

Dynamic Route Choice Modelling of the Effects of Travel Information using RP Data

Giselle de Moraes Ramos

Delft University of Technology

This thesis is one of the outcomes of the research program TRISTAM: “Traveler Response and Information Service Technology: Analysis and Modeling” funded by the NWO, Dutch Organization of Science, and TRAIL, the Netherlands Research School for Transport, Infrastructure and Logistics.



Cover illustration: Peter van Dorst

Dynamic Route Choice Modelling of the Effects of Travel Information using RP Data

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen 02 februari 2015 om 15:00 uur
door

Giselle de Moraes Ramos

Master of Science in Civil Engineering
Universidade de São Paulo
geboren te Arcoverde – Pernambuco – Brazil

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. S.P. Hoogendoorn

Samenstelling promotiecommissie :

Rector Magnificus prof. K.C.A.M. Luyben	voorzitter
Prof. dr. ir. S.P. Hoogendoorn	Technische Universiteit Delft, promotor
Dr. ir. W. Daamen	Technische Universiteit Delft
Prof. dr. ir. C.G. Chorus	Technische Universiteit Delft
Prof. dr. ir. M. Bierlaire	EPFL - École Polytechnique
Prof. dr. ir. E.C. van Berkum	Universiteit Twente
Prof. dr. H.J.P. Timmermans	Technische Universiteit Eindhoven
Dr. E.R. Frejinger	University of Montreal

This thesis is the result of a PhD study carried out from 2009 to 2014 at Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Transport & Planning.

TRAIL Thesis Series no. T2015/03, the Netherlands TRAIL Research School

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
Phone: +31 (0) 15 278 6046
E-mail: info@rsTRAIL.nl

ISBN 978-90-5584-184-4

Copyright © 2015 by Giselle M. RAMOS

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the author.

Printed in the Netherlands.

To Matthijs

“We are our choices”
- Jean-Paul Sartre

Preface

This PhD thesis is the outcome of 5 years of work, the contribution of many people and a lot of choices. Doing this PhD was only possible after I chose to move from Brazil to the Netherlands. In hindsight, it was definitely the right choice to make. At the time, however, everybody besides me may have struggled hard to understand what made the utility of moving higher than that of staying. The decision maker, however, always knows what goes in the utility function ;-)

Faith and perseverance were necessary in plenty of moments during this PhD. I thank God for providing the spiritual comfort I needed. Secondly, I would like to thank my promoter Serge Hoogendoorn and my supervisor Winnie Daamen. When I was offered a PhD position, my life changed in many ways and I am very grateful for that. I still remember my excitement when on a seemingly regular Monday afternoon mid-December 2008 we had a telephone interview. About one hour later I received an e-mail from Serge saying that after a 10 minute discussion with Winnie they had decided to offer me a PhD position. So many emotions were generated by an e-mail of just 2 lines. Serge and Winnie, on a more practical side, I am thankful for your guidance and technical support as well as for making important projects of cooperation possible. One of these projects of cooperation, which resulted in great part of the scientific contributions of this PhD thesis, was with Emma Frejinger. Emma, I really enjoyed working with you and I learned a lot from you. I am eternally grateful for your support.

This work also would not have been possible without the cooperation of TomTom and of the participants of my route choice experiment. TomTom made available the navigation devices used during the experiment and great part of the data used in the analyses presented in this thesis; the participants allowed me to track them for about 2 months, making it possible to investigate their behaviour in a real setting. Your cooperation was essential. Thanks for that.

The staff of the department of Transport & Planning of the TU Delft was of great (logistical) support and helped in many ways. So too were the participants of the research program TRISTAM. I am especially grateful to Caspar Chorus for the support in several different occasions and to my fellow PhD researchers for the interesting conversations and social activities

Especial thanks goes out to all my friends for nice conversations, jokes, laughs, gossips, drinks, nights out, etc. Without you to visit when in Brazil or to e-mail as if I was still living there, I would miss part of my identity. Without my amazing Dutch friends, who welcomed me so warmly even before I moved to the Netherlands, and who still speak in English to me – without making a face – it would have been difficult to feel at home in the Netherlands. Without my Brazilian/Portuguese friends in the Netherlands to speak Portuguese with as if I were still in Brazil, have fun together and complain about our Dutch life (and husbands), I would feel as a “peixe fora d’água”. Without my Delft (gossip) group to have drinks, talk about struggles of PhD life and also serious scientific topics (not really), I would think I was the only cool nerd person in this world. You all, my friends, are extremely special and dear to me. Having you in my life is a blessing.

Blessed I have also been with my family. Moving to the Netherlands meant that I would be even more far from you. This was – and still is – one of the most difficult parts. Thanks for everything!! Everything is really everything: talks, photos, visits, jokes, gossips and even “confusões”. I love you! Moving to the Netherlands, though, gave me a new family as a bonus. As an European family though, much less “confusões” ;-). Thanks for always being here for me and for welcoming me so well too. Your efforts to make my life in the Netherlands easier are a balm for my heart.

Last, but not least, I would like to thank my Loves, meu Amoro, meu Dearzinho. You maximize the utility of my choice irrespective of the reference point. With you, there is not regret. Being with you was the best choice I could have made. You have been fully involved in this PhD, from emotional to technical English support. You are part of a unique group of partners who actually read the PhD thesis from beginning to end. Whenever I needed, you were there. Thank you, Jurus. Thank you! (And this time, I mean it). Were this not enough, without you there would be no Bento. Meu grande amor pequeninho, you only bring the best that there is in me. You are a blessing for me. Blessing, blessed, Bento.

Giselle

Amsterdam, December 2014

Table of Contents

Preface	i
List of Figures	vii
List of Tables	ix
Notation	xi
1. Introduction	1
1.1 Research background.....	2
1.2 Conceptual framework	3
1.3 Research objective and questions	3
1.4 Research approach	5
1.5 Scientific contributions and practical relevance	5
1.5.1 Scientific contributions	5
1.5.2 Practical and policy relevance	6
1.6 Context of the research	7

1.7	Thesis outline.....	7
2.	State-of-the-art on route choice behaviour and travel information	11
2.1	Travel information.....	12
2.2	Data collection efforts and empirical findings	13
2.2.1	Use of travel information.....	13
2.2.2	Changes in route choice due to travel information	15
2.2.3	Willingness to pay for travel information.....	16
2.3	Advantages and disadvantages of different data collection methods.....	17
2.4	Conclusions	18
3.	State-of-the-art on travel behaviour theory	19
3.1	Fundamentals of travel behaviour theory	20
3.1.1	Expected utility theory.....	20
3.1.2	Prospect theory.....	21
3.1.3	Regret theory.....	23
3.2	Comparative analysis between the theories.....	23
3.2.1	Decision mechanism	24
3.2.2	Number of parameters to be estimated	24
3.2.3	Simplicity and tractability of the modelling framework.....	24
3.2.4	Ability to capture realistic behaviour.....	25
3.2.5	Overview of the comparative analysis.....	25
3.3	From behavioural theory to discrete choice models.....	25
3.4	Challenges ahead: practical issues associated to estimating random utility models based on revealed preference data.....	28
3.5	Conclusions	29
4.	Modelling framework: dynamic sequential link choice decisions with unrestricted choice set	31
4.1	Recursive Logit model for dynamic and deterministic networks.....	33
4.2	Challenges and consequences of going from RL to RL Dyn	37
4.3	Conclusions	40

5. Revealed preference data collection on travellers' reactions to travel information	41
5.1 Specifications of the data collection setup	42
5.1.1 Data collection technique(s) to be employed.....	42
5.1.2 Specifications regarding travel information.....	44
5.1.3 Duration of the data collection.....	45
5.1.4 Target group.....	45
5.1.5 Characteristics of the area to be investigated.....	46
5.1.6 Overview on the data collection design	47
5.2 The actual data collection	48
5.2.1 Period and duration of the data collection	48
5.2.2 OD zone(s) investigated.....	49
5.2.3 Selection of participants.....	50
5.2.4 Conditions of travel information provision.....	50
5.3 Data sets: characteristics and quality assessment	52
5.3.1 GPS data.....	52
5.3.2 Travel diary	53
5.3.3 Interviews.....	56
5.3.4 Traffic conditions in the network.....	56
5.3.5 Overview of the resultant data sets and quality assessment	57
5.4 Conclusions	59
6. Qualitative and quantitative analyses of travellers' reactions to travel information	61
6.1 Travellers' reactions to different timing of travel information.....	63
6.2 Travellers' reactions to different sources of travel information	65
6.3 Travellers' confidence and compliance with travel information.....	66
6.4 Willingness to pay for travel information.....	68
6.5 Changes in route choice due to travel information.....	69
6.6 Travellers' perceptions	74
6.6.1 Travellers' awareness of the route choice set	74
6.6.2 Travellers' perceptions of route reliability	75
6.7 Conclusions	79

7. Model estimation and prediction	83
7.1 Network and observations	84
7.2 Estimation results in the RL Dyn framework.....	84
7.2.1 To what extent does provision of travel information lead to (significant) changes in route choice behaviour?	85
7.2.2 How does the timing of travel information provision (pre-trip and en-route) influence route choice behaviour?	88
7.2.3 How do different sources of travel information influence route choice behaviour?	88
7.3 Comparison of model performances.....	89
7.4 Comparison of static and dynamic networks.....	90
7.4.1 Estimation results in the RL framework	90
7.4.2 Prediction results.....	92
7.5 Conclusions	94
8. Conclusions	99
8.1 Empirical findings	101
8.1.1 Empirical findings about the role of travel information	101
8.1.2 Empirical findings about travellers' perceptions	104
8.2 Methodological contributions.....	104
8.3 Conclusions	105
8.4 Implications for practice	105
8.5 Future research directions.....	106
Bibliography	109
Summary	119
Samenvatting	123
Resumo	127
About the author	131
TRAIL Thesis Series	133

List of Figures

Figure 1.1: Conceptual framework of route choice decisions.....	4
Figure 1.2 Schematic presentation of the thesis outline.....	9
Figure 4.1: Illustration of notation for dynamic and deterministic network	34
Figure 4.2: Example network.....	39
Figure 4.3: Possible states to be reached in the RL considering traveller at link k1.....	39
Figure 4.4: Possible states to be reached in the RL Dyn considering traveller at link k1 and departures at instants $t=1$ to $t=3$	39
Figure 5.1: Main cities in the Randstad.....	49
Figure 5.2: Example of the type of travel information provided by TomTom.....	51
Figure 5.3: Illustrative graphs of speed versus time and speed versus distance	52
Figure 5.4: GPS device used during the experiment.....	52
Figure 5.5: Page 1 of the travel diary	54
Figure 5.6: Page 2 of the travel diary	55
Figure 5.7: Road network studied	57
Figure 5.8: Historical versus network travel times	58
Figure 6.1: Willingness to pay for different types of travel information	70
Figure 6.2: Willingness to pay for different types of information for informed and regular travellers.....	70
Figure 6.3: Exploration rates during the experiment based on the GPS traces.....	72
Figure 6.4: Route choices in the reference and information treatments.....	72

Figure 6.5: Subjective route choice set between Delft and The Hague and home of the participants	75
Figure 6.6: Coding of the roads in the subjective route choice set	75
Figure 6.7: Objective measures of reliability of the routes in choice set 1 (CS1) in the morning (a) and afternoon peak hours (b)	77
Figure 6.8: Objective measures of reliability of the routes in choice set 2 (CS2) in the morning (a) and afternoon peak hours (b)	77
Figure 7.2: Predicted log likelihood for different models and subsamples.....	94

List of Tables

Table 5.1: Relationship between the research questions, types of data and requirements.....	43
Table 5.2: Relationship between the types of data and the data collection techniques.....	44
Table 5.3: Percentage of participants in different treatments of information provision	51
Table 6.1: Overview of the analyses performed in this chapter	62
Table 6.2: Percentage of travellers consulting travel information based on the interviews.....	63
Table 6.3: Percentage of travellers consulting travel information based on the travel diaries	64
Table 6.4: Percentage of trips in which travel information was consulted based on the travel diaries	65
Table 6.5: Percentage of use of different sources of pre-trip information for all travellers based on the travel diaries	66
Table 6.6: Percentage of use of different sources of en-route information for all travellers based on the travel diaries	66
Table 6.7: Compliance rate with travel information based on the travel diaries	66
Table 6.8: Compliance rate with different sources of pre-trip information based on the travel diaries	67

Table 6.9: Compliance rate with different sources of en-route information based on the travel diaries	67
Table 6.10: Compliance rate with travel information for informed and regular travellers based on the travel diaries	68
Table 6.11: Percentage of changes in route choices due to travel information based on the travel diaries	71
Table 6.12: Percentage of travellers willing to choose the same route in hindsight based on the travel diaries	73
Table 6.13: Percentage of changes in departure time due to travel information based on the travel diaries	73
Table 6.14: Travellers' perceptions of the routes' characteristics	79
Table 7.1: Overview of the attributes interacted with travel time in different models	85
Table 7.2: Estimation results of different model specifications in the RL Dyn framework	87
Table 7.3: Estimation results of model specification 1 (MS 1) in the RL framework	91
Table 7.4: Ratios of pairs of parameters	92

Notation

A	Vector of attributes of a prospect
$ALot$	Events or lotteries
$BLot$	Events or lotteries
C	Choice set
G	Directed connected graph defined in terms of links, nodes and time intervals
L	Set of links of a transportation network
N	Set of nodes of a transportation network
$P(.)$	Probability an alternative or path to be chosen
$R(.)$	Regret/rejoice of an alternative in relation to another
S	Set of all possible states of the world
T	Set of time intervals of a transportation network
TT	Travel time
$U(.)$	(Instantaneous) Utility of a choice
$V(.)$	Value of a prospect / Expected downstream utility

$V^+(\cdot)$	Value of a prospect for positive resultant outcomes
$V^-(\cdot)$	Value of a prospect for negative resultant outcomes
I	Identity matrix
LL ($\hat{\beta}$)	Final log-likelihood based on the estimated parameters
M	Matrix of instantaneous utilities
a	Link of the transportation network where the decision maker goes after the decision is made, i.e. next link (action) of the network
d	Destination, i.e. last link of the network to be reached by the decision maker
h	Alternative of a choice set
i	Alternative of a choice set
j	Alternative of a choice set
k	Link of the transportation network in which the decision maker currently is
n	Decision maker/individual
neg	Total number of negative outcomes
$p, p(\cdot)$	Probability of occurrence of an event, outcome or consequence
pos	Total number of positive outcomes
rp	Reference point of Prospect Theory
s	A particular state of the world
s'	State of the world reached after a decision is made, i.e. next state
t	A specific time interval
t'	Time interval reached after a decision is made
t_int	Size of the time interval
$u(\cdot)$	Instantaneous utility
$v(\cdot)$	Value function / Deterministic comportment of the utility
$w(\cdot)$	Weighting function
$w^+(\cdot)$	Weighting function for positive resultant outcomes
$w^-(\cdot)$	Weighting function for negative resultant outcomes
x	Consequence/outcome
b	Vector with elements $b_s = 0, (k,t) \neq s (d,t)$ and $b(d,t)=1$
k	Prospect, i.e. a consequence associated with a probability of occurrence

l	Prospect, i.e. a consequence associated with a probability of occurrence
m	Prospect, i.e. a consequence associated with a probability of occurrence
z	Vector with elements $z_s = e^{V(s)}$
α	Parameter of the value function of prospect theory representing the degree of diminishing sensitivity for gains
β	Vector of parameters to be estimated
$\beta_{TTInformed}$	Parameter representing the role of travel time for informed travellers
$\beta_{TTRegular}$	Parameter representing the role of travel time of travel time for regular travellers
β_{LC}	Parameter representing the role of link constant
β_{HabR}	Parameter representing the role of 2 sequential links belonging to a habitual route
β_{VMS}	Parameter representing the role of links that have a VMS panel on it
β_{TT_PreEn}	Parameter representing the role of travel time for travellers consulting pre-trip and en-route information
β_{TT_Pre}	Parameter representing the role of travel time for travellers consulting pre-trip information only
β_{TT_En}	Parameter representing the role of travel time for travellers consulting en-route information only
β_{TT_NoInfo}	Parameter representing the role of travel time for travellers who did not consult travel information at all
$\beta_{TTPreTomTom}$	Parameter representing the role of travel time for travellers consulting only TomTom as a source of pre-trip information. This was not an obstacle to also consult en-route information
$\beta_{TTPreRadio}$	Parameter representing the role of travel time for travellers consulting only radio as a source of pre-trip information. This was not an obstacle to also consult en-route information
$\beta_{TTMoreThan1SourcePreTrip}$	Parameter representing the role of travel time for travellers consulting more than one source of pre-trip information. This was not an obstacle to also consult en-route information
$\beta_{TTNoPreTrip}$	Parameter representing the role of travel time for travellers who did not consult pre-trip information at all. This was not an obstacle to consult en-route information
$\beta_{TTEnTomTom}$	Parameter representing the role of travel time for travellers consulting only TomTom as a source of en-route information. This was not an obstacle to also consult pre-trip information

$\beta_{TTEnVMS}$	Parameter representing the role of travel time for travellers consulting only VMS panels as a source of en-route information. This was not an obstacle to also consult pre-trip information
$\beta_{TTEnRadio}$	Parameter representing the role of travel time for travellers consulting only radio as a source of en-route information. This was not an obstacle to also consult pre-trip information
$\beta_{TTMoreThan1SourceEnRoute}$	Parameter representing the role of travel time for travellers consulting more than one source of en-route travel information. This was not an obstacle to consult pre-trip information
$\beta_{TTNoEnRoute}$	Parameter representing the role of travel time for travellers who did not consult en-route information at all. This was not an obstacle to consult pre-trip information
δ	Parameter of the weighting functions of Prospect Theory responsible to define the curvature of the weighting function and capture the distortion in the perception of probabilities for losses
ε	Random component of the utility function
γ	Parameter of the weighting functions of Prospect Theory responsible to define the curvature of the weighting function and capture the distortion in the perception of probabilities for gains
λ	Parameter of the value function of prospect theory representing the degree of loss aversion
μ	Scale parameter
π^+	Decision weight associated with positive resultant outcomes
π^-	Decision weight associated with negative resultant outcomes
θ	Parameter of the value function of prospect theory representing the degree of diminishing sensitivity for losses

Chapter 1

Introduction

Traffic congestion is experienced by a great number of travellers during peak hours and is directly influenced by a number of travel-related decisions, such as route and departure time choices. Congestion in the transportation network leads to travel time uncertainty and more complexity to the decision-making process. Predicting the duration of a trip, the best time to start a trip or even whether it is still feasible to leave home to engage in an activity on time, therefore, are very difficult tasks.

In order to solve problems that arise from congestion, such as delays, uncertainty and environmental effects, there has been an increasing interest to investigate the role of travel information. Travel information has the potential to influence travellers' choices, steer travellers to less congested routes and alleviate congestion (Arnott et al., 1991; Denant-Boèmont and Petiot, 2003; Bogers, 2009; Ben-Elia and Shiftan, 2010). Contrary to measures to reduce or solve congestion focusing on the construction of more road infrastructure, provision of travel information contributes to a more efficient use of the existent available road capacity. Besides that, it also overcomes issues regarding lack of space and money for investing in more road infrastructure.

However, as discussed later on in section 1.1, it is not yet fully understood how travellers' react to travel information and how it can impact behavioural changes, let alone to improve network traffic conditions. When travel information is provided, travellers have to decide whether to comply with the travel information. The complexity of the decision on whether to comply with the travel information increases, for instance, with travellers' expectations about the traffic situation, with the quality of travel information and with route choice habits. On the one hand, provision of travel information may lead to an overreaction of the system if a large proportion of travellers respond. On the other hand, the effects may be limited if travellers decide not to comply with it at all. Consequently, depending on how travellers react to travel information, different consequences can be observed in the transportation network and on the level of congestion.

1.1 Research background

Travel information has become a popular topic of academic research since the second half of the 1990s (Emmerink et al., 1996; Hato et al., 1999; Polydoropoulou and Ben-Akiva, 1999; Khattak et al., 2003; Molin and Timmermans, 2006). A main issue recognized in the literature is the lack of available data sets based on actual route choices to investigate the relationship between travel information and travellers' route choice behaviour.

A number of experiments on travellers' reactions to travel information have been reported in the literature, but are limited to either stated preference (SP) surveys in which travellers are asked which route to choose given a specific context (Abdel-Aty et al., 1997; Wardman et al., 1997; Ben-Elia and Shiftan, 2010) or interactive route choice experiments in simulation environments in which, similar to a game, travellers make consecutive route choices and their behaviour towards risk and different types and quality of travel information is investigated (Chen and Mahmassani, 1993; Koutsopoulos et al., 1994; Chen et al., 1999; Selten et al., 2007; Bogers, 2009; Ben-Elia and Shiftan, 2010). The focus on SP experiments, including travel simulators, is large because they are relatively cheap, allow more flexibility of scenarios, are more controllable and are efficient. SP experiments, however, are subject to the well-known drawback of external validity, i.e. it is questionable whether the outcomes would be valid in a real setting. This, however, is not the case of revealed preference (RP) experiments as in RP experiments travellers actually choose a route given a specific context.

RP datasets reported in the literature are either surveys in which travellers report their past route choices (Mahmassani et al., 1993; Rose et al., 2008), a combination with SP surveys (Zhang and Levinson, 2008) or based on GPS traces (Frejinger and Bierlaire, 2007; Papinski et al., 2009; Bierlaire et al., 2010). Due to the uncontrollability of RP experiments, however, issues that influence the estimation of route choice models are often observed. For instance, studies in which travellers report their past choices are hardly related to an actual network in terms of alternatives and traffic conditions. Studies based on GPS data track travellers but not much is known about whether travel information was actually accessed, the trip purpose, etc. In addition, data about the traffic conditions in the network is often missing.

Besides important data collection issues, estimation of route choice models based on RP data collected in real networks are associated with three main issues. Firstly, the sampling of alternatives is difficult (Bovy, 2009; Prato, 2009; Frejinger and Bierlaire, 2010). Secondly, alternatives may share unobserved attributes (Cascetta et al., 1996; Ben-Akiva and Bierlaire, 1999; Bekhor et al., 2001; Frejinger and Bierlaire, 2007). Thirdly, the attributes of the alternatives (e.g. travel times at the time of the observation) are in general unknown. Dealing with these three modelling issues is more complex in dynamic than static networks (Gao et al., 2008).

Outcomes of the models are impacted by the input data and ultimately this may lead to misinterpretation of results and conclusions. This thesis therefore has two main contributions: an empirical and a methodological. The empirical is to discuss the role of travel information on travellers' route choice behaviour based on RP data collected in a congested network. The methodological is to address issues associated with model estimation based on RP data in a dynamic setting.

1.2 Conceptual framework

The proposed conceptual framework primarily focuses on the role of travel information on travellers' route choice behaviour. It consists of a flowchart describing the decision making process travellers go through when making their route choice decisions (Figure 1.1).

The first main decision concerns whether the trip requires travel information. In case there is no need to consult travel information, a decision mechanism influenced by habit, experience and other attributes is used to estimate travel times. In case the trip requires travel information to be consulted, travel information needs are identified, i.e. decisions regarding providers, type and level of personalization of travel information. Outcomes of this process are the timing and sources to be consulted. In such situations, travel information is also used as a source to estimate travel times and/or update expectations.

Then, it has to be decided whether to comply with travel information. If it is decided to comply with it, the sources and timing to comply with are chosen and a route choice decision is made. Otherwise, all the other factors mentioned, including the estimated travel times, influence the decision mechanism and then a route choice is made. The experience gathered with each route choice is used to update expectations for next trip as depicted in the parallelogram "Habit, experience, etc".

1.3 Research objective and questions

The empirical investigation about the role of travel information on travellers' route choice behaviour focuses on different sources and timing of travel information provision. The main research question is:

What is the impact of travel time information on travellers' route choice behaviour?

This main research question is answered step-by-step by addressing the following research questions:

1. *To what extent does provision of travel information lead to (significant) changes in route choice behaviour? (Chapters: 2, 7)*
2. *How do different sources of travel information affect route choice behaviour? (Chapters: 2, 6, 7)*
3. *How does the timing of travel information provision (pre-trip and en-route) affect route choice behaviour? (Chapters: 2, 6, 7)*
4. *To what extent are travellers willing to comply with different sources of travel information provision (over time)? (Chapters: 2, 6)*
5. *To what extent are travellers willing to comply with different timing of travel information provision, i.e. pre-trip or en-route (over time)? (Chapters: 2, 6)*
6. *To what extent are travellers willing to pay to receive travel information (over time)? (Chapters: 2, 6).*

Although these research questions are answered throughout this thesis, a summary of the answers is presented in Chapter 8.

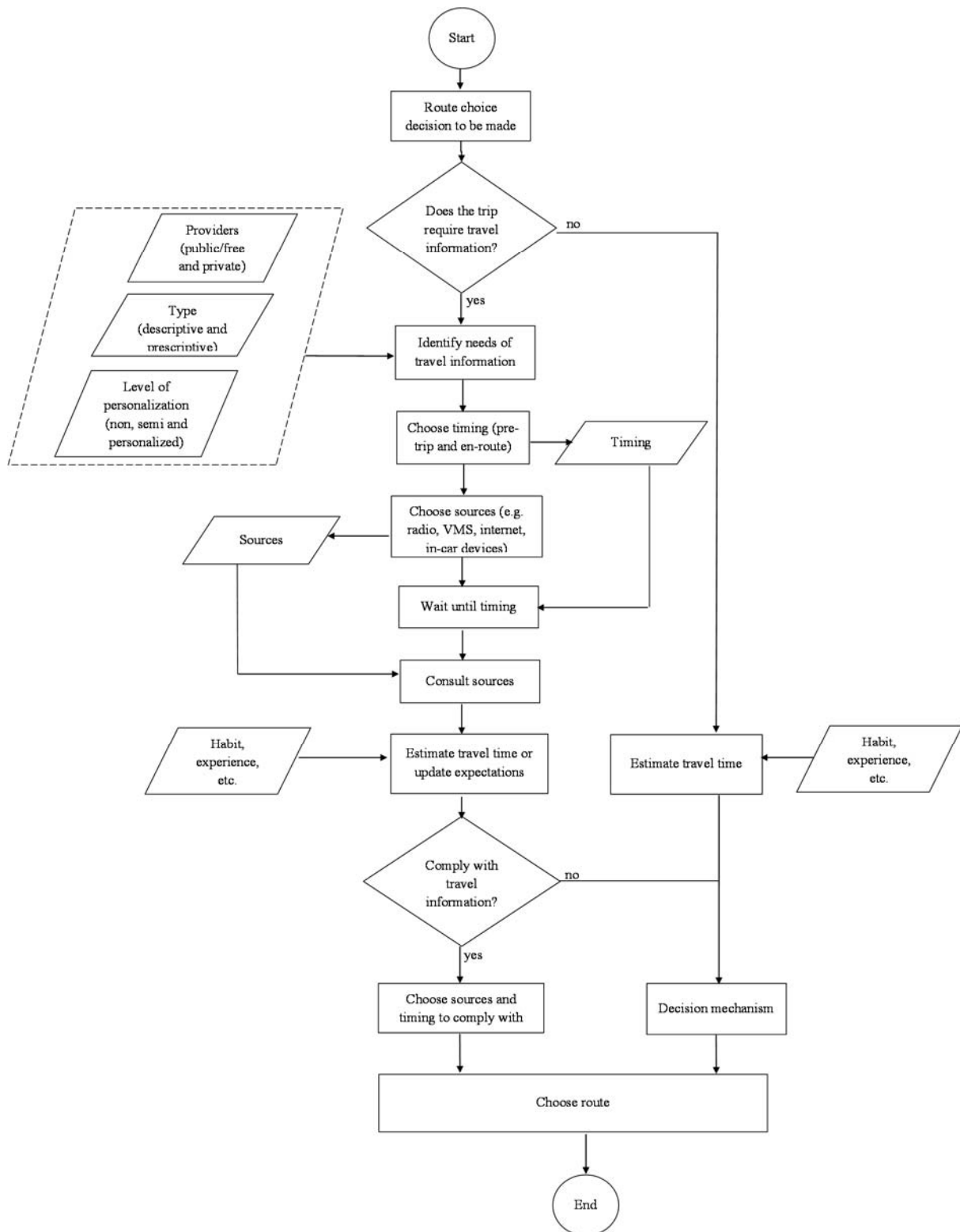


Figure 1.1: Conceptual framework of route choice decisions

1.4 Research approach

The research approach consisted in investigating travellers' route choice behaviour, and their reactions to travel time information, in a real setting while they were making their decisions. The proposed experiment consisted in using GPS devices, travel diaries and interviews to investigate the behaviour and reaction to travel information of 32 travellers for a period of 9 weeks. This resulted in a sample of 897 valid GPS traces and travel diaries, among which 374 refer to initial 3 week of data collection in which only free/public sources of travel information were available and 523 to the subsequent 6 weeks of data collection in which personalized real-time travel information was also available. Moreover, data about the traffic conditions in the network was also collected for the whole study period. A big upside of this approach is that by conducting a RP experiment, issues of external validity associated with traditional SP data collection methods are overcome.

Estimation of route choice models based on RP data collected in congested networks, however, is associated with the issue of choice set sampling. To overcome this issue, a modelling framework denominated Recursive Logit Dynamic (RL Dyn) was proposed based on the sequential link-choice decisions of the Recursive Logit (RL) (Fosgerau et al., 2013). The RL can be consistently estimated based on RP data and has the advantage of not requiring any definition of choice sets. Although both the RL and RL Dyn share the same properties, the latter is extended to a dynamic setting. Thus, it incorporates the time dimension to the route choice problem. By collecting data about the traffic conditions in the network and adopting a dynamic modelling framework which does not require choice set sampling, two of the three main modelling issues, attributes and sampling of alternatives, are addressed in a dynamic setting.

In terms of empirical contributions, extensive analyses of the relationship between travel information and travellers' route choice behaviour is done by means of quantitative and qualitative analyses of the data set(s) and model estimation. Note that despite the limited sample size, estimation results were statistically significant. In the methodological aspect, a comparative analysis between the RL and RL Dyn helps clarifying the added value of adding the time dimension to the route choice problem.

1.5 Scientific contributions and practical relevance

The contributions of this research can be divided into its scientific contributions and its practical relevance, which are described in the following.

1.5.1 Scientific contributions

The characteristics of the data collection and proposed modelling framework are the base for the main scientific contributions of this PhD thesis.

The biggest benefit of data collection based on RP experiments is to *investigate travellers' behaviour in a real-world setting*. Contrary to what happens in SP and travel simulator experiments, in RP experiments travellers are “confronted with” the consequences of their choices and not only somehow penalized by a smaller reward at the end of the experiment. Depending on the traffic conditions, travellers actually arrive later at the destination, get stuck in heavy traffic jams, miss an appointment due to congestion, etc. In particular with respect to

RP data collection conducted as part of this PhD thesis, GPS devices were used to track travellers during their trips, travel diaries were filled in after each trip and interviews were conducted in three different occasions. Besides this, data about the traffic conditions in the network for the whole study period was also collected. The combination of these data sources resulted in a very rich and unique data set that allows *overcoming many of the issues associated with the data collection and modelling processes*. Especially due to the complexity and amount of workload required, such type of RP experiments combining GPS traces with travel diaries and for which data about the traffic conditions of a real congested network is available, have not been observed in the literature.

With respect to the modelling framework, this thesis presents a *new route choice model for a dynamic network* with deterministic link attributes based on GPS data collected in a congested network for which link travel times are known. It is shown how the model of Fosgerau et al. (2013), in which no choice set generation is needed, can be specified and estimated in this context. Therefore, two of the three main issues associated with the estimation of route choice models based on RP data collected in a congested network are addressed in a dynamic setting, i.e. issues of choice set sampling and attributes of the alternatives. Adding the time dimension to the route choice problem is not a simple task and requires a careful redefinition of the state space. This consequently increases the complexity of the modelling framework and computational demand. Outcomes of this research support the added value of this approach.

The way travellers react to travel information is strongly related to how important they consider it to their route choices. The proposed methodology can be used to investigate travellers' use of different sources and timing of travel information provision makes it possible *to get more insight into the real effects of travel information provision on travellers' behaviour in order to discuss possible consequences to the network dynamics*. Besides this, it gives insights into the *relative importance of different sources* of travel information.

As the data collection comprised interviews, perceptions of the routes' characteristics could be investigated. As data about the traffic conditions in the network was also available, comparisons between travellers' perceptions with objective measures of travel time variability can be done. Although this is not the main objective of this thesis, it helps discussing the role of perceptions and possible implications on *how to define and incorporate travel time variability in route choice models*. This may have a high impact in the modelling and conclusions thereof.

1.5.2 Practical and policy relevance

Investigating travellers' behaviour in a real context while they were making their route choice decisions has a big practical relevance. For the industry and road authority it helps assessing not only the usefulness of the travel information provided from the point of view of the users, but also their willingness to pay for such services. For instance, what is the role of the variable message signs next/above the roads on traveller's route choice decisions? Are they a useful source of travel information or should they be replaced by in-car devices only? Are they well located? How about the travel information broadcasted by radio? How about the relationship between private providers of travel information and travellers' willingness to pay to use such services? Although the cost itself Benefits for the society are directly related to the way the road authorities and industry approach the derived behavioural insights.

The *detailed description of the setup of the experiment* is also part of the practical relevance. By describing and motivating each of the aspects considered for the data collection, this could be used as initial *guideline for future experiments*.

1.6 Context of the research

This thesis is part of a larger research programme called *TRISTAM: Traveller Response and Information Service Technology: Analysis and Modelling*, which focuses on investigating the impacts and consequences of Information and Communication Technologies (ICT) for the society as a whole. The goal of the TRISTAM programme is to develop, test and apply improved methods and models to assess the effects of emerging forms of teleworking and various types of personal and public sources of travel information provision for the society. In particular, four main effects are investigated in the TRISTAM programme:

- Behavioural effects. The objective is to better understand the relationship between travel information and travellers' route choice behaviour. This is the subject of this thesis.
- Accessibility effects. This project focuses on developing, testing and investigating the role of travel information as facilitator to physical travel and virtual mobility (e.g. teleconferencing).
- Spatial externalities. This research focuses on changes in the timing and location of travel activities in response to travel information and new forms of teleworking.
- Economic effects. This project focuses on the development of measures of social benefits and willingness-to-pay for travel information either as a complement or substitute to travelling.

Each of these effects has been independently investigated. Together, they contribute to a bigger and clearer picture of the potential impacts of ICT solutions to enhance the sustainability of the Randstad area¹ proposed by the TRISTAM programme.

1.7 Thesis outline

In order to answer the research questions raised in Section 1.3, this thesis is structured in nine chapters. The relationships between the chapters are depicted in Figure 1.2.

Chapter 2 presents the state-of-the-art on route choice behaviour and travel information by discussing data collection efforts and empirical findings with focus on use of travel information, behavioural changes and willingness to pay for travel information.

Chapter 3 then discusses the state-of-the-art on behaviour theories to model route choice behaviour, particularly focusing on *utility theory*, *prospect theory* and *regret theory*. Their fundamentals and developments are presented and a comparative analysis is done to help defining the behavioural rule to be used in this thesis. Following this, a link to discrete choice analysis is made and potential estimation issues are raised.

¹ The Randstad: <http://en.wikipedia.org/wiki/Conurbation> is a conurbation in the Netherlands. It consists of four largest Dutch cities (Amsterdam, Rotterdam, The Hague and Utrecht) and the surrounding areas. With a population of around 7 million habitants it is one of the largest conurbations in Europe, comparable in size to Milan or the San Francisco Bay Area and covers an area of approximately 8287 km². Wikipedia (2011) *Randstad*, [<http://en.wikipedia.org/wiki/Randstad>], page visited on July 14th, 2011.

Following both states-of-the-art, Chapter 4 introduces the proposed modelling framework. The Recursive Logit developed by Fosgerau et al. (2013) is extended to a dynamic setting by adding the time dimension to the route choice problem in a framework denominated Recursive Logit Dynamic. Although both modelling frameworks share the same properties, time-expanding the route choice problem is not a trivial task and imposes important redefinitions of the state space. The main contribution of this chapter is to elaborate on this.

The setup of the revealed preference data collection is presented in Chapter 5. It describes in detail the content of the four data sets generated by the experiment: GPS traces, travel diaries, interviews and data about the traffic conditions in the transportation network. In addition, the reasoning behind the chosen data collection setup is discussed.

Chapters 6 and 7 present the outcomes of the data collection. Qualitative and quantitative analysis of the data sets primarily focusing on travellers' reaction to travel information is described in Chapter 6. Estimation results of different model specifications are discussed in Chapter 7 with focus on travellers' sensitivity to travel time in relation to different sources and timing (pre-trip and en-route) of travel information. Furthermore, Chapter 7 compares the static and dynamic versions of the Recursive Logit by discussing their predictive power. This helps clarifying the added value of adding the time dimension to the route choice problem.

Finally, Chapter 8 presents the conclusions of this research by answering the research questions raised in section 1.3, highlighting practical and scientific implications of travellers' reactions to travel information and suggesting possible future research directions.

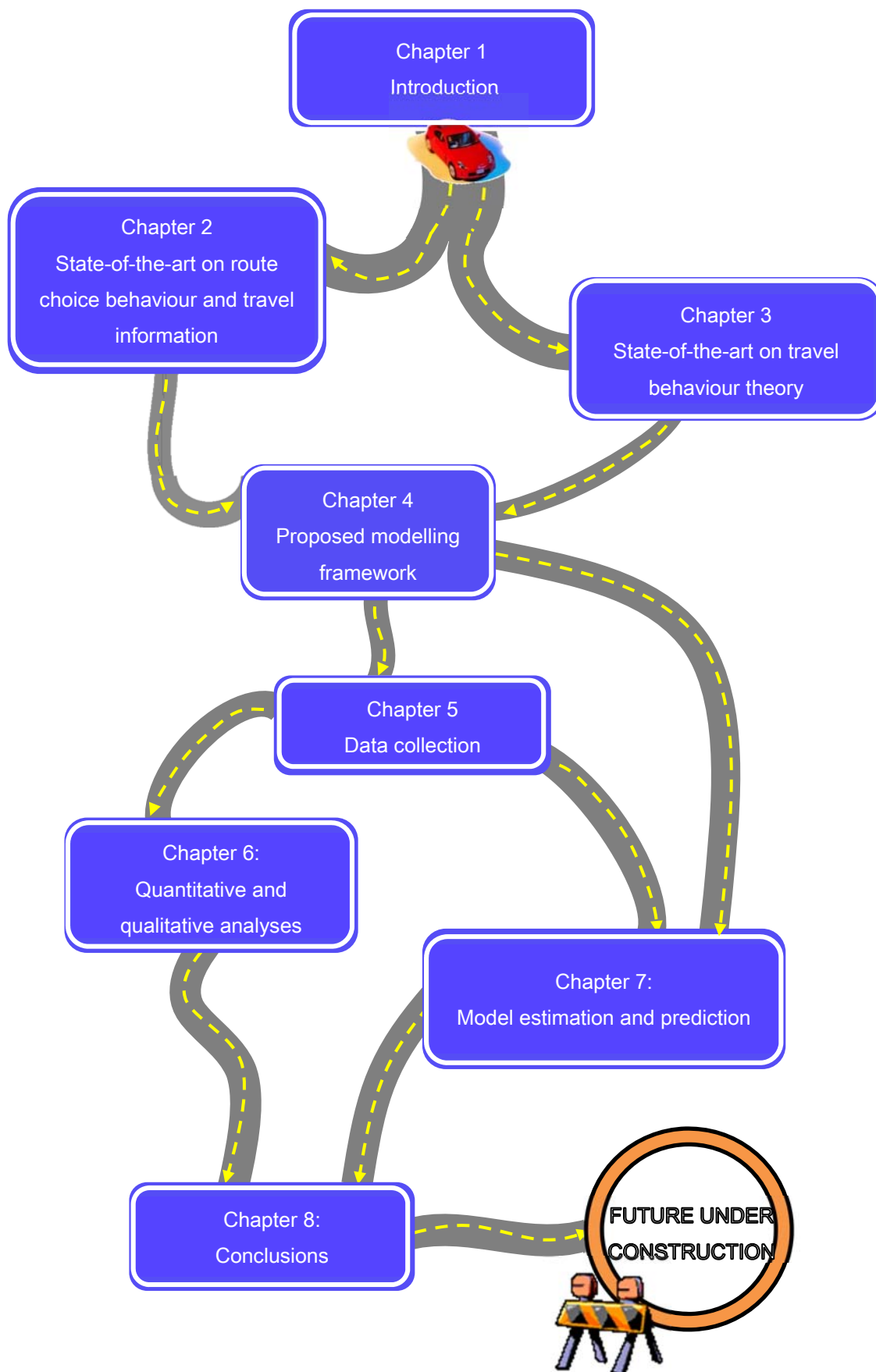


Figure 1.2 Schematic presentation of the thesis outline

Chapter 2

State-of-the-art on route choice behaviour and travel information

As discussed in Chapter 1, there has been an increased interest in the potential of travel information to help alleviate congestion and improve reliability of travel. The process through which travellers acquire and use travel information to assist their route choice decisions, however, is not fully understood (Ben-Akiva et al., 1991; Mahmassani and Jayakrishnan, 1991; Polak and Jones, 1993; Wardman et al., 1997; Chorus et al., 2007; Ben-Elia et al., 2008). This is especially relevant for situations in which multiple sources of travel information are available and travellers have to choose which source(s) to use.

Studies about travellers' reactions to travel information have employed different data collection and modelling frameworks to investigate the effects of travel information on travellers' behaviour. This chapter presents the state-of-the-art on the effects of travel information on route choice behaviour by discussing data collection efforts and its empirical findings. Part of the main contributions of this thesis come from the data collection effort and this state-of-the-art makes this contribution clearer by identifying related gaps in the literature. Main points addressed are the use of travel information, resulting behavioural changes and willingness to pay for travel information as these are related to the research questions raised in Chapter 1.

Section 2.1 introduces some basic concepts about travel information. In section 2.2, travellers' reactions to information focusing on data collection efforts and empirical findings with respect to the use, behavioural changes and willingness-to-pay for travel information are discussed. Then section 2.3 presents the advantages and disadvantages of the data collection methods. Conclusions and gaps in the literature are presented in section 2.4.

2.1 Travel information

The means of travel information provision have significantly improved in the past three decades. Until the end of 1980's the tools of travel information provision were rather low-tech systems predominantly used to help people find their way along the transportation network (Chorus, 2007). After that, technological advances led to the development of the so-called Advanced Traveller Information Systems (ATIS) with more sophisticated functionalities such as provision of expected travel times based on the current traffic situation (Boyce, 1988; Arnott et al., 1991; Polak and Jones, 1993). Nonetheless, the primary goals of providing travel information remain the same: better management of traffic flow, to enhance driving operations and to improve travellers' safety (Adler and Blue, 1998).

The level of personalization of the currently available sources of travel information include non-personalized public travel information, semi-personalized public travel information and personalized real-time travel information. Non-personalized public travel information refers to general traffic conditions in the network and comprises information regarding major congested areas *irrespective* of the interest of the traveller. This is the type of information usually provided by the radio or television. Semi-personalized public travel information, on the other hand, refers to the traffic situation of major roads *between a specific origin and destination pair of interest*. Examples of semi-personalized public travel information are the information displayed on Variable Message Signs (VMS) next to and/or above roads and the travel information provided by websites specialized in travel information provision. Finally, personalized real-time travel information refers to the current traffic situation (e.g. travel times, delays, and length of the queues) on both highways and local roads of interest for a specific origin and destination. This is the type of information provided by personal in-car navigation and applications for mobile phone devices (nomadic devices).

The level of personalization is very likely to influence travellers' behaviour not only because of the level of detail, but also as part of a strategic behaviour. For instance, as public travel information is available to everyone, travellers may act strategically and take into account how many others will comply with it before deciding themselves what to do. Conversely, if the information is personalized, travellers may act less strategically and put more importance in their own beliefs, preferences and experience rather than on the behaviour of others (Parvaneh et al., 2012).

Another important distinction regards whether the travel information is prescriptive or descriptive. While prescriptive information consists in *one* recommendation only (e.g., the fastest route), descriptive information provides a range of possibilities among which the decision maker has to choose (e.g., the travel times of several alternative routes). Behaviourally speaking, when prescriptive travel information is provided, travellers have to decide whether to comply with the recommendation by at least comparing it with their planned choice. Conversely, when descriptive travel information is provided, travellers update their expectations of travel times and congestion patterns in the network (Parvaneh et al., 2012).

The timing of travel information provision, i.e. before (pre-trip information) or during the trip (en-route information), also influences travellers' behaviour. Pre-trip information acts more on a planning level as travellers usually have the opportunity to plan not only the fastest or

less congested route to take, but also the departure time that will result in possibly less delays. As pre-trip information is gathered before the trip starts, travellers still have enough time to process the information and weigh the *pro et contra* of complying with it in case it is not aligned to their expectations. En-route information, on the other hand, acts more on the tactical level as the time to process it is more limited. This is particularly true when dealing with personalized travel information about the current traffic situation. As the traffic situation is continuously changing, the chances are high that the traveller is updated about changes in the traffic situation and has to decide “on the spot” (or within a short time-span) whether to make a diversion (from a habitual route) to another route or not. The mental workload and stress consequently increase as a decision to divert (or not) may result in much longer travel times.

Travellers confronted with too much information may become oversaturated and show some difficulty to process it (Payne et al., 1993). As a consequence, they could develop simple heuristics to determine which information to focus on. On the other hand, if too many travellers react to travel information that is not properly updated, additional fluctuation in the traffic is generated. In both situations provision of travel information would lead to even more congestion instead of solving it. Understanding how travellers react to different sources, types and timing of travel information and whether provision of (more) travel information helps alleviating congestion requires further investigation.

It has to be noted that although provision of travel information has potential to benefit both the travellers personally and the transportation system as whole, the latter is strongly related to how travellers’ react to information. It is out of the scope to discuss how an optimal system benefit could be achieved.

2.2 Data collection efforts and empirical findings

This section presents an overview on data collection efforts to investigate the relationship between travellers’ route choice behaviour and travel information. As previously mentioned, aspects of particular interest are use of travel information, behavioural changes due to travel information and willingness to pay for travel information. These aspects are discussed in sections 2.2.1 to 2.2.3.

2.2.1 Use of travel information

Since the second half of the 1990’s, the interest in the role of travel information has increased due to its potential to influence and steer travellers to less congested routes (Emmerink et al., 1996; Hato et al., 1999; Polydoropoulou and Ben-Akiva, 1999; Khattak et al., 2003; Molin and Timmermans, 2006). Initial studies about travel information traditionally focused on requirements for successful utilization of ATIS tools (Crosby et al., 1993; Koutsopoulos and Xu, 1993; Schofer et al., 1993; Ng et al., 1995). As at that time ATIS tools had just started to be used by the general public, special attention was put on requirements that would make it more attractive to them. Consequently, those studies usually engaged on discussions such as whether the travel information should be static (e.g. location of points of interest) or dynamic (e.g. travel information), qualitative (e.g. road congested) or quantitative (e.g. amount of minutes of the delay) and prescriptive or descriptive. In addition, the level of reliability of the travel information was a point of concern. In other words, focus was put on the travel information characteristics’ that would influence its utilization. A peculiarity of such studies

was that due to the novelty of the ATIS tools, the general approach consisted in guidelines on how to evaluate such tools rather than experiments about travellers' reaction to it.

From the moment that research into provision of travel information started to be conducted, attention was put on travellers' reaction to information, i.e. on its use. Use can be defined in several ways. Some drivers may reduce their level of anxiety just by being aware it exists/is available – even though they may not even consult it; others may review the information on a regular basis, but make only limited use of route guidance and not on the traffic situation; still others may accept and comply with the advice without question (Schofer et al., 1993). Each of these produce a different set of individual and social benefits, that do not necessarily help alleviate congestion.

Studies on provision of travel information have traditionally focused on two main streams: route choice behaviour of drivers as users of travel information and performance of the road network (Hato et al., 1999). While studies about drivers refer to their behaviour, the ones about road network refer to the impacts of travel information in the transportation network. This thesis primarily deals with the behaviour of drivers, but also discusses implications to network performance. A recurrent assumption of studies about drivers as users of travel information is that drivers who receive perfect travel information immediately react to the reported changes in traffic conditions. However, as discussed in the literature, the assumption that travellers immediately react to it appears to be unrealistic (Ben-Akiva et al., 1991; Mahmassani and Jayakrishnan, 1991; Polak and Jones, 1993). This is especially true for situations where multiple sources of travel information are available and travellers have to choose which source to use.

A number of experiments about the effects of travel information on travellers' behaviour have been reported in the literature. They are mostly either stated preference (SP) surveys (Abdel-Aty et al., 1997; Wardman et al., 1997; Ben-Elia and Shiftan, 2010) or interactive route choice experiments in simulation environments in which, similar to a game, travellers make (consecutive) route choices and their behaviour towards risk and different types and quality of travel information is investigated (Chen and Mahmassani, 1993; Vaughn et al., 1993; Koutsopoulos et al., 1994; Chen et al., 1999; Selten et al., 2007; Bogers, 2009; Ben-Elia and Shiftan, 2010).

Some revealed preference (RP) datasets have also been reported in the literature. These are mostly studies in which either (i) travellers report their past route choices (Mahmassani et al., 1993; Rose et al., 2008), (ii) data from Global Positioning System (GPS) devices is used (Frejinger and Bierlaire, 2007; Papinski et al., 2009; Bierlaire et al., 2010; Menghini et al., 2010) or (iii) combination with SP surveys are proposed (Zhang and Levinson, 2008). The issue with studies on past choices is that they are often hardly related to an actual network in terms of alternatives and traffic conditions. Studies using GPS data lack of knowledge/data about how travellers acquire travel information and what the trip purpose was. Finally, about studies combining RP with SP surveys, little is known about the traffic conditions in the network. Consequently, most of the available RP data sets are not entirely appropriate to investigate travellers' reactions to travel information. There are some exceptions such as the works of Shiftan et al. (2010), Joh et al. (2011) and Tseng et al. (2011), which, however, are subject to other limitations such as limited sample size and/or duration of the experiment

(Shiftan et al., 2010; Joh et al., 2011) and primary focus on departure time decisions (avoidance of morning peak hours) rather than route choices (Tseng et al., 2011).

Insights from the literature suggest that (i) less experienced drivers (in terms of frequency travelling) tend to comply more with travel information than more experienced drivers and that drivers in general can rapidly identify the accuracy of the travel information (Vaughn et al., 1993), (ii) the degree of familiarity with the road network influence the use of travel information (Bonsall, 1996), (iii) both the traffic situation and characteristics of the decision maker influences the required types and quality level of pre-trip information (Polak and Jones, 1993), (iv), the efficiency in travellers' route choice and knowledge about network conditions can be improved by provision of real-time travel information (Adler and Kalsher, 1994), (v) as experience increases, travellers are more reluctant to make use of travel information and tend to prefer routes with lower average travel times but greater travel time variance (Shiftan et al., 2010) and (vi) socio-economic characteristics and trip contexts greatly influence the process of acquiring travel information (Joh et al., 2011). In practice, however, financial results of TomTom for the first quarter of 2014, indicate that the sales of personal navigation devices in the European market decreased 12% when compared to the same period in the year of 2013 (TomTom, 2014). No specific mention is made to the services of provision of traffic information. This suggests that the interest of travellers in acquiring and making use of travel information may be diminishing.

2.2.2 Changes in route choice due to travel information

Over the past decades, the main insights into the process behind route changes also come from SP experiments. Most of these are either interactive simulator experiments (Mahmassani and Liu, 1997; Mahmassani and Liu, 1999) or home surveys (Jou et al., 1997; Jou, 2001). Findings from the literature suggest that (i) commuters are willing to change their route choice under well-developed route guidance systems (Mahmassani and Liu, 1999), (ii) commuters' route switching decisions are based on the expectation that the total travel time will decrease by a certain threshold (indifference band), which varies with the remaining travel time to the destination, subject to a minimum absolute decrease of about 1 min (Mahmassani and Liu, 1999), (iii) commuters who receive pre-trip information are more likely to switch routes than those who do not (Mahmassani and Liu, 1997), (iv) the decision to change routes and departure time is part of the same "choice package" (Caplice and Mahmassani, 1992; Mahmassani and Liu, 1999). It is as if depending on the route to be taken, travellers have to adjust their departure times (and/or vice-versa), (v) there is no consensus whether route (Neuherz et al., 2000; Petrella and Lappin, 2004) or departure time (Jou and Mahmassani, 1996) changes is the most likely adaptation due to information provision (Chorus et al., 2006).

Drawbacks of most of those data collection efforts are issues regarding their external validity² and the fact that the home surveys in general consist of single interviews, thus not taking into account behavioural changes over time. Nevertheless, the importance of such studies cannot be neglected as due to the scarcity of RP efforts, either of short or long duration, primordial insights into the role of travel information to changes in routes choices come from SP experiments.

² See section 2.3 for further details

2.2.3 Willingness to pay for travel information

Policy makers have been hesitant to adopt road pricing solutions despite the fact that economists have been in favour of the cause for many decades (Brownstone et al., 2003). One of the main reasons is the strong reluctance from the road users to accept restrictive mobility measures. Although travellers may be willing to avoid driving in congested areas in case they receive some reward (Spitsmijden Program, 2006; Tseng et al., 2011), this is generally not the case if they have to pay for driving under such conditions. A question is then raised regarding whether travellers are willing to pay to receive travel information and as a reward/compensation (possibly) avoid congestion.

As discussed in section 2.1, there is a great variety of different sources and types of travel information available. In addition, providers of travel information have become more and more specialized. The current technological developments allow, for instance, combining existent traffic information with data from millions of mobile phone users driving in the transportation network. This helps to enhance the accuracy of the travel information, which can be updated every two minutes (TomTom, 2011). While on the one hand non-personalized and semi-personalized travel information are usually provided without additional costs to the users, this is not the case for personalized real-time travel information. The latter is traditionally provided by private operators, subject to certain cost, and consists in highly specialized and personalized travel information.

The fact that travellers are faced to a great variety of sources influences the trade-off between the perceived added value of the travel information and willingness to pay it. Commuters in particular might consider this trade-off very important as on the one hand they (usually) drive in highly congested periods and on the other hand they have good knowledge about the network and traffic conditions.

Some conceptual frameworks that describe the process throughout which travellers acquire travel information provided via ATIS have been described in the literature (Ben-Akiva et al., 1991; Khattak et al., 1993; Schofer et al., 1993; Denant-Boèmont and Petiot, 2003; Ettema and Timmermans, 2006). The main drawback of most of these studies is that it consists only of a conceptual framework, thus no real data involved. Besides this, the monetary component is not directly there, i.e. the financial cost associated with the acquisition of information.

Studies directly focusing on travellers' willingness to pay to receive real-time travel information also surface the literature (Khattak et al., 1995; English et al., 1996; Hobeika et al., 1996; English et al., 1997; Polydoropoulou et al., 1997; Khattak et al., 2003). Great advantage of these studies is that most of them are based on RP surveys conducted with users of existent information systems. Disadvantage is the conduction of single interviews and lack of information regarding the frequency of use of travel information among others.

Insights from the literature suggest that (i) less experienced drivers are more unlikely to purchase information devices than more experienced drivers (Vaughn et al., 1993), (ii) travellers would only be willing to acquire an ATIS if (a) the perceived benefits outweigh the perceived costs (Schofer et al., 1993), (b) the variability of the travel conditions is high (Denant-Boèmont and Petiot, 2003) and (c) there are available comparable alternative routes (Khattak et al., 2003), (iii) current users of travel information systems would be willing to pay

for a similar type of information that was already being provided free of charge (Remer et al., 1996; Polydoropoulou et al., 1997; Wolinetz et al., 2001).

2.3 Advantages and disadvantages of different data collection methods

There are two main types of data collection methods employed to investigate how travellers react to travel information: SP and RP experiments. While in SP experiments travellers report how they would react given a specific context, in RP experiments choices are actually made. For instance, in the context of route choices, in SP experiments travellers report which route they would choose given a context, while in RP experiments the route is actually chosen.

Initial developments in the use of SP methods mainly emphasized judgemental tasks. These were tasks in which a respondent was asked to rate or rank a number of mixes of attributes associated with a particular choice context. The interest in SP methods substantially increased as of the moment it started to be applied to investigate behavioural responses related to situations in which an alternative was not currently available or it was difficult to assess its mixes of attributes (Hensher, 1994). At that time Louviere and Hensher (1983) showed how a preference experiment (i.e. a number of alternatives mixing attributes) could be extended to incorporate choice experiments in a way that an individual could choose among fixed or varying choice sets. This made it possible to estimate discrete choice models and hence predict market shares and investigate the effects of different attributes in people's choices. Since then, SP choices have been widely popular in transportation.

RP methods, although also well-known for a long time in the field of transport, have raised more attention in the past years due to the popularity of Global Positioning System devices (GPS) (Papinski et al., 2009; Bierlaire et al., 2010; Spissu and Meloni, 2011). Such technology allows the tracking of travellers during their trips and contributes to get precise information such as travel and departure times, whether and where travellers were stuck in traffic, duration and location of intermediate stops, etc. This is done without interfering in travellers' behaviour.

There are advantages and disadvantages associated with both data collection methods. Advantages of SP experiments are their efficiency, flexibility to design different scenarios (that result in adequate data to answer proposed research questions), easiness to collect multiple observations for the same person and relatively low costs. On the other hand, SP experiments are subject to the big drawback of external validity, i.e. it is questionable whether the outcomes would be valid in a real setting. In other words, it is (often) argued whether in a real situation travellers would behave as they said they would and whether the context investigated can be reproduced in the real world. Although RP experiments are not subject to drawbacks of external validity, they are very difficult to control. This is especially the case for experiments of long duration. In addition, as argued in the literature, a limitation of most RP studies is the use of cross sectional data, i.e. one observation for each individual. This is due to the fact that collecting RP data with repeated observations for the same group of people requires strong commitment from the respondents and a longer data collection period (Axhausen et al., 2002).

The use of RP and SP methods depends on the context investigated. For instance, SP methods are the only possibility for situations in which one of the alternatives is currently not

available. For other situations, in principle, RP methods would be more appropriate from a realistic point of view. However, how to control and guarantee the experimental conditions during the experiment? There is also the possibility of combining the powerful capability of SP methods with RP methods. This would help address issues/concerns about how well travellers' stated preferences about hypothetical situations correspond to their behaviour when facing such situations. The choice for one method or another (or a combination of methods), therefore should be weighed and investigated case by case.

2.4 Conclusions

The literature shows that the main insights into the relationship between travellers' route choice behaviour and travel information come from SP experiments (including travel simulators). The main advantages of SP experiments are easy collection of panel data, i.e. multiple observations for the same traveller, and flexibility of scenarios. While panel data can capture choices over time, the flexibility of scenarios allows investigating technological advances not yet implemented. A main point of criticism about SP approaches is the external validity of the data, i.e. whether in a real situation travellers would behave as they said they would and whether the scenarios reflect real situations.

RP approaches are suggested as the alternative to overcome issues regarding external validity. The drawbacks associated with RP experiments, however, may explain why they are not commonly done. For instance, if interviews about previous trips are conducted, travellers may not be entirely aware of what happened during the trip depending on the timing the interview is conducted. The possibility to relate past choices with a transportation network in terms of travel alternatives and traffic conditions is also big challenge (see section 3.4). Lack of real experimental settings with significant variation of travel time across at least two alternatives, difficulties to measure travel time data and planning and deployment costs are also limiting factors (Carrion and Levinson, 2012). In addition, to investigate behavioural changes over time longer data collection periods and strong commitment from the participants is required.

Despite the valuable insights provided by SP experiments, the literature highlights the need of more RP experiments to validate (or challenge) current findings. Especially due to the duration of the experiments, most RP data sets available are of limited use for the assessment of behavioural changes over time. This is particularly true for commuting trips. Commuters are a group of drivers of particular interest because they usually drive during (heavy) congestion periods. In such situations, the (potential) benefits of providing travel information are higher. A question then is raised regarding to what extent travellers that are already familiar to performing a specific trip react to travel information. Not only the characteristics of the travel information, such as sources, quality and timing of provision are relevant, but also the prior knowledge of traffic situation. These are aspects to be further investigated.

In order to overcome gaps in the literature regarding data collection efforts, this thesis proposes to conduct a RP data collection in a real network to investigate the route choice behaviour of commuters under provision of travel information (Chapter 5). The characteristics of the data collection should ensure that (i) the route choices can be related to a transportation network in terms of alternatives and traffic conditions, (ii) alternative routes are comparable and have significant variation of travel times and (iii) the attributes of the non-chosen alternatives can also be defined.

Chapter 3

State-of-the-art on travel behaviour theory

To estimate models of route choice behaviour, assumptions have to be made regarding the decision rule underlying travellers' decisions. The majority of existing route choice models is based on the utility maximization assumption that people act rationally to get the maximum utility (benefit) from the decision made. The first formal model, almost unchallenged for about 70 years, is the expected utility model formalized by Von Neumann and Morgenstern (1947), which has powerful and tractable axioms (de Palma et al., 2008). Nevertheless, several well-known experiments have shown that such axioms can be challenged or violated even in the context of simple choice situations. This led to the development of non-expected utility theories, which claim the ability to capture a more realistic behaviour by addressing such violations. A main question is whether travellers' route choice behaviour is better qualified as utility maximizers or should be qualified differently.

This chapter presents the state-of-the-art on behavioural theories to model travellers' route choice behaviour. Focus is put on expected utility theory, prospect theory and regret theory because the first is the main stream and the others address important violations of its axioms. For a comprehensive review on non-expected utility theories, refer to Starmer (2000).

Section 3.1 discusses the fundamentals of each theory. A comparative analysis between the theories is presented in section 3.2. Furthermore, it motivates the choice of the behavioural rule to be used in this research. Section 3.3 makes a link between behavioural theories and discrete choice analyses. In section 3.4 some important modelling issues are discussed. Finally, conclusions are then put forward in section 3.5.

This chapter is based on:

Ramos, G. M., W. Daamen and S. P. Hoogendoorn (2014). A state-of-the-art review: developments in utility theory, prospect theory and regret theory to investigate travellers' behaviour in situations involving travel time uncertainty. *Transport Reviews: A Transnational Journal*. Published on Jan. 13th, 2014, doi:10.1080/01441647.2013.856356.

3.1 Fundamentals of travel behaviour theory

In the modelling of route choice behaviour assumptions are made regarding the behavioural rule underlying travellers' decisions. The fundamentals behind *expected utility theory*, *regret theory* and *prospect theory* are respectively discussed in sections 3.1.1 to 3.1.3.

3.1.1 Expected utility theory

Main developments in *expected utility theory* date from 1947 when Von Neumann and Morgenstern provided a sound theoretical foundation for using *expected utility theory* as a guide for decision making under risk. Considering \mathbf{k} , \mathbf{l} and \mathbf{m} as prospects, where a prospect is a consequence associated with a probability of occurrence, they showed that preferences obey certain intuitively appealing axioms, namely *ordering* (divided into *completeness* and *transitivity*), *continuity* and *independence* (Starmer, 2000):

- (i) Completeness: for all \mathbf{k} , \mathbf{l} , either $\mathbf{k} \succ \mathbf{l}$, $\mathbf{k} \prec \mathbf{l}$ or $\mathbf{k} = \mathbf{l}$;
- (ii) Transitivity: for all prospects, if $\mathbf{k} \succcurlyeq \mathbf{l}$ and $\mathbf{l} \succcurlyeq \mathbf{m}$, then $\mathbf{k} \succcurlyeq \mathbf{m}$;

Axioms (i) and (ii) together imply the existence of a preference among any pair of prospects and the transitivity of this preference.

- (iii) *Continuity*: for all prospects, if $\mathbf{k} \preccurlyeq \mathbf{l} \preccurlyeq \mathbf{m}$, there is a probability P between 0 and 1 that $p \cdot \mathbf{k} + (1 - p) \cdot \mathbf{m}$ is equally preferable to \mathbf{l} alone.

Axiom (iii) implies that prospects of intermediate preference are indifferent to a specific mix between a pair of better and worse prospects. In addition, the axioms of *ordering* and *continuity* together imply that preferences over prospects are likely to be represented by a function $V(\cdot)$, which assigns real-valued index to each prospect. The function $V(\cdot)$, therefore is a representation of preferences. People are assumed to have well-defined preferences, be able to rank them and to impose minimal restrictions on their precise form. Considering two prospects \mathbf{k} and \mathbf{l} , prospect \mathbf{k} would be chosen over \mathbf{l} if, *and only if*, the value assigned to \mathbf{k} by $V(\cdot)$ is not less than that assigned to \mathbf{l} , i.e. $V(\mathbf{k}) \geq V(\mathbf{l}) \leftrightarrow \mathbf{k} \succcurlyeq \mathbf{l}$.

- (iv) *Independence*: for all prospects, if $\mathbf{k} \succcurlyeq \mathbf{l}$ then $(\mathbf{k}, p; \mathbf{m}, (1-p)) \succcurlyeq (\mathbf{l}, p; \mathbf{m}, (1-p))$.

Axiom (iv) states that preference between two prospects is independent of their common components. Therefore, in the evaluation of a prospect, their common components are perceived as if they could be cancelled (Weber and Camerer, 1987).

Based on these four axioms, expected utility theory proposes that individuals are utility maximizers, thus aiming to maximize the utility of their choice. Preferences can then be represented by:

$$V(\mathbf{k}) = \sum_q p_q \cdot U(x_q) \quad (3.1)$$

where \mathbf{k} is any prospect and $U(\cdot)$ is a utility function defined on the set of consequences x_t .

As the next sections discuss in more detail, *prospect theory* and *regret theory* are respectively derived from the relaxation of the *independence* and *transitivity* axioms. While the *independence* axiom imposes strong restrictions on the form of preferences and gives

expected utility theory its most empirical content, the *transitivity* axiom assures that there is a strict preference for one alternative over another and that this preference is transitive.

3.1.2 Prospect theory

Prospect theory, a descriptive model of decision making under risk, was developed in 1979 as a critique to *expected utility theory* (Kahneman and Tversky, 1979). The point of the critique is that changing the ways in which options are presented appears to generate predictable shifts in preference and to systematically violate the axioms of *expected utility theory*, in particular the *independence* axiom. There are three main implications. Firstly, the expected utility of a two-stage gamble (a gamble in which the outcomes are gambles) is the same as that of a single-stage gamble with the same probability distribution of final outcomes. Secondly, the expected utility of a lottery is linear in probabilities. Thirdly, the expected utility is forced to be a combination of utilities of the independent outcomes and probabilities in the prospect.

Violations of the axioms of *expected utility theory* were first shown in the well-known Allais' paradox (Allais, 1953) in which people had to choose between two lotteries $ALot_1 = (\$1M, 1)$ or $BLot_1 = (\$5M, 0.1; \$1M, 0.89; \$0, 0.01)$ and then between $ALot_2 = (\$1M, 0.11; \$0, 0.89)$ or $BLot_2 = (\$5M, 0.1; \$0, 0.9)$. A person behaving as expected utility maximizer would choose option $BLot$ in both experiments because the *independence* axiom requires that the difference between the pairs of lotteries to be ignored. Thus, the "common consequence" of a 0.89 chance of winning \$1M in the lotteries $ALot_1/BLot_1$, or a 0.89 chance of (winning) \$0 in the lotteries $ALot_2/BLot_2$ would be cancelled. Contrary to the assumption that people act rationally behind utility theory, Allais believed that people do not act rationally and as a consequence lottery $ALot_1$ would be preferred in the choice (win for certainty) and lottery $BLot_2$ in the second (very different rewards for slightly different winning probabilities). This was confirmed in the experiments and consequently violated the *independence* axiom.

Prospect theory explains choices as a two-step process: an initial phase of *editing* and a subsequent phase of *evaluation*. In the *editing phase* the options are organized and reformulated by the application of heuristics to simplify the evaluation. In the *evaluation phase* the prospect is subjectively valued. The evaluation phase is divided into two scales: a weighting function $w(p)$ and a value function $v(x)$. The weighting function $w(p)$ associates to each probability of occurrence p a decision weight ω that reflects the impact of p on the overall value of the prospect. The value function $v(x)$ reflects the subjective value of the outcome, thus measuring the deviations from the reference point. Thus, neither $v(x)$ is perceived as valuing x nor $w(p)$ as valuing p .

In 1992, advances were made to extend *prospect theory* to multinomial choice set (Tversky and Kahneman, 1992), which led to the development of *cumulative prospect theory*. This extension was necessary to evaluate situations involving uncertainty in which some of the outcomes or probabilities are unknown, which is the case for situations involving travel times. *Cumulative prospect theory* employs cumulative rather than separable decision weights, is applicable to uncertain and risky prospects and allows different weighting and value functions for gains and losses due to people's different perceptions of gains and losses. Distinct functions are associated with positive and negative outcomes, V^+ and V^- , and the overall value of a prospect is given by:

$$V = V^+ + V^- \quad (3.2)$$

Where:

$$V^+ = \sum_{q=0}^{pos} \pi_q^+ \cdot v(x_q) \quad \text{and} \quad V^- = \sum_{q=-neg}^0 \pi_q^- \cdot v(x_q) \quad (3.3)$$

where, π^+ and π^- are decision weights associated respectively with positive (*pos*) and negative (*neg*) resultant outcomes, $-neg \leq q \leq pos$, $v(\cdot)$ is the value function in CPT and x_t is the resultant outcome of an attribute.

The value function in CPT is strictly increasing and satisfies $v(x_0) = v(0) = 0$. For the other values:

$$v(x_q) = \begin{cases} (x_q - rp)^\alpha & \text{if } (x_q - rp) \geq 0 \\ -\lambda \cdot (rp - x_q)^\theta & \text{if } (x_q - rp) < 0 \end{cases} \quad (3.4)$$

where rp is the reference point against which the attribute is measured to. The other parameters of the value function, α , θ and λ , respectively reflect (i) the degree of diminishing sensitivity for gains, (ii) the degree of diminishing sensitivity for losses, i.e. the decrease on the marginal value of gains and losses with their magnitude, and (iii) the degree of loss aversion, i.e. the aggravation one experiences when incurring in losses.

The decision weights are defined by:

$$\pi_{pos}^+ = w^+(p_{pos}) \quad \text{and} \quad \pi_{-neg}^- = w^-(p_{-neg}) \quad (3.5)$$

$$\pi_q^+ = w^+(p_q + \dots + p_{pos}) - w^+(p_{q+1} + \dots + p_{pos}) \quad (3.6)$$

$$\pi_q^- = w^-(p_{-neg} + \dots + p_q) - w^-(p_{-neg} + \dots + p_{q-1}) \quad (3.7)$$

where w^+ and w^- are strictly increasing weighting functions satisfying $w^+(0) = w^-(0) = 0$ and $w^+(1) = w^-(1) = 1$ and given by:

$$w^+(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}} \quad \text{and} \quad w^-(p) = \frac{p^\delta}{(p^\delta + (1-p)^\delta)^{\frac{1}{\delta}}} \quad (3.8)$$

where the parameters γ and δ define the curvature of the weighting function and capture the distortion in the perception of probabilities for gains and losses.

The parameters of *prospect theory* are responsible for the main differences in relation to *expected utility theory*. If (i) there were no decreases in the marginal values of gains and losses with their magnitude ($\alpha = \theta = 1$), (ii) incurring in losses did not cause an aggravation of the experience ($\lambda = 1$), and (iii) there were no distortions in the perception of probabilities ($\gamma = \delta = 1$), *prospect theory* would be simplified to *expected utility theory*. Behavioural experiments, however, showed that people tend to overweight small probabilities ($p < 0.1$) and underweight moderate and high probabilities ($p \geq 0.5$) (Tversky and Kahneman, 1992). The overweight of small probabilities implies risk-seeking when offered low-probability high-reward prize. The underweight of moderate and high probabilities implies relative

insensitivity to probability differences in the middle of the range, prevalence of risk-aversion in choices between probable gains and certainty and risk-seeking in choices between probable and sure losses. These are the main behavioural features captured by *prospect theory*, but not fully explained by *expected utility theory*.

3.1.3 Regret theory

Regret theory was developed as a theory of pair-wise choice based on the assumption that preferences depend on the performance of the chosen *and* foregone alternatives (Bell, 1982; Loomes and Sugden, 1982). The possibility that the non-chosen alternative proved to be more (or less) attractive, makes people associate a feeling of regret (or rejoice) to the decision.

The main difference between *regret theory* and *expected utility theory* is the relaxation of the *transitivity* axiom. As a result, if more than two alternatives were available *regret theory* would, in principle, allow the occurrence of *intransitive* choices of the type $h \leq i$, $i \leq j$ and $j \leq h$, where h , i and j are the alternatives. This cyclical preference could lead someone to become locked in an endless chain of trades and give the possibility for a broker to operate a *money pump*, i.e. induce a person to pay a small sum of money for the privilege of making a new exchange: a person who had changed h for i would be induced to pay a small sum for the privilege of exchanging i for j then to exchange j for h and so on.

This argument is often used to support the claim that cyclical preferences are irrational. However, as argued in Loomes and Sugden (1982), choosing an alternative from a choice set with two alternatives does not imply that the same alternative should be chosen from a choice set with three (or more) alternatives. This is because the choice that yields the minimum regret can be different for each set. The *intransitivity* property, therefore, does not necessarily imply that choices must be pair wise. Generalization of *regret theory* for choices involving three or more alternatives is possible.

According to RT, an alternative i is chosen from a choice set containing i and j if and only if:

$$\sum_{s \in S} [p(s) \cdot R_{ij}(s)] > 0 \quad (3.9)$$

where s is a particular state of the world, S gives all possible states of the world, $p(s)$ is the probability of occurrence of state s and $R_{ij}(s)$ is the regret/rejoice of alternative i compared to alternative j given state s . In addition, if $R_{ij}(s) = U_j(s) - U_i(s)$, RT equals EUT.

The main behavioural difference between *regret theory* and *expected utility theory* is that while the former assumes that the benefit of the choice depends on the performance of the chosen alternative in relation to the others, the latter argues that the benefit depends solely on the chosen alternative (Chorus et al., 2008a).

3.2 Comparative analysis between the theories

The large use and acknowledged validity of the principles of *expected utility theory* to model travellers' behaviour raises questions about its advantages and disadvantages in relation to *prospect theory* and *regret theory*. For instance, does it meet requirements that the others do not? How about its potential to capture realistic traveller behaviour?

To make the advantages and disadvantages of each theory clearer, this section elaborates on the following four criteria that reflect the decision process and trade-offs between simpler and more complex modelling frameworks: (i) level of accuracy and required efforts of decision mechanisms, (ii) data requirements and model identification (calibration/estimation), (iii) level of difficulty to apply each theory and (iv) ability to capture travellers' behaviour. These are respectively discussed in sections 3.2.1 to 3.2.4. An overview of the comparative analysis is presented in section 3.2.5.

3.2.1 Decision mechanism

A variety of decision mechanisms is used to help making choices in real life situations. Some decision mechanisms are more accurate and require more effort than others. Consequently, the choice for a particular mechanism depends, for instance, on the importance of the decision, time pressure and cognitive abilities (Payne et al., 1993). This is the case with the *weighted additive rule* when compared to the *satisfying heuristic rule*.

The *weighted additive rule* is an example of a high accuracy-high effort mechanism, according to which weights are associated with attributes. The *satisfying heuristic rule*, on the other hand, is an example of a low accuracy-low effort mechanism according to which the first evaluated alternative that meets some minimum aspiration level is chosen (Simon, 1995). The *satisfying heuristic rule* argues that because people are rationally bounded, applying a rule that requires a lot of mental effort may be beyond their bounds. Due to the fact that *expected utility theory*, *prospect theory* and *regret theory* use the high accuracy-high effort decision mechanism of the *weighted additive rule*, the disadvantage applies to all three. None of these theories therefore can be considered better if based only on the decision mechanism.

3.2.2 Number of parameters to be estimated

Considering each theory in its simplest form, i.e. only one attribute to be estimated (e.g. travel time), no additional parameter has to be estimated in utility and regret based frameworks, but five additional parameters (α , θ , λ , γ , δ) and a reference point have to be estimated in prospect theoretical frameworks.

Points of additional concern for models based on *prospect theory* is the lack of consensus about the reference point in a travel behaviour context (Ramos et al., 2013) and the fact that situations in which two or more attributes are considered have not yet been investigated. The latter requires, in principle, definition of reference points and (possibly) another five additional parameters for each attribute. Therefore, concerning the number of parameters, models based on *prospect theory* are by far in disadvantage.

3.2.3 Simplicity and tractability of the modelling framework

The greatest added value of utility-based models is their simplicity. Their calculation method is very intuitive and straightforward. The same holds for *regret*-based models. This is not the case, however, for models based on *prospect theory*. Direct application of prospect theory is already quite complex to be promptly understood. A question that has not yet been raised in the literature is whether and how to take socio-economic characteristics into account in the *prospect theoretical* framework. Given its complex framework, is it really necessary? If yes, firstly it should be determined how to set reference points for social-economic characteristics such as gender and age as, in principle, no reference point appears to be applicable. The

simplicity and tractability, therefore, are also disadvantages for models based on *prospect theory* while they are similar in utility and regret based models.

3.2.4 Ability to capture realistic behaviour

The greatest advantage of *prospect theory* is its potential to capture realistic behaviour. *Prospect theory* was developed based on experiments that captured violations of the axioms of *expected utility theory* and since then has been acknowledged for that. *Regret theory* claims more simplicity than *prospect theory*, but not to better capture behaviour. This discussion, however, is theoretical. From a theoretical point of view, *prospect theory* seems to be in advantage.

From the point of view of application, in the field of economics the advantages of *prospect theory* and *regret theory* have already been acknowledged. In the field of transportation, despite recent developments in the use of *prospect theory* (Gao et al., 2010; Li and Hensher, 2011; Xu et al., 2011), more applications are necessary to value their merits (Chorus et al., 2008b, a; Ramos et al., 2011). For instance, researches on *prospect theory* usually focus on investigating its properties, especially due to the need to investigate the behavioural appeal of its parameters. The literature on *regret theory* primarily uses the random regret minimization (Chorus et al., 2008b) approach and focus on mode choices. Mode choice is a type of decision in which the trade-offs between the attributes of the alternatives is usually very clear and explicit. This is typically the type of situation in which random regret minimization performs best. From the point of view of application, the principles of *expected utility theory* appear to be in advantage.

3.2.5 Overview of the comparative analysis

Under the assumption that all aforementioned criteria weigh equally, *regret theory* and *expected utility theory* could be considered equally suitable to model travel behaviour as both are in disadvantage in only one aspect. On the other hand, *prospect theory* is in disadvantage in all aspects except for its potential ability to capture behaviour. This may explain why the literature has put more emphasis on investigating the suitability of *prospect theory* despite the apparent favouritism of *regret theory* in this comparative analysis.

Outcomes of this comparative analysis imply that the large use of the principles of *expected utility theory* is not solely due to requirements that the other theories do not meet, but to its very tractable modelling framework. A question still to be answered is whether the ability of *prospect theory* and *regret theory* to capture more realistic behaviour can be translated into more realistic models. The need of further research to validate these theories thus cannot be neglected. Both the research and practice have already acknowledged the validity of utility-based models, while more research is still necessary to value the merits of *regret theory* and especially *prospect theory*. So far, despite violations of axioms of *expected utility theory*, it still provides the best framework to model and investigate travellers' behaviour.

3.3 From behavioural theory to discrete choice models

As discussed in the previous section, theoretical comparison between the theories shows that none is dominant, i.e. scoring best for all criteria. Both *regret theory* and *expected utility theory* seem to be equally suitable to model travellers' behaviour, while *prospect theory*

appears to be in disadvantage. This allows to argue that (i) the chosen criteria should not weigh equally and (iii) the most suitable theory might depend on the criteria chosen.

The validity of the principles of *expected utility theory* has already been acknowledged by both the research and practice. On the other hand, against *prospect theory* and *regret theory* weighs the need of more experiments to have their validity fully acknowledged. Consequently, results based on *prospect theory* and *regret theory* are still subject to questionings regarding their validity. Therefore, in particular due to the acknowledged validity and tractability of utility-based models, the behavioural rule of utility maximization is the one chosen to be used in the modelling framework of this thesis.

The fact that some axioms of *expected utility theory* are challenged even in simple choice tasks, however, remains. This raises a question on how to deal with the violations of its axioms. While on the one hand non-expected approaches extend the axiomatic approaches in ways that explain violations and deviations from utility theory, non-deterministic approaches add an error term to the utility function. The error term represents uncertainties from the analyst and, as discussed in Hey and Orme (1994), depending on its structure *expected utility theory* provides satisfactory predictions of individual choice. Random utility models (RUM), which are based on the principle of utility maximization, are the most commonly used non-deterministic approach for route choice models and are traditionally estimated in the framework of discrete choice analysis (Ben-Akiva and Lerman, 1985; Train, 2009).

Discrete choice models describe the decision making as a process of choices among alternatives (Ben-Akiva and Lerman, 1985; Train, 2009). In this framework, the set of alternatives is called *choice set* and is a very important feature as it defines the set of alternatives among which the decision maker can choose. The choice set must be finite and when making choices the decision-maker has to necessarily choose one (and only one) alternative from the choice set (Train, 2009).

Incorporating a behavioural rule into a discrete choice framework starts from the recognition that preferences and perceptions may vary across individuals and are represented by specific utility functions $U_i(\cdot)$. If the decision maker does not know with certainty which event is true, the decision is called a decision under uncertainty. Conversely, if objective probabilities of occurrence are known for all events, a probability distribution is generated and a decisions under risk is faced (which are a special case of decisions under uncertainty) (de Palma et al., 2008).

Generally speaking, the expected value of a prospect x for individual n , can be expressed by:

$$\sum_i \pi_{in} \cdot U_{in}(x_i) \quad (3.10)$$

where $\pi_{in} = w_n(p_1 + \dots + p_i) - w_n(p_1 + \dots + p_{i-1})$ denotes the perceptions of probabilities as perceived by individual n and has different specifications depending on the theory applied, i.e. the definition of w .

The choice for a theory is associated with a trade-off between the flexibility of the utility and/or probability-weighting functional forms in relation to the degree of heterogeneity that can be taken into account (de Palma et al., 2008). In the RUM framework a decision maker is assumed to select the prospect x_i that maximizes the value of the prospect.

The utility function is defined by a deterministic component and a random component (the error term). The *deterministic component* is specified by the analyst and comprehends the attributes that are believed to influence the choice of the decision maker. This results in a single objective function that expresses the attraction of an alternative, or the benefit of choosing it, in terms of its attributes. The *random component*, on the other hand, represents uncertainties such as errors in the measurements, unobserved variations in taste among the decision makers and missing attributes. In other words, it reflects observational deficiencies of the analyst in the specifications of the deterministic component. The utility function can thus be expressed as:

$$U_{in} = v_{in} + \varepsilon_{in} \quad (3.11)$$

$$U_{in} = f(A_i; \beta_n) + \varepsilon_{in} \quad (3.12)$$

where U_i is the utility of choosing alternative i , v_i is the deterministic component of the utility, which depends on the vectors of attributes of an alternative (A_i) and parameters reflecting preferences and a probability weighting (β_n), ε_{in} is the random component of the utility with some standard distribution. In other words, the utility is considered to be a random variable composed of an observable $f(X_i; \beta_n)$ and an unobservable ε_{in} component.

Consider, for example, a route choice decision in which the analyst assumes that travellers' behaviour is influenced only by two attributes: travel time and availability of travel information as a dummy variable. The deterministic component of the utility would then be written as:

$$v_i = \beta_1 \cdot TT_i + \beta_2 \cdot TravelInformation_i \quad (3.13)$$

where the β values reflect the relative weight of travel time and travel information. In the RUM-setting, a parameter estimate refers to the increase or decrease in utility due to one-unit increase or decrease in an attribute's value. On the other hand, in regret-based frameworks, it refers to the increase or decrease in regret by comparing one alternative with another due to one unit increase or decrease in an attribute's value (Chorus et al., 2008b). With respect to prospect-based approaches, due to the nonlinearity of its utility function although the estimated parameters reflect an increase or decrease in utility it is not due to one-unit increase in an attribute's value.

While the specification of the deterministic part of the utility reflects the behaviour of the decision makers, the specification of the random component provides the mathematical formulation of the model. Different specifications of the random component consequently lead to different models.

According to utility theory, the decision maker chooses the alternative with the highest utility. Due to observational deficiencies from the analyst, however, the models cannot predict with certainty what alternative will be chosen. Instead, the model results in a probability of choosing an alternative.

Consider a choice among two alternatives: i and j . The probability that alternative i is chosen is given by:

$$P_n(i | C_n) = P(U_{in} \geq U_{jn}, \forall j \in C_n, j \neq i) \quad (3.14)$$

$$P(v_{in} + \varepsilon_{in} \geq v_{jn} + \varepsilon_{jn}) = P(\varepsilon_{jn} - \varepsilon_{in} \leq v_{in} - v_{jn}) = P(\varepsilon_n \leq v_{in} - v_{jn}) \quad (3.15)$$

where P_n denotes the probability of decision maker in choosing alternative i given the choice set C_n and i and j are the alternatives of the choice set.

It is possible to observe in equation 3.15 that the random term plays an important role in the probability that an alternative is chosen. Depending on its specification different model properties are defined. Many route choice models assume that the random term is Gumbel distributed type I extreme value. The difference between two extreme value variables is distributed logistic and allows viewing the disturbances as the sum of a large number of unobserved but independent components (*i.i.d.*) (Train, 2009). Under this assumption, equation 3.15 can be further developed to equation 3.16, where μ is a scale parameter³ ($\mu > 0$):

$$P_n(i | C_n) = P(U_{in} \geq U_{jn}) = \frac{1}{1 + e^{-\mu(v_{in} - v_{jn})}} = \frac{e^{\mu v_{in}}}{e^{\mu v_{in}} + e^{\mu v_{jn}}} \quad (3.16)$$

Generalizing equation 3.16 for a choice set with more than two alternatives results in the well-known probability of choosing an alternative given by the multinomial logit models (MNL):

$$P_n(i | C_n) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} \quad (3.17)$$

3.4 Challenges ahead: practical issues associated to estimating random utility models based on revealed preference data

As part of the outcomes of Chapter 2 it was decided to conduct a revealed preference (RP) experiment to investigate the relationship between travellers' behaviour and travel information. In particular, the RP experiment should be of long duration, conducted in a real network in which travellers are subject to different conditions of travel information and allow relating the choices made to a real road network in terms of alternatives and traffic conditions.

In this chapter it was defined that the behavioural rule behind travellers' decision to be adopted in this thesis should be that of utility maximization. The most traditional utility-based approach to estimate route choice models is that of discrete choice analysis (Ben-Akiva and Lerman, 1985; Train, 2009), in which the definition of choice set is a very important feature.

Considering a situation in which travellers have to make a choice among t possible alternative routes, the route to be chosen is the one from which the traveller derives the highest utility. The attributes influencing the utility of a choice are defined by the analyst taking into account

³ For a review on the role of the scale parameter see chapters 2 and 3 of Train, K. E. (2009) *Discrete choice methods with simulation*, 2nd ed, Cambridge University Press, New York.

factors that are believed to influence travellers' route choice such as travel time, travel information and habit.

Three main modelling issues, however, are associated with the estimation of route choice utility models based on RP data collected in real networks. Firstly, choice set sampling is difficult (Bovy, 2009; Prato, 2009; Frejinger and Bierlaire, 2010). Secondly, utilities may share unobserved attributes and consequently the IIA property (independence from the irrelevant alternatives) does not hold (Cascetta et al., 1996; Ben-Akiva and Bierlaire, 1999; Bekhor et al., 2001; Frejinger and Bierlaire, 2007). Thirdly, the attributes of the alternatives, such as the travel times at the time of the observation, are in general unknown.

Choice set definition is a key point in discrete choice analysis. As discussed in Prato (2009), there exists a variety of algorithms to generate choice sets, most of them based on the shortest-path rule. The main issue, however, is whether the generated choice set matches the actual choice set considered by the travellers. In order to overcome such issue, Frejinger et al. (2009) proposed a sampling approach where utilities are corrected for the sampling. Prediction issues, however, remain.

The second issue mentioned, i.e. utilities may share unobserved attributes, is also a central point in utility models, especially the multinomial logit models. In addition it is also related to the definition of the choice set. Violations of the IIA property arise when the analyst includes in the choice set alternatives that are not actually (perceived as) distinct. In the context of route choices this would be the case of considering routes with long overlaps as distinct while travellers consider them as the same route. In such situations, the attributes of the alternatives would be confounded resulting in violations of the IIA property.

Finally, the third issue put forward, i.e. the attributes of the alternatives are unknown, is a practical issue that arises from the use of RP data. As discussed in section 2.3, one of the main drawbacks in collecting RP data is the difficulty to relate the choices made to an actual network in terms of traffic conditions. In the context of route choices this has a huge impact on the definition of the travel times of the routes that were not chosen. In order to overcome such issue it is necessary that the data collection comprises not only data about travellers' route choices, but also about the traffic conditions in the network.

The level of complexity associated with the estimation of route choice models based on RP data collected in real networks is therefore very high. One of the main contributions of this PhD thesis is to overcome gaps in the literature associated with data collection efforts and modelling of travellers' behaviour. The challenge that it is portrayed to be tackled is therefore to investigate travellers' behaviour in a real network overcoming issues associated with model estimation.

3.5 Conclusions

In the modelling of travel behaviour, assumptions have to be made regarding the behavioural rule underlying the choices made. This state-of-the-art on behavioural theories presented a review on three main behavioural decision theories: *expected utility theory*, *prospect theory* and *regret theory* and some implications to model estimation. These theories were chosen because *expected utility theory* is the most used theory to model travellers' behaviour and

prospect theory and *regret theory* are among the most discussed alternatives due to their potential to capture behavioural phenomena not fully explained by *expected utility theory*.

The use, similarities and differences of each theory were discussed in the previous sections. Comparisons of their fundamentals show that under certain conditions *prospect theory* and *regret theory* equal *expected utility theory*. The added value of *prospect theory* lies within its possibility of capturing loss aversion, risk aversion and risk seeking; the added value of *regret theory* lies with its possibility of capturing regret aversion. Notwithstanding their potential to capture behaviour better than their utility counterpart, the literature on travellers' behaviour has already acknowledged the validity of utility-based models. In the context of route choice decisions under travel time uncertainty, while significant improvements have already been made with utility-based models, studies on prospect and regret based models are mostly (Ramos et al., 2014)

The superiority of one theory over the others is not claimed here. Nevertheless, this state-of-the-art helped defining the behavioural principle underlying the modelling framework of this thesis. Theoretical comparison between the theories shows that none is "dominant", i.e. scoring best for all criteria. Both *regret theory* and *utility theory* seem to be equally suitable to model travellers' behaviour, while *prospect theory* appears to be in disadvantage. This allows to argue that (i) the chosen criteria should not weigh equally and (iii) the most suitable theory might depend on the criteria chosen.

The validity of *utility theory* has already been acknowledged by both the research and practice. On the other hand, against *prospect theory* and *regret theory* weighs the need of more experiments to have their validity fully acknowledged. As a consequence, results based on these behavioural frameworks are still subject to questionings regarding their validity. Therefore, in particular due to the acknowledged validity and tractability of utility-based models, the behavioural rule of utility maximization was chosen to be used in the modelling framework of this thesis.

Three important modelling issues, however, are associated with the estimation of random utility based models based on revealed preference data: difficulty to define/sample the choice, utilities may share unobserved attributes and attributes of the foregone alternatives are in general unknown. This thesis investigates travellers' behaviour in a real network overcoming two of the three main issues associated with model estimation: sampling of alternatives and attributes of the alternatives. The adopted modelling framework that overcomes the need of choice set sampling is discussed in Chapter 4. Issues regarding the attributes of the alternative are overcome by collecting data about the traffic conditions in the network as shown in Chapter 5.

Chapter 4

Modelling framework: dynamic sequential link choice decisions with unrestricted choice set

One of the main issues associated with estimation of route choice models based on revealed preference (RP) data as proposed in this thesis (Chapter 2) is choice set definition (section 3.4). To overcome this issue, Fosgerau et al. (2013) formulated the route choice problem as a sequence of link choice decisions that can be consistently estimated based on path observations, denominated Recursive Logit (RL).

The RL was developed and validated based on a static transportation network⁴. It is similar to the Multinomial Logit Model (MNL) and mainly differs from it for being applicable to an unrestricted choice set. Although the RL is also valid in a dynamic setting, Fosgerau et al. (2013) does not elaborate on the implications and consequences of adding the time dimension to the route choice problem. As the role of real time travel information is dealt with in this thesis, it is important to elaborate on how to allow variability of traffic conditions in the network to be taken into account.

A modelling framework denominated Recursive Logit Dynamic (RL Dyn), which extends the RL to a dynamic setting, is proposed in this chapter. Although both the RL and RL Dyn share the same model properties and mathematical formulation, adding the time dimension to the route choice problem imposes important redefinitions of the state space. Such redefinitions are not trivial tasks and require careful attention. Discussing the challenges and implications of adding the time dimension to the route choice problem is the main contribution of this chapter.

⁴ The terms static transportation network and dynamic transportation network should be interpreted based on the way the travel times of the physical road infrastructure are defined.

Section 4.1 introduces the RL Dyn. It translates to a dynamic framework the main concepts behind the RL. As both the RL and RL Dyn share the same properties, significant parts of this section are taken from Fosgerau et al. (2013). This is particularly the case of the mathematical formulation and the equations which differ from Fosgerau et al. (2013) mainly by redefinition of the state space. Section 4.2, the main contribution of this chapter, discusses the modelling challenges of adding the time dimension to the route choice problem. Conclusions are drawn in section 4.3.

This chapter is based on:

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2012). Route choice model estimation in a dynamic network based on GPS data, *Proceedings of the 1st European Symposium on Quantitative Methods in Transportation Systems*, Lausanne, Switzerland.

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2014). Route choice model estimation in a dynamic network based on GPS data (*working paper*)

4.1 Recursive Logit model for dynamic and deterministic networks

The decision making process consists in making choices among alternatives. For route choice decisions this means choosing among alternative routes. Definition of the routes belonging to the choice set, therefore, is a key point.

In order to overcome the need of choice set definition and efficiently deal with model prediction, Fosgerau et al. (2013) formulated the route choice problem as a sequence of link choice decisions based on dynamic discrete choice model (Rust, 1987). As a result, it is no longer necessary to define or sample the alternatives belonging to the choice set. Instead, at each node of the transportation network the traveller chooses the utility-maximizing outgoing links. This modelling framework is denominated Recursive Logit (RL).

According to the RL all possible alternative routes are part of the choice set. One may argue that because this is a very extensive choice set, people might not be fully aware of all possible routes and that this is a deviation from driving behaviour. The target group investigated are commuters living and working in the area investigated. Therefore, it is expected that their knowledge about the area investigated is much higher than in a different population. Moreover, analyses conducted in Chapter 6 show that travellers know more alternative routes than they think they know, i.e. they drove routes that were not reported as known routes. Besides this, from a modelling perspective, definition of choice set based on RP data is a very difficult task and can lead to a huge number of alternative routes. This is particularly true if one remembers that any right/left turn may correspond to a new route and multiple destinations are considered. The methodology proposed by the RL, therefore, seems to be appropriate to deal with the problem investigated. Furthermore, the role of the model is to assign very small probabilities to unattractive alternatives. Deviations from behaviour will always occur, but for the context investigated the trade-off seems to be positive.

The RL was developed and validated based on a static and deterministic network setting. This section introduces the main concepts behind the RL and at the same time presents it in a dynamic formulation. This is done to a dynamic and deterministic network, i.e. a network where link attributes (e.g. travel times) are time dependent but deterministic within a given time interval. The dynamic representation of the RL is denominated Recursive Logit Dynamic (RL Dyn). The main challenge associated with the extension of the RL to a dynamic setting is the definition of the state space which is discussed in detail in section 4.2. As the RL and RL Dyn share the same model properties, the mathematical formulation presented in this section is the same as that of Fosgerau et al. (2013), except that it is adapted to a dynamic setting.

The RL model for static network is based on a connected graph with links (arcs) and nodes. A traveller sequentially chooses a next link and the sink node of the current. Only the location in the network matters so a state s is defined by a link k in the network. At each state a traveller observes the instantaneous utility and chooses the next link (action) a that maximizes the sum of the instantaneous utility and expected maximum utility until the destination. The RL Dyn, considers a time-expanded network over T the time intervals. This means that each link in the static network appears T times in the time-expanded network. This allows defining different link travel times for each time interval $t = 1, \dots, T$. In this setting a state is no longer defined solely by a location in the network but by a link and time pair $s = (k, t)$. In the static network,

the set of links is denoted L and the size of the state space is $|L|$. In the dynamic network size of the state space is $T|L|$.

Both the RL and RL Dyn assume that travellers behave as utility maximizers. Utilities are given by the instantaneous utility of each link, the maximum expected downstream utility to the destination and *i.i.d.* extreme value type 1 error terms. The RL Dyn defines the network as a directed connected graph $G = (L, N, T)$, where L is the set of links, N the set of nodes and T the time intervals. A path is defined as a sequence of states, $s = (k, t)$, i.e. a traveller is on link k at time interval t , where $t = 1, \dots, T$ and T is the last time interval in the time period. A time period is typically the peak hour in a given day and once defined by the analyst is constant for the whole problem.

Figure 4.1 illustrates the notation of the RL Dyn similarly to the way it is defined in Fosgerau et al. (2013). The set of all possible states to be reached from state s is $S(s)$ which has the same cardinality as the set of outgoing links $L(k)$ from link k . The decision to choose the next link a is done at the sink node of link k . Once the decision is made, the next state $s' = (a, t')$ is reached with certainty because the travel times are deterministic⁵. This means that although the topology of the transportation network defines the outgoing links of sink node k , different states are reached depending on the instant the decision is made. It is important to note that we assume that $t' > t$. That is, a traveller cannot stay in the same time interval⁶ after moving to a new location in the network. This constraint is easily satisfied by fixing the length of a time interval to the shortest link travel time in the network.

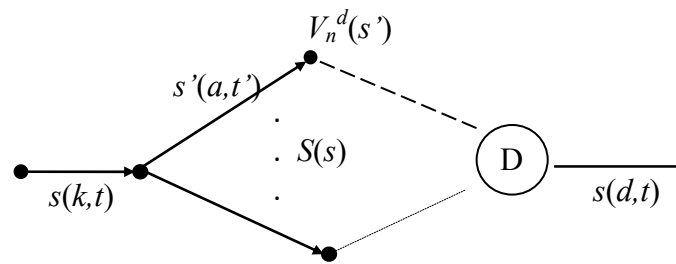


Figure 4.1: Illustration of notation for dynamic and deterministic network

*Source: Fosgerau et. al (2013, p.71)

As defined by Fosgerau et al. (2013, p.72)

a deterministic utility component $v(s' | s) = v_n(x_{n,s|s'}, \beta) < 0$ is associated with each state pair (s, s') , where $x_{n,s|s'}$ is the vector of observed characteristics (that may include characteristics of the decision maker) and β is an unknown vector of parameters to be estimated. Path attributes are assumed to be link additive. In the terminology of dynamic programming, $s(k, t)$ is a state and $s'(a, t)$ is a potential action given state s

The deterministic utilities are defined as costs and are strictly negative. The reason to define link pairs instead of node-link pairs is that the former makes it possible to include turn related attributes in $x_{n,s|s'}$, such as left and u turns.

⁵ For details about the data on traffic conditions in the network, the reader should refer to section 5.3

⁶ Discussion about the role and definition of the time intervals is presented in section 4.2

Any action results in a change of state. Note that as link-specific attributes are defined, the corresponding data also should be link-specific. The expected maximum utility from state s' to destination d is given by the function $V_n^d(s')$, where n represents the traveller.

As the graph is directed connected, for any state s there always exists a potential action s' to be taken. The network is extended by a dummy link that represents the destination. Specifically, absorbing states with zero cost are added represented by an outgoing link d from the destination node without any successors. Therefore, the network is destination and time period specific. The absorbing states $s(d,t)$ can be reached at any time and the set of all links $\tilde{L}^d = L \cup d$ is destination specific. The set of all states is $\tilde{S}^d = S \cup s(d,t)$. The deterministic part of the utility is $v_n(s'|s) < 0$, excepted for $s'=s(d,t)$ for which is $v_n(s'|s)=0$. This applies to all k that have the destination d as sink node in all time intervals. Note that as the destination is an absorbing state, virtually it costs no time to go from $s(k,t)$ to $s(d,t)$.

At any state s traveller chooses a next state s' in a dynamic process having the Markov property⁷ (Bellman, 1957a; Rust, 1987; Aguirregabiria and Mira, 2002). The next state is chosen from the set of outgoing states $S(s)$ and is given with certainty by the action taken. An instantaneous utility, $U_n(s'|s) = v_n(s'|s) + \sigma \varepsilon_n(s')$ is associated with each action and the chosen action $s' \in S(s)$. The action taken maximizes the sum of the instantaneous utility and expected downstream utility (value function). The random terms $\varepsilon_n(s')$ are assumed to be *i.i.d.* extreme value type 1 with zero mean and independent of everything else in the model.

This corresponds to an infinite horizon problem with an absorbing state, the destination state. As formulated by Fosgerau et al. (2013), the expected downstream utility is recursively defined taking into account the continuation of the choice making process by the Bellman equations⁸ (Bellman, 1957b) as follows:

$$V_n^d(k,t) = E \left[\max_{s' \in S(s)} (v_n(s'|s) + V_n^d(s') + \sigma \varepsilon_n(s')) \right] \quad \forall s \in S \quad (4.1)$$

At any state, the probability of choosing next state s' given state s is given by the MNL (Fosgerau et al. 2013).

$$P_n^d(s'|s) = \frac{e^{\frac{1}{\mu}(v_n(s'|s) + V_n^d(s'))}}{\sum_{s' \in S(s)} e^{\frac{1}{\mu}(v_n(s'|s) + V_n^d(s'))}} \quad (4.2)$$

⁷ According to the Markov property, the conditional probability distribution of future states depends only on the current state and not on the sequence of events that preceded it.

⁸ The Bellman equation defines the value of a decision problem in a given instant in terms of the payoff associated with initial choices and of the value of the remaining decision problem resultant from the initial choices. Thus, the dynamic optimization problem is split into simpler sub problems as prescribed by the Bellman's Principle of Optimality.

and the function $V^d(s)$ corresponds to the logsum (Fosgerau et al., 2013):

$$V_n^d(s) = \begin{cases} \mu \ln \sum_{s' \in S} \delta(s'|s) e^{\frac{1}{\mu}(v_n(s'|s) + V_n^d(s'))} & \forall s \in S \\ 0 & s = s^d \end{cases} \quad (4.3)$$

where $\delta(s'|s) = 1$, if state s' is reachable from s and zero otherwise. This means that the sum is over all possible states that can be reached from s . The expected utility at the destination is: $V_n(d, t) = 0$.

Observe that from now on the index n for individuals and superscript d used to clarify that value function and link choice probabilities are destination specific are suppressed.

As the travel times are deterministic, the states are also deterministic for a given departure time. Therefore, the value functions can be solved as a system of linear equations. The value function is iteratively solved for each trial value of β parameters using maximum likelihood. In order to solve the Bellman equations and find the value functions, equation 4.4 is transformed by taking the exponential and raising it to the power of $1/\mu$ (Fosgerau et al. 2013).

$$e^{\frac{1}{\mu}V(s)} = \begin{cases} \sum_{s' \in S} \delta(s'|s) e^{\frac{1}{\mu}(v(s'|s) + V(s'))} & \forall s \in S \\ 1 & s = s^d \end{cases} \quad (4.4)$$

Equation 4.4 can be written in a matrix format (Fosgerau et al., 2013). Denote \mathbf{M} ($|\tilde{S}| \times |\tilde{S}|$) as an incidence matrix of instantaneous utilities (which is known and defined based on the vectors of observed attributes/characteristics of the link pair $A_{s'|s}$ and possible characteristics of the decision maker).

$$\mathbf{M}_{s'|s} = \begin{cases} \delta(s'|s) e^{\frac{1}{\mu}v(s'|s)} & \forall s \in S \\ 0 & s = s^d \end{cases} \quad (4.5)$$

Since d has no successors, \mathbf{M} has a zero row for $s(k, t) = s(d, t)$. Denote by \mathbf{z} ($|\tilde{S}| \times 1$) a vector with elements $z_s = e^{V(s)}$ and \mathbf{b} ($|\tilde{S}| \times 1$) a vector with elements $b_s = 0$, $(k, t) \neq s(d, t)$ and $b(d, t) = 1$. Then, equation 4.5 can be written as a system of linear equations (Fosgerau et al., 2013):

$$\mathbf{z} = \mathbf{M}\mathbf{z} + \mathbf{b} \Leftrightarrow (\mathbf{I} - \mathbf{M})\mathbf{z} = \mathbf{b} \quad (4.6)$$

where \mathbf{I} is the identity matrix. As \mathbf{I} , \mathbf{M} and \mathbf{b} are known the system of linear equations can then be solved and the value function \mathbf{z} defined as long as \mathbf{M} is invertible⁹.

⁹ Discussion on whether \mathbf{M} is invertible is presented in Fosgerau, M., E. Frejinger and A. Karlström (2013) A link based network route choice model with unrestricted choice set, *Transportation Research Part B*, 56, pp. 70–80.

A path is just a sequence of states. The probabilities of choosing the next link/state for a common destination are independent of the origin and the probability of a path is given by (Fosgerau et al., 2013):

$$\mathbf{P}_d = \frac{\mathbf{M}_s \circ \mathbf{z}^T}{\mathbf{M}_s \mathbf{z}} \quad (4.7)$$

where \circ is the element-by-element product and \mathbf{M}_s is row s of matrix \mathbf{M} .

Maximum likelihood is used to estimate unknown parameters. The probability of an observation is the product of each state in the observed path. The estimation code was adapted to the dynamic case from the same code as Fosgerau et al. (2013) implemented in MATLAB and Mai (2012) further improved. The core of the optimization procedure, however, remained the same.

The RL addresses the issue of route overlapping by means of a link size attribute similar to the path size attribute of a path size logit. This, however, is not discussed or addressed in this PhD thesis. For details, refer to Fosgerau et al. (2013).

4.2 Challenges and consequences of going from RL to RL Dyn

Both the RL and RL Dyn share the same properties and definitions regarding the calculation of the probability of a path observation. In essence, therefore, the modelling frameworks are the same. The key contribution of the RL Dyn is to add the time dimension to the modelling framework. This means that while in the RL the transportation network is defined in terms of links and nodes, in the RL Dyn it is defined in terms of links, nodes *and time intervals*. Due to the fact that the time dimension is added to the route choice problem, the states in the RL Dyn are time-dependent. For instance, while in the RL the size of the state space is $|\tilde{L}| \times |\tilde{L}|$, in the RL Dyn it is at least $|\tilde{L}| \times |T| \times |\tilde{L}| \times |T|$, where T is the time intervals. As a result, while in the RL a state s is defined by link k , in the RL Dyn it is defined by a pair (k, t) meaning that a traveller is in link k at instant t . This, therefore, requires a redefinition of the state space.

Redefining the state space, however, is not a straightforward task. The key elements to define the state space in the RL Dyn are: (i) an incidence matrix defining the physical connections of the transportation network, i.e. the successor links of each link in the transportation network, (ii) data about traffic conditions in the network, i.e. the travel times to drive through each link for the whole study period and (iii) size of the time intervals. Based on these elements it is possible to define the possible states to be reached after any action. In other words, the state space is a sparse matrix representing a combination of possible next states s' that could be reached from the current state s . While in the RL the state space is static as it depends only on the incidence matrix, in the RL Dyn it varies depending on the time of the day, day of the week, week of the month, etc. the decision is made. The state space is therefore (almost) observation specific.

To understand the challenges of going from RL to RL Dyn, consider the simplified 6-link network depicted in Figure 4.2. The possible states to be reached depending on a static or dynamic definition of the network are depicted in Figure 4.3 and Figure 4.4, where the rows represent the origin states, the columns the destination states and the input corresponds to the

travel times to go from the origin to the destination state. In Figure 4.4 for instance, a traveller departing from link k_1 at $t=1$, would arrive at link k_2 at $t=2$ or at link k_3 at $t=3$; in case of departures at $t=2$, links k_2 and k_3 would be respectively reached at $t=3$, and so forth. Assume that at instant t , a traveller n is in link k_1 . In the static case, irrespective of the instant the decision is made, a traveller can only choose between links k_2 and k_3 (Figure 4.3). The instant each link will be reached will always be the same irrespective of the departure time. In the dynamic case, however, depending on the instant t the choice is made different states are reached (Figure 4.4). Observe that the possible links to be reached are always the same as this depends on the physical road infrastructure. The instant, and therefore the state, however, might change.

A key feature to be observed in the definition of the state space, therefore, is the progression in time. This means that the size of the time intervals must be carefully defined so as to ensure that each state, $s(k, t)$, is unique and can be reached solely by one action. Once this requirement is fulfilled it is possible to ensure that a traveller departing at instant t for sure will arrive at the destination at an instant t' , where $t' > t$. This can be achieved by ensuring that the time intervals are smaller or equal to the smallest travel time of a link in the network. Furthermore, the travel times of all links should be defined as multiples of the time interval. Under this condition $s'(a, t') > s(k, t)$ and consequently $(t' > t)$. The resultant state space for each time period is a sparse matrix containing entries equal to 1 for all (s, s') pairs. For $t=1, \dots, T$, the pairs (s, s') are defined by:

$$s = (t-1) \cdot |\tilde{L}| + k \quad \text{and} \quad s' = \left(t + \frac{TT(t,a)}{t_int} \right) \cdot |\tilde{L}| + a \quad (4.8)$$

where, $TT(t,a)$ is the travel time to drive through link a at instant t and t_int is the size of the time intervals. Note that as $TT(t, a) \geq t_int$ and also a multiple of t_int , the division of $TT(t,a)$ by t_int is an integer ≥ 1 .

Equation 4.8 shows that the size of the time interval plays an important role in the definition of the state space. Moreover, it indicates that despite its continuous nature, travel times are treated as a discrete variable. Discretising the time intervals has some implications for model estimation and leads to a trade-off between larger matrices resultant from small time intervals (thus more precise) and smaller matrices resultant from bigger time intervals. Although the size of the time interval does not change the modelling framework, it may lead to computational burdens such as long estimation times, lack of RAM memory and storage capacity. This is due to the fact that all link attributes, such as travel times and travel information, are defined based on the state space. As a result, the input matrices of each attribute will either be the same size or smaller than that of the state space. In other words, if a certain combination in the state space $s'|_{s=0}$, this same combination will always be equal to zero to all input matrices. As a result, the attributes of the utility function should be link (or interacted to link-specific attributes) and time-period specific. Moreover, they should be independently defined for (almost) each observation.

For the small 6-link toy network of Figure 4.2, and considering departure times between 8:00h and 9:00h and size of time interval of 1 minute, the state space is at least $(|6||60| \times |6||60|)$. Thus, it results in a matrix of at least 360×360 . If the time interval reduces to two,

i.e. 30 seconds, the size of the state space is already at least raised to the power of 2 (matrix of 720×720). This already gives an indication of the sizes of matrices one has to deal with.

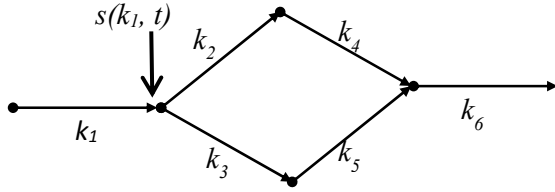


Figure 4.2: Example network

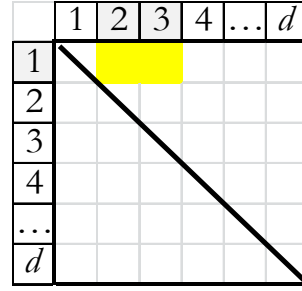


Figure 4.3: Possible states to be reached in the RL considering traveller at link k1

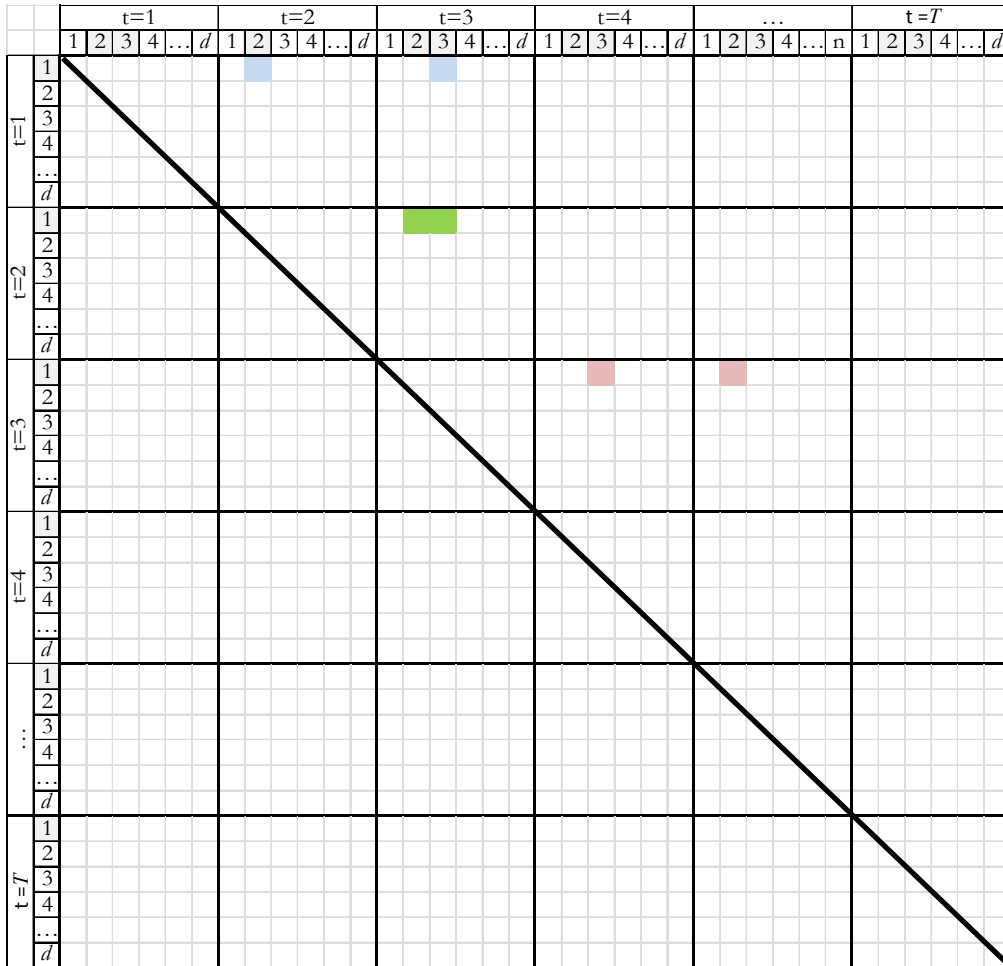


Figure 4.4: Possible states to be reached in the RL Dyn considering traveller at link k1 and departures at instants $t=1$ to $t=3$

These, however, are computational implications. More serious implications are that resultant from rounding errors. Due to the fact that the travel times are discretised, the travel times of each link should be described as multiple of the chosen time interval size. If the time interval is small, consequences may be minimal as the propagation of the error from one state to

another is small. However, the contrary happens if the time interval is big. The ideal time interval is 1 second. So, the closer the time interval is to 1 second, the better. Chapter 7 presents the details of problem dealt with in this thesis.

4.3 Conclusions

This chapter presents the model framework adopted in this PhD thesis: a dynamic link-based route choice model with unrestricted choice set. This modelling framework is denominated Recursive Logit Dynamic (RL Dyn) and is an extension of the Recursive Logit (RL) (Fosgerau et al., 2013). While the RL formulates the route choice problem as sequential link-choice decisions based on path observations without requiring any definition of choice sets, the RL Dyn extends it to a dynamic setting by adding the time dimension to the route choice problem. Therefore it addresses for the first time issues regarding choice set sampling in a dynamic setting supported by RP data collected in a congested network.

Adding the inherent variability of traffic conditions to the modelling framework is especially relevant when investigating the effects of providing real-time traffic information. Nevertheless, this substantially increases the size of the problem to be dealt with and may have more serious implications to model estimation as it requires a careful definition of the state space.

Model estimation in the RL Dyn framework is feasible because the data collection includes data about the traffic conditions in the network during the whole study period (Chapter 5). Estimation results are presented in Chapter 7. These are the first estimation results based on RP data collected in a congested network for which link travel times are known. Furthermore, it compares the predictive power of the RL and RL Dyn to make it clearer the added value of adding the time dimension to the route choice problem.

Chapter 5

Revealed preference data collection on travellers' reactions to travel information

Following the research questions (Chapter 1), the gaps associated with data collection and implications to modelling (Chapter 2) and requirements of the modelling framework (Chapter 4), this chapter presents the necessary specifications of the data collection setup, the characteristics of the actual data collection and the resulting data sets.

This chapter is structured as follows. Section 5.1 presents the specifications of the data collection setup by means of a detailed discussion of the motivation and reasoning behind each decision taken. Section 5.2 describes the characteristics of the actual data collection, i.e. what has actually been done. Section 5.3 introduces the data sets and assesses their quality. Finally, conclusions of this chapter are presented in section 5.4.

This chapter is based on:

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2012) A revealed preference study on route choices in a congested network with real-time information, *Proceedings of the 13th International Conference on Travel Behaviour Research*, Toronto, Canada.

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2014) Quantitative and qualitative analyses on route choice in a congested network with real-time information based on RP data (*working paper*)

5.1 Specifications of the data collection setup

As discussed in section 2.3, most of the data collection efforts come from stated preference (SP) experiments, which, however, are subject to questions regarding their external. As a result, there is a claim in the literature for more revealed preference (RP) experiments.

Due to the fact that RP experiments are conducted in real settings, one can investigate what travellers actually do instead of discussing hypothetical scenarios. There are two main challenges associated with the collection of data in a RP setting: relate past choices to a network in terms of alternatives and traffic conditions, i.e. availability of data about the traffic conditions in the network and find significant variation of travel times across at least two alternatives. This is more difficult to do in experiments of long duration because it requires strong commitment from the participants (Axhausen et al., 2002).

In order to answer the research questions and also to help fill in the observed gaps in the literature, a dedicated RP data collection focusing on travellers' route choice behaviour in reaction to travel information was conducted. The research questions refer to the (i) effects of different sources and timing (pre-trip and en-route) of travel information provision on route choice behaviour (over time), (ii) travellers' willingness to comply with travel information on what concerns to sources and timing over time, (iii) travellers' willingness to pay to receive travel information and (iv) changes in route choice behaviour due to travel information.

In order to answer those research questions, the data collection should comprise the following types of data (DT): (DT1) travellers' route choice, (DT2) travellers' compliance with different sources and timing of travel information provision (over time), (DT3) travellers' willingness to pay for travel information and (DT4) traffic conditions in the road network during the period of the data collection. In addition the following requirements (REQ) should be observed: (REQ1) provision of travel information via different sources and timing and (REQ2) long duration, i.e. allows investigating behavioural changes over time. Table 5.1 summarizes the relationship between the research questions, types of data and requirements to be observed.

The design of the data collection, therefore, demanded decisions regarding the following aspects: (i) data collection technique(s) to employ to properly investigate behavioural changes over time, use and compliance with different sources of travel information and willingness to pay for travel information, (ii) specifications regarding travel information, i.e. sources and timing of travel information provision and (iii) duration of the experiment. Besides this, decisions also have to be made with respect to (i) target group and (ii) characteristics of the area to be investigated. These aspects are discussed in sections 5.1.1 to 5.1.5. An overview of the decisions taken is presented in section 5.1.6.

5.1.1 Data collection technique(s) to be employed

Several RP techniques, such as interviews about past experiences, travel diaries, tracking with GPS, can be employed in order to collect data about travellers' choices. Each of these techniques has advantages and disadvantages and result in data sets with distinct characteristics. This section discusses these techniques as well as the eventual decision on the technique(s) chosen.

Table 5.1: Relationship between the research questions, types of data and requirements

Research question	Data type				Requirements	
	DT1	DT2	DT3	DT4	REQ1	REQ2
<i>Main research question:</i>						
What is the impact of travel time information on travellers' route choice behaviour?	x	x	x	x	x	x
<i>Related research questions:</i>						
1. To what extent does provision of travel information lead to (significant) changes in route choice behaviour	x	x		x	x	x
2. How do different sources of travel information affect route choice behaviour?	x	x		x	x	x
3. How does the timing (pre-trip and en-route) of travel information provision affect route choice behaviour?	x	x		x	x	x
4. To what extent are travellers willing to comply with different sources of travel information provision (over time)?		x			x	x
5. To what extent are travellers willing to comply with different timing of travel information provision, i.e., pre-trip or en-route (over time)?		x			x	x
6. To what extent are travellers willing to pay to receive travel information (over time)?			x		x	x

Interviews have traditionally been used to collect RP data about past experience(s). These may include, for instance, questions about the route and departure time chosen, whether travel information was received, whether delays occurred during the trip and reason to choose a specific route. Disadvantage of this technique is that depending on aspects such as the time elapsed between the trip and the interview or the relevance of the trip, travellers may not be entirely aware of what was exactly done. Nevertheless, in case such a technique is used to get to know travellers' perceptions over time, it can be a very powerful tool. This is because it enables one to investigate, for instance, travellers' perceptions regarding the use and reliability of travel information, travellers' willingness to pay for travel information, travellers' perceptions of routes characteristics and travellers' awareness/perceptions of route choice set. Such type of analyses greatly contributes to a better understanding of travellers' behaviour especially if such perceptions can be compared or confronted with objective measurements.

The use of travel diaries, on the other hand, given the usual short time span between the trip made and the act of filling the travel diary in, is a more accurate technique to get deeper insight into travellers' behaviour regarding a specific trip. It allows, for instance, to investigate travellers' use (the act of checking) and compliance with different sources of travel

information, perceptions about the trip made and how expectations about next/future trips. Disadvantage of this approach is that the repetitiveness of the task of filling travel diaries may tire the participants and consequently lead to lower response rates, especially in experiments of long duration. In addition, in case precise information such as departure time and travel time need to be collected, they may not be accurately registered by the travellers.

GPS devices are very useful tools to get to know with precision details such as travel and departure times, whether and where travellers were stuck in traffic during a certain trip and for how long or whether intermediate stops were made. Only by tracking travellers during their trips it is possible to get such type of data with precision. Disadvantage of such tool is that one cannot know, for instance, whether travel information was received during the trip, from which sources, whether travellers complied with it and how they perceived the trip made.

In order to answer the research questions, it was decided to combine the advantages of various techniques to investigate travellers' behaviour in reaction to travel information. The data collection should therefore comprise interviews, travel diaries and GPS traces (Table 5.2). The interviews aim at investigating travellers' perceptions over time, the travel diaries to get deeper insight into how travellers react to travel information both daily and over time and finally the GPS traces to get to know what travellers' did and to obtain precise information about the trips made.

It has to be observed that although these techniques allow one to collect data about travellers' behaviour and/or reaction to travel information, they provide only little information about the traffic conditions in the network. As a result, they do not allow one to investigate "what would have happened" in case a different choice had been made. In other words, the attributes of the non-chosen alternatives are missing. This means that data about the traffic conditions in the whole network also has to be collected.

Table 5.2: Relationship between the types of data and the data collection techniques

Type of data*	Data collection technique		
	Interviews	Travel diary	GPS
DT1: Travellers' route choice			x
DT2: Travellers' compliance with different sources and timing of travel information		x	
DT3: Travellers' willingness to pay for travel information	x		

* Data about traffic conditions in the network, DT4, also has to be collected

5.1.2 Specifications regarding travel information

As the objective of the data collection is to investigate the effects of travel information on travellers' behaviour, both the sources and timing of travel information provision, i.e. whether to provide travel information before or during the trip, have to be taken into account. In terms of sources, they should vary from public/non-personalized to private/personalized and to be

possibly reached before and/or during the trips. The distinct moments of provision of travel information allow to investigate the impact of providing both pre-trip and en-route travel information. By considering a variety of sources of travel information (and level of personalization) it is possible to compare the potential of different sources, such as radio, web variable message signals (VMS) and in-car navigation systems with real-time travel information to influence travellers' behaviour.

To investigate the effects of travel information it is important to know how travellers behave in general, i.e. irrespective of provision of travel information. The data collection should thus consist first of a reference period during which travellers drive as usual and then of a period during which personalized travel information is provided. The reference period is denominated *reference treatment*. The experimental condition is denominated *information treatment*.

In the reference treatment, as in any other trip, travellers should be able to check existent (public) sources of non-personal travel information, such as websites, radio and VMS. This is because depriving travellers from checking existing (public) sources of travel information would possibly impose changes in habits of listening to the radio, watching television and consciously ignoring VMS panels for example. Consequently, travellers' "normal/daily" behaviour would not be captured. In the information treatment, on the top of existing free/public sources, personalized real-time traffic information should also be provided.

Besides this, a reference group that will not receive personalized travel information should also be defined. The reference group is denominated *regular travellers* and the group that receives personalized travel information is denominated *informed travellers*. The existence of two treatments and two groups of travel information contributes to the achievement of three main objectives: (i) to investigate whether being part of an experiment would change the behaviour of regular travellers over time, (ii) to investigate the influence of personalized sources of travel information and (iii) to investigate whether behavioural changes are observed by comparing the behaviour of informed and regular travellers.

5.1.3 Duration of the data collection

Specifications regarding the minimum duration of RP experiments have not been found in the literature. Moreover, one of the main challenges highlighted in the conduction of RP experiments is the difficulty to engage participants in experiments of long duration (section 2.3). For instance, the RP experiments reported in the literature that required direct commitment from the participants have lasted no longer than a couple of weeks.

Nevertheless, given the focus on investigating the effects of travel information over time some aspects should be considered. For instance, the data collection should last long enough to allow one to capture (potential) behavioural changes, to allocate enough data collection time for the two treatments of travel information provision and to deal with the uncontrollability of RP experiments. Moreover, it should last at least a couple of weeks.

5.1.4 Target group

As discussed in Chapter 1, one of the main goals of this PhD thesis is to discuss the potential of travel information to (help to) alleviate congestion. In order to make this possible, the target group investigated should correspond to the largest share of travellers driving during

peak hours, thus in highly congested periods. This is the type of situation in which provision of travel information potentially helps to alleviate congestion the most. Commuters were then chosen to be investigated. Besides this, because (in theory) commuters are already familiar with the characteristics of the road network they are better able to judge the added value of making use of travel information in their decision processes.

One may argue that due to the familiarity of commuters with the transportation network they may be less susceptible to make use of travel information. This, consequently, has implications to the outcomes of the data analyses. Although this observation is correct, commuters are the largest group of travellers driving during congested periods. Consequently, if one wants to investigate the potential of travel information to alleviate congestion the target group investigated should be the one that influences congestion the most. For instance, if a traveller is not familiar with the transportation network, the use of travel information may be more associated with the need of route guidance rather than avoiding congestion. Therefore, despite the particularities inherent of having commuters as a target group, this appears to be the most appropriate target group to be investigated.

5.1.5 Characteristics of the area to be investigated

The main decision regarding where to perform the data collection concerns whether to focus on one specific origin-destination (OD) pair or not. Important factors influencing this decision are requirements concerning the existence of comparable alternative routes between the OD pairs, the level of familiarity with the routes from the point of view of the analyst (which would contribute, for example, to spot/understand unexpected behaviour resulting from incidents or disruptions), similarity of trip purposes, the availability of data regarding the traffic conditions in the network, the (possibility of) generalization of the results to bigger and/or more complex networks and the easiness to gather participants to join the experiment.

Advantage of not to focus on one specific OD pair is that the generalization of results is possible as different route characteristics and travel distances can be investigated (which is more difficult for a single OD pair). In addition, it is easier to select participants as this process is less restrictive. On the other hand, the need to get data about the traffic conditions in the network, the need to ensure similarity of trip purposes (as focus is put on commuters) and the challenges to guarantee the existence of comparable alternative routes (that are “real” alternatives) are large disadvantages. The level of complexity and workload associated with these tasks substantially increases with the amount of OD pairs. Specific OD pairs, on the other hand, have the advantage of leading to rather similar travel purposes, such as commuting trips to work or to shopping areas, enables a higher degree of familiarity of the analyst with the route choice set (and thus a better understanding of travellers’ behaviour) and facilitates the collection of data about the traffic conditions in the network.

Given the need to ensure similarities of trip purposes, the added value of a higher level of familiarity with the network from the point of view of the analyst, the workload necessary to the collection of data about the traffic conditions in the network and to make it easier to select participants and other logistical requirements (contact the participants whenever required and solve problems during the data collection), it was decided to focus not on a single OD pair, but on OD zone(s), i.e. trips between two municipalities and not between two specific locations.

Another decision to be made refers to the characteristics of the OD zone(s). A fundamental point to be observed is that the composition of traffic demand between the OD zone(s) should be representative of the population investigated. In other words, as focus is put on commuters driving in congested areas, the chosen OD zone(s) should be in a (highly) congested area and the demand of commuters between them should also be high.

Besides this, as discussed in section 2.2, travellers appear to make use of travel information only if, among other things, there is a high variability of traffic conditions in the network and there are similar/comparable available alternative routes. Therefore, the routes between the OD zone(s) should be (i) comparable in terms of their length (otherwise the routes may not be seen as alternatives) and (ii) distinct in terms of their characteristics, thus belonging to different classes/categories of roads, such as motorways and local roads, which have different levels of travel time variability.

In addition, an assumption generally made about travellers' route preferences is that travellers have a habitual route. The habitual route is likely to be the route chosen the most when no travel information is provided. In order to make the potential of travel information to influence travellers to their least preferred routes¹⁰ clearer, the OD zone(s) to be investigated should be connected by at least three comparable and distinct alternative routes. This would ensure not only the existence of more alternative routes (and not only a habitual route and a "second" choice), but possibly more travel time variability between routes.

The OD zone(s) to be investigated should therefore be in a (highly) congested area, the demand of commuters between them should be high and the OD zone(s) should be connected by at least three the comparable alternatives routes in terms of length and distinct in terms of characteristics.

5.1.6 Overview on the data collection design

The design of the data collection setup involved decisions regarding: (i) data collection techniques to be employed, (ii) specifications about travel information, (iii) duration of the experiment, (iv) target group and (iv) characteristics of the area to be investigated. The motivation behind the chosen design was presented in sections 5.1.1 to 5.1.5 and an overview of the decisions taken is presented below for the sake of completeness.

Data collection techniques

The data collection should comprise the following techniques: interviews, travel diaries and GPS traces. In addition, data about the traffic conditions in the network should be collected.

Specifications regarding travel information

Focus should be put on existent sources of public/non-personalized and private/personalized travel information such as radio, variable message signals (VMS) and in-car navigation systems with real-time travel information. Pre-trip and en-route travel information should be contemplated. Travellers should be split into *regular* and *informed* travellers. Two periods of travel information provision should be considered: one in which only public sources of travel information are available as *reference treatment* and another one in which personalized travel

¹⁰ The least preferred routes should, nevertheless, be as feasible as the habitual route.

information is also available as *information treatment*. Travel information was provided from available information providers. Travellers could consult it whenever they wanted.

Duration of the data collection

The data collection should last long enough to allow to capture behavioural changes over time and to allocate time for the two periods of travel information provision. It should not be, however, too long as this makes it difficult to engage travellers. A minimum duration of a couple of weeks should be observed.

Target group

Commuters should be investigated.

Characteristics of the area to be investigated

OD zone(s), and not an OD pair, should be investigated. The OD zone(s) should be in a (highly) congested area, the demand of commuters between them should be high and they should be connected by at least three comparable and distinct alternative routes.

Following the specifications regarding the data collection setup, next section presents the characteristics of the data collection. In other words, it is presented what was actually done and how the aforementioned specifications were taken into account.

5.2 The actual data collection

In general terms, the data collection consisted of tracking the participants with GPS devices while they performed their commuting trips by car under different timing and sources of travel information provision. A travel diary, that basically consisted of questions related to travellers' reactions to travel information, perceptions about the trip made and expectations towards the next trip, was filled in after each trip. Besides this, participants were interviewed at three different occasions: at the beginning, in the middle and at the end of the experiment. This way, travellers' perceptions was monitored during the experiment. Furthermore, data about the traffic conditions of the network during the whole study period was also collected. Sections 5.2.1 to 5.2.4 describe the setup of the data collection in terms of its period/duration, OD zones investigated, procedure to select the participants and conditions of travel information provision.

5.2.1 Period and duration of the data collection

The data was collected in the Netherlands from May 9th to July 12th, 2011, which is a period of regular traffic demand. The duration and the period were constrained by logistical requirements. The starting date was constrained by the preparations for the experiment, such as the procedure to select participants, preparation of travel diaries and interviews and purchase of GPS devices. The ending date was constrained by the beginning of the summer holidays in the Netherlands. Alternatives would be to reduce the duration of data collection or postpone the starting date of the experiment for after the summer holidays. However, as the aforementioned period was of regular traffic demand, much longer than the minimum couple of weeks and allowed time enough to investigate the two periods of travel information provision it was decided to stick to it. Despite the challenges to engage the participants to join the experiment for two months, the data collection went smooth.

5.2.2 OD zone(s) investigated

One of the requirements regarding the OD zone(s) to be investigated was that it should be in a (highly) congested area and with high the demand of commuters. According to the database of the Dutch institute of mobility research of 2008¹¹, there were about 2 million car commuters in the Netherlands. Among them, 60% lived within 19 km distance from work, 24% between 19 and 40 km and 16% over 40 km (Dutch MON database, 2008). In order to have a representative OD zone, i.e. with the largest share of commuters driving in peak hours, it was decided that the distance between the OD zones should be within 19 km. Furthermore, because this research is part of a research program focusing on the sustainability of the Randstad, which is the most congested area in the Netherlands, the area investigated should be in the Randstad (Wikipedia, 2011).

The main cities in the Randstad with high demand of traffic between them are Amersfoort, Almere, Amsterdam, Delft, Dordrecht, Gouda, Haarlem, Hilversum, Leiden, Rotterdam, The Hague, Utrecht and Zoetermeer (Figure 5.1). Due to logistical requirements, such as contact the participants whenever required and solve problems during the data collection, one of the locations was chosen to be Delft where most of this research was conducted. Other advantages of choosing Delft is that it ensures similarity of trip purposes and facilitates the selection of participants by selecting people among the staff and students of the Delft University of Technology (TU Delft).



Figure 5.1: Main cities in the Randstad

Considering that the OD zones should be within 20 km from each other and connected by at least three comparable but distinct alternative routes, the municipality of The Hague was chosen to be the other zone. The city centres of Delft and The Hague are approximately 16 km far from each other, the traffic demand between them is high and the fact that there were 13 VMS installed and operating allows investigating the role of this source of travel information.

¹¹ Preparations for the data collection started in 2010. By then the database of 2008 was the latest available.

5.2.3 Selection of participants

The procedure to select the participants consisted of contacting staff members and students of the TU Delft via e-mail and advertisement around the university. Requirements were that they lived in The Hague, commuted to Delft by car at least 2 times per week, did not have any holidays planned during the period of the experiment and were willing to join the data collection for two months. This procedure resulted in the selection of 32 participants. Each participant was introduced to the objectives of the experiment. At this occasion, the participants were interviewed about their routes' preferences, motivation for those preferences and habits regarding use of travel information among others.

All the participants were informed about what was expected from them: keep the GPS devices on during their commuting trips, fill in the travel diaries after each trip and be interviewed at two more occasions during the experiment. As a reward for joining the experiment, they were informed that at the end of the experiment three in-car navigation devices with capabilities to provide real-time travel information would be randomly draw among them.

The 32 participants who joined the experiment consisted of 44% women and 56% men. Their age ranged between 23 and 60 years old, among which 41% were between 35 and 45 years old. In terms of educational level, 3% had finished only the primary school, 6% had finished only the secondary school, 31% had finished college and 59% had university degree. Commuting frequencies varied between 2 and 5 days/week, with the majority of travellers, 61%, commuting 5 days/week. Furthermore, 19% of the participants reported to usually check pre-trip travel information and 34% to check en-route travel information.

5.2.4 Conditions of travel information provision

Travel information was provided via different sources and in different timings. The timings considered were before (pre-trip) and after (en-route) the trip started. In terms of sources, both free/public sources and personalized/private real-time travel information were considered. When dealing with pre-trip travel information the considered sources were radio, television, websites (e.g. ANWB route planner, Google maps) and real-time travel information from now on denominated TomTom¹². When dealing with en-route travel information, radio, VMS and TomTom were considered. For any source different from the ones listed, the option other was considered. When dealing with en-route travel information, radio, VMS and TomTom were considered. For sources different from the ones above, the option "other" was considered.

Regarding the treatments of travel information provision, it was decided that the two months (nine weeks) of data collection should be split into initial three weeks as *reference treatment* and subsequent six weeks as *information treatment*. This division ensures the focus on the provision of travel information and a sufficient long reference period.

In addition to this, travellers were split into two groups of travel information provision: *regular travellers* consulting only public/free sources of travel information corresponded to 20% of the participant; the remaining 80% consisted of *informed travellers* equipped with

¹² The company TomTom, specialized in provision of real-time traffic information, cooperated with this data collection. They made available the devices (type TomTom Via LIVE 120 Europe) used to provide personalized real-time traffic information. Furthermore, they made available the data about the traffic conditions in the network. (TomTom (2011) *Tom Tom Navigator*, [www.tomtom.com], page visited on July 14th, 2011.)

TomTom in the information treatment. Similar to the two treatments of travel information provision, this would allow to have a sufficient large group of regular travellers for comparisons with the informed ones, but still small enough to ensure the focus on informed travellers. Table 5.3 summarizes the number of participants under each condition of travel information provision during the experiment.

TomTom travel information consists of a recommendation regarding the fastest route between an origin and a destination, estimated trip duration, arrival time and delay (in relation to free-flow travel time). This is based on an indicated departure time (left side of Figure 5.2). Comparisons between the travel times on a suggested route (dashed line) with another alternative route (continuous line) (right side of Figure 5.2) is also possible for any given departure time.

The TomTom devices were equally set to indicate the fastest route and warn the traveller in case a new route became the fastest. In addition, the participants were told not to change the settings of the device. This way, all participants were equipped with a proper tool of travel information provision to help them to choose the best route (and departure time).

Table 5.3: Percentage of participants in different treatments of information provision

Period	Treatment	
	Reference (only public/free sources)	Information (public free sources + TomTom)
Initial 3 weeks	100% (32 participants)	0% (0 participants)
Last 6 weeks	20% (7 participants)	80% (25 participants)

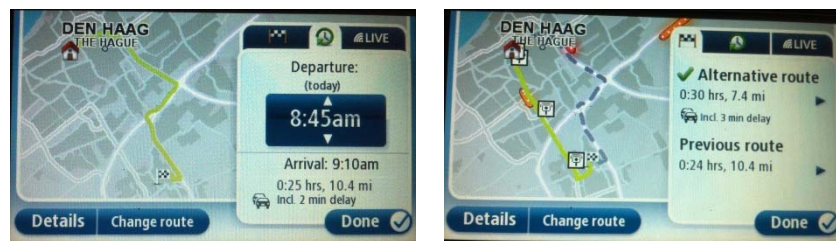


Figure 5.2: Example of the type of travel information provided by TomTom

The use of travel information was *not* imposed to the participants. Instead, at the beginning of the experiment, the participants were instructed to behave as if in a normal commuting trip. In other words, discretionary powers were given to the participants regarding when to use travel information. As a result, no change in behaviour, besides keeping the GPS device in the car and filling in the travel diaries, was imposed to the participants. As none of the participants owned TomTom devices, they were introduced to it during the experiment (this might have influenced the use of TomTom information as discussed in Chapter 6). Note that as the traffic conditions were being created as a result of the use of the transportation network, feedback information about alternative routes was not provided at the end of the trip. This applies even to the *informed travellers* who could check the expected travel times (and corresponding delays) for more than one route at the beginning of the trip.

5.3 Data sets: characteristics and quality assessment

The data collection comprised four data sets: GPS traces, travel diaries, interviews and traffic conditions in the network. Sections 5.3.1 to 5.3.4 respectively describe each data set. In section 5.3.5 an overview of the data sets and quality assessment are presented.

5.3.1 GPS data

GPS traces were used to collect with precision data about the routes chosen, the travel and departure times and whether intermediate stops were observed. It corresponds to the DT1 discussed in section 5.1. GPS data also provides information about distance and speed, thus allows one to derive relationships such as time versus speed and distance versus speed both analytically and graphically (Figure 5.3). They allow identifying, for instance, stopovers and the location where travellers got stuck in congestion (if this was the case) along the day.

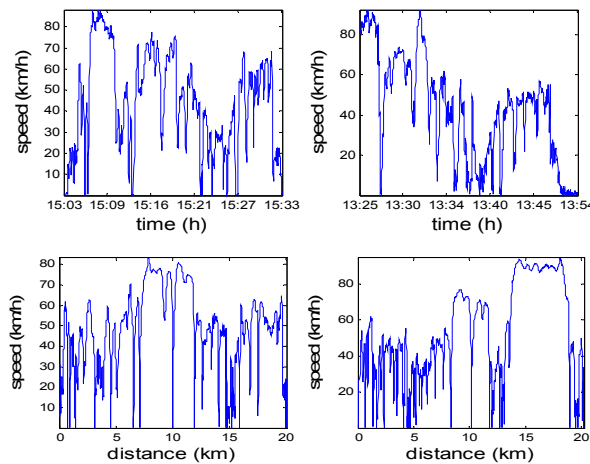


Figure 5.3: Illustrative graphs of speed versus time and speed versus distance



Figure 5.4: GPS device used during the experiment

There is a great variety of GPS models in the market and the one selected was the BT-Q1000XT from the brand Qstarz (Qstarz, 2011) due to its precision, size and price. These are small devices of approximately 8 cm, storage memory to record up to 20 consecutive days of data (considering 5 seconds interval between each logging point) and battery duration up to 42 consecutive hours (Figure 5.4). The horizontal precision of such device is around 1m and on average 8 satellites are used to define travellers' coordinates. Due to the high accuracy of the chosen GPS devices, the trips could be matched to the road network without ambiguity. These devices are thus considered to be accurate enough.

The GPS was set to automatically log every 5 seconds whenever movement was detected. This way, the participants did not have to handle the GPS. All they had to do was to keep the devices in the car.

5.3.2 Travel diary

Besides carrying the GPS devices during their commuting trips, the participants were asked to fill in a travel diary after each trip. The travel diary consists of the following five sections:

- (i) General travel information such as date of the trip, origin and destination. This make it possible to match the input of the travel diaries with the GPS traces;
- (ii) Behaviour towards pre-trip travel information, such as whether pre-trip information was checked and sources, compliance with travel information (and motivation), alignment of travel information with expectations/planned route and what was recalled from the information.
- (iii) Behaviour towards en-route travel information containing questions similar to the ones about pre-trip information.

The questions of aforementioned sections (ii) and (iii) allow one to assess travellers' reaction to different sources and timing of travel information provision.

- (iv) Feedback about the trip just made consisting of questions about the actual travel time, whether (in hindsight) the participants would have chosen the same route and departure time and whether (and what) additional travel information was needed. This set of questions allow to investigate travellers' perceptions of the trip made, get further insights into the use/need of travel information and habit.
- (v) Expectations about the next trip, such as intended route choice and departure time, expected travel time and arrival time constraints. These questions contribute to compare travellers' expectations to what actually happened in the next trip.

Figure 5.5 and Figure 5.6 depict the questions presented in the travel diary. Its structure aims not only at investigating travellers' reaction to travel information, but also to get a clearer picture of travellers' decision process by investigating the present situation, perceptions about what happened and expectations about future trips. Furthermore, the questions presented in the travel diary address the requirements of investigating the effects of different sources and timing of travel information provision, use and compliance with travel information and behavioural changes over time (data types DT2 and DT4 and requirements R1 and R2).

GENERAL INFORMATION			
1. Date:	2. Origin:	Home <input type="radio"/>	Work <input type="radio"/> _____ <input type="radio"/>
	Destination:	Home <input type="radio"/>	Work <input type="radio"/> _____ <input type="radio"/>

TRAVEL BEHAVIOUR IN RELATION TO INFORMATION RECEIVED BEFORE DEPARTURE			
3. Did you receive traffic information <i>before</i> travelling?		Yes <input type="radio"/>	No <input type="radio"/>
a. If YES	a1. Through which means?	Internet <input type="checkbox"/>	Radio <input type="checkbox"/> Navigation system <input type="checkbox"/> website: TV <input type="checkbox"/> _____ <input type="checkbox"/>
	a2. What do you remember from the information that was provided?		
	a3. Did the information favour the intended route?		Yes <input type="radio"/> No <input type="radio"/>
	a4. Did you follow the information?		Yes <input type="radio"/> No <input type="radio"/>
	a4.1. If YES	a4.1.1. Which information in particular did you follow? Internet <input type="checkbox"/> Radio <input type="checkbox"/> Navigation system <input type="checkbox"/> TV <input type="checkbox"/> _____ <input type="checkbox"/>	
	a4.2. If NO	a4.2.1. Why did you not follow the information ?	

TRAVEL INFORMATION IN RELATION TO INFORMATION RECEIVED DURING THE TRIP			
4. Did you receive traffic information <i>while</i> en-route?		Yes <input type="radio"/>	No <input type="radio"/>
b. If YES	b1. Through which means?	Navigation system <input type="checkbox"/>	Panels above or next to the roads <input type="checkbox"/> Radio <input type="checkbox"/> _____ <input type="checkbox"/>
	b2. What do you remember from the information that was provided?		
	b3. Did the information favour the intended route?		Yes <input type="radio"/> No <input type="radio"/>
	b4. Did you follow the information?		Yes <input type="radio"/> No <input type="radio"/>
	b4.1. If YES	b4.1.1. Which information in particular did you follow? Navigation system <input type="checkbox"/> Panels above or next to the roads <input type="checkbox"/> Radio <input type="checkbox"/> _____ <input type="checkbox"/>	
	b4.2. If NO	b4.2.1. Why did you not follow the information?	

Figure 5.5: Page 1 of the travel diary

YOUR FEEDBACK ABOUT THIS TRIP			
5. What was your actual travel time?			
6. Did you suffer from delays longer than expected?		Yes <input type="radio"/>	No <input type="radio"/>
7. In hindsight, would you have chosen the same route?		Yes <input type="radio"/>	No <input type="radio"/>
c. If not, which route would you have chosen?			
<i>Delft area</i>		<i>The Hague/Rijswijk area</i>	
A4 <input type="radio"/>	Van Miereveltlaan <input type="radio"/>	A12 <input type="radio"/>	A4 <input type="radio"/>
A13 <input type="radio"/>	Vrijenbanselaan <input type="radio"/>	N211/Wippolderlaan <input type="radio"/>	Jan Thijssenweg/ Westvlietweg <input type="radio"/>
Provincialeweg <input type="radio"/>		Haagweg/Rijswijkseweg <input type="radio"/>	Ypenburgse Boslaan <input type="radio"/>
		Prinses Beatrixlaan/Moerweg <input type="radio"/>	Not applicable <input type="radio"/>
		Schaapweg/Loevsteinlaan <input type="radio"/>	
8. In hindsight, would you have chosen the same departure time?		Yes <input type="radio"/>	No <input type="radio"/>
d. If not, what time do you think you should have departed?			
9. Did you depart earlier (or later) than you had initially planned because you received traffic information with respect to congestion on your route?		Yes <input type="radio"/>	No <input type="radio"/>
10. Would you have liked to have received additional traffic information for this trip?		Yes <input type="radio"/>	No <input type="radio"/>
e. If yes, could you please elaborate on the type of additional information you would like to have received (e.g. information about congestion in a specific road, construction works, accidents, etc.)?			

EXPECTATIONS ABOUT YOUR NEXT TRIP (i.e. the trip back home or to work in the next working day)			
11. Date of the next trip:			
12. Intended departure time:		13. Expected travel time	
14.: Intended route choice:			
<i>Delft area</i>		<i>The Hague/Rijswijk area</i>	
A4 <input type="radio"/>	Van Miereveltlaan <input type="radio"/>	A12 <input type="radio"/>	A4 <input type="radio"/>
A13 <input type="radio"/>	Vrijenbanselaan <input type="radio"/>	N211/Wippolderlaan <input type="radio"/>	Jan Thijssenweg/ Westvlietweg <input type="radio"/>
Provincialeweg <input type="radio"/>		Haagweg/Rijswijkseweg <input type="radio"/>	Ypenburgse Boslaan <input type="radio"/>
		Prinses Beatrixlaan/Moerweg <input type="radio"/>	Not applicable <input type="radio"/>
		Schaapweg/Loevsteinlaan <input type="radio"/>	
15. Imagine that an unforeseen event happens. As a result you are unexpectedly late or have to cancel your trip. How would this affect the activities at the destination in your next trip?			
Cancelling the trip/working from home would be fine <input type="radio"/>	I would successfully fulfil my obligations if due to a delay I arrive 30 min to 1 hour late <input type="radio"/>	10 - 20 min delay would be OK <input type="radio"/>	Punctuality (or early arrival) mandatory <input type="radio"/>

Figure 5.6: Page 2 of the travel diary

5.3.3 Interviews

Participants were interviewed at three different occasions: at the beginning, in the middle (after 5 weeks of data collection) and at the end of the experiment. The objective of the interviews was to get insights into travellers' perceptions about the use of travel information, including their willingness to pay for it, and routes' characteristics over time. In other words, it aimed at investigating whether travellers' perceptions changed over time.

At the beginning of the experiment the participants were asked about (i) perceptions regarding the fastest, most reliable and most unpredictable routes, (ii) habitual routes to and from work (and why), (iii) usual departure times, (iv) frequency of using pre-trip and en-route travel information and (v) willingness to pay to use travel information. In the middle of the experiment travellers were interviewed with respect to their reactions to the provided real-time travel information, i.e. (i) how often the route suggested by TomTom matched their expectations, (ii) how often they followed the suggested route, (iii) how satisfied they were with the travel information provided, and (iv) how reliable they considered the travel information to be. Finally, at the end of the experiment, a combination of questions presented at the beginning and in the middle of the experiment was asked again to assess changes in their perceptions.

Note that the travel diaries and interviews have complimentary objectives. The interviews aimed at investigating travellers' perceptions regarding use of travel information over time. It refers, thus, to a more subjective measurement. Conversely, by focusing on travellers' actual use of travel information, the travel diaries refer to more objective measurement. The questions asked during the interviews address: travellers' willingness to pay for travel information (DT3) and behavioural changes over time (R2).

5.3.4 Traffic conditions in the network

In order to estimate route choice models, one of the main requirements is the availability of data about the alternative routes. When dealing with RP data collected in a congested network this is a big issue. The level of difficulty increases when dealing with a dynamic network as it demands the collection of actual data for the whole study period and not only of historical data.

The modelling framework adopted in this thesis, RL Dyn, consists of sequential link-choice decisions in a dynamic network. This means that the transportation network should be defined in terms of links, nodes and time intervals. As a result, the data about the traffic conditions in the network should be collected for all links and time intervals for the whole studied period.

The investigated road network was defined based on the subjective choice set reported by the participants. During the interviews at the beginning of the experiment travellers were asked to draw on a map the routes they were aware of between homes (in The Hague) and work (in Delft). Output of the interviews resulted in the definition of a subjective route choice set combining the answers of all travellers (see section 6.6.1 for further details). The boundaries of the road network, therefore, were constrained by the most northern, southern, eastern and western routes of the subjective route choice set.

The network investigated consists of 520 unidirectional links and 200 nodes (Figure 5.7). The links represent the roads and the nodes can represent either physical intersections of the

network or points in which travellers can divert/continue straight on or start of new trips. Link travel times were retrieved from the TomTom database and consist of the travel time on each link at 1 minute intervals during the whole study period. From now on, when referring to the data from which travel times from the TomTom database was retrieved, the term *network data* will be used. Otherwise, free flow travel time is considered based on the speed limits of the roads. In Figure 5.7 links in yellow correspond to the ones for which network data is available and links in blue otherwise.

The network data allows investigating behavioural changes over time (R2) and moreover provides the traffic conditions in the road network (DT5).



Figure 5.7: Road network studied

5.3.5 Overview of the resultant data sets and quality assessment

Conduction of an RP experiment is subject to a high degree of uncontrollability. In this RP experiment, at occasions the GPS devices lost signal during the trip or before the participant arrived at the destination, data started to be logged after the trip had started, the participant did not make the regular commuting trip or the participant did not fill in the travel diary.

Due to events such as these, the resultant data sets for both GPS traces and travel diaries were smaller than what would have been in the ideal scenario of a controlled experiment. The *valid trips* consist of 897 valid GPS traces and travel diaries, among which 374 refer to the *reference* treatment and 523 to the *information* treatment. The term valid trips refers to the trips in which events such as the aforementioned *did not* happen and in which both the origin and destination of the GPS trace is within 1.5 km of home and work. The requirement of 1.5 km distance helps to ensure that the trip performed is a commuting trip and not another type of trip in the vicinities of either home or work. Especially due to the focus on investigating travellers' reaction to information, it is important to match the GPS traces with the travel diaries.

With respect to the interviews, it was logistically challenging and time consuming to personally interview each of the 32 participants in three different occasions. Nevertheless, it contributed to get more insight into their perceptions about the experiment, route choices and use of travel information. The resultant sample size of the interviews consists of 32 interviews (one for each participant) for each of the three occasions the participants were interviewed.

Despite the limited sample of 32 participants, the nine weeks of data collection allows investigating travellers' reaction to travel information over time, which is an aspect often missing in RP experiments. As discussed in Chapter 2, besides the limited number of participants in most RP experiments, only few observations per participant is usually collected. In the studies in which both the number of participants and repetitions are high, data is usually gathered in automated ways, such as by the identification of license plates, but little is known regarding the use of travel information, trip purpose, etc. The duration and strong commitment of the participants during this data collection played a very strong role in the quality of the travel diaries, interviews and GPS traces. Especially due to the commitment of the participants in properly daily filling the travel diaries, the nine weeks of data collection do not seem to have had a negative effect on the quality of the data. Instead it contributed to its richness and high quality. Furthermore, despite the limited sample size, estimation results were statistically significant

Data about the traffic conditions in the network, the network data, was obtained from the TomTom database for each link at 1 minute intervals during the whole study period. As the network data is based on the actual demand of travellers, it provides a very good estimate of travel times in the transportation network. Although network data was available for most links of the network, this was not the case for a small subset of local roads with low traffic demand. For these links, free flow travel times are considered.

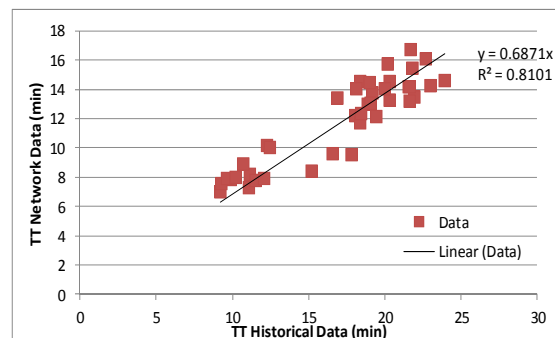


Figure 5.8: Historical versus network travel times

As the network data is raw data, TomTom still uses an algorithm to process it and report the predicted the travel times of a route to the travellers. Implication of this is that adding up the travel times of a sequence of links based on the network data does not necessarily result in the travel time of a given route. As this algorithm was not available for this research, the consistency of the network data had to be checked by means of comparison with historical data of predicted travel times for purposes of model estimation. Historical data was also retrieved from the TomTom database and consist of predicted historical travel times for each main road¹³ and for the same period of the experiment, but for the year of 2010. Thus, both the network and historical data were retrieved from the same database. The travel times of each main road resultant from the network data, i.e. adding up the travel times of each link, were compared with the historical data by means of a linear regression (Figure 5.8). Although the correction factor of 0.6871 implies that the travel times of the network data are

¹³ The main roads were defined by combining the answers of all participants regarding the roads usually driven to and from work part of the subjective route choice set

underestimated, the high coefficient of correlation ($R^2 = 0.8101$) assures the consistency of the network data. The network data can therefore be used in the analyses presented in this thesis.

5.4 Conclusions

This chapter presents the data requirements to answer the research questions raised in Chapter 1, the characteristics of the data collection, i.e. what exactly has been done, and a description of the data sets. In addition, an overview of the resultant data set followed by an assessment of their quality and issues encountered during the data collection are presented.

By conducting an RP experiment in which travellers were tracked while making their route choice decisions, it was possible to investigate their behaviour in a setting in which they immediately experienced the consequences to their choices. Furthermore, the travel diaries and interviews contribute to investigate travellers' perceptions and reactions to travel information. The uniqueness of this data collection does not lie only in the use of GPS traces, but on the fact that on top of that, travel diaries were filled in after each trip and interviews were conducted at three different occasions. Moreover, data about the traffic conditions of the network during the whole study period was collected.

Despite the limited sample of 32 participants, the nine weeks of data collection allow to investigate travellers' route choice behaviour over time. The richness and quality of the resultant data sets are aspects that should be highlighted as they definitely contribute to a better understanding of travellers' behaviour in reaction to travel information. In addition, given the lack of RP experiments such as the one described in this chapter described in the literature, the presented setup could also be used as a guideline for future research.

The data sets described in this chapter are essential inputs for the analyses carried out in this thesis. Quantitative and qualitative analyses of the data sets are presented in Chapter 6 and model estimation and prediction are put forward in Chapter 7. Combination of modelling framework and characteristic of the data collection makes the outcomes of Chapter 7 unique. They are the first ones based on a route choice model for a dynamic network with deterministic link attributes estimated with RP data collected in a congested network for which link travel times are known.

Chapter 6

Qualitative and quantitative analyses of travellers' reactions to travel information

This chapter presents quantitative and qualitative analyses of travellers' reactions to travel information based on the data sets described in Chapter 5. This is done in order to answer the research questions raised in Chapter 1.

The four data sets collected as part of this thesis are: GPS traces, travel diaries, interviews and the traffic conditions in the network, referred to as network data. The GPS traces provide information about travellers' actual choices. The travel diaries filled in after each trip allows investigating travellers' reactions to travel information. The interviews contribute to get insight into travellers' perceptions of the routes characteristics, use and willingness to pay for travel information over time. Finally, the network data provides the traffic conditions of the network, i.e. the travel times of all links of the transportation network during the period of the experiment. Contrary to most of the studies performed in the past, the analyses presented in this chapter are based on revealed preference (RP) data. The main reason to collect RP data was to investigate travellers' behaviour in a real setting in which they immediately experienced the consequences of their choices.

The analyses presented in this chapter are based on the valid commuting trips. The valid trips are the ones for which both the origin and destination of the GPS trace are within 1.5 km distance from home and work, for which there was no loss of GPS signal during the trip and the participant filled in the travel diary (section 5.3.5). The valid trips consist of 897 GPS traces and travel diaries. Among these, 374 refer to initial 3 weeks of data collection of the reference treatment and 523 to the subsequent 6 weeks of the information treatment. For the analyses based on the interviews, the sample size is of 32 x 3 (one per participant for each of the three occasions they were interviewed). The term "regular travellers" is used for travellers that were not equipped with TomTom. They were only able to check free/public sources of travel information. Otherwise, they are referred as informed travellers.

The following analyses are carried out: travellers' reactions to different sources of travel information, travellers' reactions to different timing of travel information provision (pre-trip x en-route), travellers' compliance with travel information, travellers' willingness to pay for travel information and changes in route choice behaviour due to provision of travel information. These analyses are respectively presented in sections 6.1 to 6.5. In addition, due to the importance of travellers' perceptions to the modelling of route choice behaviour, section 6.6 discusses travellers' awareness of the route choice and perceptions of route reliability. Table 6.1 depicts an overview of the analyses discussed in this chapter, required data sources and where they can be found.

Table 6.1: Overview of the analyses performed in this chapter

Analyses performed	Data Source(s)	Discussed in
Travellers' reactions to different timing of travel information provision	Travel diaries and interviews	Section 6.1
Travellers' reactions to different sources of travel information	Travel diaries	Section 6.2
Compliance with travel information	Travel diaries and interviews	Section 6.3
Willingness to pay for travel information	Interviews	Section 6.4
Changes in route choices due to travel information	GPS traces and travel diaries	Section 6.5
Travellers' awareness of route choice set	GPS traces and interviews	Section 6.6.1
Travellers' perceptions of route reliability	GPS traces, interviews and network data	Section 6.6.2

This chapter is based on:

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2012) A revealed preference study on route choices in a congested network with real-time information, *Proceedings of the 13th International Conference on Travel Behaviour Research*, Toronto, Canada.

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2014) Quantitative and qualitative analyses on route choice in a congested network with real-time information based on RP data (working paper)

6.1 Travellers' reactions to different timing of travel information

This section discusses travellers' reactions to the timing of travel information provision, i.e. travellers' behaviour of consulting travel information before (pre-trip) and/or during (en-route) the trip. The data sources used for these analyses are travel diaries and interviews.

Outcomes of the interviews suggest that (i) informed travellers tend to consult travel information more often over time, (ii) informed travellers appear to find en-route information more relevant than pre-trip information and the contrary applies to regular travellers and (iii) equipping travellers with a personalized device of travel information provision (e.g. TomTom), increases the likelihood travel information is used (Table 6.2). Observe that the percentages of regular travellers consulting travel information remained about the same. This suggests that being part of an experiment did not influence much their behaviour.

Table 6.2: Percentage of travellers consulting travel information based on the interviews

Period of the experiment	Consult pre-trip information			Consult en-route information		
	Overall	Informed travellers	Regular travellers	Overall	Informed travellers	Regular travellers
Beginning	19%	12%	43%	38%	38%	29%
Middle (5 weeks)	41%	38%	50%	72%	81%	33%
End	50%	54%	33%	66%	73%	33%

* Results refer to the participants' perceptions of "usually" checking travel information. Definition of "usually" was not presented.

The higher use of en-route information in relation to pre-trip information for informed travellers is somehow surprising as the contrary has been shown in the literature (Abdel-Aty et al., 1997; Jou, 2001). It has to be noted, however, that the outcomes of this research are based on RP data while those reported on the literature are based on SP (stated preference) data. It is then possible that travellers overestimated their use of travel information in the hypothetical situation. Outcomes of this research show that informed travellers consult travel information more often while driving (or just before starting the trip). This suggests that travellers do not tend to use travel information as a tool to plan their departure times despite having an appropriate device also for this purpose. Instead, they appear to be more interested to know the best route considering they want to depart "now". As reported by the participants, they usually consulted travel information just when entering the car and not beforehand while still at home or in the office.

Besides that, Table 6.2 shows that while the use of pre-trip information by informed travellers continuously increased along the experiment, the use of en-route information first increased and then decreased. The continuous increase on the use of pre-trip information may be due to travellers' curiosity about the traffic conditions in the transportation network triggered by the experiment. On the other hand, the level of familiarity with the TomTom devices might explain the pattern regarding the use of en-route information. TomTom information was the source of en-route travel information consulted the most. As none of the travellers previously

owned a TomTom, at first they might have been excited about the new gadget, curious to investigate the consistency of the travel information and willing to get a clearer picture of the traffic conditions in the network. As travellers became more familiar with TomTom, the interest in the new gadget may have diminished and travellers then opted to rely on their own experience about the traffic conditions. They may have thought that further travel information was not needed. This explains the reduction in the percentage of travellers consulting en-route information at the end of the experiment in comparison to the middle of the experiment.

As the outcomes of Table 6.2 are based on the interviews, they refer to travellers' perceived use of travel information. Based on the output of the travel diaries filled in after each trip, however, it is also possible to investigate whether travel information was actually consulted before and during the trip. During the interviews travellers were asked if they usually consulted travel information before or during their trips. Although a definition of "usually" was not presented, it is reasonable to assume that it corresponds to at least 40% of the trips made. Table 6.3 presents the percentage of travellers who consulted travel information in at least 40% of their trips based on the travel diaries.

Table 6.3: Percentage of travellers consulting travel information based on the travel diaries

Period of the experiment	Consult pre-trip information			Consult en-route information		
	Overall	Informed travellers	Regular travellers	Overall	Informed travellers	Regular travellers
Beginning	N/A	N/A	N/A	N/A	N/A	N/A
Middle (5 weeks)	13%	8%	33%	34%	35%	33%
End	44%	46%	33%	63%	73%	17%

* Outcomes in "Middle" consider data from "Beginning to Middle". Outcomes in "End" consider data from "Middle to End".

Comparisons between Table 6.2 and Table 6.3 show that up to the middle of the experiment (Middle – 5 weeks) travellers' overall perception of the use of both pre-trip and en-route information is overestimated, meaning that travellers believe that they are using travel information more than they actually are. At the end of the experiment, however, the perceived use is similar to the actual use. As up to the middle of the experiment the available sources of travel information were mostly public sources, this outcome might suggest some bias. Either the travellers wish to use more often what is available or they want to say they use the investments made with their taxes. It is also possible that travellers actually think that consulting travel information about 2 times per week means "usually" checking it. In practice, however, the use is rather limited. This suggests that not necessarily travellers behave the way they think they did or would do, which may have more serious implications to SP experiments.

Converting the outcomes of Table 6.3 to the number of trips in which pre-trip and en-route information were actually consulted, confirms that the percentage of trips in which travel information is consulted in the reference treatment is rather low (Table 6.4). Nevertheless,

travellers tend to consult travel information more often over time. This is especially the case for informed travellers consulting en-route information. In addition, the preference to consult en-route rather than pre-trip information further supports the argument that travellers are more interested to find an alternative solution in case their habitual route is too congested rather than looking for the best alternative route a priori. It has to be noted that as commuters was the target group investigated, they are familiar with the transportation network. Thus, it is possible that they have already incorporated the regular variability of traffic conditions to their route choice behaviour. This consequently may have influenced the way they reacted to travel information.

Table 6.4: Percentage of trips in which travel information was consulted based on the travel diaries

Treatment	Consult pre-trip information			Consult en-route information		
	Overall	Informed travellers	Regular travellers	Overall	Informed travellers	Regular travellers
Reference	15.8%	13.3%	29.3%	34.0%	33.5%	36.2%
Information	40.9%	40.3%	45.8%	66.9%	70.3%	40.7%

* Outcomes of informed and regular travellers are based on the number of trips made by each group in each information treatment. In the reference treatment it corresponds to 316 trips for informed travellers and 58 for regular travellers; in the information treatment it corresponds to 464 trips for informed travellers and 59 for regular travellers.

6.2 Travellers' reactions to different sources of travel information

The analyses regarding travellers' reactions to different sources of pre-trip and en-route travel information focus on the sources consulted during the commuting trips. The data source used is travel diaries.

Among the public sources of travel information, radio appears to be the preferred source of pre-trip information even when compared to websites on the internet (Table 6.5). Considering en-route public sources, however, radio is the preferred source only in the reference treatment while VMS is preferred in the information treatment (Table 6.6). The preference for radio might be explained by the fact that travellers tend to leave the radio on during their trips. As a result they end up listening to the news about traffic jams. With respect to the VMS, the participants reported that information displayed cannot be easily understood and that the sometimes the panels are covered by trees.

A remarkable finding is that irrespective of the timing of information provision (pre-trip or en-route), when personalized real-time travel information via TomTom is available, it turns out to be the main used source of travel information. In addition the drop in the percentage of trips in which radio information is consulted when comparing both treatments of information provision suggest that travellers ignore radio information as of the moment TomTom information is available. Travellers may be even listening to the radio, but not registering/processing that travel information may have been provided.

Table 6.5: Percentage of use of different sources of pre-trip information for all travellers based on the travel diaries

Treatment	Number of trips	Considered sources of pre-trip information				
		Internet	Radio	Television	TomTom	Other(s)
Reference	374	30.5%	62.7%	3.4%	-	6.8%
Information	523	0.9%	27.6%	0.0%	81.8%	0.5%

* The total percentage is higher than 100% because more than one source could be consulted

Table 6.6: Percentage of use of different sources of en-route information for all travellers based on the travel diaries

Treatment	Number of trips of trips	Considered sources of en-route information			
		TomTom	Radio	VMS	Other(s)
Reference	374	-	58.3%	54.3%	9.4%
Information	523	87.7%	11.4%	29.4%	0.9%

* The total percentage is higher than 100% because more than one source could be consulted

6.3 Travellers' confidence and compliance with travel information

This section discusses confidence/compliance behaviour, i.e. whether travellers actually follow the advice received. Definition of compliance was presented just before the experiment started. Travellers were informed that following the advice irrespective of having already decided to choose a specific route should be treated as being compliant with the advice. The same applied to the information displayed by VMS panels and provided via radio regarding least congested routes. These analyses are predominantly based on the *travel diaries* and supplementary on the *interviews*.

As discussed in section 6.1, en-route information is consulted more often than pre-trip information. When investigating compliance with travel information, however, the opposite holds, i.e. among the travellers who consult travel information, compliance with pre-trip information is higher than with en-route information (Table 6.7). This might be influenced by (i) the fact that changing routes while en-route would imply taking local roads, (ii) habit or (iii) travellers' misperception regarding compliance.

Table 6.7: Compliance rate with travel information based on the travel diaries

Treatment	Compliance with pre-trip	Compliance with en-route
Reference	64.4%	50.4%
Information	81.8%	75.1%

* The percentages are calculated based on the number of trips in which travel information was consulted

In the first case, taking local roads may be associated with the idea that the resultant travel time-saving is not so large, to lower speed limits and to higher amount of stops due to traffic lights and/or crossing streets. In the second case, due to habit travellers may prefer to stick to the habitual routes due to the familiarity with the environment and already feel satisfied to be aware of delays. Finally, travellers may misperceive their actual compliance when the travel information is aligned to their expectations. In such situations, although they are actually “complying with” the travel information, the travel information may be perceived as not necessary/not needed. Consequently, travellers may feel that they are not complying with it. Note that in absolute values, i.e. disregarding whether travel information was actually consulted, en-route information is consulted m

Outcomes of Table 6.8 and Table 6.9 show that the more detailed the source of the travel information, the higher the compliance rate. Besides that, compliance rates also appear to differ for informed and regular travellers (Table 6.10): informed travellers are more willing to comply with travel information than regular travellers. This suggests that provision of TomTom information has the potential to influence travellers' behaviour. It is interesting to observe that 77% of the informed travellers reported on the interviews that the route suggested by TomTom differed from their planned route less than 30% of the times. Among this group, 95% reported to have complied with the information over 70% of the times. For the one informed traveller who reported that the route suggested by TomTom differed from expectations over 70% of the times, compliance rates with travel information reduced to less than 30% of the times.

Table 6.8: Compliance rate with different sources of pre-trip information based on the travel diaries

Treatment	Considered sources of pre-trip information				
	Internet	Radio	Television	TomTom	Other(s)
Reference	25.4%	35.6%	0.0%	-	3.4%
Information	1.4%	13.6%	0.0%	66.8%	0.0%

* The percentages were calculated already considering the number of trips in which information was actually checked.

Table 6.9: Compliance rate with different sources of en-route information based on the travel diaries

Treatment	Considered sources of en-route information			
	TomTom	Radio	VMS	Other(s)
Reference	-	19.7%	29.1%	1.6%
Information	70.6%	0.9%	3.4%	0.3%

* The percentages were calculated already considering the number of trips in which information was actually checked.

Table 6.10: Compliance rate with travel information for informed and regular travellers based on the travel diaries

Treatment	Informed travellers		Regular travellers	
	Compliance with pre-trip	Compliance with en-route	Compliance with pre-trip	Compliance with en-route
Reference	52.4%	49.1%	94.1%	57.1%
Information	80.7%	79.8%	88.9%	12.5%

* The percentages were calculated already considering the number of trips in which information was actually checked

A point to be observed is that compliance rate with information depends not only on the information itself, but on travellers' expectations and willingness to change routes (may this be the case). As discussed in section 6.5, the GPS traces show that even during the information treatment most travellers tend to stick to their habitual routes and not to explore different routes. This suggests the high importance of habit for the route choice decisions of commuters. Questions concerning travellers' compliant behaviour when the travel information is not aligned to their expectations are therefore raised. If travellers comply with travel information only when it is aligned to their expectations, the potential of information to help alleviating congestion may be considerably lower than expected. As more than 75% of the informed travellers reported to be satisfied or very satisfied with TomTom and around 80% considered the travel information reliable over 70% of the times, it seems that the lower rates to consult and comply with travel information is not directly related to lack of confidence/trust in the travel information.

6.4 Willingness to pay for travel information

Outcomes of section 6.2 show that among all investigated sources of travel information, TomTom is the source used the most both pre-trip and en-route. It has to be noted, however, that TomTom information was provided without any cost to the participants. This is not what happens in a quotidian situation as on the top of purchasing costs of the equipment itself, some sort of membership is required. This section discusses travellers' willingness to pay for travel information based on the outcomes of the *interviews* conducted along the experiment.

Travellers were interviewed at two different occasions regarding their willingness to pay for travel information: at the beginning and at the end of the experiment. They were asked to rate between 1 and 5 how likely they would consider to pay for: (i) travel time of the fastest route, (ii) queue length of the fastest route, (iii) travel time of all (main) alternative routes and, (iv) queue length of all (main alternative routes). In this uniform varying scale 1 represented very unlikely to pay for travel information and 5 very likely to pay for travel information. Travellers were not informed of how much they would have to pay. Instead focus was put on whether they would consider paying at all. As travellers were interviewed at the beginning and at the end of the experiment, it is possible to investigate whether being (more) aware of the benefits of travel information influences the "purchasing" behaviour.

Outcomes regarding travellers' willingness to pay for travel information are based on the reported rates in relation to the maximum (Figure 6.1). The maximum possible outcome can

be achieved only if all travellers rate their willingness to pay for travel information to the highest level, i.e. 5. Figure 6.1 shows that overall travellers' willingness to pay for travel information increases with their experience using travel information (or stays about the same). This suggests that receiving travel information is actually more important than travellers thought at first. This is particularly the case for travel time information about the fastest route. It is interesting to observe that despite the initial higher willingness to pay for travel information about all routes, it remains about the same over time. Besides this, travellers are more willing to pay for travel time rather than queue information. In particular, travel times in all (main) routes appear to be the type that travellers are more likely to pay for. This suggests that travellers like to be in control and prefer to assess the situation and decide themselves what route to choose. Queue information about the fastest route appears to be the type that travellers are more unlikely to pay for. When queue information is provided, travellers still have to make some inference about speeds and the degree of severity of the congestion for instance (e.g. if congestion is due to a severe accident or just regular congestion). Thus, based only on the length of the queue is not possible to know the duration of a trip.

The willingness to pay for travel information also varies among travellers (Figure 6.2). For instance, informed travellers tend to be more willing to pay for travel information (irrespective of the type) over time. Conversely, the reverse pattern is (generally) observed for regular travellers. This suggests that travellers derive more satisfaction from private/personalized sources of travel information rather than from public source. This is possibly due to anticipation of the potential benefits of using TomTom in situations other than the commuting trips, which might not have been the case for regular travellers who used only public sources.

Despite the increase in travellers' willingness to pay for travel information over time, especially informed travellers, the rates are very low and travellers are at most neutral about it. For instance, with respect to the variability in the answers, with exception of travel time in all routes, more than 50% of the travellers rated their willingness to pay as 1, i.e. very unlikely to pay. For travel time in all routes, however, 30% was neutral and 22% was likely to pay for it. For the other situations, 6% was likely to pay for travel time about the fastest route, 3% for queue length in the fastest route and 15% for queue length in all routes. These outcomes, nevertheless, are expected and aligned with insights from the previous sections. If even when travel information is provided for free travellers do not consult it often, it is also expected that they would not be willing to pay for it. As mentioned by the participants, travel information seems to be more relevant for trips in which they are not familiar with or for longer trips rather than their daily commuting trips.

6.5 Changes in route choice due to travel information

This section primarily discusses the influence of travel information on changes in route choices. Besides this, as outcomes of section 6.1 suggest that travellers are more likely to use travel information to plan routes rather than departure times, a discussion concerning changes in departure time due to travel information is also provided to further support those findings. The proposed analyses are based on the travel diaries and GPS traces.

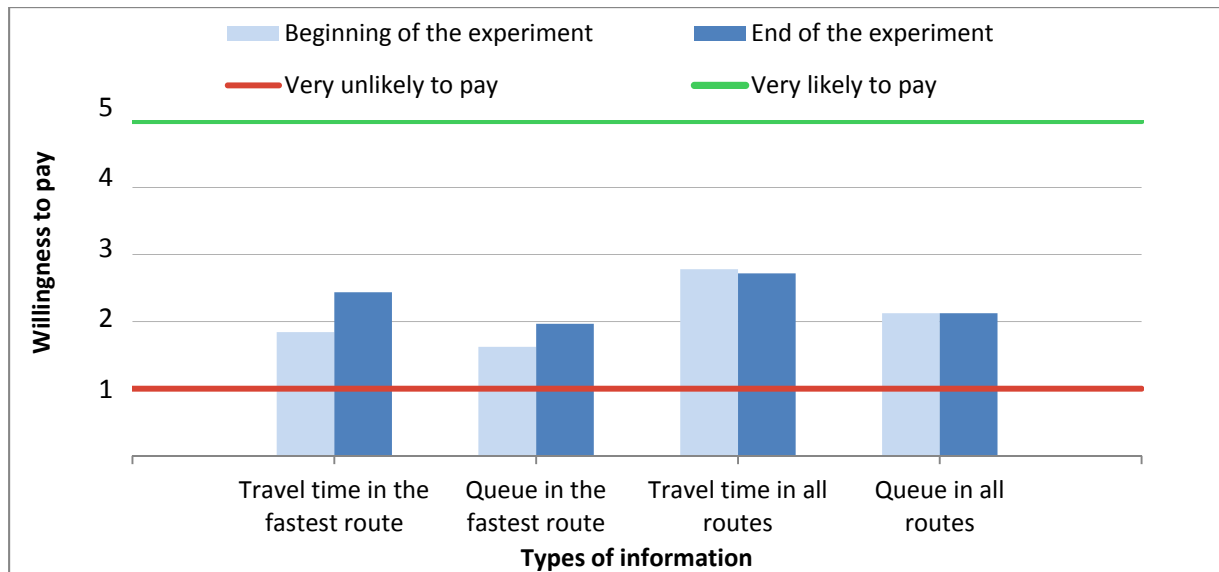


Figure 6.1: Willingness to pay for different types of travel information

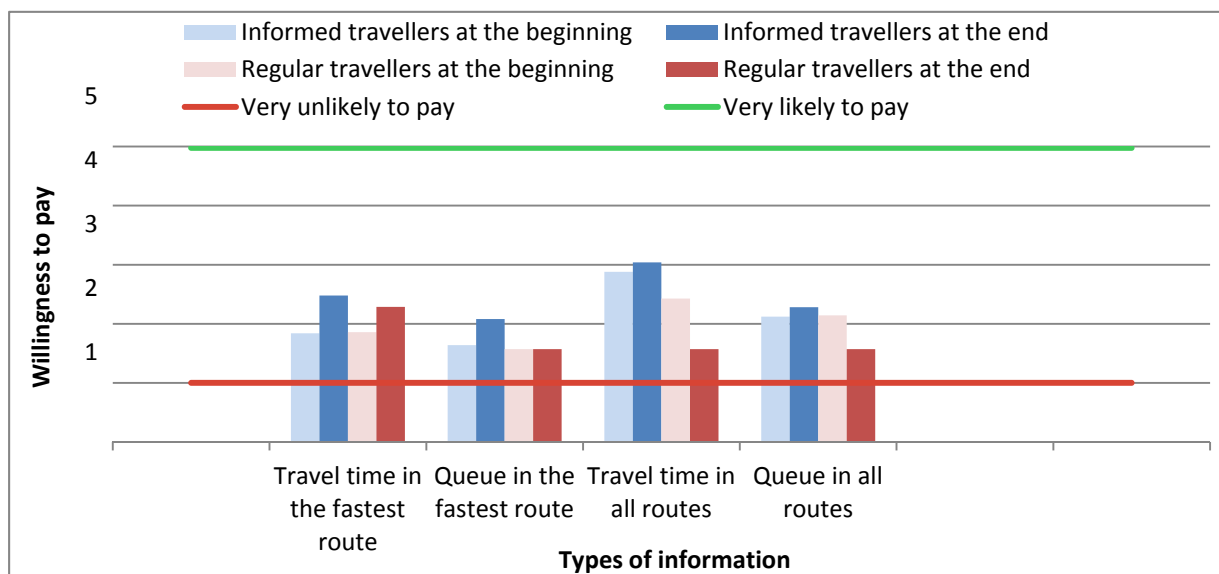


Figure 6.2: Willingness to pay for different types of information for informed and regular travellers

One of the questions of the travel diary regarded which route travellers were planning to take in the next trip. Travellers had to answer this question at the end of the current trip. The set of possible alternative routes to be chosen was defined based on a subjective route choice set discussed in detail in section 6.6.1. By comparing the planned routes reported in the travel diaries with the actual routes tracked by the GPS it is possible to verify whether the planned route matches the chosen route. A planned route was considered to match a GPS route if all links were the same.

Table 6.11 shows that travellers tend to stick to their originally planned route choices. Habit and familiarity with the transportation network are factors that may influence that.

Nevertheless, it is interesting to observe that informed travellers changed routes more often after provision of TomTom information. This suggests some potential of travel information to influence changes in route choice behaviour. It has to be pointed out that one of the travellers regularly made a small detour in the planned route. Consequently the chosen route did not fully match the planned one. If this small detour is disregarded, the percentage of route changes drops even more. For informed travellers in particular it reduces from 6.3% to 3.1%.

Table 6.11: Percentage of changes in route choices due to travel information based on the travel diaries

Treatment	Overall		Informed travellers		Regular travellers	
	Number of trips	Route Changes	Number of trips	Route Changes	Number of trips	Route Changes
Reference	374	1.1%	316	0.6%	58	3.4%
Information	523	5.7%	464	6.3%	59	1.7%

* The percentages were calculated considering the number of trips in which pre-trip information was actually checked irrespective of whether it matched expectations.

In order to further investigate the role of habit for route planning, Figure 6.3 depicts travellers' exploration rate considering the habitual routes as the reference. In other words, if a route different than the habitual one was chosen, it was considered as if exploring a different route. The habitual route was reported by the travellers at the beginning of the experiment and then confronted with the GPS traces so as to verify whether the reported habitual route was the actual habitual route. Based on Figure 6.3 it appears that travellers' route choice behaviour did not differ much with respect to the treatments of information provision. In general travellers tended to stick to their habitual routes, even during the information treatment: around 75% of the times the chosen route was the same as the habitual route either to or from work. Only few travellers showed an explorative behaviour.

As the experiment dealt with commuters, it is possible to argue that the exploration phase has already passed and that habit is more important now. Therefore, unless one is dealing with extreme conditions, such as severe congestion, accidents or extreme weather conditions, there would be little reason to choose a different route (even with the aid of information). According to the travel diaries such type of events did not occur during the experiment. The resultant congestion during the period and area investigated consisted only of regular traffic jams. Figure 6.4 depicts the routes chosen before and after provision of travel information to illustrate that travellers' route choice pattern remained about the same.

An interesting finding regards travellers' willingness to change routes after experiencing delays longer than expected. After each trip, travellers reported in the travel diaries whether they had incurred in delays longer than expected and whether in hindsight the same route would have been chosen again. A delay longer than expected was told to be discretionary to the perception of each traveller. Table 6.12 shows that even when incurring in delays longer than expected, travellers prefer to stick their choices. In other words, even if they could go back in time with extra knowledge about the consequences of their route choices, they still would stick to same route. Possible reasons for this behaviour are habit and the belief that other route(s) would result in even longer delays. As discussed in the literature, higher travel

times in habitual routes are treated as exceptions while low/normal travel times reinforce the preference for the habitual route (Bogers, 2009). The opposite holds for non-habitual routes.

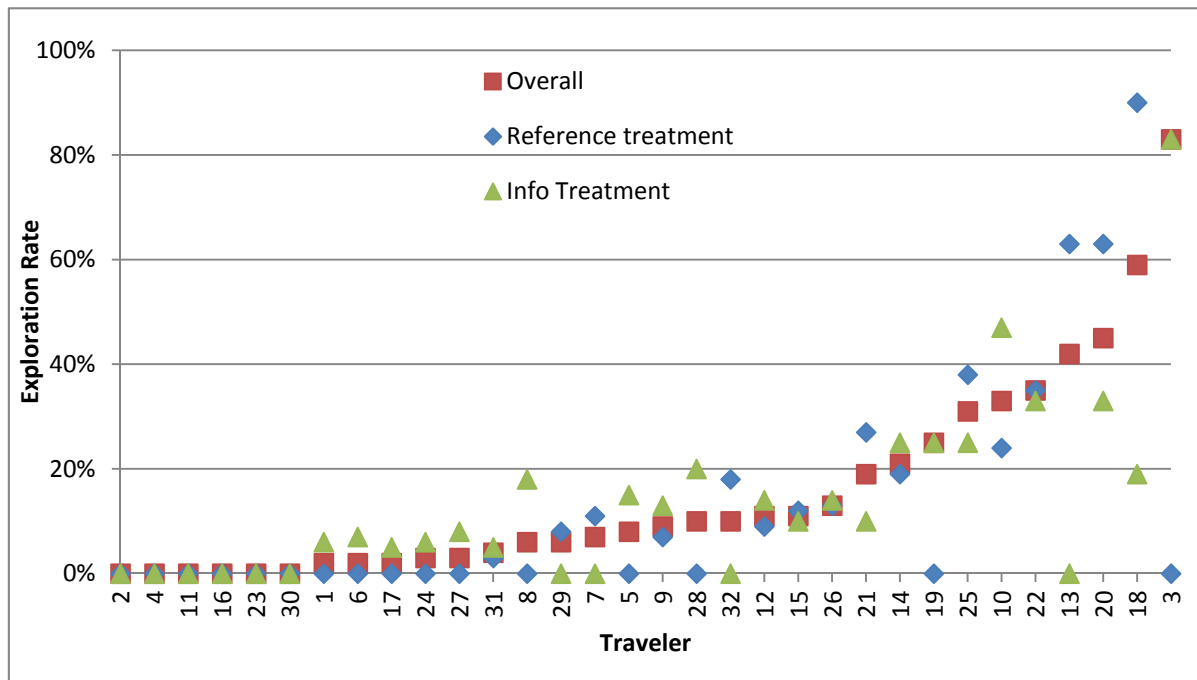


Figure 6.3: Exploration rates during the experiment based on the GPS traces

* Travellers were sorted by increasing order of exploration rate

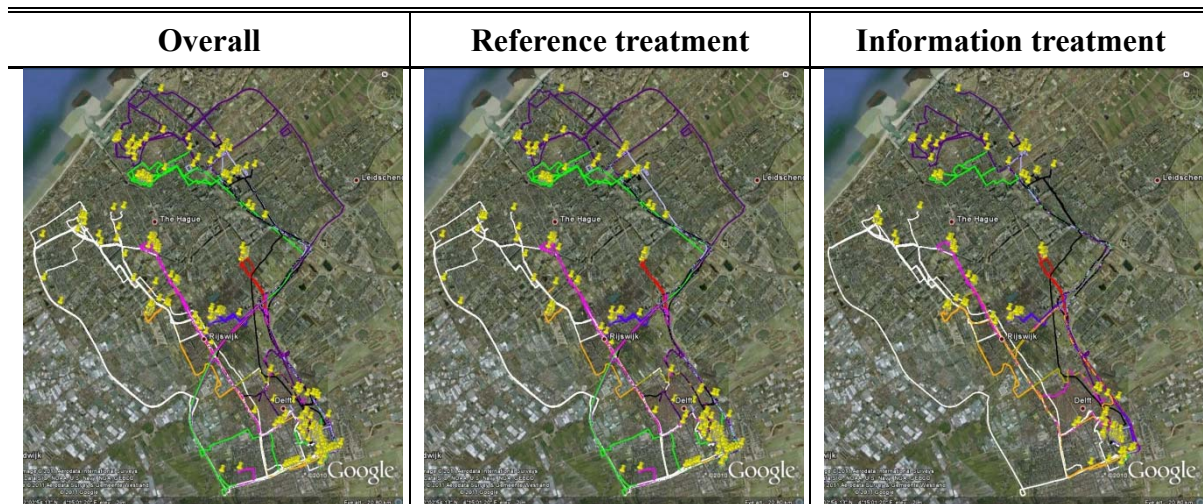


Figure 6.4: Route choices in the reference and information treatments

* Different colours correspond to different travellers

With respect to changes in departure times, section 6.1 suggests that in the context of this experiment travellers seem to be more likely to change routes rather than departure times. Travellers were asked in the travel diary whether the originally planned departure time changed due to the provision of travel information and Table 6.13 depicts the results.

Table 6.12: Percentage of travellers willing to choose the same route in hindsight based on the travel diaries

How long the delay was	Would have chosen the same route in hindsight	Would NOT have chosen the same route in hindsight
Longer than expected	10.5%	3.6%
NOT longer than expected	84.8%	1.1%

* The term "longer" was subjective to the perception of each traveller of what longer meant.

It is possible to derive from Table 6.13 that only in 6 out of 897 trips there were changes in departure time due to travel information provision. Although this is not aligned to what has been reported in the literature regarding travellers' willingness to change departure times (Vaughn et al., 1993; Mahmassani, 1997; Mahmassani and Liu, 1997), it seems very plausible in the context of commuting trips. Travellers might already have a pre-defined schedule, involving arriving and leaving work at a certain time, possibly picking up/leaving kids at school, etc. Changing schedules therefore might be associated with a high disutility and considered as an option only under extreme conditions. It seems that for the usual regular congestion travellers have already incorporated the "expected" delay as part of their routines, not being necessary to adjust their schedules.

Table 6.13: Percentage of changes in departure time due to travel information based on the travel diaries

Treatment	Overall		Informed travellers		Non-informed travellers	
	Number of trips	Changes in departure time	Number of trips	Changes in departure time	Number of trips	Changes in departure time
No information	374	1.3%	316	1.3%	58	2.0%
Information	523	0.2%	464	0.0%	59	2.0%

* The percentages were calculated already considering the number of trips in which pre-trip information was actually checked

Both changes in route and departure time choices due to provision of travel information are considerably low. This suggests that in order to change travellers' behaviour, in particular that of commuters, either the penalty associated with a late arrival or the benefit for on time/earlier arrival should be higher. For most of the regular commuting trips, however, this does not seem to be the case. For instance, only 1 out of the 32 participants had very strict working hours due to the nature of his work that required punctuality or early arrival. For the others, punctuality or early arrival was considered to be mandatory in only 12% of the trips, 10-20 minutes delay would still make it possible to fulfil the daily obligations in 49% of the trips, 30 minutes-1 hour delay would still make it possible to fulfil the daily obligations in 35% of the trips and cancelling the trip would be possible in 4% of the trips. The low percentage of trips that could be cancelled, and therefore substituted by teleworking, is very disappointing as it

suggests a limited use of information and communication technologies as a substitute for travelling.

6.6 Travellers' perceptions

As discussed in the previous sections, there appears to be some difference between what travellers actually do and how they perceive they do. The importance of travellers' perceptions to modelling and consequent behavioural implications has been largely discussed in the literature (Brownstone et al., 2003; Bonsall, 2004; Avineri, 2009; Connors and Sumalee, 2009; Peer et al., 2010) and this section address two main points: travellers' perceptions of the route choice set and of the attributes of the alternatives.

The dynamic modelling framework adopted in this thesis does not require any definition of route choice set. Moreover, it takes into account the variability of traffic conditions in the network to the modelling process. Therefore, in order to justify the choice for such modelling framework it seems logical to investigate travellers' awareness of the route choice set as well as how objective measurements of travel time variability are compared to travellers' perceptions of route reliability. These two points are respectively addressed in section 6.6.1 and section 6.6.2.

6.6.1 Travellers' awareness of the route choice set

During the interviews at the beginning of the experiment, travellers were asked to draw on a map the routes they were aware of between their homes (in The Hague) and work (in Delft). Combination of the answers of all travellers and the correspondent roads usually driven to and from work defined a subjective route choice set¹⁴. Figure 6.5 depicts the main roads of the subjective route choice set and the homes of the participants. Red dots indicate the home of the participants and the yellow drawing pin indicates the work, i.e. the TU Delft. The colours and letters next to each road represent different categories: the "A" roads in red are motorways, the "N" roads in blue are national roads (in terms of speeds it is sort of a provincial highway) and the "L" roads in yellow are local roads.

The coding of the main roads is depicted in Figure 6.6. It is possible to see 5 main roads in Delft (1 to 5) and 5 in The Hague (4 to 8). Some participants live just in the outskirts of Delft and some further in The Hague. As a result, two different subjective choice sets were derived. Choice set CS 1 is for participants living in Delft and is composed of main roads 1 to 5 (in Delft). These are the digits attributed to the roads in Delft. Choice set CS2 is for participants living in The Hague and is composed by one road in Delft (1 to 5) and one in The Hague (4 to 8). The routes belonging to CS2 are, therefore, in the range 14 to 58, i.e. the first digit belongs to the roads in Delft and the second digit to the roads in The Hague. For example, choosing road 1 in Delft and road 4 in The Hague results in route 14; choosing road 1 in Delft and road 5 in The Hague result in route 15 and so on. It has to be observed that the total amount of routes in CS2 is only 17 as not all possible road combinations were considered to be "real" alternative routes by the participants. Only the road combinations reported in the subjective choice set were considered to be real alternatives.

¹⁴ Despite the importance of the path itself, i.e. the sequence of links, and the modelling approach used in this thesis (Chapter 4), the analyses regarding travellers' perceptions considers a set of main alternative roads.

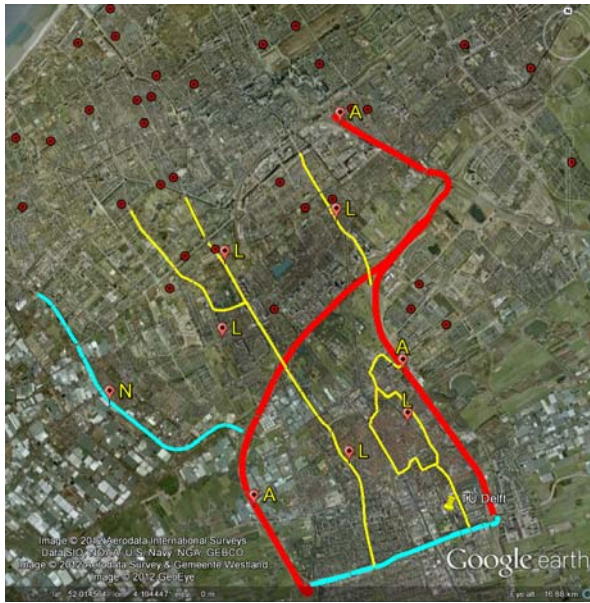


Figure 6.5: Subjective route choice set between Delft and The Hague and home of the participants

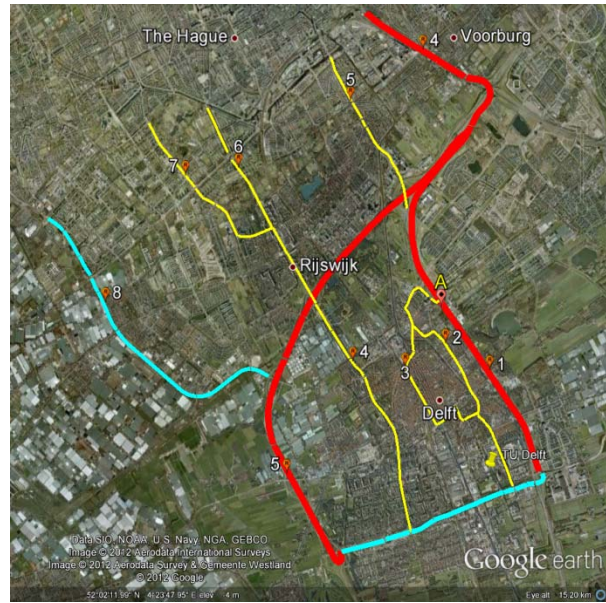


Figure 6.6: Coding of the roads in the subjective route choice set

As the participants were tracked with GPS devices, it is possible to confront the perceived awareness of the route choice set reported during the interviews at the beginning of the experiment with the actual awareness. Results show that the general route choice set provided by the participants is quite comprehensive. For instance, only 2 extra roads (in The Hague), thus in CS 2, had not been mentioned by any of the participants and were actually chosen during the reference treatment. Only minor diversions from the subjective route choice set to local routes were observed. On the individual level, i.e. comparing the outcomes of the interviews and GPS traces for each traveller, the reported route choice set was rather smaller than the observed. For instance, 12 out of 32 participants chose at least one route different from the reported during the reference treatment. The non-reported routes were in general not as straightforward as the main roads, i.e. required a considerable amount of left and right turns, and were in general rarely used. As extensively discussed in the previous sections, commuters tend to stick to the same and not to switch routes. Nevertheless, this outcome shows a big discrepancy between travellers' perceived awareness of the route choice set and their actual awareness.

As discussed in section 2.4, definition of the route choice set is one of the main modelling issues. Outcomes of this experiment suggest that even commuters who regularly drive between specific areas are not entirely aware of the route choice set. This further supports the added value of the modelling framework used in this thesis, which does not require definition of choice set and/or sampling of alternatives.

6.6.2 Travellers' perceptions of route reliability

The importance of travel time reliability for route choice decisions has been discussed in the literature at length. Measures of reliability are related to variations of the expected values and imply a notion of repetition or regularity. In addition, it also implies that different definitions of the expected value (e.g. travellers' expected travel time) might influence the outcomes. For

instance, travellers used to regularly perform a specific trip in the early hours of the morning might be surprised by the duration of the same trip during rush hours. Bates et al. (2001) propose that the concept of reliability should be associated with variations in the mean travel time that cannot be predictable by the travellers. Based on this concept, therefore, it would be inappropriate to characterize the variations of travel times (within the day) as unreliability.

A comprehensive review of the value of travel time reliability is presented in Carrion and Levinson (2012). Based on a review of empirical evidence they highlight that (i) the majority of the initial research about travel time reliability (or predictability) were mainly qualitative analysis based on questionnaires about travellers' preferences (Prashker, 1979; Chang and Stopher, 1981; Vaziri and Lam, 1983), (ii) from the moment that more quantitative studies started to be done, focus was put on stated preference experiments (Black and Towriss, 1993; Abdel-Aty et al., 1997; Cook et al., 1999; Bates et al., 2001; Hollander, 2006; Tseng et al., 2009) and (iii) there are few studies using RP data to investigate travel time reliability.

Main drawbacks usually associated with analyses of route reliability through RP experiments are the lack of a real experimental settings with significant variation of travel time across at least two alternatives and difficulties to measure travel time data (Carrion and Levinson, 2012). In addition, RP studies vary with respect to the source of travel time measurements: objective travel time distribution measured by devices such as loop detectors (Lam and Small, 2001; Small et al., 2005; Small et al., 2006) and subjective travel time distribution, i.e. travel times reported by the subjects (Peer et al., 2010).

Studies focusing on travellers' subjective value of travel time are of great importance because they deal with a key aspect for modelling, i.e. the difference between objective measurements of travel time variations from subjective travel time distributions. Due to misperceptions travellers might see worthwhile savings and predictability (low variability) that do not match the reality. This section discusses the issue of travellers' perceptions of travel time variability and its relation with route reliability based on the interviews, network data and GPS traces.

The network data was collected for the whole study period. Although for some links it was not available, this was not the case for the routes belonging to the subjective choice set. The travel times of each route of the subjective choice set were defined as the sum of the links' travel times for each direction, i.e. to and from work, and considered departure times every 5 minutes during the morning (7:00h – 10:00h) and afternoon (16:00h – 19:00h) peak hours.

At the beginning of the experiment the participants were asked to rank their known routes in terms of preference and to indicate the fastest, the most reliable and the most unpredictable route. Notion of route reliability was introduced as: "routes in which the travel times are usually about the same magnitude and you know what to expect in terms of travel times (irrespective of how fast the route is)."

The standard deviation of travel times was normalized in relation to the average travel times to measure route reliability. As the travel times of some routes were (much) higher than others, this was an appropriate way to have non-biased coefficients. The resultant measurement is a coefficient of variation. Figure 6.7 (a and b) depict objective measures of coefficient of variation and average travel times based on the network data for choice set CS1 in the morning and afternoon peak hours; outcomes of CS2 are in Figure 6.8 (a and b).

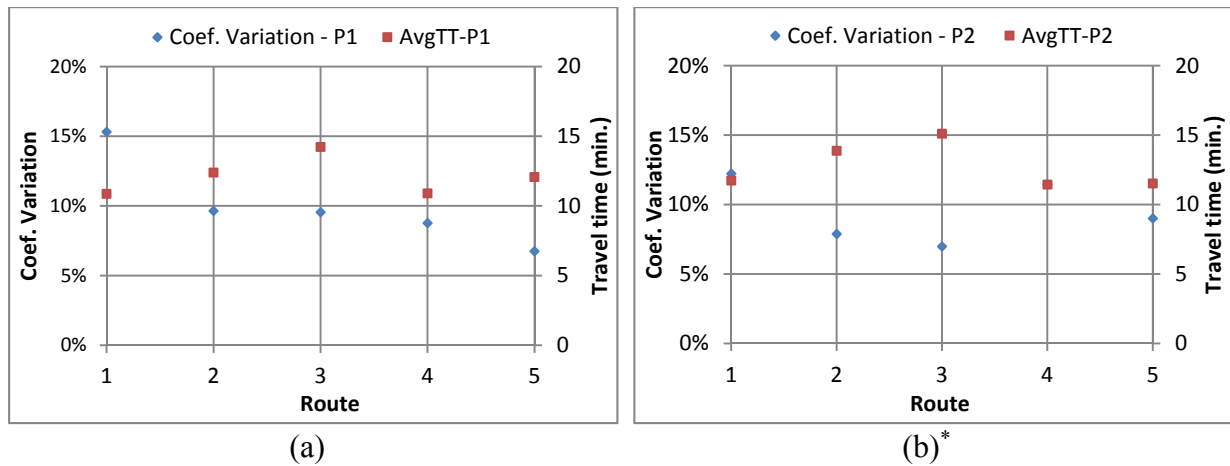


Figure 6.7: Objective measures of reliability of the routes in choice set 1 (CS1) in the morning (a) and afternoon peak hours (b)

* The coefficient of variation and average travel time of route 4 in Figure 6.7 (b) are superposed

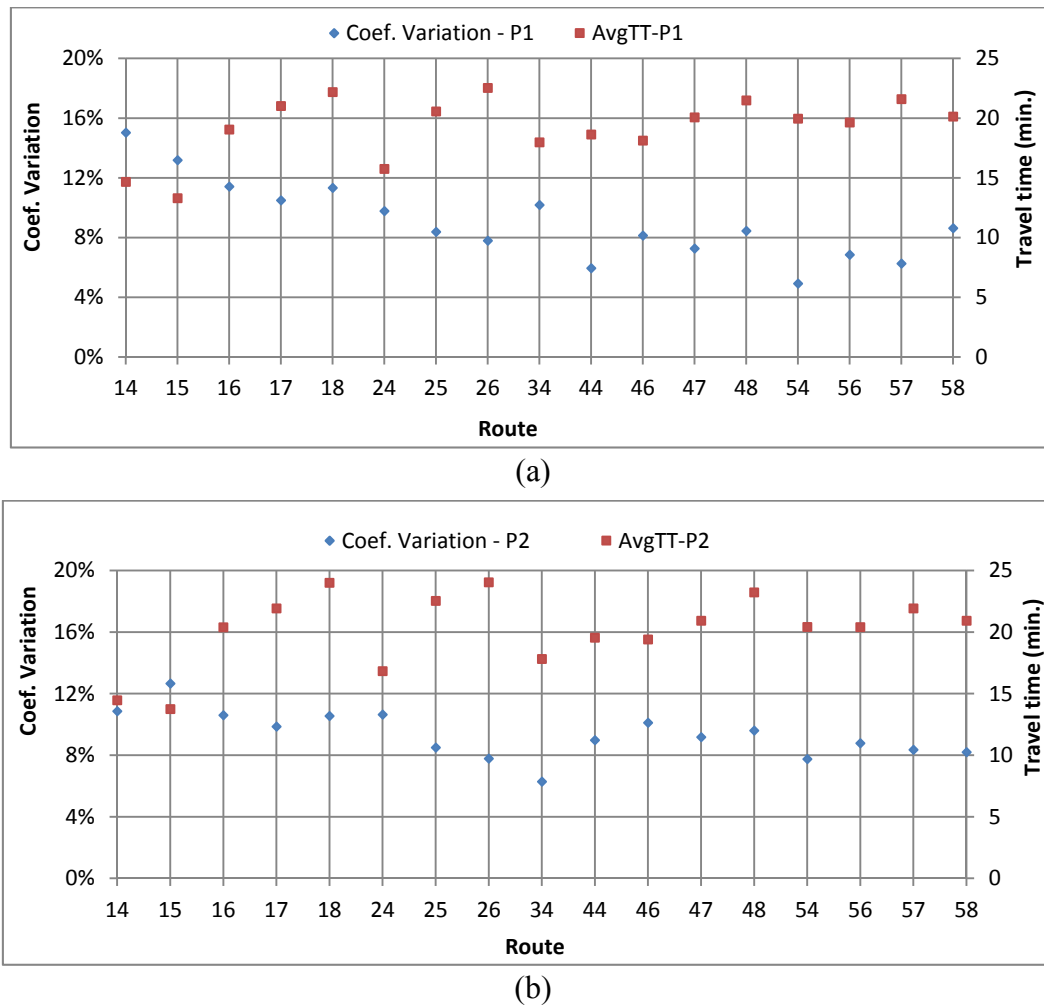


Figure 6.8: Objective measures of reliability of the routes in choice set 2 (CS2) in the morning (a) and afternoon peak hours (b)

Outcomes of the interviews conducted at the beginning of the experiment, indicate that among the routes belonging to CS1, route 1 is the habitual route and the other routes are equally less preferred. In CS2, route 14 is the habitual route and routes 46, 54 and 58 are the least preferred. Participants usually considered their habitual routes to be reliable and their least preferred route to be unreliable. By confronting the objective measures of travel time variability of the network data with travellers' perceptions, it is possible to observe that the routes perceived to be reliable (and chosen the most) are actually among the most objectively unreliable routes (although fastest routes). For instance, the GPS traces show that route 1 was chosen 73% of the times in CS 1 and route 14 in 42% of the times in CS 2. These are the main highways between Delft and The Hague and also the fastest and most straightforward routes.

Among the 32 participants who joined the experiment, 30 were assertive about which route they considered to be reliable (and 2 did not consider any route reliable) and 21 were assertive about which route they considered unreliable (and 11 did not consider any route unreliable). Table 6.14 depicts travellers' perceptions of the routes' characteristics. The first column refers to the aspect investigated. The second and third columns show travellers' perceptions of aspects that refer either to the morning or afternoon peak hours. Finally, the fourth column shows travellers' perceptions of aspects in which peak hours were not specifically investigated. For instance, the questions about the habitual routes were specific for the morning and afternoon peak hours. The other questions were generic with respect to the time of the day.

Two main aspects can be highlighted from Table 6.14: (i) travellers appear to be biased in favour of their habitual routes and (ii) the terms fastest & reliable and slowest & unreliable appear to be confound. This is aligned with findings reported in the literature regarding perception of route reliability for habitual and non-habitual routes as discussed in section 6.5, i.e. higher travel times in habitual routes are treated as exceptions while slow/normal travel times reinforce the preference for the habitual route; the opposite holds for non-habitual routes (Bogers, 2009). It is interesting to note that travellers' perceptions did not change with respect to the period of the day (morning or afternoon peak hours).

These outcomes reinforce the core assumption that the concept of reliability should be associated with variations in the mean travel time that cannot be predicted by the travellers (Bates et al., 2001). As commuters are the target group investigated, it is possible that they have already assimilated the variations of travel time. As a result, they should not be considered unpredictable. On the other hand, it is also possible that travellers assess the trade-off between faster & unreliable routes versus slower & reliable routes and opt for the former, which in the context investigated provided the best trade-off. Nevertheless, the fact that the routes chosen the most by the participants are objectively among the most unreliable but fastest cannot be contested from the data.

The findings of this research regarding the importance of route reliability for route choice decisions are somehow contrary to what has been reported in the literature. This may lead one to argue that the importance of route reliability is overestimated. Outcomes of this research suggest that when travellers are actually driving and making their route choice decisions what really matters is to arrive at the destination the fastest possible. This implies that travellers may behave as risk prone. More studies based on RP data are required to value the merits of

these findings and determine whether travellers behave as risk prone, whether the perceived travel time uncertainty is low or whether the sample size was small.

Table 6.14: Travellers' perceptions of the routes' characteristics

Aspect Investigated	Amount of travellers		
	Morning peak	Afternoon peak	Through the day
Habitual route is actually the most reliable route	2	3	-
Habitual route is actually the most unreliable route	18	17	-
Preferred route in the morning is the same as in the afternoon	-	-	25
Consider the habitual route reliable	17	17	-
Consider the habitual route to be reliable (but actually it is among the most unreliable)	8	8	-
Consider the least preferred route unreliable	-	-	8
Consider the least preferred route to be unreliable (but actually it is among the most reliable)	-	-	4
Route considered reliable is actually among the most reliable	-	-	8
Route considered unreliable is actually among the most unreliable	-	-	9
Route considered reliable is actually among the most reliable AND route considered unreliable is actually among the most unreliable	-	-	2
Route considered reliable is actually among the fastest	-	-	16
Route considered reliable is actually among the fastest AND unreliable	-	-	9
Route considered unreliable is actually among the slowest	-	-	14
Route considered unreliable is actually among the slowest AND reliable	-	-	5

6.7 Conclusions

This chapter presents some key behavioural insights into travellers' reactions to travel information based on analyses of the data sets. It primarily focuses on travellers' reactions to different timing (pre-trip and en-route) and sources of travel information provision, compliance with travel information, willingness to pay for travel information and behavioural changes due to travel information. Besides that, analyses regarding travellers' perceptions are discussed due to their importance to the modelling of travellers' behaviour.

The data sets used in the aforementioned analyses consist of GPS traces, travel diaries, interviews and data about the traffic conditions in the network, referred to as network data. The resultant data set consist of a sample of 897 valid GPS traces and travel diaries (374 observations during the initial 3 weeks of data collection of the reference treatment and 523 to subsequent 6 weeks of the information treatment) and 32 x 3 interviews (one per participant for each of the three different occasions they were interviewed).

Results of this study suggest that equipping travellers with a personalized device of travel information provision increases the likelihood that they will use it to assist their route choices. Planning the best departure time to avoid congestion seems to be less important than finding out the best route choice considering the departure time is “now”. Therefore, the importance of provision of real-time travel information cannot be neglected. As commuters were the target group investigated, their schedule may have already been adjusted taking the delays into account. This implies that in order to change the behaviour of commuters either the penalty associated with a late arrival or the benefit for on time/earlier arrival should be higher.

It has to be stressed that in order for the travel information to be effectively used, comparative alternative roads should be available and the resultant travel time savings should be high to stimulate changes from preferred habitual route. The participants reported that the travel information provided was often aligned with their and no apparent change in travellers’ route choice pattern was observed. As a result, travellers’ reactions to travel information in situations in which it diverges from their expectations still remain to be investigated. The next chapter discuss the significance to which travel information influences route choice behaviour. Remarkable finding is that even when incurring in delays longer than expected, travellers prefer to stick to their preferred planned routes. This reinforces the importance of the preferred/habitual route for route choice decisions, which is in line with the literature (Bogers, 2009; Ben-Elia and Ettema, 2011). Due to the role of habit, compliance rates with travel information that is not aligned to travellers’ expectations might be even lower (and the consequent impact on network dynamics). Further research directly focusing on the importance of habitual routes for route choice behaviour should be done. In this experiment, the habitual routes were usually highways (in general also the fastest routes). Research focusing on the role of routes’ characteristics would contribute to understand the effects of road type, i.e. local roads, provincial roads, highways, etc. on travellers’ route choice behaviour.

During the experiment travellers did not have to pay to receive personalized travel information. However, they would have to pay for it in real life. The target group investigated in this this experiment (commuters living with 19 km from work) corresponds to 60% of the car commuters in the Netherlands (Dutch MON database, 2008) and they do not appear to be willing to pay to receive travel information. This behaviour is in agreement with the patterns of use and compliance with travel information. As travellers did not often use travel information that was provided for free, it was expected that they would not be willing to pay for it. According to the participants, travel information is more relevant for trips in which they are not familiar with or for longer trips rather than the daily commuting trips. This has important implications for commercial information services and appropriate business models for such services.

Finally, regarding travellers' perceptions, outcomes of this experiment suggest that travellers are not fully aware of the route choice set. This has strong implications to the modelling of route choice behaviour based on RP. Furthermore, it supports the added value of the modelling framework used in this thesis, which does not require sampling of alternatives. Travellers' perceptions of route reliability also differ from objective (actual) measures. In particular, outcomes of this study suggest that travellers' are biased in relation to their habitual routes and consider it to be reliable when in reality they are among the most (objectively) unreliable (but fastest routes). This implies that:

- (i) Objective variations of travel time may have been already incorporate by the travellers and as such they are not considered unpredictable anymore. Therefore, they cannot be used as a good measure of reliability. One should be careful on how to define and incorporate travel time variability in route choice models as not only the objective variability, but also the perception errors should be taken into account;
- (ii) Travellers are actually interested in arriving at their destination as fast as possible, even if this implies behaving in a risk-prone fashion.

More RP experiments are necessary to validate whether travellers behave as risk prone or whether the observed behaviour is due to the fact that the perceived travel time uncertainty is low. If this is due to the latter, objective measurements of travel time variability might not be so relevant. Travellers already seem to be satisfied to know the extent of the congestion they are getting into to.

The next chapter presents estimation results having as input behavioural insights from this chapter. Due to the combination of the modelling approach (Chapter 4) and the data collection procedures (Chapter 5), the estimation results presented in Chapter 7 are unique. They are the first results based on a route choice model for a dynamic network with deterministic link attributes estimated based on RP data collected in a congested network for which link travel times are known. Furthermore, comparison between the predictive power of the RL and RL Dyn are carried out to make clear the added value of adding the time dimension to the route choice problem.

Chapter 7

Model estimation and prediction

This thesis investigates the role of travel information on travellers' route choice behaviour with the aid of a new modelling framework denominated Recursive Logit Dynamic (RL Dyn) (Chapter 4). The RL Dyn extends the Recursive Logit (RL) (Fosgerau et al., 2013) to a dynamic setting. Both modelling frameworks share the same properties and formulate the route choice problem as a sequence of link choice decisions based on dynamic discrete choice model (Rust, 1987). The added value of the RL Dyn is that it incorporates the variability of traffic conditions to the modelling framework, thus allowing the role of personalized real-time travel information to be better investigated.

This chapter focuses on model estimation and prediction. In particular, travellers' sensitivity to travel time in relation to travel information is investigated and comparisons between the RL and RL Dyn are carried out by discussing their predictive power. This helps clarifying the added value of adding the time dimension to the route choice problem time expanded.

Section 0 introduces the network and summarizes the sample used for model estimation. Section 7.2 discusses estimation results in the RL Dyn. These results are unique: the first ones based on revealed preference data collected in a congested network for which link travel times are known. Section 7.3 compares the performance of different model specifications and chooses the best one to be compared with the RL. In section 7.4 the RL and RL Dyn are compared and their predictive power discussed. Conclusions are drawn in section 7.5.

This chapter is based on:

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2012). Route choice model estimation in a dynamic network based on GPS data, *Proceedings of the 1st European Symposium on Quantitative Methods in Transportation Systems*, Lausanne, Switzerland.

Ramos, G. M., E. Frejinger, W. Daamen and S. P. Hoogendoorn (2014). Route choice model estimation in a dynamic network based on GPS data (*working paper*)

7.1 Network and observations

The sample size of the valid commuting trips is of 897 GPS traces and travel diaries. On average, a path is composed by 16 links, the minimum being 7 and the maximum 31. For estimation purposes, however, a subsample of 563 observations is taken into account. These are observations in which travellers only drove by links belonging to the road network as defined in section 5.3.5 and illustrated in Figure 5.7. Among the subsample of 563 observations, 246 observations refer to the initial 3 weeks of data collection in the reference treatment and 317 to the subsequent 6 weeks of data collection in the information treatment. It should be noted that due to the need to use a subsample of the valid trips, the number of observations in the information treatment is not much bigger than in the reference treatment.

As pointed out in section 4.2, the RL Dyn defines the modelling network as a directed connected graph $G = (L, N, T)$, where L is the set of links, N the set of nodes and T the set of time intervals. The transportation network under investigation consists of 520 links ($|L| = 520$), 200 nodes ($|N| = 200$) and 1080 time intervals ($|T| = 1080$).

The number of time intervals was defined based on the time periods investigated, i.e. morning and afternoon peak hours (from 07:00h to 10:00h and 16:00h to 19:00h), and on a 10 second size of time interval, which was the smallest travel time of a link. Around 7% of the links of the transportation network had such small travel time. By choosing the size of the time interval equal to the smallest travel time of a link, progression in time is ensured as discussed in section 4.2. The morning and afternoon peak hours were then split into 10 seconds intervals. As the data was collected at 1 minute interval, within each 1 minute interval and for each link the travel times were replicated 6 times. This resulted in $|T| = 1080$ time intervals (3 hours x 60 minutes x 6 intervals of 10 seconds).

The fact that the travel times are discretised has some computational implications and may also influence the estimation results as discussed in section 4.3. Due to the characteristics of the graph G , the size of the system of linear equations to be solved (\mathbf{M} matrix) is the order of 560.000. This has considerable computational implications as different \mathbf{M} matrices are defined for each time period. For instance, the size of each data file is around 60MB. Still, the direct solver of MATLAB can be used because the system is very sparse. Given that the duration of each time interval is 10 seconds only, significant rounding errors due to the discretization of time are not expected.

7.2 Estimation results in the RL Dyn framework

This section presents estimation results of four model specifications (MS) to answer research questions raised in Chapter 1: (i) the extent to which provision of travel information lead to (significant) changes in route choice behaviour, (ii) the role of the timing of travel information provision (pre-trip and en-route) on route choice behaviour and the role of different sources of travel information on route choice behaviour. Furthermore, it focuses on investigating the role of travel information and perceptions by means of interaction of these attributes with each other. This was done for different timing and sources of travel information, as well as for travellers subject to two different treatments of travel information provision. An overview of the attributes interacted with travel time in the different MS tested is depicted in Table 7.1. Other attributes also investigated, but not depicted in Table 7.1 are

link constant, habitual route and VMS link. Link constant is an attribute that represents the existence of decision points in the transportation network either to turn right/left or continue straight on. In practical terms it is a constant equal to one for all links and penalizes paths with many crossings. Note that a node in the transportation network does not necessarily represent a crossing as it depends on the topology of the network. As shown later on in this chapter it can represent, for instance a highway merge or exit. Habitual route is a personal characteristic. Thus, it is specific for each traveller and travelling direction, i.e. to or from work. The habitual routes were defined based on travellers' input during the interviews at the beginning of the experiment. Finally, VMS link refer to the fact of a link having a VMS on it (irrespective of travellers consulting it or not). This resulted in the investigation of 18 attributes.

Table 7.1: Overview of the attributes interacted with travel time in different models

	Considered categories				
Types of travellers	Informed travellers		Regular travellers		
Timing of travel information provision	Pre-trip	En-route	Pre-trip and En-route	No information	
Sources of pre-trip information	TomTom	Radio	More than 1 source	No pre-trip information	
Sources of en-route information	TomTom	Radio	VMS	More than 1 source	No en-route information

An overview of the estimation results is presented in Table 7.2. In the terminology used, the prefix *No* is used to indicate that travel information was not consulted. It can refer either to not consulting pre-trip, not consulting en-route information or not consulting travel information at all (both pre-trip or en-route). When travel information is consulted, a prefix *Pre* and/or *En* is used to identify the timing of the travel information provision. Furthermore, they are followed by the name of the source consulted. For the sake of simplicity, the notation of the utility functions is simplified along this section by omitting the subscripts of attributes already introduced. Discussion of estimation results is presented in sections 7.2.1 to 7.2.3.

7.2.1 To what extent does provision of travel information lead to (significant) changes in route choice behaviour?

(MS 1) 5 attributes: specific travel time attributes for informed and regular travellers, link constant, habitual route and VMS link. The deterministic component of the utility, $v(s|s)$, is defined by:

$$\begin{aligned} & \beta_{TTInformed} \cdot TTInformed_{s'|s} + \beta_{TTRegular} \cdot TTRegular_{s'|s} + \beta_{LC} \cdot LC_{s'|s} \\ & + \beta_{HabR} \cdot HabR_{(s'|s)|(s',s,s \in HabRoute)} + \beta_{VMS} \cdot VMS_{(s'|s)|(s' \in VMS)} \end{aligned} \quad (7.1)$$

where $TTInformed_{s'|s}$ and $TTRegular_{s'|s}$ are the travel times, in minutes, to go from link k to link a at instant t for informed and regular travellers respectively. $LC_{s'|s}$ and $HabR_{(s'|s)|(s',s \in HabRoute)}$ represents the habitual routes.

Model estimation in the RL(Dyn) assumes that the link utilities are additive. The utility of a path, therefore, is the sum of the link utilities. In order to define the instantaneous utility of each attribute of MS 1, (i) the travel time of each link is multiplied by a dummy variable representing informed/regular travellers, meaning that when a trip is made by an informed traveller $TTInformed_{s'|s} \neq 0$ and $TTRegular_{s'|s} = 0$ (and vice-versa), (ii) LC is defined as a constant equal to one for all possible states s' reached from s and (iii) $HabR$ is defined as a constant equal to one for all links of the habitual route of a specific observation. Note that when defining the state space of habitual routes, “ones” are put in the current link only if it is reached via a predecessor link that also belongs to the habitual route (exception is the 1st link). Strictly speaking this ensures that two consecutive links belong to the habitual route by creating a sort of “memory” of the previous choices. The \mathbf{M} matrix is defined as the sum of the instantaneous utilities of each attribute. Similar procedure is done for the other MS.

The difference in the magnitude of β_{TTInfo} and $\beta_{TTRegular}$ can be explained by the fact that the sensitivity to travel time of informed travellers is higher than that of regular travellers. For instance, informed travellers are about 30% more sensitive to travel time than regular travellers. Statistical analysis confirm that these values are significantly different from each other at 90% confidence level (t-test = -1.75).

Interpretation of β_{LC} indicates that it is worth approximately 1.4 minutes of an informed traveller and 1.7 minute of a regular traveller. Although this is somehow surprising as it indicates that an extra link is worth more than 1 minute, it can be explained by the characteristics of the transportation network. The transportation network under investigation is composed of a large number of long links and highways. In many cases a node in the network therefore does not represent a physical stop (to turn right, left or continue straight on), but a highway merge or exit. The fact that the links are long makes one extra link worth more than 1 minute of travel time.

The high significance of β_{HabR} indicates that habit plays an important role and travellers tend to repeat themselves and choose the same route based on habit. Its role is so strong that by ensuring that a link is part of the habitual route only if the previous link also is significantly increases the model fit in about 30 points. Besides this, travellers' sensitivity to travel time increases when habit is excluded from the model specifications. This suggests that travellers derive some comfort from choosing habitual routes.

Finally, with respect to β_{VMS} , quantitative analyses of the data sets indicate that travellers do not often consult the travel information displayed on the VMS panels (Chapter 6). Although estimation results suggest that links with VMS significantly influence travellers' route choice, this is not the case when separately investigating the effects of consulting and not consulting VMS travel information. Results show that despite the significance of the estimated parameters, these values are not significantly different from each other (t-test= 0.39). Therefore, it is possible to infer that the panels are very well located, but not much can be said regarding the use of travel information.

Table 7.2: Estimation results of different model specifications in the RL Dyn framework

Estimated parameters	Model Specifications			
	MS 1	MS 2	MS 3	MS 4
$\hat{\beta}_{TTInformed}$ (std. err) t-test	-0.73 (0.078) -9.34			
$\hat{\beta}_{TTRegular}$ (std. err) t-test	-0.56 (0.063) -8.91			
$\hat{\beta}_{LC}$ (std. err) t-test	-0.97 (0.037) -26.49	-0.92 (0.044) -20.98	-0.92 (0.043) -21.35	-0.93 (0.044) -21.16
$\hat{\beta}_{HabR}$ (std. err) t-test	0.57 (0.029) 19.30	0.61 (0.025) 24.20	0.61 (0.025) 24.17	0.61 (0.025) 24.12
$\hat{\beta}_{VMS}$ (std. err) t-test	0.97 (0.283) 3.42			
$\hat{\beta}_{TT_PreEn}$ (std. err) t-test		-0.62 (0.113) -5.47		
$\hat{\beta}_{TT_Pre}$ (std. err) t-test		-0.49 (0.098) -4.94		
$\hat{\beta}_{TT_En}$ (std. err) t-test		-0.66 (0.092) -7.22		
$\hat{\beta}_{TT_NoInfo}$ (std. err) t-test		-0.66 (0.074) -8.91		
$\hat{\beta}_{TTPreTomTom}$ (std. err) t-test			-0.83 (0.231) -3.61	
$\hat{\beta}_{TTPreRadio}$ (std. err) t-test			-0.32 (0.053) -6.08	
$\hat{\beta}_{TTMoreThan1SourcePreTrip}$ (std. err) t-test			-0.73 (0.166) -4.43	
$\hat{\beta}_{TTNoPreTrip}$ (std. err) t-test			-0.66 (0.060) -11.00	
$\hat{\beta}_{TTEnTomTom}$ (std. err) t-test				-0.74 (0.136) -5.42
$\hat{\beta}_{TTEnVMS}$ (std. err) t-test				-0.72 (0.227) -3.19
$\hat{\beta}_{TTEnRadio}$ (std. err) t-test				-0.40 (0.094) -4.30
$\hat{\beta}_{TTMoreThan1SourceEnRoute}$ (std. err) t-test				-0.75 (0.149) -5.02
$\hat{\beta}_{TTNoEnRoute}$ (std. err) t-test				-0.63 (0.065) -9.74
LL ($\hat{\beta}$):	-628.11	-643.88	-640.39	-641.55

7.2.2 How does the timing of travel information provision (pre-trip and en-route) influence route choice behaviour?

(MS 2) 6 attributes: specific travel time attribute for trips in which both pre-trip and en-route travel information were consulted, only pre-trip was consulted, only en-route information was consulted and travel information was not consulted at all, link constant and habitual route. This specification aims at simultaneously investigating the role of pre-trip and en-route travel information. The deterministic component of the utility is defined by:

$$\beta_{TT_PreEn} \cdot TT + \beta_{TT_Pre} \cdot TT + \beta_{TT_En} \cdot TT + \beta_{TT_NoInfo} \cdot TT + \beta_{LC} \cdot LC + \beta_{HabR} \cdot HabR \quad (7.2)$$

Outcomes of MS 2 indicate that the sensitivity to travel time of travellers consulting pre-trip information only is lower and significantly different from those who do not consult travel information at all at 85% confidence level (t-test=1.45 for β_{TT_Pre} and β_{TT_NoInfo}). As pre-trip information is provided before the trip starts, the awareness of possible delays might comfort travellers in the sense that at least they know what to expect in terms of traffic conditions. Outcomes of quantitative analyses (Chapter 6) suggest that objective variations of travel time that have already been incorporate by travellers should not be considered unpredictable. The awareness of travel time variations may comfort travellers in the sense that at least they are not surprised by the delays, they know how long it may take, etc. Results suggest that knowing this before the trip starts seems to be more important than during the trip.

7.2.3 How do different sources of travel information influence route choice behaviour?

(MS 3) 6 attributes: specific travel time attribute for travellers consulting only pre-trip TomTom information, only pre-trip radio information, more than one source of pre-trip travel information and not consulting pre-trip information, link constant and habitual route. No restriction was imposed regarding the sources consulted. In only 12 out 563 observations a source of pre-trip information different from radio or TomTom was consulted. Consulting pre-trip information is not an obstacle to also consult en-route information. The deterministic component of the utility is defined by:

$$\begin{aligned} &\beta_{TTPreTomTom} \cdot TT + \beta_{TTPreRadio} \cdot TT + \beta_{TTMoreThan1SourcePreTrip} \cdot TT + \beta_{TTNoPreTrip} \cdot TT \\ &+ \beta_{LC} \cdot LC + \beta_{HabR} \cdot HabR \end{aligned} \quad (7.3)$$

(MS 4) 7 attributes: specific travel time attribute for travellers consulting only en-route TomTom information, only en-route VMS information, only en-route radio information, more than one source of en-route information and not consulting en-route information, link constant and habitual route. Only in 4 out 563 observations a source of en-route information different from TomTom, VMS or radio was consulted. The deterministic component of the utility is defined by:

$$\begin{aligned} &\beta_{TTEnTomTom} \cdot TT + \beta_{TTVMS} \cdot TT + \beta_{TTEnRadio} \cdot TT + \beta_{TTMoreThan1SourceEnRoute} \cdot TT \\ &+ \beta_{TTNoEnRoute} \cdot TT + \beta_{LC} \cdot LC + \beta_{HabR} \cdot HabR \end{aligned} \quad (7.4)$$

Outcomes of MS 3 and MS 4 show that the sensitivity to travel information of travellers who consult very detailed information is higher than otherwise. In particular, both for pre-trip and en-route, travellers who consult TomTom are more sensitivity to travel time than those who consult radio (t-test=-2.17 for $\beta_{TTPreTomTom}$ and $\beta_{TTPreRadio}$ t-test=-2.07 for $\beta_{TTEnTomTom}$ and

$\beta_{TTEnRadio}$). Besides this, only travellers who consult radio are less sensitive to travel time variance than those who do not consult travel information at all (the other parameter combination are not significantly different from each other). The travel information provided via radio is very generic as (in general) only queue lengths of some major roads are provided. Nevertheless, travellers can infer the degree of severity of congestion from the information provided. This may provide some sort of comfort for being aware of whether the congestion is a regular congestion or not. For the other sources, the information provided and experience appear to have similar effect on travellers' sensitivity to travel time.

7.3 Comparison of model performances

The previous section tested different model specifications to discuss the role of travel information on route choice behaviour. This section compares the performance of the estimated models by conducting statistical tests.

In general, there are two main tests to evaluate model performances: likelihood ratio test and non-nested hypothesis test. The likelihood ratio test is a statistical test to compare model specifications in which one is a (un)restricted version of the other. In other words, there is a null model and there is an alternative model, and the null model is a special case of the alternative model. In case the two models to be compared do not have this restricted-unrestricted relationship, a non-nested hypothesis test is the usual alternative.

Although the non-nested hypothesis test makes it possible to compare rival model specifications, such type of test cannot be performed in the RL Dyn (or RL) framework. As discussed in Fosgerau et al. (2013), model estimation in the RL framework requires definition of initial values of the parameter to be estimated. As result, neither the initial null log likelihood nor the adjusted rho-square are available. Without these values it is not possible to test the hypothesis that the model with the lower adjusted rho-square is the true model. Besides this, a composite model is not always possible to be built as in our case it would result counting travel time at least twice. Therefore, only the performance of models in which one is a (un)restricted version of another can be compared with the aid of likelihood ratio tests.

The likelihood ratio test expresses how many more times the data is likely to be under one model specification rather than under the other model specification. Depending on the value of the likelihood ratio in relation to a critical value, one can decide whether to reject the null model in favour of the alternative model. The log likelihood ratio test checks whether $-2(LLR-LLU) > \chi^2_{((1-\alpha),df)}$, where LLR is the final log likelihood of the restricted model, LLU is the final log likelihood of the unrestricted model and $\chi^2_{((1-\alpha),df)}$ is the critical chi square value for significance level α and df degrees of freedom.

The restricted version of each model assumes that travel time is similarly perceived irrespective of the type of traveller, timing and sources of travel information provision. The way this was considered for the MS tested is depicted in each row of Table 7.1. The likelihood ratio tests indicate that MS 1 is significantly better than the others and consequently fits the data better. When comparing it with a simplified version that disregards that informed and regular travellers value travel time differently, there is a significant improvement in the model fit at 0.1 significance level. Similarly, MS 3 is significant at 0.05.

This means that it is possible to reject the null hypotheses that the sensitivity of travellers to travel time is the same irrespective of the source of pre-trip travel information consulted. MS 2 and MS 4, however, fail even at 0.15 when compared to their respective restricted versions. This means that it is neither possible to reject the null hypothesis that travellers' sensitivity to travel time is the same irrespective of the timing (pre-trip or/and en-route) of travel information provision nor that it is the same with respect to the sources of en-route information.

7.4 Comparison of static and dynamic networks

As already discussed, the RL Dyn is computationally too intense. For a small size network such as the one of this thesis, the size of the system of linear equations to be solved is the order of 560.000. Besides this, as the input files are also very big, it demands a lot of storage and RAM memory from the computers.

Statistical comparison between the RL and RL Dyn is not possible because they are based on different definitions of the modelling network/state space. In order to investigate the added value of adding the time dimension to the route choice problem, therefore, it is necessary to look into predictions. A path is defined as a sequence of states and the probability of a path observation is the product of the probability of each state. While in the RL a state is defined by a link, in the RL Dyn it is defined by a pair link-time interval. In other words, the network in the RL Dyn is time-expanded. The question is whether the RL Dyn leads to more accurate predictions despite its much bigger state space. Besides this, what are the behavioural implications? Sections 7.4.1 and 7.4.2 discuss these two aspects.

7.4.1 Estimation results in the RL framework

Among the model specifications tested, MS 1 fits the data better and for this reason is chosen for comparisons between the RL and RL Dyn. As the RL does not consider the variability of traffic conditions in the network, the travel times are treated as if static irrespective of the departure times. As a result, the way of defining the travel times appears to be an important aspect when applying the RL. Should the travel times be averaged over the day, be time-period specific (morning and afternoon peak hours), be specific for the day of the week and time-period, etc.? For investigation purposes, the following three measures of travel time were considered as travel time of a link:

- AVG Day: Average travel times of the whole day based on all days of the experiment;
- AVG Peak: Average travel times based either on the morning or afternoon peak hours of all days of the experiment. Two static state spaces, one for the morning peak and another for the afternoon peak, are defined;
- AVG Date & Peak: Average travel times based either on the morning or afternoon peak hours for each day of the experiment. For each observation there is a correspondent state space defined depending on the date and time-period of the commuting trip.

Estimation results in the RL framework are depicted in Table 7.3. The estimated parameters are significant and have the expected sign. Besides this, all of them are significantly different from each other with exception of the pair habit-VMS. Moreover, with exception of β_{VMS} , the magnitude of each parameter does not vary much within the different measures of travel time. Interpretation of estimates indicates that informed travellers are about 2.4 times more

sensitive to travel time than regular travellers; a link constant is worth approximately 0.75 minute; and habit is worth approximately 0.4 minutes for informed travellers and 0.9 minute for regular travellers.

Table 7.3: Estimation results of model specification 1 (MS 1) in the RL framework

Estimated Parameters	AVG Day	AVG Peak	AVG Date & Peak
$\hat{\beta}_{TTInformed}$ (std. err)	-1.65 (0.162)	-1.64 (0.166)	-1.62 (0.143)
t-test	-10.14	-9.89	-11.37
$\hat{\beta}_{TTRegular}$ (std. err)	-0.69 (0.032)	-0.70 (0.032)	-0.68 (0.032)
t-test	-21.79	-22.22	-21.41
$\hat{\beta}_{LC}$ (std. err)	-1.25 (0.104)	-1.23 (0.102)	-1.24 (0.100)
t-test	-12.07	-11.97	-12.31
$\hat{\beta}_{HabR}$ (std. err)	0.60 (0.044)	0.61 (0.044)	0.63 (0.042)
t-test	13.88	13.93	14.76
$\hat{\beta}_{VMS}$ (std. err)	0.83 (0.140)	0.72 (0.142)	1.02 (0.190)
t-test	5.94	5.11	7.35
$LL(\hat{\beta})$	-708.43	-710.83	-713.66

A logical question that follows is how the estimation results in the RL framework compares to those of the RL Dyn. In other words, what the behavioural implications are. Table 7.4 presents the elasticities of the parameters in relation to travel time for regular and informed travellers. Comparison of elasticities indicates a big difference between the RL and RL Dyn. For instance, while in the RL Dyn informed travellers are about 30% more sensitive to travel time than regular travellers, they are about 140% more sensitive in the RL framework. The relationship between travel time and link constant also leads to different conclusions. While in the RL a link constant is worth less than 1 minute for informed travellers, the contrary applies in the RL Dyn. Behavioural implications are therefore quite different as in one case it appears that an extra link is more negatively perceived than 1 extra travel time minute. These results indicate that incorporating (or not) travel time variability to the modelling framework indeed plays a big role.

A key question to be answered is which of the two modelling frameworks better represents travellers' route choice behaviour. Moreover, does adding the time dimension leads to better predictions? The next section answers these questions by discussing the likelihood of path observation through cross-validation techniques as well as the predictive power of both modelling frameworks.

Table 7.4: Ratios of pairs of parameters

Parameter ratios	RL Dyn	RL (AVG Day)	RL (AVG Peak)	RL (AVG Date & Peak)
$\hat{\beta}_{TTInformed} / \hat{\beta}_{TTRegular}$	1.30	2.39	2.33	2.40
$\hat{\beta}_{LC} / \hat{\beta}_{TTRegular}$	1.72	1.82	1.75	1.82
$\hat{\beta}_{HabR} / \hat{\beta}_{TTRegular}$	-1.01	-0.88	-0.86	-0.93
$\hat{\beta}_{VMS} / \hat{\beta}_{TTRegular}$	-1.72	-1.20	-1.03	-1.52
$\hat{\beta}_{TTRegular} / \hat{\beta}_{TTInformed}$	0.77	0.42	0.43	0.42
$\hat{\beta}_{LC} / \hat{\beta}_{TTInformed}$	1.32	0.76	0.75	0.76
$\hat{\beta}_{HabR} / \hat{\beta}_{TTInformed}$	-0.78	-0.37	-0.37	-0.39
$\hat{\beta}_{VMS} / \hat{\beta}_{TTInformed}$	-1.32	-0.50	-0.44	-0.63

7.4.2 Prediction results

In order to investigate the predictive power of both modelling frameworks, a cross-validation technique of repeated random sub-sampling was performed. MS 1 was re-estimated based on 80% of the sample taking stratification for informed and regular travellers into account. The remaining 20% of the sample was used for model prediction. This means that all models were re-estimated based on 450 random observations and the estimation results were used to predict the probability of the remaining 113 path observations. Note that all models were applied to the same samples. This procedure was repeated 20 times to validate the results. Although we do not present these estimation results, the estimated parameters were significant and had the expected sign for all subsamples and models estimated. Among the 520 links of the transportation network, circa 36% are connected to one link only, 48% to 2 links and 17% to 3 links. These percentages disregard the possibility of U-turns. Moreover, on average, a path observation is composed by 16 links/states, the minimum being 7 and the maximum 31.

On average, a realized path has 62.3% probability of being correctly predicted in the RL Dyn as opposed to 57.4% in the RL AVG Day, 57.1% in the RL AVG Peak and 57.2% in the RL AVG Date & Peak. These, however, are average values and just give an indication of the performance of each modelling framework. For instance, the standard deviation of the probabilities to which a path is correctly predicted is about 35%, which indicates the existence of outliers. These outliers substantially influence the average predictive power. A thorough investigation shows that in about 50% of the cases the probability of a path observation in the RL Dyn framework was higher than in any of the RL frameworks. The question, however, is not solely if the predictive power is high (or small), but in which situations one is better off/safer using the simpler approach of the RL or the more complex of the RL Dyn. Situations in which there was a big difference in the predictive power of both modelling frameworks were investigated. This considered outliers for which the differences (delta) in the predicted probability of a path observation was $|\text{delta}| > 20\%$. In such situations, RL Dyn predicted better on average in 18 cases and worse in 6 cases (for each of the 20 random samples). Detailed

investigation of the outliers shows that the situations in which the RL Dyn outperformed the RL involve high variability of traffic conditions and the contrary applies otherwise.

In order to assess the best modelling framework and consequently identify which better represents behaviour, the log likelihood of a path was computed. The log likelihood of a path observation was defined as $\sum_{n=1}^N \ln(\Pr(obs_n))$, where N is the total number of path observations.

Figure 7.1 depicts the predicted log likelihood for each model and each of the 20 random samples of 113 observations. Results are ordered based on decreasing values of predicted log likelihood in the RL Dyn framework.

Figure 7.1 shows that the predicted log likelihood in the RL Dyn is overall much higher than that in the RL framework (irrespective of the measure of travel time). In particular, in 18 out of 20 trials the predicted log likelihood was higher in the dynamic case. These outcomes indicate that despite the much bigger state space of the RL Dyn, adding the time dimension to the modelling framework substantially improves the predicted log likelihood of a path. This supports the added value of having the network time expanded and consequently compensates the computational demand. Furthermore, the higher predicted log likelihood of the RL Dyn indicates that the estimated parameter values under this framework are the ones that better represents behaviour.

Despite the different measures of travel time within the RL framework, the predicted log likelihoods among them are very similar. Therefore, it does not seem to matter what average measure of travel time is used in the static version. This has positive implications for model estimation and prediction in the RL framework: average travel times are the easiest measure of travel time to be collected and the results indicate that this is a good way of defining travel times in the static case. The data collection procedure, therefore, can be substantially simplified by using historical data. Although this raises questions regarding the frequency to which new expectations are updated or created, which is a point to be further investigated, it makes clear that irrespective of the measure of travel time, the dynamic version is clearly better for predictions.

Finally, it is important to highlight that although model estimation in the RL Dyn is too computationally demanding, estimation is done only once. Afterwards, the model is used for predictions. Based on a windows 64-bit operating system, processor Intel ® core™ i5-2400 and 12GB of memory RAM, it takes around 11 minutes to predict the probability of 113 path observations considering the characteristics of transportation network investigated. Conversely, model prediction in the RL takes around 5 seconds only. The longer computational time for model prediction in the RL Dyn, however, is due to the size of the input data files. This requires time to read the files (and also storage capacity) from the computers. Nevertheless, the computational complexity for model prediction is essentially the same: disregarding the computational time to read the input data, model prediction in the RL Dyn takes around 90 seconds. Thus, one is dealing with computational/memory issues that can be overcome by using a server with high memory capacity.

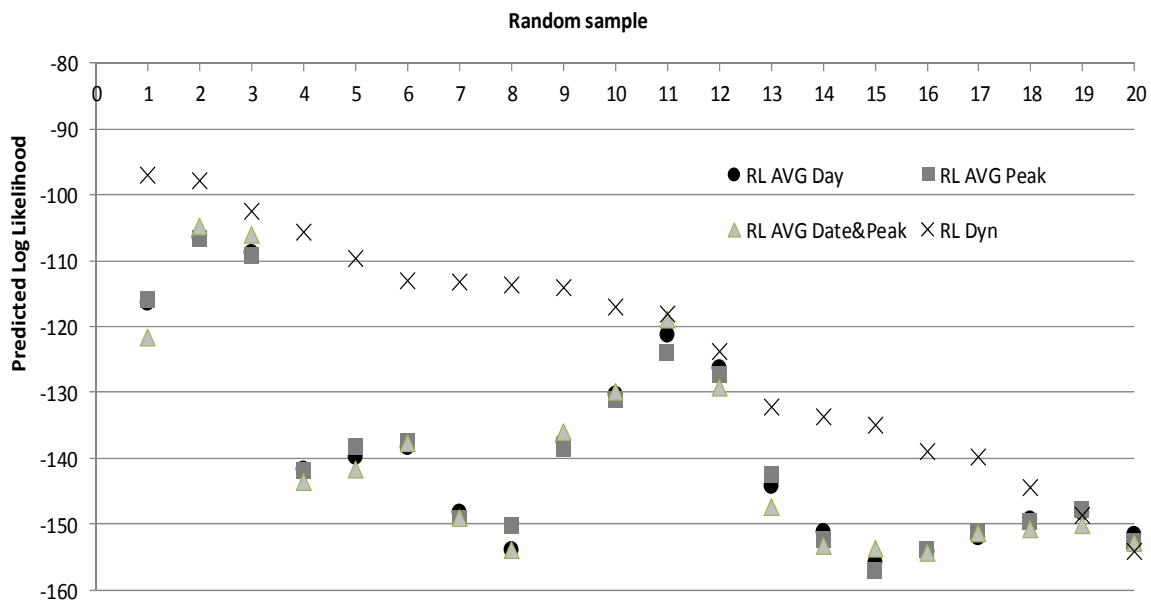


Figure 7.1: Predicted log likelihood for different models and subsamples

7.5 Conclusions

This chapter further investigates the role of travel information on travellers' route choice behaviour by means of model estimation and prediction. To the best of our knowledge, these are the first estimation results based on revealed preference data collected in a congested network for which link travel times are known. Model estimation is done in a new modelling framework denominated Recursive Logit Dynamic (RL Dyn), which builds up from the Recursive Logit (RL) (Fosgerau et al., 2013). Furthermore, discussion about their predictive power helps clarifying the added value of adding the time dimension to the route choice problem. While the RL formulates the route choice problem as sequential link-choice decisions based on path observations without requiring any definition of choice sets, the RL Dyn extends it to a dynamic network by adding the time dimension to the route choice problem.

Different model specifications were tested to answer the following research questions:

- To what extent does provision of travel information lead to (significant) changes in route choice behaviour?
- How do different sources of travel information influence route choice behaviour?
- How does the timing of travel information provision (pre-trip and en-route) influence route choice behaviour?

Regarding changes in behaviour, estimation results show that informed and regular travellers value travel time differently. For instance, informed travellers are about 30% more sensitive to travel time than regular travellers. Statistical analyses confirm that these values are significantly different from each other at 90% confidence level ($t\text{-test} = -1.75$).

With respect to different sources of travel information, estimation results indicate that the more detailed they are, the more travellers are sensitive to it. In particular, travellers who consult TomTom are more sensitive to travel time than those who consult radio both for pre-trip and en-route information (t-test=-2.17 for $\beta_{ITPreTomTom}$ and $\beta_{ITPreRadio}$; t-test=-2.07 for $\beta_{ITEnTomTom}$ and $\beta_{ITEnRadio}$). Conclusions regarding the role of variable message signs (VMS) should be drawn carefully. Although estimation results indicate that links with VMS significantly influences route choice behaviour, it appears that consulting or not consulting VMS similarly influence travellers' decisions. This suggests that the VMS panels are well located, but not much can be said regarding the use of travel information.

Finally, with respect to the timing of travel information provision, estimation results show that travellers who consult pre-trip information only are less sensitive to travel time than travellers who do not consult information at all. However, not much can be said regarding the role of en-route information. As pre-trip information is provided at the beginning of the trip, being aware of the traffic conditions may provide some comfort to travellers. As reported by the participants in the travel diaries, they tend to stick to their habitual routes even when encountering longer than expected delays. Estimation results confirm this: habit turned out to be a highly significant parameter in all model specifications.

Among the different model specifications tested, the one the better fits the data is MS1 that considers 5 attributes: specific travel time attributes for informed and regular travellers, link constant, habitual route and VMS link. In order words, such model specification should be preferred to model travellers' behaviour. Estimation results of such model specification indicate that regular and informed travellers value travel time differently and that habit has a positive significant strong effect in route choice decisions. Besides this, although not much can be said regarding the use of travel information displayed in VMS panels, it appears that travellers regularly drive through such links indicating that the panels are well located.

As MS 1 better fits the data, such model specification was used to compare the RL and RL Dyn. The RL and RL Dyn define the network differently and because of this a statistical comparison between them is not possible. Comparisons, however, can be done by discussing behavioural implications from ratios of pair of parameters and their predictive power. Outcomes of the ratio of pair of parameters show that while the sensitivity to travel time of informed travellers is about 30% higher than that of regular travellers in the RL Dyn, it increases to 140% in the RL framework. Interpretation of a link constant is also quite different: in the RL framework it is worth less than 1 minute and the contrary applies in the RL Dyn framework. Investigation of the predicted log likelihood of a path observation show that despite the much bigger state space of the RL Dyn, its predictive power is substantially higher than that of the RL framework. This was the case for 18 out 20 trials each with 113 random samples. These outcomes, therefore, indicate that adding the time dimension to the route choice problem compensates the higher complexity and computational demand for model estimation. Furthermore, this implies that the ratio of parameters of the RL Dyn provide more appropriate behavioural interpretation of travellers' behaviour. More experiments, nevertheless, should be performed to value the merits of these findings and to further investigate to what extent the physical characteristics of the transportation network and/or variability of traffic conditions influence the outcomes.

It is important to highlight that results indicate that under high variability of traffic conditions, the predictive power of the RL Dyn is much higher than that of the RL. These are the type of situations in which provision of real-time traffic information is likely to be most useful due to its potential to help travellers avoiding congestion by choosing a different route. Therefore, the RL Dyn seems to be the most appropriate to such type of investigations. Nevertheless, the RL is a very good approach for situations of lower variability of traffic conditions or for which dynamic traffic data is not available. The fact that different definitions of travel time measures within the RL framework similarly impact the predicted log likelihood has positive implications from the point of view of data collection: it suggests that the use of historical data is a good starting point.

Finally, in terms of computational demand, for the network under investigation and model specifications tested, the maximum estimation time in the RL framework was about 10 minutes as opposed to more than 8 hours in the RL Dyn. This, however, has to be done only once in order to obtain the estimates. Afterwards, the model is used only for predictions. Based on a windows 64-bit operating system, processor Intel ® core™ i5-2400 and 12GB of memory RAM it takes around 11 minutes to predict the probability of 113 path observations considering the characteristics of transportation network investigated. Conversely, model prediction in the RL takes around 5 seconds only. The longer computational time for model prediction in the RL Dyn is due to the big data files that have to be read. In terms of computational complexity for model prediction, however, both frameworks are essentially the same: model prediction in the RL Dyn takes around 90 seconds when the time to read the input data is not taken into account. Thus, one is dealing with computational/memory issues that can be overcome by using a server with high memory capacity.

The application of the RL Dyn still faces some challenges, such as the possibility of considering travel time as a stochastic attribute and the need to correct for panel data. Having several observations for the same traveller violates the assumption that the error term is *i.i.d* and requires correction for panel data. As discussed, for instance, in Daly and Hess (2010) and Train (2009), naïve estimates, that is, estimates that are based on the incorrect assumption that the error term is *i.i.d.*, will still provide consistent estimation results, but biased standard errors, as long as the correlation in errors is related to the observed attributes.

Louviere and Woodworth (1983) discuss that a very conservative and simple way of checking the significance of naïve estimates is to multiply the standard error by the squared root of the number of repetitions. The number of trips made by the participants during the experiment, however, was not the same. For instance, the average number of valid trips per traveller was 16, but some travellers had more than 40 and some less than 5. Based on the squared root of the average number of repetitions, $|t\text{-test}| > 5.94$ are significant at 85% confidence level, $|t\text{-test}| > 6.69$ are significant at 90% confidence level and $|t\text{-test}| > 7.76$ are significant at 95% confidence level. This means that with exception of β_{MS} , all attributes of MS 1 are statistically significant.

Limitations in logit-based models arise when the model attempts to allow tastes to vary with respect to unobserved attributes or purely randomly. Since according to Train (2009) the logit formula seems to be fairly robust to misspecifications, such models might still be able to capture average tastes fairly well even when tastes are random. Applications for forecasting and identification of market shares of new products, for instance, might lead to biased/wrong

conclusions. Options to properly correct for panel data are re-sampling methods such as the “Bootstrap” and the “Jack-knife” which try to measure the variation of the estimates as the data changes. Model estimation in the RL Dyn framework, however, is too computationally demanding to use such procedures. Depending on the number of parameters to be estimated, the estimation time takes more than 8 hours. Considering the number of re-samples needed, the estimation time substantially increases. Although there are other alternatives to correct for panel data, these were also not applied for several other reasons. Correction for panel data, therefore, is left for future research.

Besides this, in terms of empirical contributions regarding the role of travel information on travellers’ route choice behaviour, future research should contemplate a more detailed network (and hence more observations) and try to include the information provided itself. This would allow identifying what the informed links were. It is acknowledged, however, that this is a rather challenging task as it depends of the availability of data from several information providers.

The next chapter summarizes the main findings of this PhD thesis. Moreover, it presents conclusions, implications for practice and future research directions.

Chapter 8

Conclusions

Traffic congestion is worldwide experienced by a great number of travellers driving during peak hours. In order to try to solve problems that arise from congestion, such as delays, uncertainty and environmental issues, there has been an increased interest in the role of travel information. How travellers' react to travel information and its consequences regarding the level of congestion in the transportation network raised the main research question "What is the impact of travel time information on travellers' route choice behaviour?", which was addressed and answered in this PhD thesis.

In order to answer the main (and other complementary) research question(s), a dedicated revealed preference (RP) experiment was conducted. It consisted of tracking commuters to and from work with GPS devices under different treatments of travel information provision. Travellers were then investigated in a real setting, while making their route choice decisions and immediately experiencing the consequences of their choices. On top of that, travel diaries were filled in after each trip, interviews were conducted at three different occasions and data about the traffic conditions in the network was collected. The characteristics of the data collection made it unique and resulted in a very rich data set.

The RP experiment was conducted during nine weeks, from May 9th to July 12th, 2011, in the Netherlands. Thirty two travellers were tracked and their behaviour in relation to travel information was investigated. During the initial three weeks of data collection, travellers were allowed to consult available public sources of travel information; during the subsequent six weeks, on top of the public sources of travel information, 80% of the travellers were equipped with a personalized tool of provision of real-time traffic information via TomTom (type TomTom VIA Live 120 Europe). The initial three weeks of data collection was denominated reference treatment and the subsequent six weeks was denominated information treatment. Travellers equipped with TomTom were labelled informed travellers. Otherwise, they were labelled regular travellers. The resulting sample size consists of 897 valid GPS

traces¹⁵ and travel diaries (374 observations during the reference treatment and 523 during the information treatment) and 32 x 3 interviews (one per participant for each of the three different occasions they were interviewed).

The fact that travellers' behaviour was investigated in a real setting has three main implications for the estimation of route choice utility models. Firstly, choice set sampling is difficult (Bovy, 2009; Prato, 2009; Frejinger and Bierlaire, 2010). Secondly, utilities may share unobserved attributes and consequently the IIA property (independence from the irrelevant alternatives) does not hold making the Multinomial Logit (MNL) models unsuitable to be directly applied (e.g. Cascetta et al., 1996; Ben-Akiva and Bierlaire, 1999; Bekhor et al., 2001; Frejinger and Bierlaire, 2007). Thirdly, the attributes of the alternatives (e.g. travel times at the time of the observation) are in general unknown. Dealing with these three issues is more complex in dynamic than static networks (Gao et al., 2008). This thesis addresses two of these three main modelling issues: choice set sampling and attributes of the alternatives.

In order to overcome issues regarding choice set definition, model estimation was done in a new modelling framework denominated Recursive Logit Dynamic (RL Dyn). The RL Dyn is based on the Recursive Logit (RL) (Fosgerau et al., 2013), which formulates the route choice problem as a sequence of link choice decisions based on dynamic discrete choice model (Rust, 1987). The RL, therefore, does not require choice definition. The RL Dyn goes one step further by incorporating the variability of traffic conditions to the modelling framework. As a result, it is a proper framework to estimate discrete choice random utility models based on RP data and to investigate the role of real-time travel information. The issue of attributes of the alternatives was overcome by collecting data about the traffic conditions in the network during the whole study period.

This thesis, therefore, has two main contributions: an empirical and a methodological. The empirical contribution is to investigate the role of travel information on travellers' route choice behaviour based on RP data. The methodological is to address for the first time issues regarding choice set sampling in a dynamic setting supported by real data about the traffic conditions in the network. Moreover, comparisons between the RL Dyn and RL are carried out to clarify the added value of adding the time dimension to the route choice problem.

The main findings of this PhD thesis are summarized in section 8.1, methodological contributions in section 8.2 and conclusions in section 8.3. The conclusions drawn are based on the main findings, which, by definition, are not specific for a case but have a generic character for similar situations. Then, implications for practice are discussed in section 8.4. Finalizing, possible future research directions are presented in section 8.5.

¹⁵ For estimation purposes a subsample of 563 observations was considered

8.1 Empirical findings

The main empirical findings of this thesis are split in two categories: role of travel information on route choice behaviour and travellers' perceptions. Each of them is respectively discussed in sections 8.1.1 and 8.1.2. Furthermore, methodological contribution is discussed in section

8.1.1 Empirical findings about the role of travel information

The role of travel information on travellers' behaviour was addressed both from model estimation and qualitative and quantitative perspectives. Focus was put on travellers' reactions to different sources and timing (pre-trip and en-route) of travel information provision, compliance with travel information, willingness to pay for travel information and behavioural changes due to travel information. This was done to answer the following research questions raised in Chapter 1:

To what extent does provision of travel information lead to (significant) changes in route choice behaviour?

Estimation results show that informed and regular travellers value travel time differently. For instance, informed travellers are about 30% more sensitive to travel time than regular travellers. Statistical analyses confirm that these values are significantly different from each other at 90% confidence level (t-test = -1.75).

It is important to highlight the importance of habit in commuters' behaviour. It turned out to be a highly significant parameter in all model specifications tested ($\hat{\beta}_{HabR} \sim 0.60$ and t-test ~ 24). Participants also reported that even when incurring delays longer than expected, they prefer to stick to their preferred planned routes and in hindsight would still choose the same route. This implies that in order to change the behaviour of commuters either the penalty associated with a late arrival or the benefit for on time/earlier arrival should be or perceived as higher.

How do different sources of travel information affect route choice behaviour?

Estimation results indicate that the more detailed the travel information is, the more travellers are sensitive to it. Both for pre-trip and en-route information, travellers who consult TomTom are significantly more sensitive to travel time than those who consult radio ($\hat{\beta}_{TTPreTomTom} = -0.83$, $\hat{\beta}_{TTPreRadio} = -0.32$, $\hat{\beta}_{TTEnTomTom} = -0.74$ and $\hat{\beta}_{TTEnRadio} = -0.40$) at 0.05. Conclusions regarding the role of VMS should be drawn carefully. Although estimation results indicate that road links with VMS panels significantly influences route choice behaviour, it appears that consulting or not consulting VMS similarly influence travellers' decisions. This suggests that the VMS panels are well located, but not much can be said regarding the use of travel information.

Outcomes of the quantitative and qualitative analyses indicate that equipping travellers with TomTom devices increases the likelihood that they will use travel information to assist their route choices. Irrespective of the timing of travel information provision (pre-trip or en-route), when TomTom information is available, it turns out to be the main used source. Among the public sources of travel information, radio appears to be the preferred source of pre-trip

information both in the information and reference treatments. The use of en-route public information, however, depends on the treatment of information provision: radio is preferred during the reference treatment and VMS is preferred otherwise.

How does the timing of travel information provision (pre-trip and en-route) affect route choice behaviour?

With respect to the timing of travel information provision (i.e. pre-trip or en-route), estimation results suggest that travellers who only consult pre-trip information are significantly less sensitive to travel time than travellers who do not consult information at all ($\hat{\beta}_{TT_Pre} = -0.49$ and $\hat{\beta}_{TT_NoInfo} = -0.66$; t-test for $\hat{\beta}_{TT_Pre}$ and $\hat{\beta}_{TT_NoInfo}$) at 0.15. Not much can be said regarding the role of en-route information. As pre-trip information is provided at the beginning of the trip, being aware of the traffic conditions may provide some sort of comfort to travellers.

Outcomes of the quantitative and qualitative analyses suggest that finding out the best route choice considering the departure time is “now” appears to be more important than planning an “ideal” departure time to avoid congestion. As reported by the participants, they usually consulted travel information just when entering the car and not beforehand while still at home or in the office. The higher use of en-route information in relation to pre-trip information for informed travellers is somehow surprising as the contrary has been shown in the literature (Abdel-Aty et al., 1997; Jou, 2001). It has to be noted, however, that the outcomes of this research are based on RP data while those reported on the literature are based on SP (stated preference) data. It is then possible that travellers overestimated their use of travel information in the hypothetical situation.

To what extent are travellers willing to comply with different sources of travel information provision (over time)?

Definition of compliance was presented just before the experiment started. Travellers were informed that following the advice irrespective of having already decided to choose a specific route should be treated as being compliant with the advice. The same applied to the information displayed by VMS panels and provided via radio regarding least congested routes.

Outcomes of the quantitative and qualitative analyses show that (in general) the more detailed the source of the travel information is, the higher the compliance rates are. Besides this, compliance rates also appear to differ for informed and regular travellers: informed travellers are more willing to comply with travel information than regular travellers. It is interesting to observe that 77% of the informed travellers reported on the interviews that the route suggested by TomTom differed from their planned route less than 30% of the times. Among this group, 95% reported to have complied with the information over 70% of the times. For the one informed traveller who reported that the route suggested by TomTom differed from expectations over 70% of the times, compliance rates with travel information reduced to less than 30% of the times.

To what extent are travellers willing to comply with different timing of travel information provision, i.e. pre-trip or en-route (over time)?

Outcomes of the quantitative and qualitative analyses show that travellers are more willing to comply with pre-trip, rather than en-route, information. This might be influenced by (i) the fact that changing routes while en-route would imply taking local roads, (ii) habit or (iii) travellers' misperception regarding compliance.

Taking local roads may be associated with the idea that the resultant travel time-saving is not so large, to lower speed limits and to higher amount of stops due to traffic lights and/or crossing streets. Habit also plays an important role as extensively discussed in this thesis. Due to habit travellers may prefer to stick to the habitual routes due to the familiarity with the environment and already feel satisfied to be aware of delays. Finally, travellers may misperceive their actual compliance when the travel information is aligned to their expectations. In such situations, although they are actually "complying with" the travel information, the travel information may be perceived as not necessary/not needed. Consequently, travellers may feel that they are not complying with it. For instance, the information favoured travellers' planned routes in about 88% of the trips in which travellers complied either with pre-trip or en-route information.

To what extent are travellers willing to pay to receive travel information (over time)?

Travellers' willingness to pay for travel information increases with their experience using travel information (or stays about the same). In particular, travellers are more willing to pay for travel time rather than queue information and travel times in all (main) routes appear to be the type that travellers are more likely to pay for. This suggests that travellers like to be in control and prefer to assess the situation and decide themselves what route to choose.

The willingness to pay for travel information also varies among travellers. For instance, informed travellers tend to be more willing to pay for travel information (irrespective of the type) over time. This suggests that travellers derive more satisfaction from private/personalized sources of travel information rather than from public source. This is possibly due to anticipation of the potential benefits of using TomTom in situations other than the commuting trips, which might not have been the case for regular travellers who used only public sources.

Despite the increase in travellers' willingness to pay for travel information over time, especially informed travellers, without much surprise, the rates are very low and travellers are at most neutral about it. This behaviour is in agreement with the patterns of use and compliance with travel information. As travellers did not often use travel information that was provided for free, it was expected that they would not be willing to pay for it. According to the participants, travel information is more relevant for trips in which they are not familiar with or for longer trips rather than the daily commuting trips. For instance, with exception of travel time in all routes, more than 50% of the travellers reported to be very unlikely to pay for travel information. For travel time in all routes, 30% of the participants were neutral and 22% were likely to pay for it.

8.1.2 Empirical findings about travellers' perceptions

Due to its importance for modelling, another aspect investigated was travellers' perceptions. In particular, investigation of travellers' perceptions of the routes' characteristics and of the route choice set. Outcomes of this PhD thesis suggest that travellers are not fully aware of the route choice set and are biased in favour of their preferred habitual routes.

Regarding the choice set, the general route choice set provided by the participants is quite comprehensive. For instance, only two roads had not been mentioned by any of the participants and were actually chosen during the reference treatment. Nevertheless, investigation on the individual level showed that 12 out of 32 participants chose at least one route different from the reported during the reference treatment. This shows a big discrepancy between travellers' perceived awareness of the route choice set and their actual awareness. The non-reported routes were in general not as straightforward as the main ones, i.e. required a considerable amount of left and right turns, and were rarely used. As extensively discussed during this thesis, commuters tend to stick to the same route.

Regarding travellers' perceptions of the routes' characteristics, outcomes show bias in favour of travellers' habitual routes. Besides this, the terms fastest & reliable and slowest & unreliable seem to be confounded. For instance, travellers consider their habitual routes to be reliable when in reality they are among the most (objectively) unreliable (but fastest routes). The contrary applies to non-habitual routes. This is aligned with findings reported in the literature regarding perception of route reliability for habitual and non-habitual routes, i.e. higher travel times in habitual routes are treated as exceptions while slow/normal travel times reinforce the preference for the habitual route; the opposite holds for non-habitual routes (Bogers, 2009). Travellers' perceptions did not change with respect to the period of the day (morning or afternoon peak hours).

8.2 Methodological contributions

The main scientific contribution of this PhD thesis is the proposed modelling framework, RL Dyn. In a dynamic setting, the RL Dyn overcomes the important modelling issue of choice set definition based on RP data collected in congested networks. Outcomes of this research indicate that the proposed modelling framework turned out to be (predictively) valid and appropriate to the context investigated.

Investigations about the predicted log likelihood of a path observation show that despite the much bigger state space of the RL Dyn, its predictive power is substantially higher than that of the RL. This was the case for 18 out of 20 trials each with 113 random samples. These outcomes, therefore, indicate that adding the time dimension to the route choice problem compensates the higher complexity and computational demand for model estimation. Furthermore, this implies that the ratio of parameters of the RL Dyn provide more appropriate behavioural interpretation of travellers' behaviour.

It is important to highlight that results indicate that under high variability of traffic conditions, the predictive power of the RL Dyn is much higher than that of the RL. These are the type of situations in which provision of real-time traffic information is likely to be most useful due to its potential to help travellers avoiding congestion by choosing a different route. This reinforces the pertinence of the RL Dyn for the type of investigations done in this PhD thesis.

8.3 Conclusions

The target group investigated in this research consist of commuters living within 19 km distance from work. According to the Dutch institute of mobility research, among the 2 million car commuters in the Netherlands 60% live within 19 km distance from work (Dutch MON database, 2008). This means that the target group investigated corresponds to the largest group of travellers driving during peak hours in the Netherlands. Main findings of this research indicate that despite the potential effects of travel information, the impact on the route choice behaviour of commuters is smaller than expected. Although informed travellers appear to be more sensitive to travel time variance than regular travellers, the role of habit in commuters' route choice decisions is very strong. In particular, even when incurring in delays longer than expected, they prefer to stick to their preferred planned routes. Based on these findings it can be concluded that the role of travel information to alleviate congestion, and the consequent effects on network performance, resultant from commuting trips is limited. This means that it is not expected that travellers will change their route choices due to provision of travel information unless one is dealing with extreme conditions. Similar outcomes are expected for other cases in which commuters living within 19km from work are investigated in a real setting under regular traffic conditions.

Based on the fact that travellers are not fully aware of the choice set, it is possible to conclude that the modelling approach used in this thesis, which does not require definition of choice set, is appropriate for the situation investigated. This is further supported by the fact that the RL Dyn turned out to be (predictively) valid.

The fact that travellers appear to be biased in favour of their habitual routes reinforces the importance of perceptions for route choice decisions. Although this aspect was not tackled in this research from a modelling perspective, as it would lead to an even higher degree of complexity of the modelling framework, discussing it contributes to increase awareness about the problem.

8.4 Implications for practice

Main conclusions of this thesis highlight the role of habit on the route choice behaviour of commuters. In particular, the role of habit/experience in commuters' route choice decisions appears to be more important than travel information. As a result, they are not willing to pay to receive travel information about their commuting trips. As reported during the experiment, travel information appears to be more useful for a different context rather than for the daily commuting trip. This has strong implications for commercial information services and appropriate business models.

Private providers of travel information, therefore, should deal with commuters in a more tailor-made way. For instance, in-car devices could already "know" the most likely trip to occur given a certain day of the week and period of the day and provide a "green light/OK" as push notification in the event of regular traffic conditions in the preferred habitual route. A warning would be given otherwise. Commuters then would not need to actively check travel information or alternative routes. Besides this, they would have a feeling of reassurance/confirmation that the choice made was good. A tailor-made approach could help increase the interest of travellers in travel information during their commuting trips.

The panels of traffic information on/next to the roads are another point of attention. Outcomes of this research suggest that they are well located, but not much can be said regarding the use of traffic information. Something that could be improved is the way of displaying the traffic information. The participants often reported that it was difficult to understand the information provided. Some sort of educational advertisement could help making the content clearer.

Outcomes of this research indicate that among different sources of travel information, commuters are more willing to use personalized sources of travel information, namely TomTom. In particular, irrespective of the timing of travel information provision, TomTom is a chosen source in more than 80% of the times in which information is consulted. Conversely, the VMS panels are a chosen source in about 30% of the trips¹⁶. One can then argue that market parties, such as TomTom are more appropriate to perform traffic management and current VMS should be dismantled. Although the data indicates that commuters rely more on in-car information, there is an important issue of data privacy. By transferring this responsibility to market parties, all travellers would be subject to have their data under control of a market party. Otherwise, they would not have access to travel information. Therefore, it is important to carefully evaluate the *pro et contra* of such measure.

The fact that travellers appear to be biased in favour of their habitual routes by considering it reliable when actually it is among the most unpredictable (but fastest) routes further supports the importance of perceptions for route choice decisions. This suggests that (i) one should be careful on how to define and incorporate travel time variability in route choice models as not only the objective variability, but also the perception errors should be taken into account and (ii) travellers may be actually interested in arriving at their destination as fast as possible, even if this implies behaving in a risk-prone fashion.

A final implication regards the use of the RL Dyn in bigger networks. Model estimation in the RL Dyn is too computationally demanding. The high computational time is mostly due to the need to define specific time-period input matrices. Considering the characteristics of the network and sample size under investigation, 80 different travel time matrices were needed for the sample size of 563 observations. The size of each file was about 60MB. Besides this, hundreds of other smaller input matrices (~1MB) had to be defined. As these are mostly observation-specific attributes, they resulted in at least more 563 files for each attribute investigated. This requires time to read the files, but also storage and RAM capacity from the computers. This means that one is dealing with computational memory issues that can potentially be overcome by using servers with high memory capacity.

8.5 Future research directions

Future research directions regarding travellers' route choice behaviour should contemplate (i) the role of travel information in non-commuting trips, (ii) the role of travel information that is not aligned to travellers' expectations and (iii) the role of habitual routes. Outcomes of this research show that commuters behave in a particular way, especially concerning habit. Due to the role of habit, even when incurring in delays longer than expected, commuters prefer to stick to their preferred planned habitual routes. Besides this, as during the experiment the travel information was often aligned with their expectations, their reaction in situations in

¹⁶ Observe that total is higher than 100% because multiple sources can be consulted in the same trip.

which the travel information diverges from expectations still remains to be investigated. Compliance rates with travel information that is not aligned to travellers' expectations might then be even lower. The same, consequently, might apply to impacts on network dynamics.

Regarding travellers' perceptions, outcomes of this PhD research suggest that the importance of route reliability for route choice decisions is contrary to what has been reported in the literature. This may lead one to argue that the importance of route reliability is overestimated. Besides this, it seems that when travellers are actually driving and making their route choice decisions, what really matters is to arrive at the destination the fastest possible. This implies that travellers may behave as risk prone. More RP experiments are required to value the merits of these findings and determine whether (i) travellers behave as risk prone, (ii) the perceived travel time uncertainty is low or (iii) the sample size was small.

Finally, future research regarding the modelling framework should try to incorporate stochastic travel times, a more detailed network (and hence more observations), implement correction for panel data and perceptions via random error terms. Besides this, it should also try to include the information provided itself so as to allow identifying what the informed links were. It is acknowledged, however, that this is a rather challenging task as it depends of the availability of data from several information providers.

Bibliography

Abdel-Aty, M. A., R. Kitamura and P. Jovanis (1997) Using stated preference data for studying the effect of advanced traffic information on drivers' route choice, *Transportation Research Part C*, 5, pp. 39-50.

Adler, J. L. and V. J. Blue (1998) Toward the design of intelligent traveler information systems, *Transportation Research Part C*, 6, pp. 157 - 172.

Adler, J. L. and M. J. Kalsher (1994) Human factors studies to investigate driver behavior under in-vehicle information systems: an interactive microcomputer simulation approach, *38th Annual Meeting of the Human Factors and Ergonomics Society*, Massachusetts, USA, pp. 461-465.

Aguirregabiria, V. and P. Mira (2002) Swapping the nested fixed point algorithm: a class of estimators for discrete Markov decision models, *Econometrica*, 70(4), pp. 1519-1543.

Allais, M. (1953) Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'Ecole Americaine, *Econometrica*, 21(4), pp. 503-546.

Arnott, R., A. de Palma and R. Lindsey (1991) Does providing information to drivers reduce traffic congestion?, *Transportation Research Part A*, 25(5), pp. 309-318.

Avineri, E. (2009) The fuzzy meaning of reference-based perceptions in travel choice modeling, *88th Annual Meeting of the Transportation Research Board*, Washington DC, USA.

Axhausen, K. W., A. Zimmermann, S. Schönfelder, G. Rindsfuser and T. Haupt (2002) Observing the rhythms of daily life: a six-week travel diary, *Transportation*, 29(2), pp. 95-124.

Bates, J., J. Polak, P. Jones and A. Cook (2001) The valuation of reliability for personal travel, *Transportation Research Part E*, 37, pp. 191-229.

- Bekhor, S., M. Ben-Akiva and M. Ramming (2001) Estimating route choice models for large urban networks, *Proceedings of the 9th World Conference on Transport Research*, Seoul, Korea.
- Bell, D. E. (1982) Regret in decision making under uncertainty, *Operations Research*, 30(5), pp. 961-981.
- Bellman, R. (1957a) A Markovian decision process, *Journal of Mathematics and Mechanics*, 6.
- Bellman, R. (1957b) A theory of the economics of time, *The Economic Journal*, 81(324), pp. 828-846.
- Ben-Akiva, M. and M. Bierlaire (1999) Discrete choice methods and their applications to short-term travel decisions, in: *Handbook of Transportation Science*, ed Hall, R., Kluwer Academic Publishers, pp. 5-34.
- Ben-Akiva, M., A. de Palma and I. Kayasi (1991) Dynamic network models and driver information system, *Transportation Research Part A*, 25(251-266).
- Ben-Akiva, M. and S. R. Lerman (1985) *Discrete choice analysis: theory and application to travel demand*, The MIT Press, Cambridge Massachusetts.
- Ben-Elia, E., I. Erev and Y. Shiftan (2008) The combined effect of information and experience on drivers' route choice behavior, *Transportation*, 35, pp. 165-177.
- Ben-Elia, E. and D. Ettema (2011) Rewarding rush-hour avoidance: a study of commuters' travel behavior, *Transportation Research Part A* 45(7), pp. 567-582.
- Ben-Elia, E. and Y. Shiftan (2010) Which road do I take? A learning-based model of route-choice behavior with real-time information, *Transportation Research Part A*, 44, pp. 249-264.
- Bierlaire, M., J. Chen and J. Newman (2010) *Modeling route choice behavior from smartphone GPS data*, Ecole Polytechnique Fédérale de Lausanne, Report TRANSP-OR 101016, Lausanne.
- Black, I. G. and J. G. Towriss (1993) *Demand effects of travel time variability*, Department of Transport, Her Majesty's Stationery Office, London, UK.
- Bogers, E. A. I. (2009) Traffic information and learning in day-to-day route choice, T2009/5, June 2009, *TRAIL Thesis Series, the Netherlands*.
- Bonsall, P. (2004) Traveller behavior: decision-making in an unpredictable world, *Journal of Intelligent Transportation Systems*, 8(1), pp. 45-60.
- Bonsall, P. W. (1996) Analysing and modelling the influence of roadside variable message displays on drivers' route choice, *Proceedings of the 7th World Conference on Transport Research*, Sydney, Australia, pp. 11-26.
- Bovy, P. H. L. (2009) On modelling route choice sets in transportation networks: a synthesis, *Transport Reviews*, 29(1), pp. 43-68.
- Boyce, D. E. (1988) Route guidance systems for improving urban travel and location choices, *Transportation Research Part A*, 22(4), pp. 275-281.

- Brownstone, D., A. Ghosh, T. F. Golob, C. Kazimi and D. V. Amelsfort (2003) Drivers willingness-to-pay to reduce travel time: evidence from the San Diego I-15 congestion pricing project, *Transportation Research Part A*, 37, pp. 373-387.
- Caplice, C. and H. S. Mahmassani (1992) Aspects of commuting behavior: preferred arrival time, use of information and switching propensity, *Transportation Research Part A*, 26, pp. 409-418.
- Carrion, C. and D. Levinson (2012) Value of travel time reliability: a review of current evidence, *Transportation Research Part A* 46, pp. 720-741.
- Cascetta, E., A. Nuzzolo, F. Russo and A. Vitetta (1996) A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks, *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Lyon, France.
- Chang, X. and P. Stopher (1981) Defining the perceived attributes of travel modes for urban work trips, *Transportation Planning and Technology*, 7, pp. 55-65.
- Chen, P. and H. S. Mahmassani (1993) A dynamic interactive simulator for the study of commuter behavior under real time traffic information supply strategies, *Transportation Research Record*, 1413, pp. 12-21.
- Chen, P. S.-T., K. K. Srinivasan and H. S. Mahmassani (1999) Effect of information quality on compliance behavior of commuters under real-time traffic information, *Transportation Research Record*, 1676, pp. 53-60.
- Chorus, C. (2007) Traveler response to information, T2007/2, February 2007, *TRAIL Thesis Series, the Netherlands*.
- Chorus, C., T. Arentze, H. J. P. Timmermans, Z. Sun, E. Molin and B. Van Wee (2007) Behavioral aspects of travel information: models and experiments, *Proceedings of the European Transport Conference*, Leeuwenhorst Conference Centre, the Netherlands.
- Chorus, C. G., T. A. Arentze and H. J. P. Timmermans (2008a) A comparison of Regret-Minimization and Utility Maximization in the context of travel mode-choices, *87th Annual Meeting of the Transportation Research Board*, Washington DC, USA.
- Chorus, C. G., T. A. Arentze and H. J. P. Timmermans (2008b) A Random Regret-Minimization model of travel choice, *Transportation Research Part B*, 42, pp. 1-18.
- Chorus, C. G., E. J. E. Molin and B. Van Wee (2006) Travel information as an instrument to change car drivers' travel choices: a literature review, *European Journal of Transport and Infrastructure Research*, pp. 335 - 364.
- Connors, R. D. and A. Sumalee (2009) A network equilibrium model with travellers' perception of stochastic travel times, *Transportation Research Part B*, 43, pp. 614-624.
- Cook, A., P. Jones, J. Bates and J. Polak (1999) Improved methods of representing travel time reliability in SP experiments, *Proceedings of the 27th European Transport Forum: 27th PTRC Annual Meeting*, London, UK.

- Crosby, P., J. Spyridakis, J. Ramey, M. Haselkorn and W. Barheld (1993) A primer on usability testing for developers of traveler information systems, *Transportation Research Part C*, 1(2), pp. 143-157.
- Daly, A. J. and S. Hess (2010) Simple approaches for random utility modelling with panel data, *European Transport Conference*, Glasgow.
- de Palma, A., M. Ben-Akiva, D. Brownstone, C. Holt, T. Magnac, D. McFadden, P. Moffatt, N. Picard, K. Train, P. Wakker and J. Walker (2008) Risk, uncertainty and discrete choice models, *Marketing Letters*, 19, pp. 269-285.
- Denant-Boèmont, L. and R. Petiot (2003) Information value and sequential decision-making in a transport setting: an experimental study, *Transportation Research Part B*, 37, pp. 365-386.
- Dutch MON database (2008) *Mobility Onderzoek Nederlands*, [http://www.rijkswaterstaat.nl/wegen/innovatie_en_onderzoek/betere_doorstroming/mobiliteitsonderzoek_nederland/], page visited on August 20th, 2010.
- Emmerink, R. H. M., P. Nijkamp, P. Rietveld and J. N. Van Ommeren (1996) Variable message signs and radio traffic information: an integrated empirical analysis of drivers' route choice behavior, *Transportation Research Part A*, 30(2), pp. 135-153.
- Englisher, L., S. Bergman, S. Pepin and A. Wilson (1997) Promoting advanced traveler information systems among cellular and land-line phone users, *Transportation Research Record*, 1588, pp. 63-69.
- Englisher, L., R. Juster, S. Bregman, D. Kosses and A. Wilson (1996) Consumer response to the SmarTraveler ATIS: findings from the second year evaluation, *Proceedings of the 1996 Annual Meeting of ITS America*, Houston, Texas.
- Ettema, D. and H. Timmermans (2006) Costs of travel time uncertainty and benefits of travel time information: conceptual model and numerical examples, *Transportation Research Part C*, 14, pp. 335-350.
- Fosgerau, M., E. Frejinger and A. Karlström (2013) A link based network route choice model with unrestricted choice set, *Transportation Research Part B*, 56, pp. 70-80.
- Frejinger, E. and M. Bierlaire (2007) Capturing correlation with subnetworks in route choice models, *Transportation Research Part B*, 41, pp. 363-378.
- Frejinger, E. and M. Bierlaire (2010) On path generation algorithms for route choice models, in: *Choice Modelling: The State-of-the-Art and the State-of-Practice*, eds Hess, S. and A. Daly, Emerald Group Publishing Limited, pp. 307-315.
- Frejinger, E., M. Bierlaire and M. Ben-Akiva (2009) Sampling of alternatives for route choice modeling, *Transportation Research Part B*.
- Gao, S., E. Frejinger and M. Ben-Akiva (2008) Adaptive route choice models in stochastic time dependent networks, *Transportation Research Record*, 2085, pp. 136-143.
- Gao, S., E. Frejinger and M. Ben-Akiva (2010) Adaptive route choices in risky traffic networks: a Prospect Theory approach, *Transportation Research Part C*, 18, pp. 727-740.

- Hato, E., M. Taniguchi, Y. Sugie, M. Kuwahara and H. Morita (1999) Incorporating an information acquisition process into a route choice model with multiple information sources, *Transportation Research Part C*, 7(2), pp. 109-129.
- Hensher, D. A. (1994) Stated preference analysis of travel choices: the state of practice, *Transportation*, 21, pp. 107-133.
- Hey, J. D. and C. D. Orme (1994) Investigating generalisations of expected utility theory using experimental data, *Econometrica*, 62, pp. 1291-1326.
- Hobeika, A., R. Sivanandan, K. Jehanian and M. Ameen (1996) Advanced Traveler Information System users needs in I-95 Northeast Corridor, *Transportation Research Record*, 1537, pp. 55-62.
- Hollander, Y. (2006) Direct versus indirect models for the effects of unreliability, *Transportation Research Part A*, 40, pp. 699-711.
- Joh, C.-H., B. Lee, M. Bin, T. Arentze and H. Timmermans (2011) Exploring the use of travel information - identifying contextual market segmentation in Seoul, Korea, *Journal of Transport Geography*, pp. 1245-1251.
- Jou, R. C. (2001) Modeling the impact of pre-trip information on commuter departure time and route choice, *Transportation Research Part B*, 35, pp. 887-902.
- Jou, R. C., T. Y. Hu and C. W. Lin (1997) Empirical results from Taiwan and their implications for advanced travelers' pre-trip information systems, *Transportation Research Record*, 1607, pp. 126-133.
- Jou, R. C. and H. S. Mahmassani (1996) Comparability and transferability of commuter behavior characteristics between cities: departure time and route switching decisions, *Transportation Research Record*, 1556, pp. 119-130.
- Kahneman, D. and A. Tversky (1979) Prospect Theory: an analysis of decision under risk *Econometrica*, Vol. 47(2), pp. 263-291.
- Khattak, A., J. Schofer and F. Koppelman (1995) Effect of traffic reports on commuters route and departure time changes, *Journal of Advanced Transportation*, 29(2), pp. 193-212.
- Khattak, A. J., J. L. Schofer and F. S. Koppelman (1993) Commuters' enroute diversion and return decisions: analysis and implications for advanced traveler information systems, *Transportation Research Part A* 27(2), pp. 101-111.
- Khattak, A. J., Y. Yim and L. S. Prokopy (2003) Willingness to pay for travel information, *Transportation Research Part C*, 11(2), pp. 137-159.
- Koutsopoulos, H. N., T. Lotan and Q. Yang (1994) A driving simulator and its application for modeling route choice in the presence of information, *Transportation Research Part C*, 2(2), pp. 91-107.
- Koutsopoulos, H. N. and H. Xu (1993) An information discounting routing strategy for advanced traveler information systems, *Transportation Research Part C*, 1(3), pp. 249-264.
- Lam, T. and K. Small (2001) The value of time and reliability: measurements from a value pricing experiment, *Transportation Research Part E*, 37, pp. 235-251.

- Li, Z. and D. Hensher (2011) Prospect theoretic contributions in understanding traveller behaviour: a review and some comments, *Transport Reviews: A Transnational Transdisciplinary Journal*, 31(1), pp. 97-115.
- Loomes, G. and R. Sugden (1982) Regret Theory: an alternative theory of rational choice under uncertainty, *The Economic Journal*, 92(368), pp. 805-824.
- Louviere, J. J. and D. A. Hensher (1983) Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event, *Journal of Consumer Research*, 10(3), pp. 348-381.
- Mahmassani, H. S. (1997) Dynamics of commuter behaviour: recent research and continuing challenges, in: *Understanding Travel Behaviour in an Era of Change*, eds Stopher, P. and M. Lee-Gosselin, Pergamon Press, Oxford, pp. 279-313.
- Mahmassani, H. S. and R. Jayakrishnan (1991) System performance and user response under real-time information in a congested traffic corridor, *Transportation Research Part A*, 25(5), pp. 293-307.
- Mahmassani, H. S., T. Joseph and R. C. Jou (1993) Survey approach for study of urban commuter choice dynamics, *Transportation Research Record*, 1412, pp. 80-89.
- Mahmassani, H. S. and Y. Liu (1999) Dynamics of commuting behaviour under advanced traveller information systems, *Transportation Research Part C*, 7, pp. 91-107.
- Mahmassani, H. S. and Y. S. Liu (1997) Models of user pre-trip and en-route switching decisions in response to real-time information, *Proceedings of the 8th IFAC Symposium on Transportation Systems*, Chania, Greece.
- Mai, T. (2012) Revisiting optimization algorithm for maximum likelihood estimation, *Master thesis*.
- Menghini, G., N. Carrasco, N. Schüssler and K. W. Axhausen (2010) Route choice of cyclists in Zurich, *Transportation Research Part A*, 44(9), pp. 754-765.
- Molin, E. J. E. and H. J. P. Timmermans (2006) Traveler expectations and willingness-to-pay for web-enabled public transport information services, *Transportation Research Part C*, 14 (1), pp. 57-67.
- Neuherz, M., V. Patz, T. Pischner and H. Keller (2000) User acceptance and impacts of new multimodal traffic information systems in Bayerninfo, *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin, Italy.
- Ng, L., W. Barfield and F. Mannering (1995) A survey-based methodology to determine information requirements for advanced traveler information systems, *Transportation Research Part C*, 3(2), pp. 113-127.
- Papinski, D., D. M. Scott and S. T. Doherty (2009) Exploring the route choice decision-making process: a comparison of planned and observed routes obtained using person-based GPS, *Transportation Research Part F*, 12, pp. 347-358.
- Parvaneh, Z., T. Arentze and H. Timmermans (2012) Understanding travelers' behavior in provision of travel information: a Bayesian belief approach, *Procedia - Social and Behavioral Sciences*, 54, pp. 251 - 260.

- Payne, J. W., J. R. Bettman and E. J. Johnson (1993) *The adaptive decision maker*, Cambridge University Press, Cambridge.
- Peer, S., Koster and E. Verhoef (2010) The perception of travel time variability, *Proceedings of the 4th International Symposium on Transportation Network Reliability*, Minneapolis, USA.
- Petrella, M. and J. Lappin (2004) Comparative analysis of customer response to online traffic information in two cities: Los Angeles, California and Seattle, Washington, *Transportation Research Record*, 1886, pp. 10-18.
- Polak, J. and P. Jones (1993) The acquisition of pre-trip information: a stated preference approach, *Transportation*, 20(2), pp. 179-198.
- Polydoropoulou, A. and M. Ben-Akiva (1999) The effect of advanced traveller information systems (ATIS) on travelers' behavior, in: *Behavioural and network impacts of driver information systems*, Aldershot, Ashgate, pp. 317-352.
- Polydoropoulou, A., D. Gopinath and M. Ben-Akiva (1997) Willingness to pay for advanced traveler information systems: SmarTraveler case study., *Transportation Research Record*, 1588, pp. 1-9.
- Prashker, J. N. (1979) Direct analysis of the perceived importance of attributes of reliability of travel modes in urban travel, *Transportation*, 8, pp. 329-346.
- Prato, C. G. (2009) Route choice modeling: past, present and future research directions, *Journal of Choice Modelling*, 2(1), pp. 65-100.
- Qstarz (2011) *GPS recorder*, [www.qstarz.com], page visited on July 14th, 2011.
- Ramos, G. M., W. Daamen and S. P. Hoogendoorn (2011) Expected Utility Theory, Prospect Theory, and Regret Theory compared for prediction of route choice behavior, *Transportation Research Record*, 2230, pp. 19-28.
- Ramos, G. M., W. Daamen and S. P. Hoogendoorn (2013) Modelling travellers' heterogeneous route choice behaviour as prospect maximizers, *The Journal of Choice Modelling*, 6, pp. 17-33.
- Ramos, G. M., W. Daamen and S. P. Hoogendoorn (2014) A state-of-the-art review: developments on utility theory, prospect theory and regret theory to investigate travellers' behaviour in situations involving travel time uncertainty, *Transport Reviews: A Transnational Journal*, doi: 10.1080/01441647.2013.856356.
- Remer, M., T. Atherton and W. Gardner (1996) TravLink operational test results, *Proceedings of the 3rd World Congress on Intelligent Transport Systems*, Orlando, Florida.
- Rose, J. M., M. C. J. Bliemer, D. A. Hensher and A. T. Collins (2008) Designing efficient stated choice experiments in the presence of reference alternatives, *Transportation Research Part B*, 42, pp. 395-406.
- Rust, J. (1987) Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher, *Econometrica*, 55(5), pp. 999-1033.

- Schofer, J. L., A. Khattak and F. S. Koppelman (1993) Behavioral issues in the design and evaluation of advanced traveler information systems, *Transportation Research Part C*, 1(2), pp. 107-117.
- Selten, R., T. Chmura, T. Pitz, S. Kube and M. Schreckenberg (2007) Commuters route choice behaviour, *Games and Economic Behavior*, 58, pp. 394-406.
- Shiftan, Y., S. Bekhor and G. Albert (2010) Route choice behaviour with pre-trip travel time information, *IET Intelligent Transport Systems*, doi: 10.1049/iet-its.2010.0062.
- Simon, H. A. (1995) A Behavioral model of rational choice, *Quarterly Journal of Economics*, 69(1), pp. 99-118.
- Small, K., C. Winston and J. Yan (2006) Differentiated road pricing, express lanes, and carpools: exploiting heterogeneous preferences in policy design, *Brookings-Wharton Papers on Urban Affairs*, 7, pp. 53-96.
- Small, K. A., C. Winston and J. Yan (2005) Uncovering the distribution of motorists' preferences for travel time and reliability, *Econometrica*, 73(4), pp. 1367-1382.
- Spissu, E. and I. Meloni (2011) A behavioral analysis of daily route choice using GPS-based data, *90th Annual Meeting of the Transportation Research Board*, Washington DC, USA.
- Spitsmijden Program (2006) *Spitsmijden*, [www.spitsmijden.nl/], page visited on June 12th, 2012.
- Starmer, C. (2000) Developments in Non-Expected Utility Theory: the hunt for a descriptive theory of choice under risk, *Journal of Economic Literature*, 38(2), pp. 332-382
- TomTom (2011) *Tom Tom Navigator*, [www.tomtom.com], page visited on July 14th, 2011.
- TomTom (2014) *Quarterly financial results*, [http://corporate.tomtom.com/presentations.cfm], page visited on July 20th, 2014.
- Train, K. E. (2009) *Discrete choice methods with simulation*, 2nd ed, Cambridge University Press, New York.
- Tseng, Y., E. Verhoef, G. De Jong, M. Kouwenhoven and T. Van der Hoorn (2009) A pilot study into the perception of unreliability of travel times using in-depth interviews, *Journal of Choice Modelling*, 2, pp. 8-28.
- Tseng, Y. Y., J. Knockaert and E. T. Verhoef (2011) A revealed-preference study of behavioural impacts of real-time traffic information, *Transportation Research Part C*, 30, doi: 10.1016/j.trc.2011.11.006, pp. 196-209.
- Tversky, A. and D. Kahneman (1992) Advances in Prospect Theory: cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5(4), pp. 297-323.
- Vaughn, K. M., M. A. Abdel-Aty, R. Kitamura, P. P. Jovanis and H. Yang (1993) Experimental analysis and modeling of sequential route choice under ATIS in a simple traffic network, *Proceedings of the 72nd Annual Meeting of the Transportation Research Board*, Washington DC, USA.
- Vaziri, M. and T. N. Lam (1983) Perceived factors affecting driver route decisions, *Journal of Transportation Engineering*, 109, pp. 297-311.

Von Neumann, J. and O. Morgenstern (1947) *Theory of Games and economic behavior*, Princeton University Press, Princeton.

Wardman, M., P. W. Bonsall and J. D. Shires (1997) Driver response to variable message signs: a stated preference investigation, *Transportation Research Part C*, 5(6), pp. 389-405.

Weber, M. and C. Camerer (1987) Recent developments in modelling preferences under risk, *OR Spektrum*, 9, pp. 129-151.

Wikipedia (2011) *Randstad*, [<http://en.wikipedia.org/wiki/Randstad>], page visited on July 14th, 2011.

Wolinetz, L., A. Khattak and Y. B. Yim (2001) Why will some individuals pay for travel information when it can be free? Analysis of a Bay Area travel survey, *Transportation Research Record*, 1759, pp. 9-18.

Xu, H., J. Zhou and W. Xu (2011) A decision-making rule for modeling travelers' route choice behavior based on Cumulative Prospect Theory, *Transportation Research Part C*, 19, pp. 218 – 228.

Zhang, L. and D. Levinson (2008) Determinants of route choice and value of traveler information: a field experiment, *Transportation Research Record*, 2086, pp. 81-92.

Summary

Traffic congestion is experienced by a great number of travellers during peak hours and is directly influenced by travel-related decisions, such as route choice decisions. In order to minimize problems that arise from congestion, such as delays, uncertainty and environmental effects, there has been an increased interest to investigate the role of travel information on travellers' route choice behaviour. When travel information is provided, travellers have to decide whether to comply with the travel information. The complexity of such a decision increases, for instance, with travellers' expectations about the traffic situation, the quality and reliability of travel information and route choice habits.

A review of the literature shows that estimation of route choice utility models based on stated preference data is the most traditional way to investigate the relationship between travellers' route choice behaviour and travel information. Stated preference experiments, however, are subject to the main drawback of external validity, i.e. it is arguable whether the experimental design and its results are valid and can be replicated in a real setting. Although revealed preference experiments are not subject to issues regarding external validity, there are three main modelling issues associated to the estimation of route choice utility models based on revealed preference data collected in real networks: choice set sampling is difficult, utilities may share unobserved attributes and the attributes of the alternatives (e.g. travel times) are in general unknown.

This thesis has two main contributions: an empirical and a methodological. The empirical contribution is to investigate the role of travel information on travellers' route choice behaviour based on revealed preference data collected in a real congested network. The methodological is to address for the first time issues regarding choice set sampling and attributes of the alternatives in a dynamic setting supported by real data about the traffic conditions in the network.

A revealed preference experiment making use of GPS devices, travel diaries and interviews was conducted to investigate the route choice behaviour of 32 travellers for a period of 9 weeks in the Netherlands. Travellers were subject to different sources and timing of travel

information provision. The target group that was investigated consisted of commuters living within 19 kilometres distance from work, who, according to the Dutch institute of mobility research, correspond to 60% of the 2 million car commuters in the Netherlands (Dutch MON database, 2008). This means that the investigated target group corresponds to the largest group of travellers driving during peak hours in the Netherlands. The sample resulted in 897 valid GPS traces and travel diaries, among which 374 refer to the first 3 weeks of data collection in which only free/public sources of travel information were available and 523 correspond to the subsequent 6 weeks of data collection in which personalized real-time travel information was also available. Moreover, data about the traffic conditions in the network was also collected for the whole study period at 1 minute intervals. The group subject only to free/public sources of travel information can be categorized as regular travellers. Otherwise, they are labelled informed travellers.

Model estimation was done based on a new modelling framework denominated Recursive Logit Dynamic (RL Dyn), which is based on the sequential link-choice decisions of the Recursive Logit (RL) (Fosgerau et al., 2013). The RL can be consistently estimated with revealed preference data and has the advantage of not requiring of choice sets to be defined. Although both the RL and RL Dyn share the same properties, the latter is extended to a dynamic setting. In other words, while in the static formulation of the RL the network is defined in term of links and nodes, in the dynamic one it is defined in term of links, nodes and time intervals. This way, the time dimension is incorporated to the route choice problem and not only the route to chosen matters, but also the moment the route is chosen.

This thesis shows the pertinence of using the RL Dyn for the proposed investigations. In particular, combination of the interviews conducted at the beginning of the experiment with the GPS traces indicate that travellers are not fully aware of the route choice set. Furthermore, the (predictively) validity of the RL Dyn and a substantially higher predictive power than its static counterpart (RL) were demonstrated. This indicates that adding the time dimension to the route choice problem compensates the higher complexity and computational demand for model estimation.

Regarding the role of different sources of travel information, estimation results indicate that the more detailed the travel information is, the more travellers are sensitive to it. Both for pre-trip and en-route information, travellers who consult TomTom are significantly more sensitive to travel time than those who consult radio. Besides this, although estimation results indicate that road links with variable message signs (VMS) panels significantly influence route choice behaviour. It appears that consulting or not consulting the information displayed on them similarly influence travellers' decisions. This suggests that the VMS panels are well located, but not much can be said regarding the use of travel information. Besides this, equipping travellers with TomTom devices increases the likelihood that they will use travel information to assist their route choices.

With respect to the timing of providing travel information (i.e. pre-trip or en-route), estimation results suggest that travellers who only consult pre-trip information are significantly less sensitive to travel time than travellers who do not consult information at all. As pre-trip information is provided at the beginning of the trip, being aware of the traffic conditions may provide some sort of comfort to travellers. Besides this, outcomes of the quantitative and qualitative analyses suggest that finding out the best route choice considering

the departure time is “now” appears to be more important than planning an “ideal” departure time to avoid congestion. As reported by the participants, they usually consulted travel information just when entering the car and not beforehand while still at home or in the office.

Regarding compliance with travel information, outcomes of the quantitative and qualitative analyses show that (in general) the more detailed the source of the travel information is, the higher the compliance rates are. Besides this, compliance rates appear to differ for informed and regular travellers: informed travellers are more willing to comply with travel information than regular travellers. In addition, travellers are more willing to comply with pre-trip, rather than en-route, information.

Despite the potential effects of travel information, the impact on the route choice behaviour of commuters is smaller than expected. Although informed travellers are about 30% more sensitive to travel time than regular travellers, the importance of habit to commuters’ route choice decisions is very strong. Habit turned out to be a highly significant parameter in all model specifications tested. In particular, even when incurring delays longer than expected, travellers tend to stick to their preferred planned routes. This suggests that in order to change the behaviour of commuters either the penalty associated to a late arrival or the benefit for on time/earlier arrival should be or perceived as higher. Based on these findings the (overall) conclusion can be drawn that the role of travel information to alleviate congestion, and the consequent effects on network performance, resultant from commuting trips is limited. This means that it is not expected that travellers will change their route choices due to provision of travel information unless one is dealing with extreme conditions. Similar outcomes are expected for other situations in which commuters living within 19 kilometres from work are investigated in a real setting under regular traffic conditions.

Samenvatting

Congestie wordt tijdens de spits door een groot aantal reizigers ervaren en wordt direct beïnvloed door reisgerelateerde beslissingen zoals routekeuzebeslissingen. Om problemen die ontstaan door congestie, zoals vertragingen, onzekerheid en milieueffecten, te minimaliseren, is er toenemende interesse geweest in het onderzoeken van de rol die reisinformatie speelt bij het routekeuzegedrag van reizigers. Indien reisinformatie beschikbaar is moeten de reizigers beslissen of ze de reisinformatie opvolgen. De complexiteit van een dergelijke beslissing neemt toe met, bijvoorbeeld, de verwachtingen die reizigers hebben ten aanzien van de verkeerssituatie, de kwaliteit en betrouwbaarheid van de reisinformatie en de gewoontes op het gebied van routekeuze.

Literatuuronderzoek laat zien dat de meest gebruikelijke manier om de relatie tussen het routekeuzegedrag van reizigers en reisinformatie te onderzoeken is het schatten van nutsmodellen voor routekeuze op basis van *stated preference data*. *Stated preference* experimenten kennen daarentegen het belangrijke nadeel van externe validiteit, dat wil zeggen het is te betwisten of het ontwerp van het experiment en de resultaten geldig zijn en overgenomen kunnen worden naar een werkelijke situatie. Hoewel *revealed preference* experimenten niet onderhevig zijn aan kwesties van externe validiteit zijn er drie modelleringskwesties verbonden aan het schatten van nutsmodellen voor routekeuze gebaseerd op *revealed preference* data welke zijn verzameld in werkelijke netwerken: het opstellen van een keuzeset is lastig, nutsfuncties delen mogelijk niet-geobserveerde kenmerken en de kenmerken van de alternatieven (bijvoorbeeld reistijden) zijn doorgaans onbekend.

De twee belangrijkste bijdragen van dit proefschrift zijn een empirische en een methodologische bijdrage. De empirische bijdrage bestaat uit het onderzoeken van de rol die reisinformatie speelt bij het routekeuzegedrag van reizigers gebaseerd op *revealed preference* data welke zijn verzameld in een werkelijk netwerk met congestie. De methodologische bijdrage is het voor de eerste keer aan de orde stellen van kwesties op het gebied van het opstellen van een keuzeset en attributen van de alternatieven in een dynamische setting ondersteuning door werkelijke data van de verkeerscondities in het netwerk.

Een *revealed preference* experiment met behulp van GPS-apparatuur, reisverslagen en interviews is uitgevoerd om het routekeuzegedrag van 32 reizigers te onderzoeken voor de duur van 9 weken in Nederland. De reizigers werden voorzien van reisinformatie van verschillende bronnen en op verschillende momenten. De onderzochte doelgroep bestond uit forenzen die binnen 19 kilometer van hun werk wonen, net als 60% van de 2 miljoen forenzen die in Nederland met de auto reizen (NederlandseDutch MON database, 2008). Dit betekent dat de onderzochte doelgroep overeenkomt met de grootste groep reizigers die tijdens de spits met de auto rijdt in Nederland. Deze steekproef heeft geresulteerd in 897 geldige GPS-sporen en reisverslagen. 374 van deze sporen hebben betrekking op de eerste drie weken van dataverzameling waarin enkel gratis/publieke bronnen van reisinformatie beschikbaar waren en 523 sporen horen bij de daaropvolgende 6 weken van dataverzameling waarin ook persoonlijke real time reisinformatie beschikbaar was. Gedurende de hele onderzoeksperiode zijn er bovendien data verzameld over de verkeerscondities in het netwerk, met intervallen van 1 minuut. De groep die enkel gratis/publieke bronnen van reisinformatie ontvangen behoren tot de groep reguliere reizigers. Het overige deel behoort tot de groep geïnformeerde reizigers.

Een modelschatting is gedaan op basis van een nieuw modelleringskader genaamd *Recursive Logit Dynamic (RL Dyn)*, welke is gebaseerd op de opeenvolgende beslissingen voor de keuze voor een wegdeel van *Recursive Logit (RL)* (Fosgerau et al., 2013). De RL kan goed geschat worden met *revealed preference data* en kent het voordeel dat er geen afbakening van keuzesets vereist is. Hoewel RL en RL Dyn dezelfde kenmerken hebben, is de laatste uitgebreid naar een dynamische situatie. Op deze manier wordt de dimensie tijd opgenomen in het routekeuze probleem.

Dit proefschrift laat de relevantie zien van het gebruiken van de RL Dyn voor het voorgestelde onderzoek. De interviews die zijn uitgevoerd aan het begin van het experiment in combinatie met de GPS-sporen laten in het bijzonder zien dat reizigers niet geheel op de hoogte zijn van het routekeuzeset. Bovendien zijn de (voorspelbare) validiteit van de RL Dyn en een substantieel hogere voorspellingskracht dan zijn statische tegenhanger (RL) aangetoond. Dit laat zien dat het toevoegen van de dimensie tijd aan het routekeuze probleem compenseert voor de hogere complexiteit en rekenkundige vraag voor het schatten van het model.

Ten aanzien van de rol van verschillende bronnen van reisinformatie laten de schattingsresultaten zien dat hoe gedetailleerder de reisinformatie is, hoe gevoeliger de reizigers zijn voor deze informatie. Zowel voor informatie voorafgaand aan als tijdens de reis geldt dat de reizigers die TomTom raadplegen significant meer gevoelig zijn voor reistijd dan degenen die de radio raadplegen. Daarnaast blijkt dat, hoewel schattingsresultaten laten zien dat wegdelen met variabele informatiepanelen (VMS) het routekeuzegedrag significant beïnvloeden. Het wel of niet raadplegen van de informatie die op de panelen wordt vertoond beïnvloedt de beslissingen van de reizigers op dezelfde manier. Dit suggereert dat de VMS panelen zich op de juiste locatie bevinden, maar er kan weinig gezegd worden over het gebruik van de reisinformatie. Het uitrusten van reizigers met TomTom apparatuur maakt het waarschijnlijker dat zij reisinformatie gaan gebruiken als hulp bij hun routekeuzes.

Met betrekking tot het moment van het verschaffen van reisinformatie (namelijk voorafgaand of tijdens de reis) suggereren schattingsresultaten dat reizigers die enkel voorafgaand aan de

reis informatie raadplegen significant minder gevoelig zijn voor reisinformatie dan reizigers die helemaal geen reisinformatie raadplegen. Aangezien informatie wordt verschaft aan het begin van de reis, biedt het op de hoogte zijn van de verkeerscondities de reizigers mogelijk een mate van comfort. Daarnaast suggereren de uitkomsten van de kwantitatieve en kwalitatieve analyses dat het vinden van de beste routekeuze uitgaande van een vertrek “per direct” belangrijker is om congestie te voorkomen dan het plannen van een “ideaal” vertrektijdstip. De deelnemers gaven aan dat zij reisinformatie meestal raadpleegden bij het betreden van de auto en niet van te voren wanneer ze nog thuis of op kantoor waren.

Ten aanzien van het naleven van verkeersinformatie laten de uitkomsten van de kwantitatieve en kwalitatieve analyses zien dat (in het algemeen geldt dat) hoe gedetailleerder de bron van reisinformatie, hoe hoger de mate van naleving ervan is. De mate van naleving blijkt bovendien te verschillen voor geïnformeerde en reguliere reizigers: geïnformeerde reizigers zijn meer bereid om reisinformatie na te leven dan reguliere reizigers. Reizigers zijn daarnaast meer bereid om informatie voorafgaand aan een reis dan tijdens een reis na te leven.

Ondanks de potentie van reisinformatie is het effect van deze informatie op het routekeuzegedrag van forenzen kleiner dan verwacht. Hoewel geïnformeerde reizigers ongeveer 30% gevoeliger zijn voor reistijden dan reguliere reizigers, is het belang van de gewoonte erg sterk bij routekeuzebeslissingen van forenzen. De gewoonte bleek een sterk significante parameter te zijn binnen alle geteste modelspecificaties. Zelfs wanneer de vertragingen groter zijn dan verwacht, blijven de reizigers op de geprefereerde en geplande routes. Dit suggereert dat het nadeel behorende bij een late aankomst of het voordeel voor een vroege aankomst (of een aankomst op tijd), groter moet zijn of als groter beleefd moet worden om het gedrag van de forenzen te veranderen. Op grond van deze bevindingen kan de (totaal)conclusie worden getrokken dat de rol van reisinformatie bij het verlichten van congestie, en de voortvloeiende effecten op de netwerkprestatie, voor woon-werkverkeer beperkt is. Dit betekent dat niet verwacht wordt dat reizigers hun routekeuzes zullen veranderen door hen te voorzien van reisinformatie tenzij zich extreme omstandigheden voordoen. Vergelijkbare uitkomsten worden verwacht voor andere situaties waarin onderzoek wordt gedaan onder forenzen die binnen 19 kilometer van hun werk wonen, in een werkelijke situatie onder reguliere verkeerscondities.

Resumo

Uma grande parte dos motoristas¹⁷ dirigindo em horário de pico diariamente vivencia congestionamento de trânsito. Decisões como, por exemplo, a rota a ser escolhida na viagem pretendida influenciam diretamente o nível de congestionamento. Com o intuito de minimizar os problemas derivados do congestionamento de trânsito tais como atrasos, incertezas e impactos ambientais, mais pesquisas têm sido feitas para investigar o papel que a informação de trânsito tem na escolha das rotas pelos motoristas. Quando informação de trânsito é fornecida, os motoristas têm que decidir se vão seguir a recomendação e a complexidade desse tipo de decisão aumenta, por exemplo, com as expectativas dos motoristas quanto à situação do trânsito, com a qualidade e nível de confiabilidade da informação de trânsito e com o hábito/preferência por alguma rota.

Uma revisão da literatura mostra que o modo mais tradicional de investigar como informações de trânsito influenciam o comportamento dos motoristas quanto à escolha de rotas é estimando modelos de utilidade com dados de preferência declarada¹⁸. Experimentos de preferência declarada, entretanto, têm a grande desvantagem de terem sua validade externa questionada. Em outras palavras é discutível se a configuração do experimento e os resultados também são aplicáveis (ou podem ser replicados) em um ambiente real. Embora experimentos de preferência revelada¹⁹ não estejam sujeitos a esse tipo de questionamento, três grandes dificuldades estão associadas com a estimativa de modelos de utilidade de escolhas de rotas baseado em dados de preferência revelada coletados em uma rede rodoviária real: dificuldade de definição do conjunto de possíveis rotas alternativas, as utilidades de cada alternativa podem compartilhar características não observadas pelo analista e falta de dados sobre as alternativas que não foram escolhidas (por exemplo, duração da viagem nas rotas que não foram escolhidas).

¹⁷ No contexto dessa pesquisa, motorista é qualquer pessoa dirigindo um automóvel.

¹⁸ Em experimentos de preferência declarada, as pessoas dizem como se comportariam em uma determinada situação.

¹⁹ Em experimentos de preferência revelada, as pessoas dizem como de fato se comportaram ou observa-se como as pessoas se comportaram em uma determinada situação real.

Esta tese de doutorado tem duas contribuições principais: uma empírica e uma metodológica. A contribuição empírica é investigar como o fornecimento de informação de trânsito influencia o comportamento dos motoristas na escolha de suas rotas com base em dados de preferência revelada coletados em uma rede rodoviária real. A contribuição metodológica é abordar pela primeira vez dificuldades associadas com a definição do conjunto de possíveis rotas alternativas com base em uma definição dinâmica da rede rodoviária²⁰ e substanciada por dados reais sobre a situação do trânsito em toda rede rodoviária investigada.

Um experimento de preferência revelada fazendo uso GPS, diários de viagem e entrevistas foi conduzido durante nove semanas para investigar o comportamento de escolha de rotas de 32 motoristas na Holanda. Os motoristas receberam diferentes tratamentos quanto às fontes e momentos de fornecimento de informações de trânsito. De acordo com o instituto holandês de pesquisa de mobilidade (Dutch MON database, 2008), 60% dos 2 milhões de motoristas de automóveis nos Países Baixos mora a uma distância de até 19 quilômetros do trabalho. De modo a focar no grupo que corresponde à maioria dos motoristas dirigindo em horário de pico, os motoristas que fizeram parte deste experimento foram selecionados dentro desse grupo.

A coleta de dados resultou em uma amostra válida de 897 observações de GPS e diários de viagem, dos quais 374 referem-se às primeiras 3 semanas de coleta de dados em que apenas fontes públicas/gratuitas de informação de trânsito estavam disponíveis e 523 referem-se às 6 semanas subsequentes em que informações de trânsito personalizadas quanto ao tempo real da viagem também estavam disponíveis. Dados sobre as condições do trânsito em toda rede rodoviária foi coletado durante o período do experimento em intervalos de 1 minuto. O grupo de motoristas que recebeu informações de trânsito apenas de fontes públicas/gratuitas foi rotulado de motoristas regulares, e os que também receberam informação personalizada foram rotulados motoristas informados.

O modelo foi estimado em uma nova metodologia denominada *Recursive Logit Dinamic* (RL Dyn), que é baseado em decisões sequenciais em cada via da rede rodoviária, ou seja, no final de cada via o motorista decide que outra via escolher, conforme proposto pelo *Recursive Logit* (RL) (Fosgerau et al., 2013). O RL pode ser estimado de forma consistente com dados de preferência revelada e tem a vantagem de não necessitar definição do conjunto de possíveis rotas alternativas. Embora o RL e RL Dyn compartilhem as mesmas propriedades, este último foi ampliado para uma formulação dinâmica. Dessa forma, o RL Dyn incorpora a dimensão temporal ao problema de escolha de rota. Não apenas o caminho é importante, mas o caminho depende do momento em que a escolha é feita.

A pertinência da aplicação do RL Dyn ao problema estudado foi demonstrada nesta tese de doutorado. Comparando as entrevistas realizadas no início do experimento com os dados do GPS é possível verificar que os motoristas não estão plenamente conscientes de todas as rotas que eles conhecem, ou seja, eles dirigem por mais caminhos do que eles reportaram conhecer. Definição de quais rotas incluir no conjunto de possíveis alternativas de fato mostra-se um problema na prática e o RL Dyn supera esse problema por não requerer definição do conjunto

²⁰ Em uma definição estática da rede rodoviária, a rede de transportes é definida por vias e interseções. Em uma definição dinâmica, a dimensão temporal também é considerada. Dessa forma é possível considerar como trânsito varia ao longo do dia. Dessa forma, a rota a ser escolhida também depende do momento em que a escolha é feita.

de possíveis rotas alternativas. Importante ressaltar que o poder preditivo do RL Dyn mostrou-se substancialmente maior do que o do estático RL. Isso indica a pertinência em adicionar a dimensão temporal ao problema de escolha de rotas, apesar do maior nível de complexidade metodológica e demanda computacional para estimar modelos.

Resultados da estimativa do modelo quanto ao papel de diferentes fontes de informação de trânsito indicam que quanto mais detalhada a informação, mais os motoristas são sensíveis a ela. Tanto para informações fornecidas antes e durante as viagens, motoristas que consultam informação personalizada sobre a situação real do trânsito (via TomTom) são significativamente mais sensíveis ao tempo de viagem do que aqueles que consultam informação via rádio. Outrossim, equipar os motoristas com TomTom aumenta a probabilidade de que informações de trânsito serão consultadas antes ou durante a viagem. Além disso, apesar dos resultados indicarem que vias com painéis informando a situação do trânsito significativamente influenciam o comportamento quanto a escolha de rota, isso não se deve ao fato dos motoristas estarem interessados em consultar a situação do trânsito nesses painéis. Os resultados mostram que consultar ou não consultar as informações de trânsito dos painéis influenciam as decisões dos motoristas de forma semelhante. Isto sugere que os painéis estão bem localizados, mas não se pode inferir muito sobre o uso das informações de trânsito exibidas.

Em relação ao momento em que a informação de trânsito é fornecida, ou seja, antes ou durante a viagem, os resultados da estimativa do modelo sugerem que os motoristas que consultam informações apenas antes da viagem são significativamente menos sensíveis ao tempo da viagem do que os motoristas que não consultam nenhuma informação. Devido ao fato que a informação de trânsito é fornecida antes da viagem, os motoristas podem se sentir mais seguros/confortáveis por saber o que esperar quanto à situação do trânsito. Além disso, os resultados das análises quantitativas e qualitativas sugerem que escolher a melhor rota considerando que o horário de partida é “agora” parece ser mais importante do que planejar o momento de ideal (de menor congestionamento) para iniciar a viagem. Conforme relatado pelos participantes, informações sobre o trânsito são normalmente consultadas no momento de entrar no carro e não com antecedência enquanto ainda em casa ou no escritório.

No que diz respeito a seguir as informações de trânsito, os resultados das análises quantitativas e qualitativas mostram que (em geral) quanto mais detalhada a fonte de informação, maior a porcentagem de motoristas seguindo a informação. Além disso, a porcentagem difere entre motoristas informados e regulares: motoristas informados estão mais dispostos a seguir as informações de trânsito do que de motoristas regulares. Adicionalmente, os motoristas estão mais dispostos a seguir informações fornecidas antes da viagem começar do que durante.

Não obstante os possíveis efeitos que o fornecimento de informação de trânsito tem nos motoristas, o seu impacto é menor do que o esperado. Embora motoristas informadas sejam cerca de 30% mais sensíveis ao tempo de viagem do que motoristas regulares, a importância de hábito nas escolhas de rota é muito grande. Hábito mostrou-se um parâmetro altamente significativo em todas as especificações do modelo testadas. Até mesmo quando incorrem em atrasos maiores do que o esperado, os motoristas tendem a se ater às suas rotas preferenciais. Isto sugere que para mudar o comportamento dos motoristas, ou a penalidade associada a atrasos ou o benefício associado a chegadas antecipadas devem ser considerados maiores.

Baseado nestes resultados pode-se concluir que o papel da informação de trânsito para aliviar congestionamentos é limitado. Isso significa que ele não é esperado que os motoristas mudem suas rotas devido ao fornecimento de informações de trânsito a não ser que estejam lidando com situações extremas, como uma via bloqueada, acidentes, etc. Resultados semelhantes são esperados para outros cenários em que os motoristas que morem a até 19 quilômetros de distância do trabalho sejam investigados em um ambiente real sob condições de trânsito regulares.

About the author

Born in Arcoverde (Pernambuco – Brazil), Giselle graduated as a Civil Engineer at Universidade Federal de Pernambuco (Pernambuco – Brazil) in 2001 and two years later obtained her MSc degree in Civil Engineering with specialization in the field of Transportation at Universidade de São Paulo (São Paulo – Brazil).

She then worked for a consultancy company and a governmental agency in Brazil before moving to the Netherlands in 2009 to conduct her PhD research at the TU Delft. During her PhD Giselle investigated travellers' route choice behaviour focusing on dynamic route choice modelling of the effects of travel information. The research was funded by the Dutch Organization of Science (NWO) under the program TRISTAM (Traveller Response and Information Service Technology: Analysis and Modelling).



During her PhD, Giselle cooperated with KTH Royal Institute of Technology (Stockholm – Sweden) and TomTom among others. Besides this, she was also involved in giving lectures to MSc and bachelor students of the TU Delft.

Publications:

Journals

Ramos, G.M., W. Daamen and S.P. Hoogendoorn (2014). A state-of-the-art review: developments in utility theory, prospect theory and regret theory to investigate travellers' behaviour in situations involving travel time uncertainty. *Transport Reviews: A Transnational Transdisciplinary Journal*, DOI: 10.1080/01441647.2013.856356.

Ramos, G.M., W. Daamen and S.P. Hoogendoorn (2013). Modelling travellers' heterogeneous route choice behaviour as prospect maximizers. *The Journal of Choice Modelling* 6 (2013) 17–33. <http://dx.doi.org/10.1016/j.jocm.2013.04.002>

Ramos, G.M., W. Daamen and S.P. Hoogendoorn (2011). Expected utility theory, prospect theory, and regret theory compared for prediction of route choice behavior. *Transportation Research Record*, Vol. 2230, pp. 19-28. (TUD)

Peer reviewed conference papers

Ramos, G.M., E. Frejinger, W. Daamen and S.P. Hoogendoorn (2012). Commuters and traffic information: A revealed preference study on route choice behavior. Presented at the 12th TRAIL Congress. Delft: TRAIL Research School. (TUD).

Ramos, G.M., E. Frejinger, W. Daamen and S.P. Hoogendoorn (2012). Route choice model estimation in a dynamic network based on GPS data. Presented at the 1st European Symposium on Quantitative Methods in Transportation Systems. September 4th - 7th, 2012 - Lausanne - Switzerland.

Ramos, G.M., E. Frejinger, W. Daamen and S.P. Hoogendoorn (2012). A revealed preference study on route choices in a congested network with real-time information. Presented at the 13th International Conference on Travel Behaviour Research. July 15th - 20th, 2012 - Toronto - Canada.

Ramos, G.M., W. Daamen and S.P. Hoogendoorn (2011). Modelling travellers' heterogeneous route choice behaviour as prospect maximizers. In s.n. (Ed.), Presented at the 2nd International Choice Modelling Conference 2011 (pp. 1-15). Leeds: ICMC. (TUD)

Ramos, G.M., W. Daamen and S.P. Hoogendoorn (2011). Comparative analysis on purpose of expected utility theory, prospect theory and regret theory to predict route choice behavior. In s.n. (Ed.), Presented at the 90th TRB Annual Meeting. Washington DC: Transportation Research Board. (TUD)

Ramos, G.M., W. Daamen and S.P. Hoogendoorn (2010). Towards a better understanding of the reference point in a travel behaviour context. In s.n. (Ed.), Presented at the 11th TRAIL Congress. Delft: TRAIL Research School. (TUD)

Reports

Ramos, G.M., W. Daamen and S.P. Hoogendoorn (2011). A comparative analysis between Expected Utility theory, Prospect theory and Regret theory to evaluate travellers' behaviour. Report. Delft: Delft University of Technology. (TUD)

Working papers

Ramos, G.M., E. Frejinger, W. Daamen and S.P. Hoogendoorn (2014) Quantitative and qualitative analyses on route choice in a congested network with real-time information based on RP data.

Ramos, G.M., E. Frejinger, W. Daamen and S.P. Hoogendoorn (2014). Route choice model estimation in a dynamic network based on GPS data.

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 100 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.

Ramos, G. M., *Dynamic Route Choice Modelling of the Effects of Travel Information using RP Data*, T2015/3, February 2015, TRAIL Thesis Series, the Netherlands

Sierzychula, W.S., *Development and Early Adoption of Electric Vehicles: Understanding the tempest*, T2015/2, January 2015, TRAIL Thesis Series, the Netherlands

Vianen, T. van, *Simulation-integrated Design of Dry Bulk Terminals*, T2015/1, January 2015, TRAIL Thesis Series, the Netherlands

Risto, M., *Cooperative In-Vehicle Advice: A study into drivers' ability and willingness to follow tactical driver advice*, T2014/10, December 2014, TRAIL Thesis Series, the Netherlands

Djukic, T., *Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management*, T2014/9, November 2014, TRAIL Thesis Series, the Netherlands

Chen, C., *Task Complexity and Time Pressure: Impacts on activity-travel choices*, T2014/8, November 2014, TRAIL Thesis Series, the Netherlands

Wang, Y., *Optimal Trajectory Planning and Train Scheduling for Railway Systems*, T2014/7, November 2014, TRAIL Thesis Series, the Netherlands

Wang, M., *Generic Model Predictive Control Framework for Advanced Driver Assistance Systems*, T2014/6, October 2014, TRAIL Thesis Series, the Netherlands

Kecman, P., *Models for Predictive Railway Traffic Management*, T2014/5, October 2014, TRAIL Thesis Series, the Netherlands

Davarynejad, M., *Deploying Evolutionary Metaheuristics for Global Optimization*, T2014/4, June 2014, TRAIL Thesis Series, the Netherlands

Li, J., *Characteristics of Chinese Driver Behavior*, T2014/3, June 2014, TRAIL Thesis Series, the Netherlands

Mouter, N., *Cost-Benefit Analysis in Practice: A study of the way Cost-Benefit Analysis is perceived by key actors in the Dutch appraisal practice for spatial-infrastructure projects*, T2014/2, June 2014, TRAIL Thesis Series, the Netherlands

Ohazulike, A., *Road Pricing mechanism: A game theoretic and multi-level approach*, T2014/1, January 2014, TRAIL Thesis Series, the Netherlands

Cranenburgh, S. van, *Vacation Travel Behaviour in a Very Different Future*, T2013/12, November 2013, TRAIL Thesis Series, the Netherlands

Samsura, D.A.A., *Games and the City: Applying game-theoretical approaches to land and property development analysis*, T2013/11, November 2013, TRAIL Thesis Series, the Netherlands

Huijts, N., *Sustainable Energy Technology Acceptance: A psychological perspective*, T2013/10, September 2013, TRAIL Thesis Series, the Netherlands

Zhang, Mo, *A Freight Transport Model for Integrated Network, Service, and Policy Design*, T2013/9, August 2013, TRAIL Thesis Series, the Netherlands

Wijnen, R., *Decision Support for Collaborative Airport Planning*, T2013/8, April 2013, TRAIL Thesis Series, the Netherlands

Wageningen-Kessels, F.L.M. van, *Multi-Class Continuum Traffic Flow Models: Analysis and simulation methods*, T2013/7, March 2013, TRAIL Thesis Series, the Netherlands

Taneja, P., *The Flexible Port*, T2013/6, March 2013, TRAIL Thesis Series, the Netherlands

Yuan, Y., *Lagrangian Multi-Class Traffic State Estimation*, T2013/5, March 2013, TRAIL Thesis Series, the Netherlands

Schreiter, Th., *Vehicle-Class Specific Control of Freeway Traffic*, T2013/4, March 2013, TRAIL Thesis Series, the Netherlands

Zaerpour, N., *Efficient Management of Compact Storage Systems*, T2013/3, February 2013, TRAIL Thesis Series, the Netherlands

Huibregtse, O.L., *Robust Model-Based Optimization of Evacuation Guidance*, T2013/2, February 2013, TRAIL Thesis Series, the Netherlands