealed preference data, including route choices, mode choices, and shopping destination choices. The model comparison results show that *i*) there is great potential for improving model performance (in terms of model fits) by incorporating relative thinking into the RRM models and some of the improvements in model fits are quite substantial; *ii*) the new models are not always superior to their original counterparts, they perform equally well as or even worse than their originals in some cases; *iii*) when the level-based RRM model and the range-based RRM model are compared, no consistent results are found in terms of which model always performs better than the other. These findings suggest that there is no "perpetual winner" among these models; The answer to the question of which model performs better is actually data set-specific.

In Chapter 4, we introduce a new loss aversion model. This new model was adapted from the specification of the μRRM model. The main merits of this model, compared with most previous loss aversion models, are *i*) it has a smooth function and *ii*) it can straightforwardly show the degree of loss aversion of multiple attributes, which is done by estimating an attribute-specific parameter μ_m (as the one in the μRRM model). The robustness of the new loss aversion model is tested using three empirical datasets. Specifically, it is compared with the linear-additive RUM model and a piecewise-linear model which is commonly used to capture loss aversion. Model estimation results show that the new loss aversion model performs as well as the RUM model when there is no loss aversion in behaviour; it fits the data just as well as the piecewise-linear model when only loss aversion presents in behaviour; and it is outperformed by the piecewise-linear model when an opposite behavioural pattern—gains are valued more importantly than losses—exists in behaviour. To conclude, the new loss aversion model is an effective modelling tool to model loss aversion, but the convexity of its loss aversion function may sometimes jeopardise the model performance in terms of model fit.

Chapter 5 provides a new tool for collecting stated preference choice data. Specifically, it introduces an innovative method of constructing experimental designs into the choice modelling field. Currently, the mainstream design method is the so-called efficient designs, which focus on the reliable recovery of model parameters for a given model in a statistically efficient way. However, our method is from a different perspective: it focuses on how to generate data with the maximum model discrimination capability. Like many other recent choice modelling studies, this thesis put forward several new model specifications. The main business of such research is not to find the most reliable parameters for a specific model, but to select the model that best describes the underlying choice behaviour from a set of competing models. Thus, there is a need to propose an experimental design method suitable for this aim. In this thesis, we propose an innovative method of constructing experimental designs called *discriminatory designs*, which aim to generate experiments that lead to the data with the maximum model discrimination capability, thereby yielding the result that clearly favours the most plausible model over other competing models based on a given, limited number of choice observations. To construct such designs, we apply Bayes' theorem and the Bayesian Information Criterion approximation. The robustness of discriminatory designs is tested on both synthetic data and empirical data. The test results show that an elaborate discriminatory design can discriminate between different competing models. Specifically, for synthetic data, our method can generate larger model fit differences, yielding the result that clearly favours the "true" model (i.e., true DGP given the data), and for empirical data, if there is the "true" model given data (among all competing models), using discriminatory designs can maximize the likelihood of discriminating it from other competing models.

In conclusion, this thesis provides a profound understanding of reference dependence in travel behaviour and enriches the toolkit for choice modellers to effectively model and analyse choice behaviour. The empirical and methodological contributions of this thesis have important implications for travel demand forecasts and transport policy development.

About the author

Bing Huang was born on March 27th, 1990 in Zhengzhou, China. She obtained a Bachelor's degree in Traffic Engineering from Dalian Jiaotong University in 2013 and a Master's degree in Transportation planning and Management from University of Shanghai for Science and Technology in 2016. In September 2016, supported by the China Scholarship Council, Bing joined the Transport & Logistics group of TU Delft and started her PhD journey. Bing's PhD research focuses on modelling and analysing people's choice behaviour primarily in Transportation. During her PhD, Bing has conducted several empirical and methodological research topics. For empirical research, Bing specifically looked at people's choice behaviour in the face of emerging technologies, such as the general public's acceptance of automated vehicles and consumers' preferences for vehicle-to-grid technology. For methodological research, Bing developed several new choice models that incorporate behavioural and psychological elements (reference dependence, loss aversion and relative thinking) and proposed a novel method of constructing experimental designs, called discriminatory designs. Besides research, Bing is also a sports enthusiast; she likes Hip-hop dance, mountain biking, and skiing.

Journal publications and conference contributions

Huang, B., van Cranenburgh, S., & Chorus, C. G. (2020). Death by automation: Differences in weighting of fatalities caused by automated and conventional vehicles. European journal of transport and infrastructure research, 20(3), 71-86.

Huang, B., Meijssen, A. G., Annema, J. A., & Lukszo, Z. (2021). Are electric vehicle drivers willing to participate in vehicle-to-grid contracts? A context-dependent stated choice experiment. Energy Policy, 156, 112410.

Huang, B. (2022). Incorporating relative thinking into the Random Regret Minimization model framework. Submitted to the Journal of Choice modelling.

Huang, B. (2022). Optimal experimental designs for discriminating between choice models. Submitted to Transportation Research Part B.

Huang, B., van Cranenburgh, S., & Chorus, C. G. (2019). Choice Experimental designs optimized for discriminating between different choice models. Contribution to the Sixth International Choice Modelling Conference, 2019, Kobe, Japan.

Huang, B., van Cranenburgh, S., & Chorus, C. G. (2018). Death by automation? Differences in weighting of fatalities caused by automated and conventional vehicles. Contribution to the Fifteenth International Conference on Travel Behaviour Research, 2018, Santa Barbara, United States.

Huang, B., van Cranenburgh, S., & Chorus, C. G. "It's the relativity, stupid!" Testing Weber's law in utility-based and regret-based models of travel behaviour. Contribution to the sixth Symposium of the European Association for Research in Transportation, 2017, Haifa, Israel.

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 275 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Huang, B., *Assessing Reference Dependence in Travel Choice Behaviour,* T2022/6, May 2022, TRAIL Thesis Series, the Netherlands

Reggiani, G., *A Multiscale View on Bikeability of Urban Networks*, T2022/5, May 2022, TRAIL Thesis Series, the Netherlands

Paul, J., *Online Grocery Operations in Omni-channel Retailing: opportunities and challenges,* T2022/4, March 2022, TRAIL Thesis Series, the Netherlands

Liu, M., *Cooperative Urban Driving Strategies at Signalized Intersections,* T2022/3, January 2022, TRAIL Thesis Series, the Netherlands

Feng, Y., *Pedestrian Wayfinding and Evacuation in Virtual Reality*, T2022/2, January 2022, TRAIL Thesis Series, the Netherlands

Scheepmaker, G.M., *Energy-efficient Train Timetabling,* T2022/1, January 2022, TRAIL Thesis Series, the Netherlands

Bhoopalam, A., *Truck Platooning: planning and behaviour,* T2021/32, December 2021, TRAIL Thesis Series, the Netherlands

Hartleb, J., *Public Transport and Passengers: optimization models that consider travel demand,* T2021/31, TRAIL Thesis Series, the Netherlands

Azadeh, K., *Robotized Warehouses: design and performance analysis,* T2021/30, TRAIL Thesis Series, the Netherlands

Chen, N., *Coordination Strategies of Connected and Automated Vehicles near On-ramp Bottlenecks on Motorways,* T2021/29, December 2021, TRAIL Thesis Series, the Netherlands

Onstein, A.T.C., *Factors influencing Physical Distribution Structure Design*, T2021/28, December 2021, TRAIL Thesis Series, the Netherlands

Olde Kalter, M.-J. T., *Dynamics in Mode Choice Behaviour,* T2021/27, November 2021, TRAIL Thesis Series, the Netherlands

Los, J., *Solving Large-Scale Dynamic Collaborative Vehicle Routing Problems: an Auction-Based Multi-Agent Approach,* T2021/26, November 2021, TRAIL Thesis Series, the Netherlands

Khakdaman, M., *On the Demand for Flexible and Responsive Freight Transportation Services,* T2021/25, September 2021, TRAIL Thesis Series, the Netherlands

Wierbos, M.J., *Macroscopic Characteristics of Bicycle Traffic Flow: a bird's-eye view of cycling,* T2021/24, September 2021, TRAIL Thesis Series, the Netherlands

Qu, W., *Synchronization Control of Perturbed Passenger and Freight Operations,* T2021/23, July 2021, TRAIL Thesis Series, the Netherlands

Nguyen, T.T., *Highway Traffic Congestion Patterns: Feature Extraction and Pattern Retrieval,* T2021/22, July 2021, TRAIL Thesis Series, the Netherlands

Pudāne, B., *Time Use and Travel Behaviour with Automated Vehicles,* T2021/21, July 2021, TRAIL Thesis Series, the Netherlands

Gent, P. van, *Your Car Knows Best,* T2021/20, July 2021, TRAIL Thesis Series, the Netherlands

Wang, Y., *Modeling Human Spatial Behavior through Big Mobility Data,* T2021/19, June 2021, TRAIL Thesis Series, the Netherlands

Coevering, P. van de, *The Interplay between Land Use, Travel Behaviour and Attitudes: a quest for causality,* T2021/18, June 2021, TRAIL Thesis Series, the Netherlands

Landman, R., *Operational Control Solutions for Traffic Management on a Network Level,* T2021/17, June 2021, TRAIL Thesis Series, the Netherlands

Zomer, L.-B., *Unravelling Urban Wayfinding: Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists,* T2021/16, May 2021, TRAIL Thesis Series, the Netherlands

Núñez Velasco, J.P., *Should I Stop or Should I Cross? Interactions between vulnerable road users and automated vehicles,* T2021/15, May 2021, TRAIL Thesis Series, the Netherlands

Duivenvoorden, K., *Speed Up to Safe Interactions: The effects of intersection design and road users' behaviour on the interaction between cyclists and car drivers,* T2021/14, April 2021, TRAIL Thesis Series, the Netherlands

Nagalur Subraveti, H.H.S., *Lane-Specific Traffic Flow Control,* T2021/13, March 2021, TRAIL Thesis Series, the Netherlands

Beirigo, B.A., *Dynamic Fleet Management for Autonomous Vehicles: Learning- and optimization-based strategies,* T2021/12, March 2021, TRAIL Thesis Series, the Netherlands

Zhang, B., *Taking Back the Wheel: Transition of control from automated cars and trucks to manual driving,* T2021/11, February 2021, TRAIL Thesis Series, the Netherlands

Boelhouwer, A., *Exploring, Developing and Evaluating In-Car HMI to Support Appropriate use of Automated Cars,* T2021/10, January 2021, TRAIL Thesis Series, the Netherlands