Route Choice Behaviour under Uncertainty in Public Transport Networks
Stated and Revealed Preference Analyses

Shelat, S.

DOI
10.4233/uuid:abdccd21-e390-45f5-b7a3-985f9a8a682e

Publication date
2023

Document Version
Final published version

Citation (APA)
Shelat, S. (2023). *Route Choice Behaviour under Uncertainty in Public Transport Networks: Stated and Revealed Preference Analyses*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:abdccd21-e390-45f5-b7a3-985f9a8a682e

TRAIL

Summary

Arguably, nearly all real-world decisions, including travel choices, are inherently associated with subjective uncertainty where decision-makers' personal evaluations play a significant role. In public transport networks, uncertainty due to waiting time and, recently, the COVID-19 pandemic possibly induce the most frustration and anxiety. Therefore, with the overarching aim of making public transport a viable and satisfying option, this thesis is dedicated to modelling and analysing the impact of such pervasive uncertainty on public transport travellers' route choice behaviour.

About the Author

Sanmay Shelat conducted his PhD research at the department of Transport and Planning, Delft University of Technology. He holds bachelor's and master's degrees in Civil Engineering (Nirma University) and Transport and Planning (TU Delft), and currently works as a Data Scientist.

TRAIL Research School ISBN 978-90-5584-327-5

Erasmus University Rotterdam · Radboud University · rijksuniversiteit groningen

TU Delft Delft University of Technology · UNIVERSITY OF TWENTE. · TU/e Technische Universiteit Eindhoven University of Technology

TRAIL THESIS SERIES T2023/8

Sanmay Shelat Route Choice Behaviour under Uncertainty in Public Transport Networks

Route Choice Behaviour under Uncertainty in Public Transport Networks

Stated and Revealed Preference Analyses

Sanmay Shelat

Route Choice Behaviour under Uncertainty in Public Transport Networks

Stated and Revealed Preference Analyses

Sanmay Shelat

Delft University of Technology

This research has been funded by the My-TRAC (My Travel Companion) project (EU-H2020 grant number 777640) and the Amsterdam Institute for Advanced Metropolitan Solutions

Cover photos and composition: Sanmay Shelat

Route Choice Behaviour under Uncertainty in Public Transport Networks

Stated and Revealed Preference Analyses

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Thursday 29 June 2023 at 10:00 o' clock

by

Sanmay SHELAT
Master of Science in Civil Engineering, Delft University of Technology, Netherlands
born in Vadodara, India

This dissertation has been approved by the promotors and copromotor.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof. dr. ir. J.W.C. van Lint	Delft University of Technology, promotor
Prof. dr. O. Cats	Delft University of Technology, promotor
Dr. ir. N. van Oort	Delft University of Technology, copromotor

Independent members:

Prof. dr. ir. C.G. Chorus	Delft University of Technology
Prof. dr. D.J. Graham	Imperial College London, United Kingdom
Prof. dr. I. Erev	Technion – Israel Institute of Technology, Israel
Dr. J.-D. Schmöcker	Kyoto University, Japan
Prof. dr. ir. S.P. Hoogendoorn	Delft University of Technology (reserve member)

TRAIL Thesis Series no. T2023/8, the Netherlands Research School TRAIL

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

ISBN: 978-90-5584-327-5

Printed in the Netherlands

Acknowledgements

In my application for the Transport and Planning master's programme at Delft, I remember specifically noting that I wanted to research travel behaviour. Although my master thesis took me in a different direction, as luck would have it, near its end this PhD project on choice modelling was offered to me. I am truly fortunate to have a research topic that, through good (acceptances, conference trips!) and bad (frustrating reviewers, the pandemic) times, drove me to wake up curious every single day. But no person is an island (as I also sorely realised during the lockdowns) and my PhD journey wouldn't have been as wonderful without the support of those around me.

To my amazing supervision team: Hans, Oded, and Niels; I am extremely thankful for your constant guidance and empathetic support. Working with your unique approaches has been an incredible experience and I hope that I have managed to imbibe at least some of your qualities. Hans, thank you for creating such a fabulous atmosphere for research in Dittlab (and bringing together such wonderful people). I also appreciate how welcome I have always felt to openly speak my mind in your office, which I have definitely taken advantage of whenever we discussed My-TRAC! And although I have often been wary of the number of editing notes you would have on my papers, I can confidently say that they have influenced my writing style the most. Niels, thank you for always reminding me to balance my work and life—I always walked away from our meetings a little calmer. Your networking skills are inspiring and encouraged me to collaborate with colleagues in the Smart Public Transport Lab. Finally, Oded, thank you for being a mentor to me since I did the honour's project with you as a master's degree student. You have been my go-to person for any professional advice and I can honestly say that I am in awe of your structured and methodical approach to all topics. You have made me a better researcher.

Before I move on to thanking a fairly long list of people, I would like to take a moment to extend my heartfelt appreciation to the committee members for accepting our invitation and assessing my thesis. Each of your research is truly inspiring and it is an honour to have members with such fantastic and diverse backgrounds on my defence committee.

Being a part of the Transport and Planning department has been amazing. Here, Dittlab (room 4.02) was my home. I cannot thank Panchamy, Ding, Tin, Leonie (who had started a couple months earlier than me), and Simeon enough for how welcome you made me feel when I joined right after I finished my master thesis. Thank you, Panchamy for being a great (opposite) desk neighbour and for being my Indian connection in the lab (so I could always intervene with my Gujarati perspective whenever you described something about India!!). Leonie, who I missed sorely as my PhD partner (and side desk neighbour) later, thank you for always having a positive outlook on things (even on My-TRAC stuff!) and being ever-ready to teach us Dutch. Tin and Ding, thank you for being such cheerful presences in the lab. (Special thanks to Tin for helping out with all my questions about the Dittlab backend system!). Guopeng, Zahra, and Ehab who joined the lab after me—thank you for being such great additions to the group. Guopeng and Zahra, thank you for regularly coming to the lab during the semi-lockdowns—it was so good not to be working alone. Ehab, thank you for being so easy going and hanging out with us PhD folk! Thank you all for the fun conversations, philosophical discussions, rants, etc. and making Dittlab the wonderful place to work it is. I would also like to thank Viktoriya, Salil, Ali, the visiting researchers (Bobin, Kai, and others), and the newer PhDs. Special thanks to Kristel for being a great addition to the (quickly dwindling!) My-TRAC team and for our coffee chats in the work-from-home era after she left the lab.

Outside Dittlab, I would like to especially thank Malvika for always reaching out with a friendly conversation (especially for keeping in touch after I left for Hamburg), encouraging me to wrap up the final bits of the thesis, and our research collaboration; Maria and Konstanze for great advice—whether it's on my research or my job search; and to Danique for reaching out to collaborate on a paper. Thank you also to Subodh, Menno, Nejc, Arjan, Peyman, Jishnu, Giulia, Yan, Lara, Roy, Alexandra, and others for being great colleagues. To visiting researchers, Arek and Carlos, thank you for our research collaborations and discussions. I would be amiss if I didn't thank the department secretariat, Moreen, Dehlaila, Priscilla, Marije (and others) for all the administrative support over the years, and Conchita from TRAIL for making sure the very last stage of the PhD was as smooth as possible. Finally, special thanks to Sander van Cranenburgh from the TPM faculty for being on my go-no-go committee, collaborating on the pandemic paper, and inviting me to tutor on a very interesting student assignment.

I was also lucky to have great master project/thesis students from TIL and T&P; thank you all for being my first students! On this, I would also like to express gratitude to Winnie again for her exceptional guidance as my master thesis supervisor. I have tried to emulate you in my own supervisory roles.

This PhD was sponsored by the My-TRAC project (a European Union Horizon 2020 initiative) and the Amsterdam Institute for Advanced Metropolitan Solutions. I am thankful for their funding grants. I also thank HTM and GVB for making anonymised public transport data available for research and the TU Delft Transport Institute for funding the pandemic data collection. The 3-year My-TRAC project was a significant aspect of this PhD and despite some frustrations, I will fondly remember the fun times and the interesting work we did. Special thanks to Leonie for making the meetings bearable (and making me laugh so much during them!) and being a fellow sightseer on our trips. Thank you also to Viktoriya, Ehab, and Kristel (and also Ding for a little bit!) for collaborating on the project from the TU Delft end. I would like to also thank all My-TRAC partners but in particular Eleni Mantouka, Konstantinos

Mavromatis, Alexandros Papacharalampous, and Ismini Strompou with whom I closely collaborated on the deliverables I was leading.

The very final bits of this thesis were written on weekends after I shifted to Hamburg to work as a Data Scientist for FREENOW. I would like to thank, Devangana, Ines, Vikram, Alessandro, Bruno, Gil, Sam, Jan, Marcus, Priscila, Ali, Shreyash, Arpita (amongst others) for being fantastic colleagues/friends and making my transition to the industry and Hamburg so smooth. Special thanks to Prachi, who is both a colleague and a neighbour, for all the dinners and work gossip!

Given how bad I am at staying in touch with friends, I should thank Shiksha and Parshwa for sticking with me for more than 10-15 years (even though we are so far)! Dugar, I am finally getting the same title as you (although I'll still be absolutely useless if someone asks for a doctor)! Parshwa, thanks for all the long phone calls and discussions on automobiles (and, of course, for the amazing coffee and popcorn in Detroit; we should come visit you again soon!). Thank you both for all your support throughout this journey.

Finally, I would like to thank Rathore and my parents but I find that it is often most difficult to thank those closest to us. There are simply too many things I am grateful for and (after thinking about it for ~~hours~~ days) I have concluded that my words will simply not do justice. But know that I wouldn't be here without your constant support, encouragement, and love.

Sanmay
Jersey City, May 2023

Contents

Subjective Waiting Time Uncertainty

COVID-19 related Uncertainty

Conclusions and Recommendations

References

Summary

Samenvatting

About the Author

List of Publications

TRAIL Thesis Series

Introduction

1 Research motivation

Arguably, nearly all real-world decisions are made under uncertainty. The anatomy of any choice consists of the alternatives, their properties, and the set of outcomes possible for each alternative under various eventualities. These eventualities are the realizations of uncertainties in the choice situations. Consider somebody deciding whether to go for a long bicycle ride tomorrow. The cyclist's options may be whether to go cycling or not. Their outcome, in terms of joy derived from the ride, depends on the weather tomorrow, which is clearly uncertain. The uncertainty can be defined by the various weather 'events' possible; for instance, it could be sunny or rainy. Additionally, the source of uncertainty could also be regarding the availability of options (whether they could even go cycling tomorrow or not), the decisions of others (whether it will be too crowded on the cycle paths), or the impact of outcomes (whether they will enjoy cycling as much as they think).

Uncertainty, when represented by a set of events, is often classified on the basis of whether objective probabilities exist for these events (Knight, 1921). Uncertainty is objective when probabilities for different events are available to and trusted by decision-makers, either because they are integrated in the choice situation itself (e.g., in a casino roulette game) or because there is a consensus regarding them (e.g., trusted expert opinion of striking oil). Most decisions, however, are inherently associated with subjective uncertainty where the decision-makers' personal beliefs regarding the likelihood of events play a much larger role. Such personal beliefs may be informed by a number of factors, including available information, experience, trust, decisions of others, and so on. The pervasive nature of such uncertainty in everyday decision-making means that often intuition reigns and decisions lack retrospection (Gilboa, 2012).

Uncertainty has a significant impact on decisions because of its intricate relationship with emotion and affect. Not only do the emotional and affective contexts (i.e., state when making a decision) influence the perception of uncertainty (Loewenstein and Lerner, 2003) but the reverse is also true (Anderson et al., 2019). Uncertainty is largely associated with negative affective responses and has been strongly linked with anxiety (Carleton, 2016) and stress (Buhr and Dugas, 2002).

Travelling is one such sequence of decisions where uncertainty is present, largely subjective, and usually dealt with not by carefully weighting probabilities but using intuition. Deciding whether to travel, to go by car or bus, and to take the freeway or urban roads are all associated with uncertainty. A large number of studies have demonstrated the importance of travel time reliability on various decisions, including mode and route choice (Bates et al., 2001; Börjesson et al., 2012; Carrion and Levinson, 2012; de Palma and Picard, 2005; Small et al., 1999; Small et al., 2005). While travel time is typically the main source of uncertainty in this domain, others may arise with newer forms of mobility, such as availability of shared vehicles (Hsu et al., 2016), or following security events (Elias et al., 2013; Holguín-Veras et al., 2003) and disease outbreaks (Lau et al., 2003; Wang, 2014) which may cause travellers to feel uncertain about their safety.

Public transport is different from personal modes of transportation, such as the car, in that it (typically) has fixed schedules, serves discrete locations, and is shared. Because of these properties, the two main sources of uncertainty associated with travelling, reliability and safety, are particularly important for public transport passengers. In fact, several authors have identified these aspects as being the most basic of public transport travellers' needs and potential dissatisfiers (Allen et al., 2019; van Hagen, 2011). With respect to reliability and impact on overall satisfaction, waiting time is perhaps the most important aspect of travelling with public transport (Abenoza et al., 2018). Waiting is inherently frustrating (Maister, 1985) and its inevitability in public transport and travellers' apparent lack of control can induce stress and dissatisfaction (Cantwell et al., 2009).

While there are numerous safety aspects related to public transport due to its shared nature, such as at public transport stations and on vehicles (and arguably during access and egress trips as well; e.g., walking home from the station at night), with the outbreak of the COVID-19 pandemic in early March 2020, the risk of contracting or spreading the virus on public transport rose to prominence in travellers' and authorities' minds alike (Shelat et al., 2022). Like previous epidemics (Lau et al., 2003; Rubin et al., 2009), this pandemic too induced anxiety and public transport avoidance (de Haas et al., 2020; Gerhold, 2020). This has been compounded by unprecedentedly sustained uncertainty as different parts of the world suffered seemingly endless infection waves.

Keeping in mind the overarching aim of reducing dependence on unsustainable motorized individual transport modes by offering public transport as a viable and satisfying alternative, this thesis analyses the impact of pervasive uncertainty on route choice behaviour within public transport networks. That is, we analyse the choice of (a combination of) public transport lines or vehicles in order to obtain insights into traveller preferences. Given their prominent impacts, the focus is on behaviour under waiting time uncertainty and in the context of the uncertainty presented by COVID-19.

2 Research gaps

Although a number of studies have demonstrated and discussed the importance of travel time uncertainty for route choice decisions, potential for improving the external validity of these behavioural models remains. This is particularly true in the context of waiting time uncertainty in public transport. This thesis addresses different research gaps (RG) related to this potential by focussing on two important dimensions: the source of behavioural data and the type of uncertainty analysed. Additionally, it also undertakes new research avenues that result from the impact of the COVID-19 pandemic on public transport travel behaviour.

2.1 Source of behavioural data

The impact of travel time uncertainty on travellers' choices has been analysed using behavioural data obtained from (i) stated choices, that is, between hypothetical alternatives in surveys or laboratory experiments, or from (ii) reported or passively observed choices in real-world trips. While the former mode of data collection offers more experimental control and has traditionally been easier to collect, it is likely to suffer from hypothetical bias. Hypothetical bias is a catch-all term for suspected deviations of reported behaviours or values in stated preference studies from those potentially evident in the market. For behavioural studies, this bias likely arises due to varying degrees of: (i) absence or misalignment of consequences of respondents' choices, that is, improper incentives, (ii) removal from real-world contexts under which the choice would normally be made, or (iii) strategic responses aimed at influencing the perceived policy outcome of the study (Haghani et al., 2021b). Readers are referred to Hensher (2010) and Haghani et al. (2021a, 2021b) for a detailed review.

Although, due to the above factors, stated choice surveys typically offer the lowest level of realism, the vast majority of behavioural studies in the domain of travel time uncertainty have used observations from such surveys (e.g., Small et al. (1999), Schwanen and Ettema (2009), Tilahun and Levinson (2010)). Laboratory experiments, while also collecting hypothetical choices, reduce hypothetical bias by emulating the choice context via simulation, enabling learning by giving feedback for every choice made, and/or awarding choice-based incentives or penalties. A few laboratory experiments have analysed travel behaviour under uncertainty but they have largely focussed on road traffic networks and learning mechanisms (e.g., Ramos et al. (2011), Ben-Elia and Avineri (2015)). Moreover, laboratory experiments incentivising revelation of true preferences by linking rewards to choices are not common in studies of travel behaviour.

Revealed preferences have the least hypothetical bias; yet only a handful of studies have studied revealed preferences related to travel time uncertainty and have focussed on road users. These studies have used reported perceptions of travel time (e.g., Peer et al. (2014)), active observations obtained by asking drivers to carry out specific real-world trips (e.g., Carrion and Levinson (2013)¸ Dixit et al. (2019)), or passive observations of choices between tolled, reliable routes and untolled, unreliable routes (e.g., Lam and Small (2001), Brownstone and Small (2005)). The latter—passive choice observations or naturalistic data—have the highest realism (Haghani et al., 2021b) but have usually been the most difficult to collect.

One of the key difficulties with using revealed preferences is the lack of experimental control: researchers not only have to identify situations with sufficient variation of the variable of interest across alternatives (Carrion and Levinson, 2012), but they also have to infer which alternatives are considered by the decision-makers. For active observations, one might simply ask the participant which alternatives they considered; and for passive observations, the set of all chosen alternatives in a given choice situation may be assumed to be its considered choice set. However, the self-reporting in the former is subject to various errors (Hoogendoorn-Lanser, 2005) and the latter precludes the possibility to include non-chosen but considered alternatives (Raveau, 2017). As such both of these direct-identification methods lack transferability as similar insights may not be available for other public transport networks (Ton et al., 2018). Because of this a number of approaches have been put forward for route choice set generation but none are completely satisfactory. These approaches either make strong, simplistic behavioural assumptions—such as those in shortest-path methodologies—or depend on the researchers' (or experts') judgements regarding travel behaviour—such as those in link-labelling or uncalibrated constrained enumeration methodologies (Bovy, 2009; Prato, 2009). Thus, a route choice set generation methodology that produces transferable parameters, requires minimal assumptions regarding traveller behaviour, and can be calibrated using revealed choice observations is missing (RG1). Moreover, transferable behavioural insights regarding the consideration set formation need to be derived (RG2).

Another difficulty in collecting naturalistic data for route choice studies has been the effort and expense related to observing choices and the values of attributes of different alternatives (Carrion and Levinson, 2012). This has become considerably easier in public transport networks with automatic fare collection (AFC) systems (e.g., those using smart cards) which collect data that can then be used to trace travellers' routing choices and the attribute values of various alternatives. Although a number of studies have used such data to study route choice behaviour in public transport networks (e.g., Guo and Wilson (2011), Jánošíková et al. (2014), Yap et al. (2018)) only one previous study, (Leahy et al., 2016), has specifically analysed the effect of travel time uncertainty. However, it examines the impact of total travel time uncertainty and only uses a sample of AFC transactions over the time period of their study. Thus, an explicit analysis of the impact of waiting time uncertainty on route choice behaviour in public transport networks remains to be conducted (RG3). Note that all revealed preference studies on this topic have analysed the impact of *objective* travel time uncertainty. Empirical distributions of travel time are summarised through various statistical measures that can be broadly classified into statistical ranges and buffer times (van Lint et al., 2008). While some studies have even compared different statistical indicators for total travel time (e.g., Bogers et al. (2008), Alemazkoor et al. (2015), Leahy et al. (2016); Tilahun and Levinson (2010)), there is no consensus regarding the best representation and specific insights for waiting time uncertainty in public transport networks are missing (RG4).

RG1 A route choice set generation methodology that can be calibrated using revealed choice observations using minimal behavioural assumptions to produce transferable behavioural parameters

RG2 Transferable behavioural insights regarding considered route choice set formation from a real-world public transport network

RG3 An explicit analysis of the impact of waiting time uncertainty on route choice behaviour in public transport networks

RG4 Comparison of different statistical representations of waiting time uncertainty in route choice models for public transport networks

2.2 Type of uncertainty analysed

Travellers in public transport networks rely on published schedules or frequencies to draw their expectations regarding waiting times. Since vehicles are not always on time, waiting times are associated with subjective uncertainty which travellers may gauge based on past experiences with the system, their own personal characteristics (e.g., pessimism), and other contextual variables (e.g., bad weather) (Cheng and Tsai, 2014; Kugler et al., 2012). While real-time information can offset some of this uncertainty, it is also liable to be distorted by travellers' personal, subjective beliefs. Although waiting time uncertainty is subjective, the existing body of literature has largely studied its impact on route choice using objective probabilities (or probability distributions) that are normally not available to travellers.

The type of uncertainty—(objective) risk, simulated uncertainty, or natural ambiguity—under which decisions have been observed is closely related to the data collection method. Stated choice experiments have typically presented hypothetical route choices with objective distributions of travel times. A number of presentation methods have been used for such distributions, such as displaying an array of equiprobable travel times (e.g., Swierstra et al. (2017)) or indicating the probability of a fixed delay (e.g., Schwanen and Ettema (2009)), yet conveying objective probabilities to respondents who are not used to receiving such information remains difficult (Bates et al., 2001; Carrion and Levinson, 2012).

With laboratory experiments, a number of studies have observed behaviour under simulated or artificial uncertainty. Typically in such experiments respondents face repeated choices between routes for which imperfect or no information regarding travel times is provided and the true travel times of the chosen (and sometimes also non-chosen) routes is revealed after the choice has been made to allow respondents to develop likelihood beliefs just as travellers do in the real-world (e.g., de Moraes Ramos et al. (2013), Avineri and Prashker (2006), Chorus et al. (2008)). However, the uncertainty simulated is usually unrelated to that in real-world public transport networks but is instead tuned to fulfil analysis requirements such as studying a given range of values or estimating a specific parameter.

Finally, route choice decisions under natural ambiguity—subjective uncertainty (also known as ambiguity) for natural (i.e., not artificial but real-world) events (Baillon et al., 2018)—have only been studied using revealed preferences from the field experiments mentioned in the previous section. However, as discussed in that section, although the travellers made choices under natural ambiguity, the analyses on these choices have assumed the uncertainty to be objective that is, as if travellers had access to the probability distributions (or statistical measures thereof) of travel times (e.g., Small et al. (2005)).

Thus, a method to analyse route choice decisions made under natural ambiguity without assuming the availability of objective probability distributions is missing (RG5). This is especially true for waiting time related uncertainty in public transport networks because studies where decisions are at least observed under some kind of uncertainty (i.e., laboratory and field experiments) are mainly car traffic focussed. Such a method would result in a concrete evaluation of travellers' subjective beliefs regarding waiting time uncertainty for real-world public transport networks as well as describe heterogeneity in these beliefs (RG6).

RG5 An experimentation method to observe and analyse public transport route choice decisions under natural waiting time ambiguity

RG6 Evaluation of travellers' subjective beliefs regarding waiting time uncertainty (and heterogeneity therein) in real-world public transport networks

2.3 COVID-19 and travel behaviour

Since the coronavirus causing COVID-19 spreads through airborne means, exposure is typically through proximity to infectious persons. As such, transmissions are more successful with closer proximity and longer exposure times to infectious persons (Hu et al., 2020; Prather et al., 2020). This is why public transport, which involves moving a large number of people in dense, enclosed spaces, poses a transmission risk for COVID-19[1]. As a result of governmental and individual interventions to avoid such transmission, the COVID-19 pandemic (2020–) has resulted in significant disruption to public transport ridership (Shelat et al., 2022). Similar impacts have been found in previous (respiratory) disease outbreaks, such as SARS (2002–2004), MERS (2012), and swine flu (2009). Moreover, survey responses during these epidemics indicated that people perceived public transport avoidance as an effective risk mitigation strategy. Given the relatively wider and more sustained impact of COVID-19, travellers in this pandemic are also likely to change their behaviour and focus on factors contributing to disease transmission.

In the context of public transport travel during (both past and ongoing) disease outbreaks, research has largely focussed on analysing (i) aggregate statistics of ridership drops and mode shifts (e.g., Bucsky (2020); Teixeira and Lopes (2020)), (ii) travel/activity pattern changes (e.g., Beck and Hensher (2020); de Haas et al. (2020); Engle et al. (2020); Kim et al. (2017a); Lau et al. (2003)), and (iii) Likert scale-based measures of transmission risk perception (e.g., Dryhurst et al. (2020); Gerhold (2020); Rubin et al. (2009)). Even though—as a result of its unprecedented scale—a large body of literature on COVID-19 exists, only a handful of studies (e.g., Aaditya and Rahul (2021); Aghabayk et al. (2021); Cho and Park (2021)) have directly modelled travel choices or analysed how public transport travellers trade-off various transmission risk determinants with travel time attributes (RG7)[2].

In particular, travellers can be expected to change their valuations for on-board crowding and in-vehicle times; factors that are directly related to exposure risk. Furthermore, at a macroscopic level, a higher proportion of infectious (i.e., capable of spreading the disease) people in the population increases the probability of exposure; and at the microscopic level, weaker immunity makes one more susceptible to the disease, increasing the likelihood of transmission. Thus, in addition to trip-specific factors, travellers may also take into account the general COVID-19 situation in their region as well as the susceptibility of themselves and their loved ones.

As with any new disease, in its early stages, there was a great deal of uncertainty around COVID-19. Although consensus regarding factors causing the disease and the risk it posed grew amongst the medical community, perceptions in the general population remained divergent (Smail et al., 2021); a problem further exacerbated by misinformation spread via social media (Bridgman et al., 2020). Since beliefs regarding the risks imposed by the pandemic (and attitudes towards these risks) are so varied, it would be pertinent to measure heterogeneity in public transport travellers' behaviour, particularly with respect to factors related to the disease, such as on-board crowding (RG8). Assessing how different travellers will behave under different COVID-19 situations is not only useful in communicating and designing transmission-preventing policies but is also critical for eventually bringing back public transport ridership levels.

RG7 A choice analysis of travellers' trade-offs in public transport between COVID-19 transmission risk determinants and travel time attributes

RG8 Evaluation of heterogeneity in public transport travellers' choice behaviour in context of the COVID-19 pandemic

[1] Although this risk can be attenuated by better respiratory protective measures, such as masks and HVAC filters.
[2] It should be noted that as this is a very active domain of research, the number of studies is expected to rise.

3 Research objective, scope, and questions

The objective and scope of this thesis follows from the research motivation and gaps in literature outlined above. The overarching objective is to **model and analyse the impact of the pervasive uncertainty on public transport travellers' route choice behaviour**. To do this, we use observations from stated choices experiments and passively collected, naturalistic data from automatic fare collection systems. The focus is on analysing the impact of waiting time uncertainty and the changes in behaviour brought about by the uncertainty presented by the COVID-19 pandemic. Although the case studies presented in the thesis use observations from travellers in the Netherlands, the methodologies and, to a limited extent, findings are applicable to developed public transport networks around the world.

To achieve this objective, we aim to answer the following four research questions (RQ). First, to model travel behaviour from passively observed route choices (which have no hypothetical bias), choice sets must be inferred. While several choice set generation methods have been proposed in the literature, they are not completely satisfactory, leading to the need for a new parsimonious and transferable methodology.

RQ1 How to infer route choice sets from passively observed choices using minimal assumptions and producing transferable behavioural insights? (Addressing RG1, RG2)

The next two questions are about analysing the impact of waiting time uncertainty on travellers' route choice behaviour. First, with the conventional assumption that travellers are aware of the empirical distribution of waiting time (i.e., objective uncertainty), and then relaxing this assumption to evaluate travellers' subjective beliefs regarding waiting time uncertainty and subsequently their route choices under the natural ambiguity that exists in the real-world.

RQ2 What is the impact of waiting time uncertainty, expressed as different statistical representations of its historical values, on route choice behaviour? (RG3, RG4)

RQ3 How to evaluate subjective beliefs regarding waiting time for route choices made under the ambiguity that is naturally present in the real-world? (RG5, RG6)

Finally, the COVID-19 pandemic (2020–) has imposed new uncertainties everywhere, potentially also resulting in changes in how travellers perceive various components of public transport travel. This leads us to our final research question:

RQ4 What are the impacts of COVID-19 transmission risk determinants on public transport travellers' route choice behaviour? (RG7, RG8)

4 Research approach

To fulfil the research objective effectively, this thesis addresses each of the three groups of research gaps identified previously. First, for source of behavioural data, a mix of stated and revealed preferences is used. The former have been collected from (online) stated choice experiments with travellers in the Dutch railways, while the latter have been obtained from passively collected smart card observations from the public transport services of The Hague and Amsterdam. Second, behaviour is modelled under different assumptions of uncertainty. The uncertainty under which the choice has been made depends on the source of the behavioural data. For stated choices, this depends on the type of experiment whereas for smart card observations, choices have been made under the natural ambiguity that exists in the system. The behavioural analysis can then either ignore the uncertainty, include it as if it were objective, or account for its subjective nature. Finally, research gaps related to the new uncertainties brought about by the COVID-19 pandemic are directly addressed by the last research question. Next, we discuss the approach taken for each research question.

The first question (RQ1) seeks to overcome shortcomings of current choice set generation methods in order to empower the use of passively collected data for route choice analysis in public transport networks. To do this, we model consideration set formation using a non-compensatory decision heuristic (elimination-by-aspects) as is common in marketing literature (Hauser, 2014). This heuristic is calibrated for a given public transport network in two parts: (i) a constrained enumeration of feasible routes using the network topology and schedule (and two logical constraints), and (ii) a brute force calculation of the parameters in the elimination-by-aspects model (i.e., attribute ranking and thresholds). The model parameters are calculated to optimize the balance between coverage—the proportion of observed routes that have been generated (i.e., recall)—and efficiency—the proportion of generated routes that have been observed (i.e., precision). The calibration methodology is used to generate route choice sets from smart card observations in The Hague tram and bus networks. To develop our understanding of the methodology, we examine the intermediate analysis as well as final results in detail.

To answer the remaining research questions, a series of choice models are estimated. All choice analyses in this thesis are underpinned by the random utility maximization paradigm as is conventional in transportation modelling. The paradigm assumes that decision-makers make choices that maximize utility, which is a latent variable consisting of systematic and random components. In particular, we employ multinomial logit models which assume Gumbel distributed random components. Where panel information is available, taste heterogeneity is analysed with latent class choice models which are useful for identifying specific segments of travellers. Both models are introduced in the methodology sections of chapters (e.g., section 5.3) employing them.

In the next two research questions (RQ2, RQ3), the impact of waiting time uncertainty is studied. First, route choices made under natural ambiguity are analysed as if empirical distributions of waiting time were available. We use the tram and bus networks of The Hague as a case study, combining smart card, automatic vehicle location, and general transit feed specification (GTFS) data of the networks. Before choice analyses can be performed, the following steps are required: data preparation (including transfer inference), choice set identification, and attribute assignment. Special attention is paid to analysing waiting time which is considered separately for the origin and transfer stops. Uncertainty in waiting time is characterised by the reduced-form approach (Börjesson et al., 2012) which divides the uncertainty into regular and irregular deviations from scheduled values. We then estimate choice models with various representations of irregular deviations and compare attribute ratios and model fits. Finally, model validation is performed using a *k*-fold procedure.

Next, we propose a method to quantify travellers' evaluation of waiting time uncertainty under natural ambiguity. In order to guide the scope of the analysis, first a theoretical framework of decision-making under uncertainty is sketched from which we focus on quantifying the effects of (i) travellers' attitudes and perceptions and (ii) situational contexts. To do this, observations from a realistic choice situation (occurring in many real-world public transport networks) are used to quantify subjective uncertainty beliefs in terms of a certainty equivalent: a risk-less value for any situation with uncertain outcomes—a gamble—such that the decision-maker is indifferent between the risk-less value and the gamble. We employ this method in two case studies.

For the first case study, the choice situation is contextualised for the Dutch railways within a stated choice experiment. The experiment is carefully designed so that respondents' choices reflect their subjective beliefs regarding the real-world network and so that we are able to measure a potentially wide range of beliefs. The choice analysis includes estimating multinomial logit models, with and without the certainty equivalent term, to assess the impact of explicitly accounting for waiting time uncertainty (also using *k*-fold cross-validation for model fit); and a latent class choice model to capture decision-maker heterogeneity. The latent class choice model is followed by a posterior analysis of its class membership whereby we analyse the distribution of attitudinal characteristics in the three classes.

The second case study analyses smart card observations (in combination with vehicle location and GTFS data) from Amsterdam's tram and bus networks. Unlike the controlled experiment above, in this naturalistic study, observability of attributes is a key concern. Therefore, we define a subset of the smart card data wherein all key choice attributes are completely observed. Following data preparation similar to that in RQ2, choice sets are enumerated using techniques (e.g., topological representation, attribute assignment) comparable to those developed for RQ1 and filtered based on design considerations analogous to the case study above. The choice analysis is performed with multinomial logit models which explicitly take into account differences not present in the controlled experiment and is also followed by a *k*-fold cross-validation.

Finally, for the last research question (RQ4), another online survey is conducted with travellers of the Dutch railways. The survey, distributed at the end of May 2020, includes a stated choice experiment and collects travellers' socio-demographics, mobility choices, and (Likert scale) pandemic-related attitudes and opinions. In the experiment design, emphasis is placed on carefully communicating on-board crowding for route alternatives and contextual information regarding COVID-19. As we are specifically interested in the heterogeneity in traveller behaviour under pandemic-related uncertainties, we estimate latent classes of behaviour and perform a posterior analysis to derive class profiles of the collected Likert scale measures. This is followed by an in-depth comparison of these classes.

5 Main contributions

By addressing the research gaps discussed in section 2, a number of scientific contributions with societal relevance have been made.

5.1 Scientific contributions

The primary scientific contributions of this thesis are:

Development of an assumption-parsimonious and transferable route choice generation methodology for public transport networks (chapter 2; RG1): A route choice set generation methodology is developed where consideration set formation is modelled as an elimination-by-aspects process and the parameters of this model are calibrated with passively observed choices. The non-compensatory decision heuristic used as the model is well-aligned to the actual cognitive process and requires minimal behavioural assumptions on the part of the modeller. The calibration results in actionable insights and the estimated parameters can be used for choice set identification in similar public transport systems.

Analysis of the impact of waiting time uncertainty on route choices in public transport networks using revealed preferences (chapters 3 and 4; RG 3,): The impact of waiting time uncertainty on route choice is explicitly estimated in discrete choice models using observations from passively collected smart card data (in combination with vehicle location data, network topology, and schedule information). Using revealed preferences overcomes hypothetical bias related shortcomings of previous stated preference experiments on the impact of reliability.

Suitability comparison of different statistical representations of waiting time uncertainty in route choice models for public transport networks (chapter 3; RG 4): Statistical summaries of historical values of travel times have been typically used to represent reliability in route choice models. Using vehicle location data and schedule information, different measures of types statistical range and buffer times are calculated for waiting time. Their suitability in representing travellers' perceptions of the uncertainty is then assessed by using these measures in discrete choice models and comparing their model fits (and the face validity of other coefficient ratios).

Development of an experimentation method to observe and analyse public transport route choice behaviour under natural waiting time ambiguity (chapter 4; RG 5): An experimentation method involving a realistic route choice situation is developed to enable quantification of travellers' subjective beliefs regarding waiting time uncertainty and the impact of situational contexts thereon in terms of a certainty equivalent. The choice situation occurs fairly commonly in public transport networks, allowing us to generalize findings beyond the scope of the specific situation. Furthermore, the experimentation method lends itself to both controlled and natural route choice experiments as demonstrated in the case studies which use stated and passively observed choices, respectively.

Analysis of the (heterogeneity in the) impact of COVID-19 transmission risk determinants on public transport route choice behaviour (chapter 5; RG7, 8): A choice analysis of the trade-offs between COVID-19 transmission risk determinants, particularly, on-board crowding, exposure duration, and prevalent infection rate, and travel time attributes is performed. The analysis is conducted in the early stages of the pandemic for travellers of the Dutch railways. Latent classes of travellers are estimated and their crowding valuations and propensities to avoid public transport given a prevalent infection rate are compared. Furthermore, class profiles composed of socio-demographics and pandemic-related attitudes and opinions are derived.

5.2 Societal relevance

This thesis makes contributions that will help planners, operators, and policy-makers improve public transport services and make it a more attractive alternative to other modes. This is achieved by enabling better forecasting through improved models, offering improved behavioural assessment through better experimentation methods, and providing actionable insights arising from our case studies:

Improved models: By explicitly including waiting time uncertainty and using revealed preference data sources, we estimate public transport route choice models with better performance and reduced hypothetical bias. Furthermore, explicit choice analysis of travel behaviour under the new uncertainties presented by the COVID-19 pandemic has led to updated crowding valuations and infection rate based propensity-to-travel estimates. Planners can incorporate this information into demand estimation and assignment models by including the suggested attributes (e.g., waiting time irregular deviations, certainty equivalent) or they can directly use the estimated coefficient ratios (and comparisons to previous appraisals either pre-pandemic or without the said attributes).

Better experimentation methods: The proposed route choice set generation methodology enables experiments with revealed choices when direct identification of choice sets is not possible or suitable. While the method is demonstrated on a large dataset of passively observed choices, it can also be used for limited actively observed choices as demonstrated in Ton et al. (2020). Planners can use this methodology to shift towards using behavioural data sources with smaller hypothetical bias for their choice analyses.

With the proposed experimentation method for capturing travellers' subjective beliefs regarding waiting time, planners can explicitly account for uncertainty. Moreover, since waiting time uncertainty is critically linked to public transport travellers' satisfaction (Abenoza et al., 2018), the snapshots of certainty equivalents captured by this method can be used by operators as an indicator to analyse which situational or environmental variables cause higher uncertainty perception; or test within randomized experiments which measures are effective in lowering anxiety.

Actionable insights: The models and case studies in the thesis give insights that public transport providers can act upon directly. The parameters of the proposed route choice set generation methodology give insights into which routes are not likely to be considered. This can add to the learnings obtained from traditional fully-compensatory models. Moreover, journey planner applications can use these insights to provide travellers with useful mode/route recommendations.

Such applications can also use our analysis of waiting time uncertainty perceptions to nudge travellers into making choices that are more optimal for them and the network. Moreover, as noted above, operators can use this experiment (which can be automated when smart card data is used) to understand and remedy situations that generate a higher perception of uncertainty in travellers.

Finally, while we note the rapidly changing nature of the pandemic, our study contains important insights about travellers' response to COVID-19 risk determinants for policy-makers. The updated choice parameters can be used to re-plan supply and identified latent clusters of behaviour can help tailor marketing campaigns that balance bringing travellers back to public transport and educating them about the need for respiratory protective measures. Moreover, even if the choice parameters themselves become outdated they are integral in understanding the evolution of travel behaviour and can support proactive action in the next one.

6 Research Context

This thesis was part of and supported by My-TRAC (my travel companion), a European Union Horizon 2020 project[3] with a consortium of academic, consultancy, and operator partners in the

[3] Support from ancilliary funding sources (Amsterdam Institute for Metropolitan Solutions, Transport Institute of TU Delft) and data providers (HTM, GVB) is also acknowledged.

Netherlands, Greece, Spain, and Portugal. The project aim was to develop a '[mobile] application for seamless transport and an ecosystem of models and algorithms for public transport—public transport user choice simulation, data analytics and affective computing'[4]. A substantial portion of the research in this thesis was conducted in the context of project tasks relating to the development of a choice modelling framework that extends the state-of-the-art in simulating and predicting traveller behaviour and providing recommendations to public transport operators. This thesis contributes to these tasks and the overall aim by (i) focussing on understanding (heterogeneity in) behaviour under uncertainty so that this companion application can assist people with making more 'rational' choices and (ii) developing methodologies that can be used to draw inferences from passively collected data that can be useful for operator planning.

7 Outline

Figure 1 shows an overview of the thesis, arranging the research contributions (chapters 2–5) into three parts and dividing each analysis based on the source of behavioural data. As discussed above, two types of behavioural data have been collected and analysed in this thesis: (i) stated choices and (ii) smart card observations.

The first part, which consists of chapter 2, focusses on enabling the use of passively observed route choices for behavioural analysis. Since, such naturalistic data has no hypothetical bias, we are interested in using it but are obstructed by the lack of experimental control, particularly the inability to observe considered alternatives. Therefore, to answer RQ1, a novel choice set generation methodology is proposed in this chapter. The proposed methodology uses a non-compensatory decision model—in line with theoretical models of choice set formation—and calibrates it using smart card data. This chapter is an edited version of the following article:

> **Shelat, S.**, Cats, O., van Oort, N., van Lint, J.W.C. Calibrating Route Choice Sets for an Urban Public Transport Network using Smart Card Data. 6th International Conference on Models and Technologies for Intelligent Transportation Systems (2019). *(chapter 2)*

The second part contains chapters 3 and 4 which answer RQs 2 and 3, respectively, on the impact of waiting time uncertainty on route choice behaviour. First, as has been done for travel time uncertainty in the literature, the conventional assumption that objective probabilities of waiting time are available to travellers is adopted. However, smart card observations are used for the analysis giving us important advantages over previous studies. One, the choices studied are made under natural ambiguity (even if they are not analysed as such), and two, model performance with different statistical representations of empirical waiting time distributions can be compared.

Next, a method to assess travellers' route choice behaviour under natural ambiguity is proposed so that the above assumption can be relaxed. The method can provide snapshots of travellers' evaluations of waiting time uncertainty in real-world public transport systems using, both, stated choice experiments and passively collected smart card data. As shown in Figure 1, this is demonstrated with two case studies using these two sources of data, respectively. The

[4] http://www.my-trac.eu/about/

chapters in this part of the thesis are based on edited versions of the following articles and conference presentation:

> **Shelat, S.**, Cats, O., van Oort, N., van Lint, J.W.C. Evaluating the impact of waiting time reliability on route choice using smart card data. Transportmetrica A: Transport Science (2022). *(chapter 3)*
> **Shelat, S.**, Cats, O., van Lint, J.W.C. Quantifying travellers' evaluation of waiting time uncertainty in public transport networks. Travel Behaviour and Society (2021). *(chapter 4)*
> **Shelat, S.**, Dixit, M., Cats, O., van Oort, N., van Lint, J.W.C. What does smart card data reveal about subjective beliefs regarding waiting time uncertainty? 8th International Symposium on Transport Network Reliability (2021). *(chapter 4)*

The third part, comprising of chapter 5, takes a look at travel behaviour under the new uncertainties brought about by the COVID-19 pandemic. To answer RQ4, stated route choice observations are analysed—specifically, with respect to factors affecting the risk of COVID-19 transmission. Moreover, the heterogeneity in behaviour is explored with a posterior analysis of various covariates in order to understand how the variation in beliefs regarding the pandemic affect behaviour in public transport network. This chapter is an edited version of the following article:

> **Shelat, S.**, Cats, O., van Cranenburgh, S. Traveller Behaviour in Public Transport in the Early Stages of the COVID-19 Pandemic in the Netherlands. Transportation Research Part A: Policy and Practice (2022). *(chapter 5)*

Finally, in chapter 6, overall conclusions of the thesis are drawn from findings and discussions in the preceding chapters. This is followed by a discussion of policy recommendations and ideas for further research.

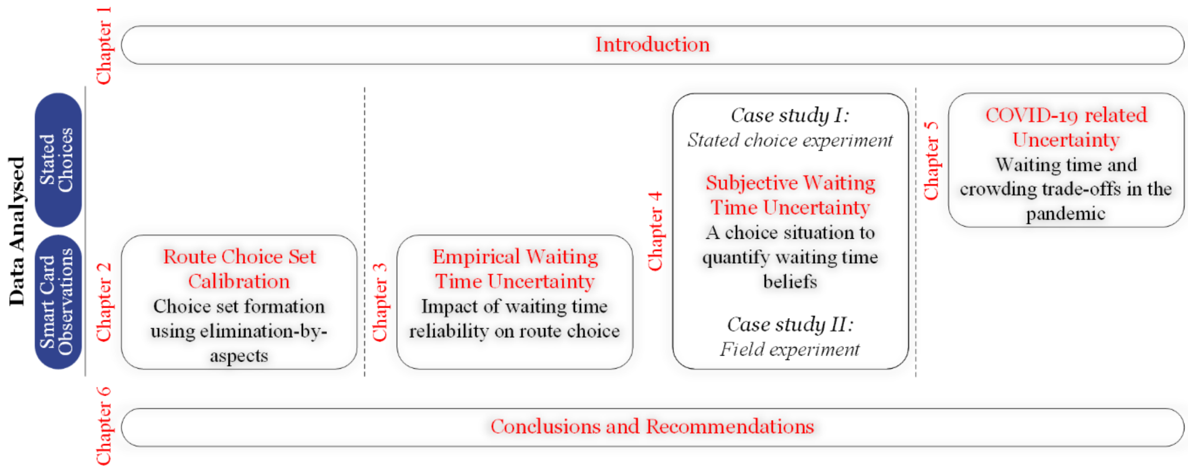


Figure 1: Thesis outline

Route Choice Set Calibration

Identifying the set of routes from which travellers choose is a crucial first step in estimation and application of route choice models. However, it is a typically difficult exercise when using passively collected data, such as from smart cards, because while chosen routes are observed, the ones considered are not. Approaches proposed in literature are not completely satisfactory, either lacking transferability or requiring strong assumptions regarding traveller behaviour.

Therefore, in this chapter, we propose a novel methodology wherein a non-compensatory decision model—elimination-by-aspects—is applied to alternative sets that are generated by constrained enumeration. The decision model is calibrated using observed route choice behaviour from smart card data. In addition to the choice sets, the calibration also returns key insights regarding choice set formation. We first present the proposed methodology followed by a demonstration using data from The Hague as a case study. The chapter concludes with an outline of the contributions and results, important remarks on limitations, and notes on potential avenues for further research.

This chapter is an edited version of the following article:
Shelat, S., Cats, O., van Oort, N., van Lint, J.W.C. Calibrating Route Choice Sets for an Urban Public Transport Network using Smart Card Data. 6th International Conference on Models and Technologies for Intelligent Transportation Systems (2019).

1 Introduction

There is widespread agreement in the marketing field that consumption choices occur in a two-stage process whereby consumers first form a consideration choice set and then make the final choice from this set (Hauser, 2014). Choice set composition and size can affect the ultimate decision in a number of ways (Prato and Bekhor, 2007), with the most obvious being that the exclusion of an alternative from the choice set means that it cannot be selected. Identification of choice sets may be straightforward when the number of alternatives is limited but it becomes more difficult as this number increases. In these cases, correctly identifying the choice set is important not only for real-world application of estimated choice models but also for the estimation of choice models from revealed preferences where choices are observed but choice sets are not.

Public transportation provides vital, sustainable transportation in many regions, making their planning, maintenance and operation a priority for authorities. In order to provide an appropriate level of service, understanding traveller behaviour to correctly model network flows has become increasingly important. Amongst other traveller decisions, route choices have a significant impact on network flows. Therefore, for both, estimation and application of route choice models, identification of route choice sets is a crucial step (Prato and Bekhor, 2007).

However, identifying route choice sets for origin-destination (OD) pairs in a network is a non-trivial task for several reasons. First, due to the combinatorial nature of the problem, the number of available and attractive routes is usually large. Second, public transport characteristics, such as fixed routes, schedules, and headways, which are usually time-dependent, add to the complexity of the task. Finally, the existence of different forms of travel costs, for instance, transferring or in-vehicle time, mean that traveller preferences have to be taken into account when identifying route choice sets.

Given the importance and complexity of route choice set identification, several studies in transportation literature have either entirely focussed on or have employed some form of route choice set identification methodology. These methodologies can be broadly divided into: (i) direct identification of choice sets and (ii) choice set generation methodologies (CSGMs).

Direct identification of choice sets may be based on reporting or observations of non-selected and selected alternatives, respectively. In the former, surveyed travellers are asked to report alternatives to their chosen route that they did not select but considered. This method has the obvious advantage that researchers do not have to guess what travellers have in their mind and the consideration set is known at the individual level. However, this reporting is subject to a number of errors (e.g., forgetfulness) and, as suggested in (Hoogendoorn-Lanser, 2005), is ‘at best a subset of the true choice set’. Furthermore, such interview techniques are time consuming and difficult to implement when choice sets are required for network-wide analysis.

For network-wide identification of choice sets, observations of selected alternatives offer a more suitable data source. In this method, the sets of unique routes observed are assumed to be the choice set for the respective OD pairs. The argument is that, if such data is collected over a long period of time, it should include all routes considered by travellers. Practically, this is facilitated by the creation of large data sources as an increasing number of public transport services turn to automatic fare collection (AFC) technologies. As a result, several studies using smart card data employ this method for the identification of choice sets. However, this technique precludes the possibility of taking into account why some non-selected but feasible routes are never chosen (Raveau, 2017). Moreover, the transferability of behaviour parameters, estimated with choice sets thus obtained, is precarious because matching choice set generation methodologies are not available for other public transport networks (Ton et al., 2018).

Some drawbacks of direct identification of route choice sets can be overcome by using CSGMs. The aim with this approach is to develop a generic algorithm, that satisfies requirements associated with the purpose of the choice set (Bovy, 2009), for identification of route alternatives. Thus, route CSGMs are suitable for network-wide application and by nature more transferable than direct identification techniques. These methods are typically classified into: (i) deterministic and stochastic shortest path, (ii) constrained enumeration, and (iii) probabilistic approaches (Bovy, 2009; Prato, 2009). Below, the most important approaches are discussed and the comparison of their performance is reviewed (for a complete literature review see Bovy (2009); Prato (2009)).

Shortest path-based methodologies, which compose the largest group of models, search for optimal routes in the network and assume them to be the route choice set. Variations are based on the link impedances optimized, route constraints, and other search criteria (Bovy, 2009; Prato, 2009). Approaches in this category that are based on either purely topological criteria or use only travel time have the drawback that the choice sets do not reflect traveller preferences. On the other hand, methods that do have some degree of behavioural sophistication, such as the link-labelling approach (Ben-Akiva et al., 1984), are criticised for their dependence on analyst judgments to make assumptions regarding traveller behaviour for the definition of objective functions (Bovy, 2009; Guo and Wilson, 2011; Prato, 2009). Furthermore, shortest path methods tend to produce more homogenous routes and are, therefore, typically unable to reproduce all observed routes.

Unlike the above approaches, constrained enumeration methodologies are based on rules other than minimum cost paths. Since these methods aim to generate all possible routes between OD pairs whilst being constrained by some rules, they usually perform better in terms of reproducing observed routes than the shortest path CSGMs (Bovy, 2009). Constraints used to reduce the number of irrelevant routes generated may be based on logic or common sense, feasibility, degree of choice set heterogeneity, or behavioural preferences (Hoogendoorn-Lanser, 2005; Prato and Bekhor, 2006). The disadvantages of this approach include the high computational effort required for route enumeration and the fact that, here too, the method depends on the definition of behavioural constraints which have been typically based on the expertise of analysts. Despite this drawback, in a comparison of various (uncalibrated) route CSGMs, a branch-and-bound based enumeration with threshold-based behavioural constraints performed better than other shortest path approaches on all the validation criteria considered (Bekhor and Prato, 2009; Prato and Bekhor, 2006).

From the common disadvantages of the above approaches, it is clear that calibration of behavioural parameters is an important aspect of route CSGMs. Yet, while studies often validate their models against observed (selected) or reported (non-selected) route alternatives, calibration is rarely performed. A Scopus search[5] for studies that perform such calibration returned only five relevant studies, including two studies that considered public transport modes (Bovy and Fiorenzo-Catalano, 2007; Hoogendoorn-Lanser et al., 2007). The latter studies use trial-and-error methods to calibrate their models on the basis of analyst judgments and, observed

[5]Search term: (TITLE-ABS-KEY (calibr*) AND TITLE-ABS-KEY ((route OR path)) AND TITLE-ABS-KEY (("choice set*" OR "consideration set*"))); Access date: 8 February 2019

and reported route alternatives. However, a shortcoming of these studies is that sample sizes of the data used are relatively small in comparison to the networks considered (which may be at least in part due to data collection difficulties).

Given the importance of identifying route choice sets in public transport networks and the drawbacks of existing studies, we propose a methodology that adopts an intuitive and accepted behavioural model of choice set formation and includes calibration of parameters of the same using smart card data. The proposed CSGM takes a constrained enumeration approach similar to those used (and proven to perform well) in (Hoogendoorn-Lanser, 2005; Prato and Bekhor, 2006). However, the methodology developed here avoids (almost completely) the need for any subjective assumptions regarding traveller preferences by delaying the application of behavioural constraints until after all logical and feasible routes have been generated. Instead of assumptions, behavioural constraints are directly obtained from AFC data, the increasing availability of which makes it possible to more easily collect network-wide route choice observations. Moreover, the constraints, which are based on a non-compensatory decision model, offer an intuitive insight into travellers' choice set formation preferences (section 2). To demonstrate the methodology, it is applied to the urban public transport network of The Hague, Netherlands.

In the next section, the behavioural model used for choice set formation is discussed. Section 3 describes the choice set generation methodology which is applied to the urban public transport network of The Hague, Netherlands. Results are discussed in section 5 followed by a summary of the contributions, key insights from the case study, and potential paths for future research.

2 Behavioural models of choice set formation

When a large number of alternatives are involved, consumers are likely to apply heuristic decision rules, rather than perform a comprehensive evaluation, when forming their considered choice set. These choice set formation heuristics are usually more reasonable because of the relatively high (cognitive and explicit) costs of complete evaluations (Hauser, 2014). Therefore, since the number of route alternatives available in transportation networks is typically large, travellers can be reasonably expected to use such heuristics to identify their choice sets (Bovy, 2009; Prato, 2009).

While complete evaluations are typically compensatory in nature, heuristics involve non-compensatory decision rules. Compensatory models take into account trade-offs between alternative attributes whereas non-compensatory models only apply constraints on individual attributes. A number of non-compensatory decision models have been proposed in literature, such as: (i) disjunctive, (ii) conjunctive, (ii) lexicographic, and (iv) elimination-by-aspects (Hawkins and Mothersbaugh, 2010). Some of these have been used in the route choice set generation literature.

Disjunctive and conjunctive rules both set minimum thresholds for all important attributes. The former accepts alternatives that comply with at least one requirement while the latter needs all attribute thresholds to be met. Most route CSGMs that apply detour thresholds to different attributes (e.g., Hoogendoorn-Lanser (2005); Prato and Bekhor (2006)) are applying conjunctive rules. In these studies, thresholds are usually set as multiplicative factors (greater than one but not necessarily integers) of the attribute value of the best performing alternative (for that particular attribute).

Under lexicographic decision-making, first, attributes are ranked by importance; then alternatives are selected on the basis of their performance of the top-ranking attribute. In case of a tie, the performance on the second-best attribute is checked, and so on. Since this method does not set thresholds, desired choice set sizes need to be defined for their formation (lexicographic and conjunctive decision rules become the same if choice set size is defined for both) (Hauser, 2014). The link-labelling route CSGM (Ben-Akiva et al., 1984), which assumes that travellers optimize paths for different attributes, is an example of this category.

Elimination-by-aspects (EBA) combines attribute ranking and setting of thresholds. Although the original version (Tversky, 1972) was proposed as a probabilistic rule, most applications for choice set formation use a deterministic version (Hauser, 2014). For the choice set formation, first the most important attribute is selected and alternatives that do not meet its threshold are eliminated. This is repeated until all attribute thresholds have been checked although in another version, elimination stops once the required choice set size has been achieved (Hauser, 2014). Based on the literature review conducted here, no route CSGM could be found that uses this behavioural model. A possible reason could be that in the absence of a calibration method, because this model combines ranking and setting thresholds, researchers are required to make more assumptions regarding traveller behaviour.

This study assumes deterministic EBA as the behavioural model for route choice set formation. In the version employed here, no assumptions are made regarding choice set size. Deterministic EBA implies that the choice set for an OD pair at a given time is the same for all travellers. Attribute values are obtained from the general transit feed specification (GTFS) data. Therefore, attributes included in the process are limited to those observable in this data.

The output of the methodology proposed here are route choice sets per OD pairs and time periods. Each alternative in the route choice set is defined uniquely by the sequence of alighting stations and the common lines (lines passing through the same sequence of stations) connecting the respective stations. Although common lines are thus accounted for, issues concerning partial route overlap are assumed to be handled at the next stage of choice modelling. In addition to the route choice sets, calibration of the choice set formation model returns two insights regarding traveller behaviour: (i) the importance ranking of attributes and (ii) the acceptable detour threshold for each attribute.

3 Methodology

To give structure to the complexity of route choice set identification, a hierarchy of route choice sets (for a given OD pair and time period) is proposed in (Bovy and Stern, 1994) and presented from traveller and researcher perspectives in (Hoogendoorn-Lanser, 2005). Similar to those, for the methodology presented here, the following hierarchy is used (Figure 2, right hand side): (i) complete network containing the universal set of all possible paths from origins to destinations, (ii) logical routes per OD pair, (iii) feasible routes per OD pair for different times (OD-T), (iv) considered routes per OD-T, and finally (v) chosen routes. Here, the consideration route set is obtained from the generated-feasible and observed route sets. The following sub-sections describe the steps in the proposed methodology (Figure 2, centre) that take some inputs and produce the desired outputs (Figure 2, left hand side), by progressively moving down the hierarchy.

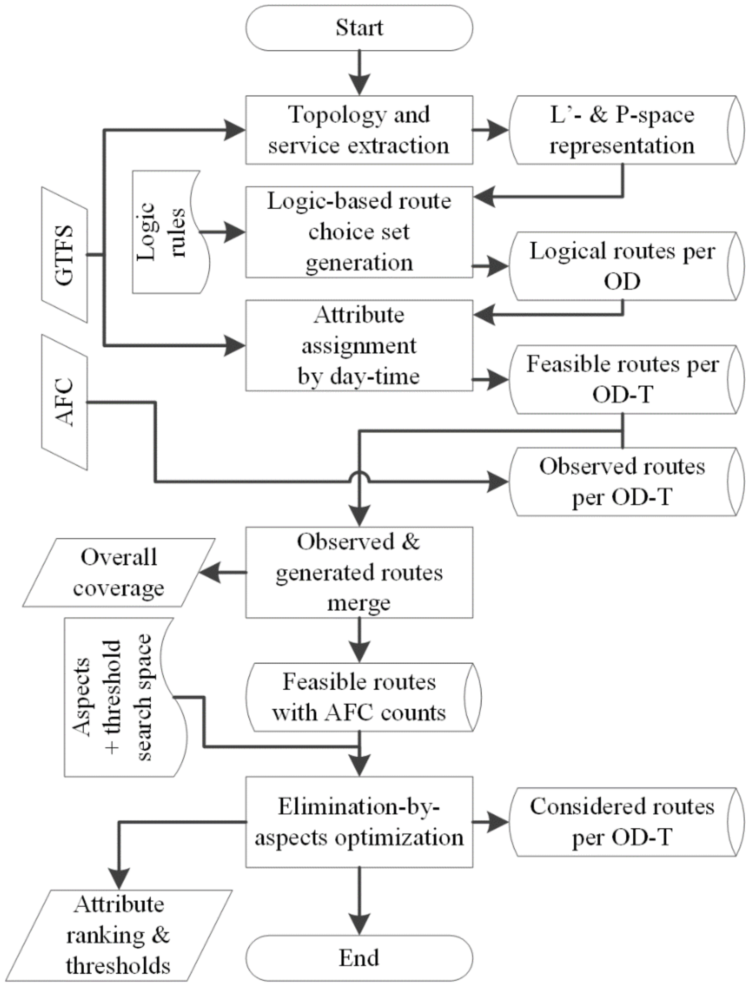


Figure 2: Algorithm for proposed route choice set generation methodology

3.1 Inputs

Two main data sources are required for the route CSGM proposed here: (i) GTFS and (ii) AFC. GTFS data contains information regarding the service layer of the network and its properties. These define public transport lines connecting different sequences of stops in the network, the in-vehicle travel time (time taken by a vehicle) between OD pairs, and the frequencies (vehicle departures per hour) of each line. Although, a frequency-based system is assumed in this CSGM, line frequencies are allowed to be time-dependent. AFC data is used to generate the set of selected route alternatives. Ideally, for each observation, the data should contain information regarding the sequence of stops (i.e., origin, transfers, and destination), the lines used between each stop, and boarding times. It should be noted, however, that for data from nearly all AFC systems, at the least, transfer inference will be required.

In addition to the above data sources, rules regarding which routes are logical, and which aspects and thresholds are considered by the calibration process are also inputs to the methodology. However, in the current implementation, these inputs are defined as part of the methodology.

3.2 Network representation

Proper network representation is key to the computational efficiency of route generation. For the topological constrained enumeration methodology used in this study, the public transport network is represented in the P-space which explicitly represents the service layer: nodes are stops while links are (groups of) public transport lines that provide direct (transfer-less) connections to other stops. Although urban public transport networks may use schedules, this study assumes a frequency-based system. Thus, time is not included in the graph representing the network.

Stops in the representation are defined by the ‘parent station’ field in the GTFS data. Moreover, different lines are grouped together as one connection in P-space if they pass through the same sequence of stops, that is, they are common lines. Each cell in the P-space adjacency matrix contains information about the connections between the origin and destination stops. For each connection, this information consists of the common lines and the stops they traverse (pass without alighting) through for this connection.

Since the generation methodology also considers transfers which require walking to another stop, walkable links are stored as a binary adjacency matrix of all stops. To avoid generating too many irrelevant alternatives, a conservative threshold of 200 Euclidean metres is set as the maximum acceptable walking distance.

3.3 Constrained enumeration

Enumeration

The enumeration methodology applied here uses a one-to-all, breadth-first search algorithm, similar to that used in (Shelat and Cats, 2017). The methodology is applied to the P-space graph representation of the network as defined above.

First, a stop is selected as the origin and the vertex root of the search tree. The stops it is directly connected to in the P-space graph of the network become destinations; this is the first level of the search tree. The information contained in the connections (lines and traversed stops) are stored for the respective OD pairs. For the next level of the search tree, the following become intermediate origins: transfer stops (stops connected by more than one line) amongst the neighbouring stops and stops accessible by walking from the neighbouring stops. Then the stops directly connected to these intermediate origins become destinations and connection information is stored, retrospectively from the origin stop (vertex root), for the respective OD pairs. For the next level, intermediate origins are selected in the same way as above, and the process is repeated up to a desired depth of the search tree. This way, all route alternatives between the origin stop and others are generated and stored. The procedure can then be repeated with another stop as the origin.

Constraints

Obviously, simply enumerating this universal set of routes would be unending. To prune the search tree, the depth is constrained by assuming that travellers accept a maximum of two transfers. This behavioural assumption should be reasonable for most urban public transport networks. Additionally, to ensure that only logical routes are produced, two rules are used as breadth-wise constraints to the enumeration. (i) No loops—traversing through, alighting at, or walking to stops previously traversed through or boarded from is not allowed. (ii) No transferring between common lines—alighting at a stop which is connected by the same set of lines as the previous connection is not allowed. Since travellers may want to shift time spent waiting for a particular line downstream, transferring to stops with a subset of the previous connection's lines is permitted. In this case, this subset of lines is removed from the previous connection to ensure that transfers do not occur between the same lines. In the current implementation, it is assumed that travellers do not shift their waiting times by walking to another stop, hence, walking to a stop connected by a subset of the lines as the previous connection is not allowed. These logical constraints are only applied after the first level of the search tree.

3.4 Attribute assignment

In this step, route alternatives, that were generated from an unweighted graph, are assigned attribute values. This is required to remove infeasible routes as well as for the consideration set formation in section 3.6.

Attribute values

The following route attributes are observable from the GTFS data, and therefore included in the study: (i) waiting time, (ii) in-vehicle time, and (iii) number of transfers. Currently, the waiting and in-vehicle times over different legs of the route are not considered separately and only the total values are used.

Expected waiting time for connections between two stops is calculated as the inverse of the sum of frequencies of the connecting lines, implicitly assuming them to be evenly spaced as well as assuming uniform arrivals of travellers at stops. The time-dependent nature of public transport line frequencies is taken into account by assigning them separately for each hour of the day in weekdays and weekends, respectively. Routes that become infeasible at a certain time (because a link has zero frequency) are eliminated from the feasible choice set for the respective time period. Values of the other attributes are time-independent. Although there might be small time-dependent fluctuations in the planned in-vehicle times, they are ignored for the sake of computational efficiency.

As discussed in section 2, the consideration set formation model employed here uses the EBA behavioural model, which requires setting threshold constraints to different attributes. These thresholds are some factors of the attribute values of the alternatives (in the same time period) that perform best on the respective attributes. In preparation for the calibration step, these factors are calculated for each attribute in all the alternatives. Since waiting and in-vehicle times are more continuous in nature, multiplicative factors are employed, whereas for number of transfers an additive factor is used.

Dominated alternatives

Once attribute values have been assigned, alternatives that are state-wise dominated, that is, perform worse on all attributes, by others (in the same time period) are removed. It is rarely disputed that choosing such a dominated alternative is irrational. Although the existence of dominated alternatives in the choice set may have a decoy effect (see (Puto et al., 1982)), such effects are rarely modelled in the route choice context.

3.5 Generated and observed routes merging

The calibration uses generated-feasible routes as well as those observed from AFC data. This step merges these two route sets on the basis of the sequence of stops boarded, lines used, the hour of the first boarding, and the final destination stop. For the calibration, only those observed routes that were also generated are considered. Given that the constraints assumed during route enumeration are not very restrictive, discarding observations that are not generated should not affect the final calibration too much. In case, the overall coverage does turn out to be low, it may make sense to check the AFC data for issues such as improper transfer inference.

3.6 Elimination-by-aspects calibration

In EBA, travellers are assumed to rank and set threshold cut-offs for attributes. In order to deduce these preferences, the generated feasible route alternatives may be compared with the observed ones. For such a comparison, two indicators are commonly used in literature, albeit for validation purposes rather than calibration: (i) coverage—the proportion of observed routes that have been generated, and (ii) efficiency—the proportion of generated routes that are observed.

With respect to calibration, clearly, the likeliest combination of choice set formation preferences is one that maximize both coverage and efficiency; that is, reproduces as many observed routes as possible while not generating too many irrelevant alternatives. Thus, to derive EBA preferences, an optimization problem that maximizes these indicators is setup. First, however, small modifications to the above indicators are proposed.

Indicators

In their simplest form, coverage and efficiency do not take into account demand across OD pairs and weigh each route alternative as the same. For example, if there is an OD pair with only one trip, it would still have an effect on the choice set calibration even though there is little behaviour to be observed. To this end, the coverage indicator is modified by simply adding demand weights per route. Efficiency is changed more fundamentally by making it a proportion of routes not observed (but in the generated feasible choice set), rather than a proportion of generated routes, to avoid asymmetric demand weighting in the definition. These indicators are defined below.

Let N be the set of stops in the network under consideration; and $\mathrm{R}_{ij}^{\mathrm{f}}$ the set of generated-feasible routes between OD pairs $i, j \in \mathrm{N}$, $\mathrm{R}_{ij}^{\mathrm{o}}$ the set of observed routes therein $\left(\mathrm{R}_{ij}^{\mathrm{o}} \subseteq \mathrm{R}_{ij}^{\mathrm{f}}\right)$, and $\mathrm{R}_{ij}^{\mathrm{c}}$ be the calibrated choice set, such that $\mathrm{R}_{ij}^{\mathrm{c}} \subseteq \mathrm{R}_{ij}^{\mathrm{f}}$, for a given combination of EBA preferences. Then, Table 1 gives the four possible sets (and notations) of route alternatives that result when comparing the observed and calibrated choice sets. Finally, let q_{ij} be the total of all demand on routes $\mathrm{R}_{ij}^{\mathrm{o}}$, and $q_{ij}^{\mathrm{TP}}, q_{ij}^{\mathrm{FN}}$ be the total demand for route sets $\mathrm{R}_{ij}^{\mathrm{TP}}, \mathrm{R}_{ij}^{\mathrm{FN}}$, respectively. Then, coverage and efficiency are defined in this study as:

$$\text{coverage} = \frac{\sum_{i,j} q_{ij}^{\mathrm{TP}}}{\sum_{i,j} q_{ij}^{\mathrm{TP}} + q_{ij}^{\mathrm{FN}}} \qquad [1]$$

$$\text{efficiency} = \frac{\sum_{i,j} \left|\mathrm{R}_{ij}^{\mathrm{TN}}\right| \cdot q_{ij}}{\sum_{i,j} \left(\left|\mathrm{R}_{ij}^{\mathrm{FP}}\right| + \left|\mathrm{R}_{ij}^{\mathrm{TN}}\right|\right) \cdot q_{ij}} \qquad [2]$$

where $|\bullet|$ denotes set size. Since coverage and efficiency move in opposite directions (as one increases the other decreases), to achieve a balance between coverage and efficiency (in the absence of other requirements), the following optimization indicator is minimized for each attribute, a:

$$x_a = \text{abs}\left(\text{coverage}_a - \text{efficiency}_a\right) \qquad [3]$$

Algorithm

The EBA based analysis conducted here considers only a few aspects (i.e., attributes). Moreover, it is reasonable to expect that the potential thresholds are close to the respective smallest values (i.e., 1 for waiting time and in-vehicle time ratios, and 0 for difference in number of transfers). Therefore, to deduce EBA preferences, a brute force algorithm may be feasibly employed. The algorithm to calculate indicator values for different attribute rankings (Figure 3) works as follows: all possible attribute permutations are listed; for a given permutation, different thresholds from the pre-defined search space are tried to find the minimum indicator value for the first attribute; before repeating this for the next attribute, routes that do not comply with the previously found threshold(s) are eliminated; this is repeated until all attribute thresholds (and indicator values) for the permutation have been found; and the process is repeated for the next permutation.

It should be noted that a key difference from other threshold based CSGMs is the sequential elimination of routes. Thus, for each permutation we have a number of optimization indicator values associated with each attribute in it. The performance of a permutation, *p*, is assessed by calculating the natural logarithm of the product of attribute optimization indicator values in that permutation:

$$x^{\mathrm{p}} = \ln\left(\prod_a x_a^{\mathrm{p}}\right) \qquad [4]$$

Since the optimization indicator has to be minimized, the permutation with the lowest value is considered optimal.

Table 1: Comparison between calibrated and observed choice sets

		In calibrated choice set?	
		Yes	No
In observed choice	Yes	True Positives $R_{ij}^{TP} = R_{ij}^{c} \cap R_{ij}^{o}$	False Negatives $R_{ij}^{FN} = \{R_{ij}^{c} \cup R_{ij}^{o}\} - R_{ij}^{c}$
	No	False Positives $R_{ij}^{FP} = \{R_{ij}^{c} \cup R_{ij}^{o}\} - R_{ij}^{o}$	True Negatives $R_{ij}^{TN} = R_{ij}^{f} - \{R_{ij}^{c} \cup R_{ij}^{o}\}$

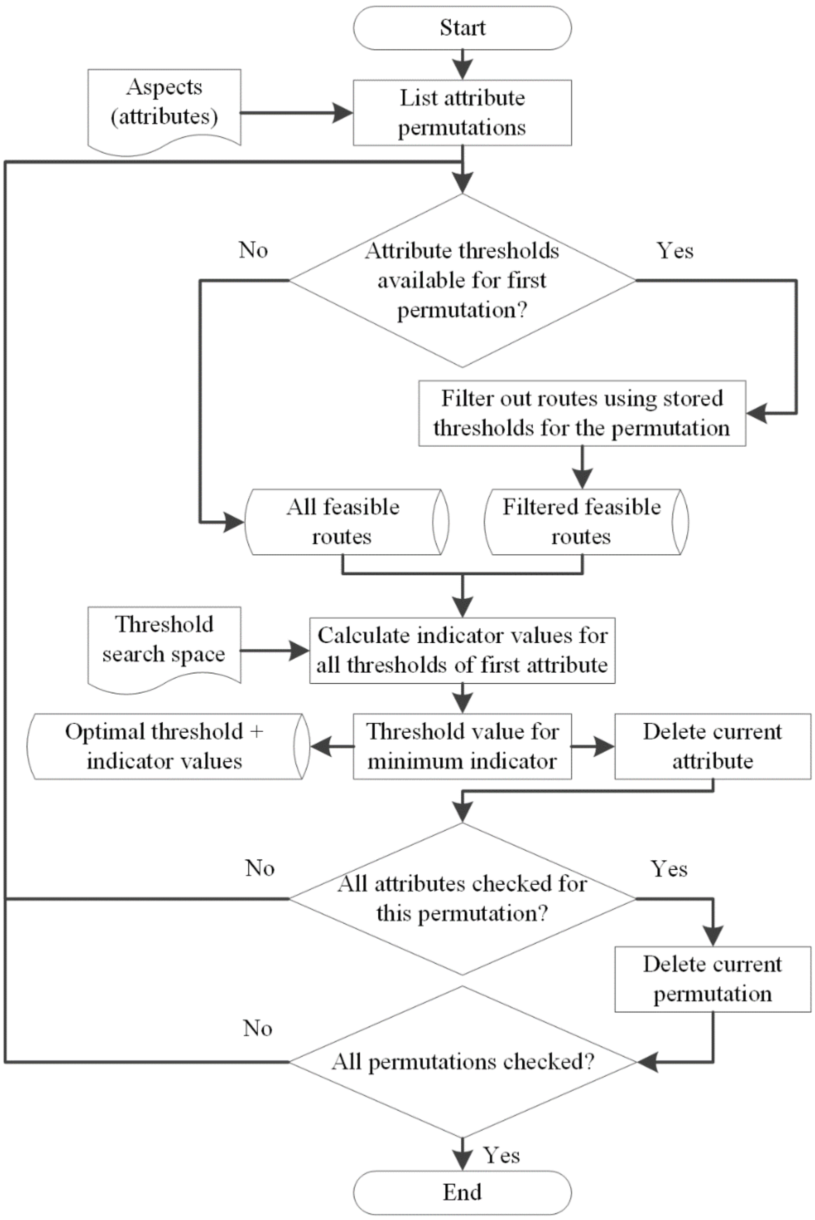


Figure 3: Algorithm for elimination-by-aspects calibration

4 Case study analysis

To demonstrate the above methodology and obtain attribute ordering and threshold preferences of travellers, the urban public transport network of The Hague, Netherlands (Figure 4) is used as a case study. The network consists of both tram and bus lines which mainly serve The Hague but also connect to the neighbouring cities of Zoetermeer and Delft. The case study uses smart card data from March 2015 and the corresponding GTFS data for the analysis. The network then consisted of 12 tram and 8 bus bidirectional lines serving a total of 459 stations (as defined under ‘parent stations’ in the transit feed).

The AFC system on both trams and buses requires travellers to check-in and out with the *OV-chipkaart*, (the national public transport smart card; for more details see (Van Oort et al., 2015a)) every time they board and alight a vehicle; thus, potentially allowing full observation of chosen routes. Moreover, since, a large number of travellers in the network use smart cards for fare payment a significant amount of data is available for analysis. The data, made available by the operator, is pre-processed such that individual smart card transactions (check-ins and outs) are already chained to approximately 5.9 million journeys from origin to destination stations. Out of these, the case study, which only includes trips in weekday extended morning peak hours (0600h to 1100h), makes use of about 1.5 million journeys.

The pre-determined journeys used in the smart card dataset have been inferred using a simplistic rule based on maximum transfer time (35 minutes (Yap et al., 2017)). Such inference methods typically lead to an overestimation of routes with more transfers and can leave seemingly illogical trips in the data. A full and robust (against disruptions) transfer inference algorithm as given in (Yap et al., 2017) can solve these issues. However, this is not done and misidentified journeys are directly filtered out when they do not match with the generated feasible routes. This seems to have a relatively low impact for the time period selected for the analysis as the overall coverage of the generated-feasible routes is nearly 85 percent of the observed routes.

Figure 5 compares the logical, feasible, and identified route choice set size distributions. As one would expect, logical choice sets are typically large (median size: 58 routes). A sharp decline in the sizes for the feasible set (median size: 9 routes) is brought about mainly by the state-wise dominancy elimination rule, although some routes are also removed due to service unavailability in certain time periods.

For the EBA calibration, three attributes, waiting time, in-vehicle time, and number of transfers, are considered. Based on experience and with an eye on computational efficiency, the threshold search space for the former two is defined between 1 and 2 with a step size of 0.025, while all possible values (0 to 2) are tried for the latter. Note that, if an intuition for these values is not available, one could simply try a larger search space.

5 Results and discussion

Performance of the six permutations (Figure 6) indicates a clear preference in attribute ranking. Similar to findings for fully compensatory route choice models in literature, people rank number of transfers as the most important parameter followed by waiting and in-vehicle time, respectively.

For all permutations, constraints on individual attributes are quite restrictive: for waiting and in-vehicle time most multiplicative thresholds lie between 1 and 1.1 (meaning that only a 10 percent increase is acceptable), while for transfers, routes with even a single extra transfer are unacceptable in the choice set. These thresholds are lower than those assumed in CSGM studies assuming a conjunctive model for consideration set formation. For instance, for road traffic, the threshold used for travel time is 1.5 in (Prato and Bekhor, 2006). Moreover, because of low thresholds, the choice sets sizes are also small (Figure 5) with a median size of only 2 routes. gives the threshold values obtained for individual attributes.

To assess the performance of the calibration the overall coverage of the EBA model can be calculated as the product of the coverage values obtained sequentially for each attribute (Table 2). The overall coverage for this case study is 63.9% which on its own is a somewhat moderate performance, but one that may be expected because, in an effort to be more transferable, the model trades-off coverage for an increase in efficiency.

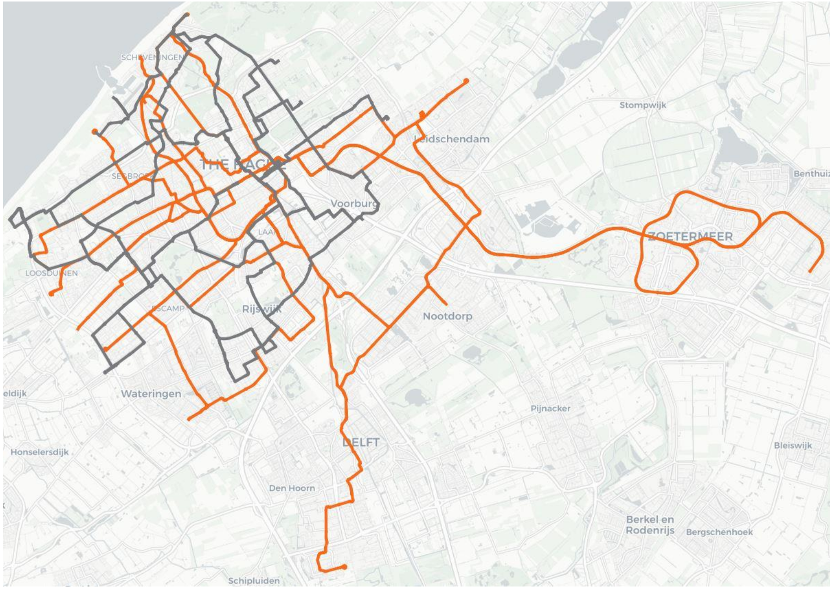


Figure 4: The Hague tram (orange) and bus (gray) networks in March 2015

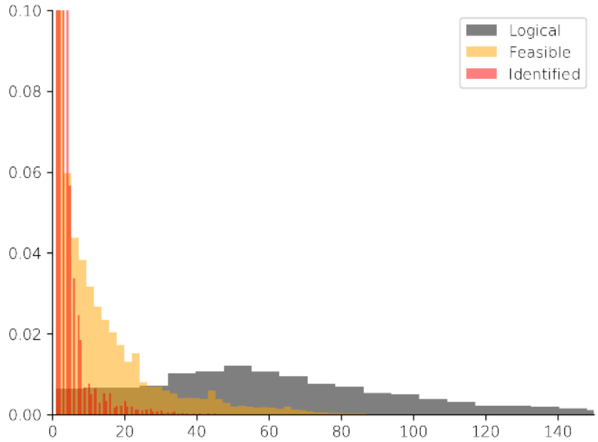


Figure 5: Comparison of choice set size distributions (normalized) of logical, feasible, and identified route choice sets (top- and right- censored for better focus)

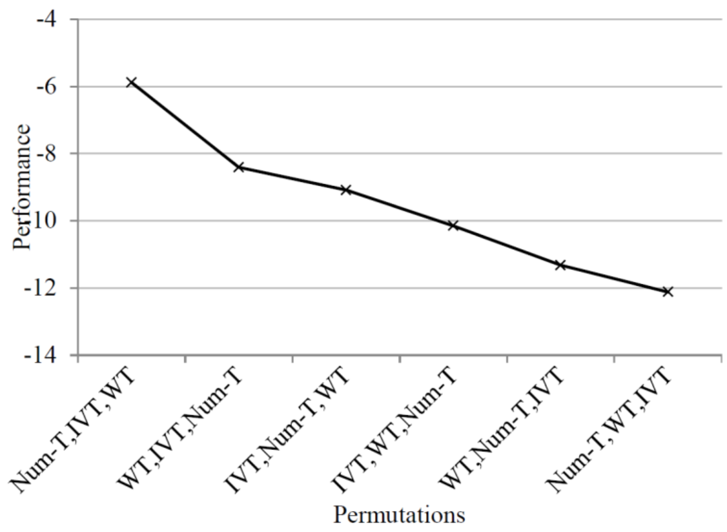


Figure 6: Performance for different attribute rankings (Num-T: number of transfers, WT: waiting time, IVT: in-vehicle time) (lower is better)

Table 2: Optimal attribute ranking and thresholds

Rank	Attribute	Threshold[a]	Sequential Coverage
1	Number of transfers	0	99.3%
2	Waiting time	1.1	82.0%
3	In-vehicle time	1.1	78.4%

[a]Threshold accuracy for waiting and in-vehicle time = 0.025

Figure 7 takes a deeper look into the values of individual thresholds for the optimal permutation. It can be seen that for the first attribute—number of transfers—at 99.3 percent, coverage is already extremely high with no extra transfers; a clear reflection of travellers' dislike for transferring. Thus, any increase in the transfer threshold only decreases efficiency thereby increasing the indicator value. For waiting time too, the initial coverage is quite high, meaning that improvements in coverage tend to be quite slow. Accepting twice the least possible waiting time only increases coverage from 79.9 to 93.3 percent. On the other hand, efficiency quickly decreases by approximately 40 percentage points. Although the initial slope for coverage is slightly higher than efficiency, the overall change in the latter is higher for in-vehicle time too. The fact that the initial value of coverage is more moderated for this attribute could be because the values are calculated after the feasible choice set has been filtered based on the thresholds of the previous two attributes. Finally, it should be noted that for all three attributes, the optimal indicator values are unambiguous.

A possible explanation for the restrictive constraints may be a combination of the following statistical observations and hypothesis. The statistical observations are (Figure 8) (i) OD pairs with a high demand tend to be nearby (in terms of in-vehicle time) and (ii) OD pairs that are farther away tend to have more feasible routes generated by the CSGM. The hypothesis is that (iii) travellers are either able to evaluate alternatives better or have a lower threshold acceptance for OD pairs that are nearby. From statistical observation (i), the hypothesis in (iii), and the definition given in Equation [1], it can be seen why the coverage values are already quite high at low thresholds. This increase in coverage is mainly due to the highly used routes between OD pairs that are close to one another. On the other hand, statistical observation (ii) and the definition in Equation [2] explain why non-selected alternatives from farther away OD pairs might play a larger role in the value of efficiency. This potential disconnect might cause a decrease in efficiency that is not sufficiently balanced by the increase in coverage, leading to smaller thresholds. The larger slopes of efficiency in comparison to coverage (Figure 7) seem to indicate that this is indeed the case here.

6 Conclusion

Route choice set identification for public transport networks is a vital but complicated task. Identifying the correct route choice sets are crucial for both, estimation and application, of route choice models. However, approaches developed and commonly employed in literature either lack transferability (observation-driven methods) or require strong assumptions regarding traveller behaviour (uncalibrated CSGMs).

Given this key scientific gap, and the context of increasing availability of smart card data for public transport networks, this research makes two crucial contributions. First, a choice set generation methodology is proposed which uses elimination-by-aspects as the consideration set formation model. This model adds more behavioural dimensions over those used previously by taking into account attribute ranking as well as threshold constraints. Second, rather than making subjective assumptions regarding traveller preferences, the elimination-by-aspects model is calibrated using revealed behaviour observations from smart card data. The proposed methodology can be used to identify choice sets for estimating route choice models from revealed preferences as well as to predict alternative shares on the basis of available choice parameters.

Application of the proposed methodology to the urban public transport network of The Hague revealed that the number of transfers is the most important attribute for travellers, followed by waiting time, and in-vehicle time. Furthermore, the thresholds obtained for individual attributes are quite restrictive indicating that travellers make more optimal choices than previously assumed. Although the overall coverage for the EBA model is on the lower side, it makes up for this by being a more transferable model rather than a network-specific one. While coverage and efficiency are weighted equally here, one may want to tune this trade-off. For instance, Ton et al. (2020) assign a higher weight to coverage because they had a relatively small number of observations for their choice model. The ideal trade-off can also be learnt by jointly evaluating the CSGM with the performance of the subsequent choice model.

An important limitation in the current implementation of the model is the assumption that the public transport services are frequency-based. Based on this, waiting times are calculated from the headways of individual lines under the assumption of evenly spaced arrivals of public transport lines and uniformly distributed traveller arrivals at stops. These assumptions may not hold outside rush hours or for non-urban networks where line frequencies are often lower, or when lines are explicitly synchronised to reduce transfer waiting time. To overcome issues arising from the assumption of a frequency-based system, future implementations may consider using the following: (i) a schedule-based network which includes time in its representation and (ii) more complex traveller arrival models. Further improvements to the model could include taking into account that travellers behave differently for OD pairs that are relatively near, as hypothesised in the discussion of the case study results. Finally, future research could focus on using the calibration procedure proposed here, for the comparison of different behavioural models of route choice set formation.

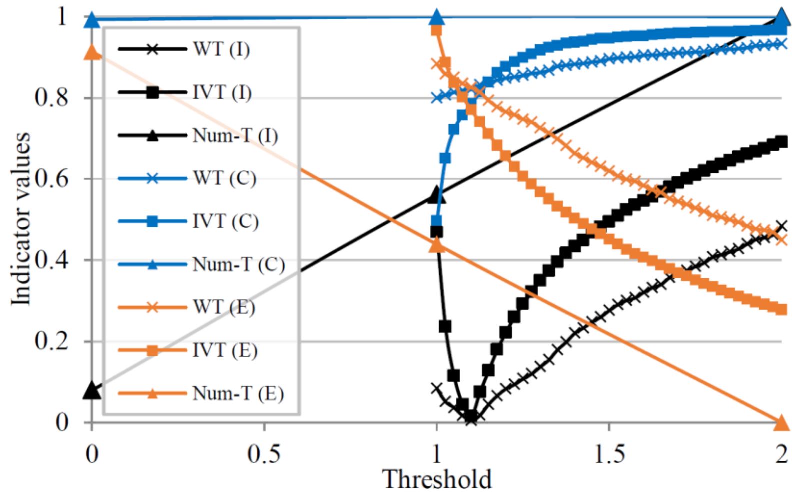


Figure 7: Coverage (C), efficiency (E), and optimization indicator (I) values (y-axes) by threshold values (x-axis) of different attributes for the optimal permutation

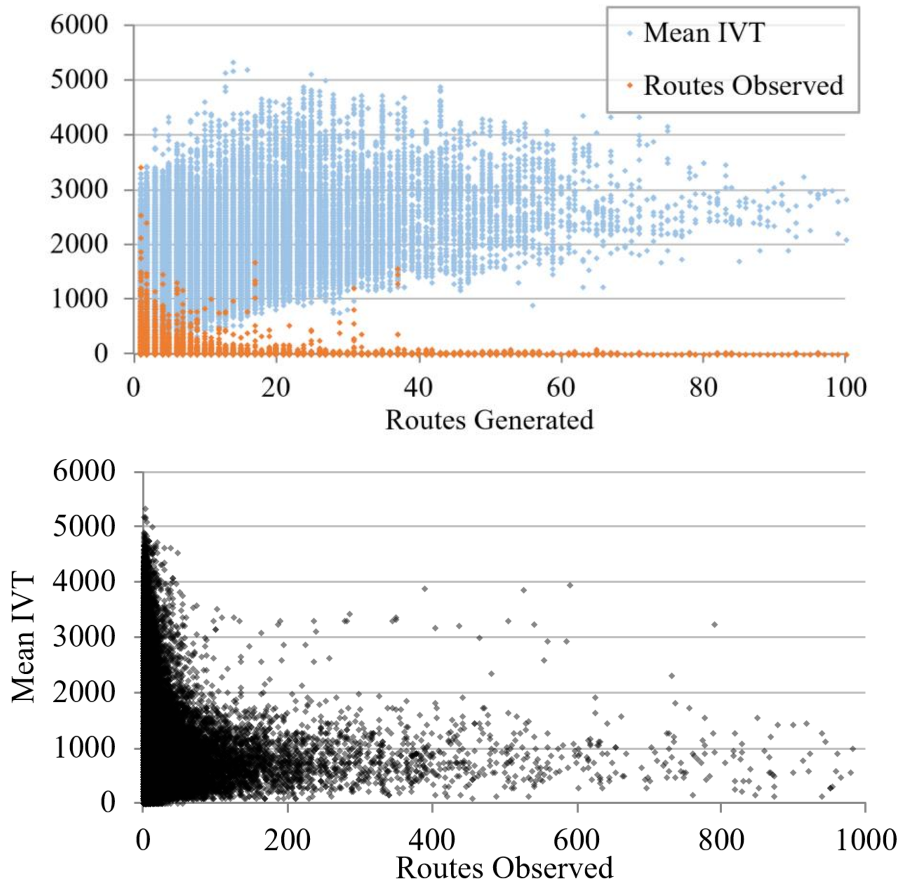


Figure 8: Comparison between the number of generated-feasible routes, number of observed routes and average in-vehicle times (in seconds) per origin, destination, and time period (x-axis right-censored for better focus)

Empirical Waiting Time Uncertainty

The effect of travel time reliability on route choice behaviour has been studied extensively—typically as the impact of an objective representation of uncertainty on hypothetical choices. In this chapter, we overcome disadvantages that are inherent to stated choice experiments, such as hypothetical bias and difficulty in uncertainty presentation, by using revealed preferences from smart card data. We analyse the effect of waiting time uncertainty as regular and irregular deviations from scheduled values. Since passively collected data is used, we are also able to examine a number of statistical indicators for the latter.

Using data from The Hague, first choice sets suitable to analyse the effect of waiting time uncertainty are obtained. (We do not employ the methodology proposed in the previous chapter but adopt one that is more commonly used.) Next, we assign attributes to the choices and inspect, in particular, a range of indicators for irregular waiting time deviations. Multinomial logit choice models for morning peak and off-peak hours, with and without reliability coefficients are estimated and validated, followed by a detailed discussion of the estimated models. Finally, we conclude with a summary of our contributions, main results, and a discussion on the limitations of the analysis.

This chapter is an edited version of the following article:
Shelat, S., Cats, O., van Oort, N., van Lint, J.W.C. Evaluating the impact of waiting time reliability on route choice using smart card data. Transportmetrica A: Transport Science (2022).

1 Introduction

The impact of travel time reliability on route choice behaviour has received much attention in literature. The vast majority of studies have used stated preferences collected through surveys or simulation experiments (Carrion and Levinson, 2012; Li et al., 2010). Due to the costs associated with data collection, only a few have analysed revealed preferences and almost all of these have focussed on car traffic. Fortunately, however, an increasing number of public transport networks are integrated with automatic fare collection (AFC) systems, which enable analysts to glean revealed preferences from this passively collected data. Moreover, automatic vehicle location (AVL) data, provides detailed information on the realised supply characteristics (e.g., travel time, waiting time) of public transport services. In this chapter, we combine the revealed preferences from AFC data with reliability information derived from AVL data to establish the role of waiting time reliability in public transport route choice.

Passively collecting revealed preferences offers two important advantages over simulation experiments and conventional stated preferences. First, and most importantly, revealed preferences are free from hypothetical biases that are prevalent in the other methods since observed choices have consequences and are made in real-world situations. Second, passively collecting revealed preferences obviates the need to explicitly convey reliability information, which has proven to be difficult (Bates et al., 2001; Carrion and Levinson, 2012). Furthermore, unlike revealed preference questionnaires that enquire about past behaviour, passive data collection allows us to gather significantly more observations. However, since revealed preferences inherently lack experimental control, the amount of information obtained per observation is likely to be significantly lower than stated preferences or simulation experiments. The lack of experimental control could also lead to issues in determining the direction of causality. Privacy regulations may also restrict the amount and type of data that can be used for analysis. Finally, passively collected data typically require significant processing effort and require assumptions regarding attributes that cannot be observed (e.g., Lam and Small (2001); Luo et al. (2018)).

While more complete reviews of studies analysing the effect of travel time reliability on travel behaviour are available elsewhere (Carrion and Levinson, 2012; Li et al., 2010), here, we briefly outline those using passively collected data for their analysis. For car traffic, such studies have largely made use of road-pricing experiments in the United States (Alemazkoor et al., 2015; Carrion and Levinson, 2013; Lam and Small, 2001). This setting offers researchers a unique opportunity to observe choices between a free but (potentially) congested road and a tolled but (almost certainly) uncongested road, and thus, estimate the value of reliability. Travel times and choices are obtained via loop detectors or GPS devices and toll transponders, respectively. In recent years, with the introduction of AFC systems, researchers have used this passively collected data to analyse travel behaviour in public transport systems, often in combination with AVL data. Amongst these studies, for many, general route choice behaviour is the primary aim of the analysis (e.g., Jánošíková et al. (2014); Kim et al. (2019)). Other studies focus on specific aspects such as, on-board crowding (Hörcher et al., 2017; Yap et al., 2018), transfer inconveniences (Guo and Wilson, 2011), route choice variability (Kim et al., 2017b; Kurauchi et al., 2014), or strategic behaviour (Nassir et al., 2018; Schmöcker et al., 2013). To the best of our knowledge, only Leahy et al. (2016) have used such data to analyse the impact of travel time reliability on route choice behaviour in public transport networks. Our study is different from theirs in two important ways: (i) whereas they evaluate the role of the *total* trip travel time reliability we focus on and are able to specifically estimate perceptions related to waiting time reliability; and (ii) they use a sample of AFC transactions over the time period under study while we have access to the full population dataset.

In general for the service industry, waiting time is a critical component of satisfaction (Maister, 1985). For public transportation too, it has been consistently shown that waiting time has a large impact on behaviour. Unreliability in this important component of travel time may lead to frustration and anxiety amongst travellers. We note that unreliability is inherently an uncertain (Knight, 1921) attribute—the true distribution of waiting times is unknown to the travellers. Instead, travellers may have their own subjective distributions (Dixit et al., 2019; Meng et al., 2018) that they use to make decisions. However, in line with the majority of previous research on this topic (Carrion and Levinson, 2012), we model waiting time unreliability as if it were a known risk. We compare a number of empirical measures of unreliability in our choice analysis. Furthermore, the effects of waiting time unreliability are modelled separately for origin and transfer stations, for different public transport modes, and for morning peak and off-peak hours.

Next, we briefly present our case study: the urban public transport network of The Hague, followed by the methodology outlining data preparation, choice set identification, attribute extraction, and finally choice analysis. We then discuss the estimated choice models and conclude with a summary of the main results and suggest avenues for future work.

2 Case study

For our analysis, we use smart card data from the urban public transport system in The Hague, the third largest city in the Netherlands. The network contains 12 tram and 8 bus lines connecting 499 aggregated stops (operator-defined ‘parent stations’ in the GTFS data) in The Hague and neighbouring suburbs and towns (Figure 9). About 90% (Yap et al., 2018) of the trips in the network are paid for through a smart card based AFC system which requires travellers paying with the smart card to interact with the system (i.e., check-in and check-out) upon boarding and alighting a vehicle. Travellers in the network can check scheduled departure times at all tram and bus stops, at most of which, real-time information is also available. For those using mobile internet, real-time information for the entire network is always available.

The urban public transport operator in The Hague, HTM, provided us with processed and anonymised AFC and AVL data from March 2015. The provided AFC data consists of journeys constructed by linking individual AFC transactions using a time-based transfer inference method wherein a trip with the same smart card identifier is included in the same journey as the previous trip if the boarding time of the second trip is within 35 minutes of the alighting time of the first. Since the data we received does not contain these unique identifiers, we are unable to follow individual cards across different journeys. Furthermore, since no information about the type of card or discounts applied is available, segmentation based on such variables is not possible.

The data consists of about 5.9 million inferred journeys and information on boarding and alighting time, stop, and line for each trip in every journey is available. In this study, we analyse and contrast route choice behaviour in weekday morning peak (06:00–09:00) and off-peak (09:00–16:00) hours. After filtering the datasets accordingly and applying the transfer inference procedure (described next), we are left with 1.02 million and 2.63 million journeys for the morning peak and off-peak hours, respectively.

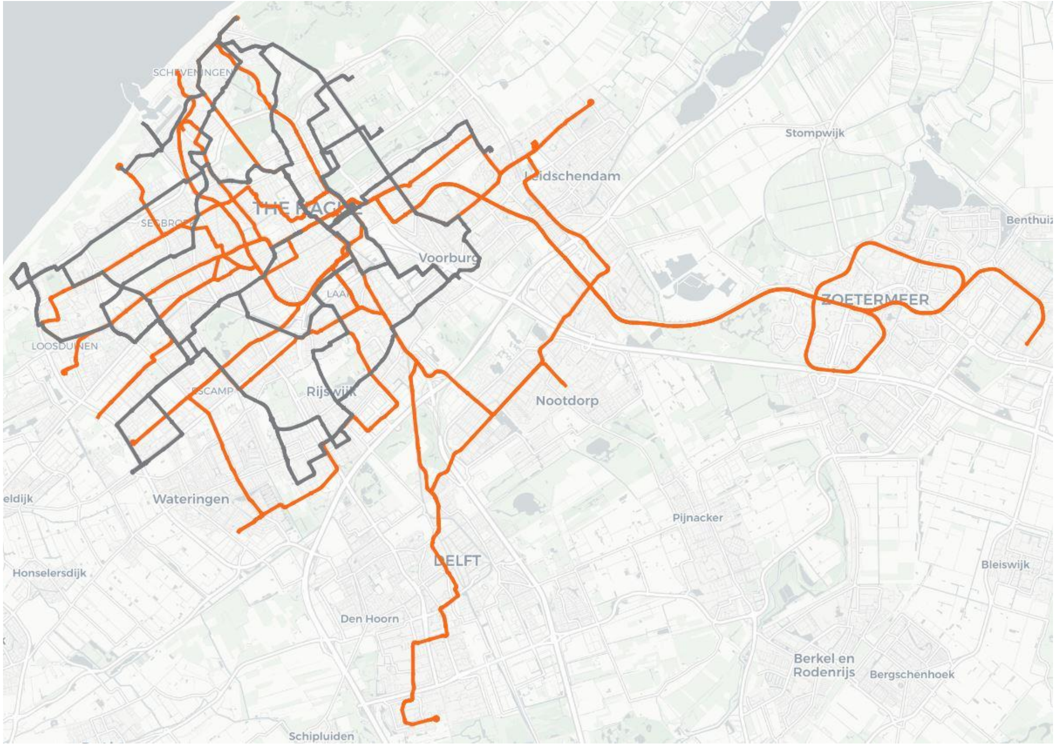


Figure 9: The Hague tram (orange) and bus (gray) networks in March 2015

3 Methodology

3.1 Data Preparation

Three data sources are used: (i) automatic fare collection, (ii) automatic vehicle location, and (iii) general transit feed specification (GTFS) data. While AFC data is the source for behavioural observations, the latter two, being data on the realized and scheduled operations, respectively, provide information on travel time characteristics including waiting time reliability.

To analyse route choice behaviour, complete journeys—as sequences of trips (i.e., rides on a single vehicle) without intervening trip-generating activities—have to be known. Although, the AFC data provided by the operator already consists of journeys, these are inferred by a time-based algorithms which tend to over-estimate the number of transfers (Yap et al., 2017). The time-based algorithm considers two trips to be part of the same journey if the time between the last check-out and next check-in is less than 35 minutes. Thus, trip-generating activities of shorter duration are ignored, leading to an overestimation of the number of transfers. This is particularly true for an urban public transport system where services have a relatively higher frequency and transfer times are rarely greater than the threshold. Therefore, we apply our own transfer inference algorithm to each journey to check whether the trips linked together by the operator indeed constitute one journey. For this, the AFC and AVL datasets were merged using the technique detailed in Luo et al. (2018).

The transfer inference algorithm is composed of one spatial and one temporal rule. The spatial rule ensures that the alighting stop of one trip and the boarding stop of the next are within 400 Euclidean metres (Yap et al., 2017) of one another. This places an upper bound on the distance travellers will walk to transfer. The temporal rule checks whether, after alighting, the first plausible service (or an earlier one) of the line used in the next trip is boarded. When this is the case, it is unlikely that a traveller will have performed an intermediate trip-generating activity. If the boarding stop is the same as the alighting stop of the previous trip, then the first plausible service is the same as the first service. That is, if the alighting and boarding stops for two consecutive trips are the same, then the first plausible service is the first vehicle of the line (actually) used in the second trip to arrive at the stop after the traveller has alighted. If the two stops are different, then it is the first service (of the line used in the next trip) after adding the time required by most people to walk between the two stations. For this, Euclidean distances and a walking speed of 0.66 m/s (Hänseler et al., 2016) are used. A slower walking speed is used to ensure that the first plausible service is feasible for the majority (in this case 97.5%) of the population. Note that since it is rare for vehicles in the network to be too crowded for passengers to board, we do not account for the possibility that a traveller does not board the first plausible vehicle due to overcrowding. As an example, to understand how the transfer inference algorithm works, consider an operator-inferred journey where a traveller goes from stop A to D, first taking line 1 from A to B, then walking from B to C, and finally taking line 2 from C to D. The traveller alights line 1 at 10:03 and boards line 2 at 10:20. Line 2 departs from C every 10 minutes (i.e., at 10:00, 10:10, 10:20, etc.) and the distance between stops B and C is 200m. The transfer thus passes the spatial rule as the distance is less than the threshold of 400m. Based on the assumed walking speed of 0.66m/s, the traveller should have arrived at stop C by about 10:05. Thus, the first plausible service of line 2 at stop C is the one at 10:10. Since the traveller instead boarded the service departing at 10:20, it fails the temporal rule and the two trips (A to B and C to D) are separated to two journeys.

Journeys containing transfer between the same lines (about 0.47% of the data) are not removed as long as they pass the above temporal criterion. This is done to accommodate travellers affected by planned and unplanned short-turning, stop-skipping or dead-heading. To avoid considering such routes as different alternatives, wherever the temporal criterion is passed for such transfers, the trips involved are merged into one (thus removing the extra transfer to the same line).

Applying this transfer inference algorithm reduces the number of journeys with at least one transfer from 18.8% of all journeys in the dataset provided by the operator to 10.8%. Journeys with more than one transfer make up about 0.5% of the total.

3.2 Choice Set Identification

A number of choice set generation methodologies have been proposed in literature. These approaches typically either require enumerating shortest paths or relying on a variety of behavioural assumptions (see Bovy (2009); Prato (2009) for an overview). However, studies analysing travel behaviour using AFC data have typically identified the route choice set for each origin-destination (OD) pair directly from the set of observed routes (Kim et al., 2019; Leahy et al., 2016; Yap et al., 2018). A potential disadvantage of this method is that we do not include in our analysis routes that are feasible but are only used for a few trips or not at all. However, since the data used for this study covers the entire network and covers a reasonably long period of time, this disadvantage is fairly low and the direct identification method is also suitable for our analysis.

The following filtering rules are applied to identify OD pairs (and associated route alternatives) for choice analysis: (i) each OD pair must have at least 2 route alternatives, (ii) each OD pair must have at least 200 trips between them, and (iii) each route alternative must make up at least 10% of the observations of its OD pair. While the first rule ensures that we are able to observe trade-offs between alternatives, the lower limits on the number of observations ensure that there is sufficient information to estimate behavioural parameters as well as to eliminate unusual observations that do not take place regularly. The set of eligible OD pairs is filtered iteratively using these rules until a stable set is obtained.

For the choice analysis, routes have to be uniquely defined such that travellers can be reasonably expected to perceive one route to be different from another. To this end, route alternatives are defined by the sequence of modes used and the boarding, intermediate and alighting stops. We consider lines (of the same mode) along a common corridor (i.e., traversing the same sequence of stops) to be perceived equivalently by travellers. Moreover, routes that have different transfer stops but are otherwise similar are distinguished as separate alternatives. This is to account for the fact that different stops may be associated with different waiting time reliability characteristics due to various reasons, such as planned transfer coordination.

Given these filtering rules, for the morning peak, we are left with 39 OD pairs and 30,606 inferred journeys suitable for choice analysis while for the off-peak, we have 105 OD pairs and 85,952 inferred journeys. As shown in Figure 10, most OD pairs in both time periods have only 2 alternatives.

3.3 Choice Attributes

Once all eligible OD pairs and associated choice sets are obtained, attribute values for the alternatives are assigned. The following attributes are used for the choice analysis: for each leg of the route, its (i) mode, (ii) in-vehicle time, and (iii) waiting time (and its components: scheduled waiting time, and regular and irregular deviations), and for the route alternative as a whole, its (iv) path size factor (indicating degree of overlap with other alternatives) and the (v) number of transfers. Attributes (i) and (v)—mode used in each leg and number of transfers—are directly known from the route definition. Travel time attributes are aggregated over each hour per day of the week (e.g., Mondays, 0900h–1000h) to account for the fact that travel time attributes may vary over time. Unfortunately, the data available does not include fare (or discount) information. However, for alternatives of a given OD pair, the price difference is typically in the order of cents. Moreover, paying by smart card may make it even more difficult for passengers to internalize costs. Having said that, we acknowledge that any fare effects will instead show up in the coefficients of other attributes, particularly, in-vehicle time, number of transfers, and mode, which are associated with price differences. We describe each choice attribute and its calculation below.

Modes and Number of Transfers

The mode used (denoted by m in Equation [7]) helps to understand how travel time components are weighed for different modes by travellers while the number of transfers (n^{trans}) is used to evaluate the transfer penalty, that is, the additional disutility beyond the transfer waiting time. Amongst the journeys eligible for choice analysis, all have a maximum of one transfer although the vast majority of observations are direct tram trips. Barring two OD pairs in the off-peak hours, whenever a bus-based option is available, it competes against a tram alternative. However, the overall proportion of observations where the choice between bus and tram is observed is also quite low: only 5 (5.12% of observations) and 11 (9.86% of observations) such OD pairs are available in the peak and off-peak hours, respectively. This may be expected given the different functions the two networks perform in a wheel & spoke-like network where the tram lines radiate out of the centre and the bus lines provide peripheral connections.

In-vehicle Time

Since the focus here is on waiting time reliability, only scheduled in-vehicle times (t_m^{ivt}) are used in the analysis. In-vehicle times are separately calculated for each stop pair and line combination by taking the mean of the scheduled values over the aggregation period. For common corridors, the in-vehicle times are assigned by taking the median of the in-vehicle times of the individual lines in the corridors.

Waiting Time

We use the reduced-form approach to include reliability effects in the route choice model. This method directly introduces statistical measures of travel times in the utility function (Börjesson et al., 2012). Usually, centrality and dispersion of realised travel times are used (Alemazkoor et al., 2015; Carrion and Levinson, 2012) with the aim of estimating how expectations of travel times (centrality measures such as mean or median) are traded-off against dispersion parameters (such as standard deviation or skew). However, unlike road traffic, since schedules exist for public transport networks, they may loom large in the decision process. Therefore, based on van Oort et al. (2015b) and van Oort (2016), we consider the following waiting time components to quantify the effect of reliability on route choice behaviour: (i) scheduled waiting time ($t_{0,m}^{\mathrm{wt}}$); and (ii) regular deviations ($t_{\mathrm{d},m}^{\mathrm{wt}}$) and (iii) irregular deviations ($t_{\mathrm{r},m}^{\mathrm{wt}}$; where r is replaced by the indicator name) of the realised waiting times from the schedule.

Regular deviations are calculated as the difference between the median of realised and scheduled values. For irregular deviations, a few studies on road traffic have compared indicators (Alemazkoor et al., 2015; Bogers et al., 2008; Bogers et al., 2006) but there is little consensus regarding which dispersion measures best represent the perception of travel time unreliability. van Lint et al. (2008) categorize these measures into (i) statistical ranges, (ii) buffer times, (iii) tardy trip measures, and (iv) probabilistic measures. The first two categories are commonly included in travel behaviour studies: statistical ranges (for instance, variance or standard deviation) measure the variation around the central value while buffer times indicate the extent of worse case scenarios, usually through the difference between the 90$^{\text{th}}$ or 95$^{\text{th}}$ percentile travel time and the median. In our analysis, we tested absolute and normalized formulations of these two commonly used dispersion measures (Table 3) to evaluate which representations best explain observed behaviour.

The effects of waiting times at origin (t^{owt}) and transfer (t^{twt}) stops are considered separately. We assume that travellers arrive uniformly at origin stops and subsequently consider half of the headway time (of relevant lines) as the origin waiting time. Previous studies have found that for headways up to 10–12 minutes, passengers tend to indeed arrive randomly (Ansari Esfeh et al., 2020; Fan and Machemehl, 2009; van Oort, 2011). Figure 11, top-left shows the distribution of calculated waiting times at origin (the headways are double these values). It can be seen that although our assumption is reasonable for most alternatives, for some, the headways are slightly higher. For these, we may be considering a slightly higher waiting time than reality as we can expect travellers to begin coordinating their arrivals to departure times (or be ‘less random’ Fan and Machemehl (2009)). Furthermore, we assume that travellers take the first vehicle available to them since denied boarding is rare in The Hague (Yap et al., 2018). Transfer waiting times, on the other hand, can be fully observed as the time between departures of the lines used to and from the transfer stop. However, for alternatives with common corridors, the expected transfer waiting time also depends on the lines used from the previous stop; and therefore ultimately on the assumptions made regarding traveller arrivals at the origin. For example, to obtain the waiting time at the first transfer stop, the proportion of travellers using each line at the origin stop (under given assumptions) is used to weight the feasible transfer times between the lines from the origin to the transfer stop and from the transfer stop onward.

Figure 11 shows the distribution of scheduled waiting times, and regular and irregular deviations for the route alternatives eligible for analysis. (Since it is ultimately selected for the choice models in section 4, only standard deviation is shown here.) While the origin waiting times are generally low, transfer waiting times show a stark difference between the two time periods. Lower transfer waiting times in the peak hours may be a result of increased transfer coordination by the operator. Characteristic of urban public transport networks, regular deviations are distributed around zero. Most values fall within a narrow deviation of 0.5 minutes for both time periods, indicating a fairly reliable service on average—although again a wider spread can be observed in the off-peak hours at transfer stops. On average, irregular deviations seem to be slightly higher in the peak hours presumably due to disturbances caused by the higher number of travellers on the public transport network and heavier road traffic.

High correlation between regular and irregular deviation indicators may affect choice analysis. Unlike dispersion indicators calculated for total travel times (for example in (Alemazkoor et al., 2015; Lam and Small, 2001)), a small negative correlation (0 to -0.3) is found between absolute values and median scheduled waiting times. Naturally, therefore, when the indicators are normalized with the median value, the magnitude of negative correlation is stronger (-0.5 to -0.8). For routes that include transfers, the variability of waiting times at transfer stops may depend on that for the origin stops (which is essentially the variation in headways) (Bates et al., 2001). For eligible routes in our analysis, we find only a small positive correlation (0 to 0.2) in the peak hours, while in the off-peak hours dispersion indicators for the origin and transfer stops of a route appear to be unrelated.

Path Size Factor

In order to account for overlap between the available alternatives, the path size factor for each route is calculated. We define the degree of overlap as the number of links shared with other alternatives and use the simplest form of the path size factor (Hoogendoorn-Lanser et al., 2005). Specifically, if route k of an OD pair traverses over links $l \in L_k$, and the number of alternatives using link l is n_l, then the route's path size factor, p_k, is given by Equation [5] ($|L_k|$ indicates the number of links in L_k). The path size factor lies between 0 and 1, with higher values corresponding to lesser overlap. As described in the following sub-section, the natural logarithm of the path size factor enters the systematic utility under multinomial logit.

$$p_k = \frac{1}{|\mathrm{L}_k|} \sum_{l \in \mathrm{L}_k} \frac{1}{n_l} \quad [5]$$

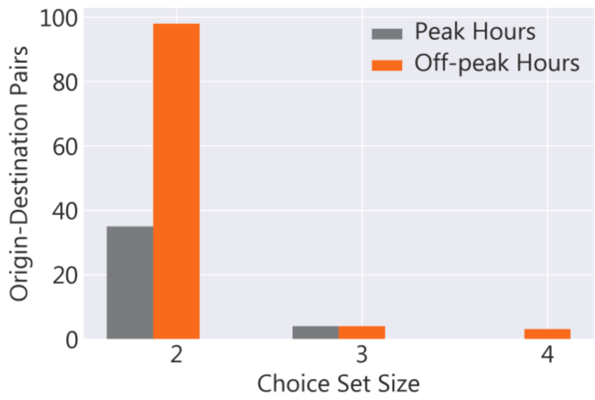


Figure 10: Choice set size distribution for the morning peak and off-peak hours

Table 3: Waiting time dispersion measures considered

Category	Statistical measure	Formula
Statistical range	Standard deviation	$\sqrt{\frac{\sum (x_i - \mu)^2}{N}}$
	Normalized variance (van Lint et al., 2008)	$\frac{p_{90} - p_{10}}{p_{50}}$
Buffer times	Reliability buffer time	$p_{95} - p_{50}$
	Reliability buffer index	$\frac{p_{95} - p_{50}}{p_{50}}$
	Normalized skew (van Lint et al., 2008)	$\frac{p_{90} - p_{50}}{p_{50} - p_{10}}$

For a given time period aggregate (hour-day of week), x_i is the ith realised value, μ is the mean of realised values, N is the number of realisations, p_m is the mth percentile of the realised values

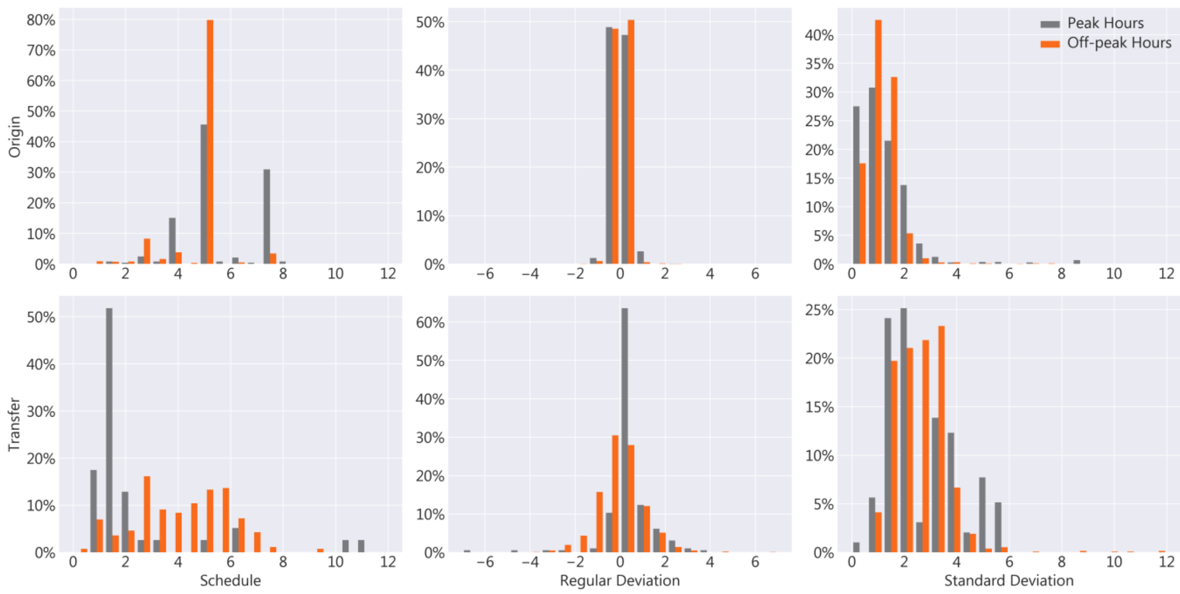


Figure 11: Scheduled, regular deviations, and irregular deviations for origin and transfer waiting times of route alternatives in the choice analysis (in minutes)

3.4 Choice Analysis

The effect of waiting time reliability on route choice behaviour is assessed under the conventional random utility maximisation (RUM) paradigm. Since the data provided does not contain individual identifiers, we treat each observation as independent (although in reality choices made by the same person may be correlated) and employ multinomial logit (MNL) models. Within the RUM paradigm, the utility of an alternative a, U_a, is composed of systematic (V_a) and random (ε) components. The systematic part is the product of the vector of taste preferences (β) and the vector of alternative attributes (x_a). The MNL model assumes that the random components are i.i.d. Gumbel distributed, which gives the probability of choosing alternative i from I alternatives as the following:

$$U_a = V_a + \varepsilon;\ V_a = \beta \cdot x_a;\ P_{ni} = \frac{e^{V_i}}{\sum_{a=1}^{I} e^{V_a}} \quad [6]$$

In the most generic form of the model, all attributes associated with separate legs of the route are mode-specific as shown in Equation [7] (symbol descriptions can be found in section 3.3). To examine the impact of including reliability parameters, we estimate one set of models with only planned values and another which also incorporates parameters for regular and irregular deviations. We also estimated models with quadratic terms for in-vehicle times and interaction effects between irregular deviations and in-vehicle times but this resulted in less generalizable models; that is, for values outside of the ranges observed here, the effects would be in the wrong direction. The effect of transfer distances was excluded because of the relatively small magnitude. Choice model parameters are estimated using PandasBiogeme (Bierlaire, 2018).

$$\begin{aligned} V = {} & \beta^{\text{trans}} \cdot n^{\text{trans}} + \beta^{\text{ps}} \cdot \ln(p) + \sum_{m\in \mathrm{M}} \beta_m^{\text{ivt}} \cdot t_{0,m}^{\text{ivt}} + \\ & \sum_{m\in \mathrm{M}} \beta_{0,m}^{\text{owt}} \cdot t_{0,m}^{\text{owt}} + \beta_{0,m}^{\text{twt}} \cdot t_{0,m}^{\text{twt}} + \\ & \sum_{m\in \mathrm{M}} \beta_{\mathrm{d},m}^{\text{owt}} \cdot t_{\mathrm{d},m}^{\text{owt}} + \beta_{\mathrm{d},m}^{\text{twt}} \cdot t_{\mathrm{d},m}^{\text{twt}} + \\ & \sum_{m\in \mathrm{M}} \beta_{\mathrm{r},m}^{\text{owt}} \cdot t_{\mathrm{r},m}^{\text{owt}} + \beta_{\mathrm{r},m}^{\text{twt}} \cdot t_{\mathrm{r},m}^{\text{twt}} \end{aligned} \quad [7]$$

We also separately validate the off-peak hour models (with and without reliability parameters), using a k-fold procedure where each fold is assigned observations from a unique OD pair. In one iteration of this procedure, we keep the observations of an OD pair as the test dataset; estimate the choice model using observations from the remaining OD pairs; and calculate the likelihood that this estimated model predicts the test data. This is then repeated until all OD pairs have been kept as the test dataset. After all iterations have been completed for a particular model, all prediction likelihoods are aggregated by taking their product. The two models can then be compared using the likelihood ratio test.

4 Results and Discussion

Table 4 shows the estimated route choice models, with and without reliability parameters for morning peak and off-peak hours. Estimated coefficients are scaled relative to that for in-vehicle times in trams to easily compare models. To arrive at the final models, statistically insignificant ($p > 0.1$) coefficients are removed (i.e., fixed to zero) one-by-one followed by reestimation until all coefficients retained in the model are significant ($p \leq 0.1$). Other exceptional conditions for keeping or removing parameters are discussed below.

Most parameters are statistically significant although, given the relatively low number of trips with transfers to buses, mode-specific coefficients for transfer waiting times were difficult to estimate with sufficient confidence. As discussed previously, we tried different irregular deviation measures in the choice models with reliability parameters. Likelihood ratio tests confirmed that in both time periods, all dispersion measures led to a better fit than models with only scheduled travel times ($p < 0.001$). For each time period, (with the exception of normalized variance in peak hours) fairly similar model fits were found and values for transfer penalties and waiting-to-in-vehicle time ratios were comparable. Using both indicator types (statistical range and buffer times) together, did not improve the model significantly and led to counterintuitive results.

Other revealed preference studies have also reported similar results. In their analyses of choices between a free and tolled route, Lam and Small (2001) find that reliability buffer time and standard deviation perform almost the same in terms of model fit while Carrion and Levinson (2013) find that using the mean and standard deviation gives the best fit followed closely by the combination of median and reliability buffer time. The fact that we find comparable models with different indicators may be because, similar to the abovementioned studies, in each time period the indicator types are generally highly correlated. Ultimately, we choose to present results with standard deviation as the dispersion indicator in both time periods. As noted below, the coefficients for the off-peak hours seem more in line with expectations; therefore, we choose the indicators that fit best for that time period. Since standard deviation is also most commonly used to calculate reliability ratios in literature, using this indicator also enables comparison. In the following, unless noted otherwise, we discuss parameters of the models with reliability parameters.

Unexpected effect of irregular deviations in peak hours

Some coefficients for origin waiting time in peak hours are unexpected. The scheduled waiting to in-vehicle ratio (in the model with reliability parameters) at the origin for trams is 0.81, indicating that travellers value in-vehicle time more than waiting time, while the opposite effect has been typically found. We also confirmed (by means of a t-test) that the two parameters are indeed different with a statistical confidence of >99%. Furthermore, irregular deviations—both statistical ranges and buffer times as indicators—are found to have a positive effect. Leahy et al. (2016), who analyse smart card data from London, also find higher travel time standard deviation to have a positive effect on the utility of one of the mode combinations studied and suggest that this is indicative of risk-seeking behaviour. We, however, submit that this anomaly in our results may have arisen because crowding and dispersion measures are correlated in the long run. That is, as more travellers choose a particular line it becomes more unreliable because of delays due to greater boarding and alighting times. If in reality unreliability has a small effect on travellers' choices, the estimation procedure would find that travellers tend to choose the unreliable alternative because of this underlying long-run relationship. This phenomenon would be particularly in effect in peak hours because concentrated demand then would lead to higher crowding which causes the unreliability.

Table 4: Estimation results

		Peak Hours (06:00-09:00)						Off-peak Hours (09:00-16:00)					
		Scheduled values only			With reliability param.			Scheduled values only			With reliability param.		
	Initial LL	-22420.31			-22420.32			-62909.21			-62909.21		
	Final LL	-20964.06			-20938.99			-57663.89			-57544.00		
	Parameter	**Coeff.**	**p-val**	**Scaled**	**Coeff.**	**p-val**	**Scaled**	**Coeff.**	**p-val**	**Scaled**	**Coeff.**	**p-val**	**Scaled**
Number of transfers	β^{trans}	-1.960	0.000	10.103	-1.520	0.000	7.958	-2.420	0.000	14.405	-2.060	0.000	12.118
Path size factor	β^{ps}	0.373	0.004	-1.923	0.234	0.076	-1.225	-0.208	0.017	1.238	-0.117	0.183	0.688
Scheduled in-vehicle times	$\beta^{ivt}_{0,bus}$	-0.086	0.004	0.445	-0.079	0.016	0.414	-0.186	0.000	1.107	-0.184	0.000	1.082
	$\beta^{ivt}_{0,tram}$	-0.194	0.000	1	-0.191	0.000	1	-0.168	0.000	1	-0.170	0.000	1
Scheduled waiting times	$\beta^{owt}_{0,bus}$	-0.338	0.000	1.742	-0.337	0.000	1.764	-0.281	0.000	1.673	-0.287	0.000	1.688
	$\beta^{owt}_{0,tram}$	-0.167	0.000	0.861	-0.155	0.000	0.812	-0.204	0.000	1.214	-0.262	0.000	1.541
	$\beta^{twt}_{0,bus}$	-0.223	0.000	1.149	-0.227	0.000	1.188	0	(fixed)	0	0	(fixed)	0
	$\beta^{twt}_{0,tram}$	-0.179	0.021	0.923	-0.244	0.010	1.277	-0.104	0.000	0.619	-0.216	0.000	1.271
Regular deviations	$\beta^{owt}_{d,bus}$				0	(fixed)	0				-0.135	0.013	0.794
	$\beta^{owt}_{d,tram}$				-0.130	0.000	0.681				-0.269	0.000	1.582
	$\beta^{twt}_{d,bus}$				0	(fixed)	0				0	(fixed)	0
	$\beta^{twt}_{d,tram}$				0	(fixed)	0				-0.247	0.000	1.453
Irregular deviations (standard deviation)	$\beta^{owt}_{std,bus}$				0.141	0.035	-0.738				-0.323	0.000	1.900
	$\beta^{owt}_{std,tram}$				0.053	0.000	-0.280				-0.053	0.000	0.314
	$\beta^{twt}_{std,bus}$				0	(fixed)	0				0	(fixed)	0
	$\beta^{twt}_{std,tram}$				-0.169	0.000	0.885				0	(fixed)	0

Waiting times

Other parameters in the study are generally in line with expectations. In the off-peak hours at origin stops, for both, trams and buses, 1 minute of waiting time is about 1.55 in-vehicle minutes in the respective modes. This is comparable to previous revealed preferences studies which have found similar values in urban public transport networks with ratios ranging between 1.5–1.7 (Kim et al., 2019; Nassir et al., 2018; Yap et al., 2018). Transfer waiting time to in-vehicle time ratios for trams in the off-peak hours is about 1.27, indicating that waiting time at transfers has a smaller impact on route choice than that at origin. While most studies lump origin and transfer waiting times together, results from Guo and Wilson (2011), who focus on transfer inconvenience, also indicate that the latter has a lower effect. The lower weight attached to transfer waiting time compared to origin waiting time may be indicative of travellers adopting a strategy-based decision rule where the transfer waiting time would depend on the line boarded at the origin. Another possible reason could be that travellers value waiting time costs closest to them more whilst discounting future losses. That is, the anticipated disutility of waiting at a transfer stop is lower or the traveller pays less attention to it simply because it is further in the future. Since the route observations suitable for choice analysis do not contain journeys with more than one transfer, we are unable to check if the impact of waiting time at subsequent transfers would be even lower.

Transfer penalties

We estimate transfer penalties of more than 8 and 12 in-tram minutes in the peak and off-peak hours, respectively. In comparison, other revealed preference studies for public transport networks have reported values ranging from as low as 3–5 minutes (Guo and Wilson, 2011; Yap et al., 2018) to values in the vicinity of ours (Nassir et al., 2018) to extremely high penalties of 0.5–2 hours (Han, 1987; Jánošíková et al., 2014; Kim et al., 2019). A higher transfer penalty in the off-peak hours could be a result of travellers having more modest time constraints than commuters in the morning peak or because travellers are more worried about missing a transfer as the frequencies are slightly lower in this time period.

Effect of overlapping alternatives

Since the logarithm of the path size factor is itself negative, a positive coefficient indicates a penalty that corrects for correlation due to overlapping routes. Negative coefficients have also been found for public transport networks which have been interpreted as overlapping routes adding robustness thus making them more attractive (Hoogendoorn-Lanser et al., 2005). We find positive and negative coefficients for the peak and off-peak hours, respectively. The negative sign for the off-peak hours may, again, be indicative of travellers seeking more robustness in light of slightly higher headways in this time period.

Tram bonus

Previous research (both stated and revealed preference studies) has indicated that travellers in urban networks find each minute on a bus to be equivalent to 1.2–1.67 minutes on a tram (Axhausen et al., 2001; Bunschoten, 2012; Yap et al., 2018). We too find a consistent tram bonus on the lower end of this range. Although in Table 4, peak hour in-bus coefficients are much lower than those for in-tram, closer inspection reveals that, correspondingly, origin waiting times for trams are weighted much lower than those for buses. Using mode-agnostic (generic) coefficients for waiting times reveals a bus to tram in-vehicle time ratio of 1.2, indicating a small preference for trams. For the off-peak hours, using either mode-specific or generic coefficients for waiting times, one minute in a bus is perceived as approximately 1.1

minutes in a tram. Furthermore, in the off-peak hours waiting for buses is valued slightly worse than trams. Since many stops in The Hague serve both trams and buses, it is unlikely that this is caused by different waiting conditions. A more likely reason may be that travellers perceive waiting for trams to be less uncertain than for buses. Note that this would not necessarily be captured in the schedule deviation terms as this perception may be independent of empirical values.

Regular and irregular deviations in waiting time

Regular deviations for trams in the off-peak hours are evaluated as 1.03 and 1.14 times the scheduled values at origin and transfer stops, respectively. This may indicate that travellers do not really consider these values separately but rather internalize the distribution of actual waiting times in their decision making. In contrast, regular deviations for buses are weighted about half of their scheduled waiting times. A possible reason for this is that of the observations suitable for choice analysis, the majority of those choosing a bus at the origin (in the off-peak hours) are along a corridor served by two bus lines and a tram line. Since the two bus lines are along a common corridor, they form a single choice alternative with a net higher frequency which may have reduced the importance of regular deviations.

We calculate the reliability ratio to compare our dispersion parameter estimates against a fairly large number of studies who also calculate this indicator albeit typically for total travel time rather than only waiting time. The reliability ratio is defined as the 'marginal rate of substitution between average travel time and travel time variability' (Li et al., 2010) and can be calculated as the ratio of the coefficients of time variability to the coefficient of average time in the utility function. A smaller value indicates a weaker impact of travel time variability (irregular deviations in our case) relative to the impact of average travel times (scheduled waiting times in our case). In the morning peak, at transfer stops, we find a reliability ratio of 0.69 for trams. In the off-peak hours at origin stops, reliability ratios for trams and buses are found to be 0.20 and 1.12, respectively.

Since previous studies calculate reliability ratios for total travel time there is little empirical evidence that can be directly compared. However, in comparison to these values, our findings are overall in agreement. Literature reviews (Carrion and Levinson, 2012 (Figure 3); Li et al., 2010) have found a wide range of reported reliability ratios from 0.1 to 3.3. These studies focussing mostly on car traffic contain results from both stated and revealed preference studies. In their meta-analysis, Carrion and Levinson (2012) do not find the type of data (stated or revealed preferences) to have a significant effect on the value of reliability. However, amongst the studies they reviewed, those that had carried out analysis on both stated and revealed preferences, reported that estimates from the latter were typically higher (e.g., Ghosh (2001); Small et al. (2005)). In contrast, Bates et al. (2001) argue that protest responses in stated preference experiments may lead to higher value of reliability ratios for public transport. Recent empirical evidence concurs with these expectations: studying crowding valuations using smart card data, Yap et al. (2018) conclude that stated preferences tend to overestimate values. Leahy et al. (2016), whose study using smart card data is the closest to ours, find reliability ratios for the London Underground to be below 0.6 in two model specifications (higher ratios were found for light and heavy rail modes), in line with our results.

The fact that our estimates find that travellers do not react too strongly to irregular deviations, means that these measures contribute fairly little to improving model performance; especially, given the limited disturbances (indicating a reliable service) in the public transport network in The Hague.

Model validation

Model validation using a *k*-fold procedure returned an aggregate log-likelihood of -57725.09 and -57799.80 for off-peak models with and without reliability parameters, respectively. Both of these have a poorer fit in comparison with models estimated with the entire dataset (Table 4). However, this is expected with models estimated from cross-validation folds. Moreover, the performance gap between them is smaller although the likelihood ratio test confirms that the model with reliability parameters is still better ($p < 0.001$). Analysis of predicted probabilities for observed choices did not reveal any obvious patterns between difference in performance of the two models and reliability attributes of the alternatives. Thus, the model with reliability parameters did not, as such, describe behaviour better for OD pairs with greater differences in reliability. Given the generally small deviations from schedule in the public transport network of The Hague and the relatively small coefficients for reliability, the fact that reliability parameters add little predictive power is not surprising. Validation tests for the peak hour models found aggregate log-likelihoods of -21535.88 and -21412.92 for models with and without reliability parameters, respectively, indicating that the model with reliability parameters was overfitting the data. This could possibly support our hypothesis that the unexpected signs of irregular deviation coefficients in the peak hours models were caused by data resulting from a specific situation rather than describing the underlying behaviour.

5 Conclusion

In this study, we evaluate the impact of waiting time reliability on route choice behaviour in public transport networks. Unlike the majority of studies on this topic, rather than stated preference surveys or laboratory experiments, we use revealed preferences derived from passively collected AFC data. While several studies have used smart card data for analysing choice behaviour, to the authors' best knowledge, this is the first study explicitly analysing the impact of waiting time reliability using all AFC transactions in the network. As a case study, we analyse the urban public transport network of The Hague in the Netherlands. Different models are estimated for peak and off-peak hours; and, as far as possible, separate coefficients are estimated for different modes, and origin and transfer stops. Furthermore, we used both statistical range and buffer time type waiting time dispersion measures to evaluate which best represents travellers' perceptions of reliability.

All tested dispersion measures performed nearly equally well although in comparison with most previous studies, we find a relatively small effect of unreliability on route choice behaviour. Reliability ratios estimated with standard deviation were in the range of 0.20–1.12. In addition to the possibility that our values are lower because we use passively collected revealed preferences rather than stated preferences, the fact that public transport in The Hague is overall quite reliable may have also contributed to smaller reliability coefficients. In other networks where travellers have to regularly face delays, we may find more risk averse behaviour. This may be investigated further in future studies, particularly because a number of studies (e.g., Bordagaray et al. (2014); Soza-Parra et al. (2019)) have found that reliability is usually on of the most important stated satisfaction determinants. Differences in behaviour between the two time periods are also found, arising mainly from travellers being wary of missing transfers in the off-peak hours. Further, small differences are found in travel time weights for different modes and origin/transfer stops.

This study has several limitations stemming from the nature of passively collected data. Due to privacy regulations, we do not have unique identifiers that link different journeys in the data. Typically, assuming each choice observation to be independent (as we do here) leads to poorer model fit. Moreover, we also cannot comment on the nature of the heterogeneity in taste preferences of travellers. Although the AFC system used in The Hague allows us to record complete trips, origin waiting times are not observed, forcing us to make an overarching assumption regarding passenger arrivals at stops. This may have led to overestimation of waiting times and subsequently underestimation of the impact of waiting time. Furthermore, the absence of fare data from this period meant that we could not include it in our model. However, we do not expect this particular limitation to have a significant impact on our model.

Passively collected revealed preferences have a number of advantages over stated preferences, in particular the absence of hypothetical bias and the need to convey probabilistic information. However, by their nature, such data lacks experimental control. Thus, we did not necessarily observe all trip types (with/without transfer) of all modes (trams/buses) in equal numbers, which may have resulted in some of our coefficients being insignificant. More importantly, the lack of control over causality may have led to anomalies such as travellers preferring more unreliable lines. Experimental setups that can disentangle causally linked variables are an interesting avenue for future research into using such revealed preferences.

Finally, we reiterate an important assumption made in this study (and similar revealed preferences-based studies in literature): although we use observations of choices made in real-life, our analysis is made under the assumption that travellers are able to internalize and integrate empirical measures (such as median or standard deviation) of waiting time distributions into their decision-making process. The idea is that, on average, these measures should represent travellers' perception of travel time but, clearly, this assumption will not always hold true. Even experienced travellers, who may be fairly aware of waiting time distributions, would be influenced by personal beliefs and subjective probability weighting. Future work may also want to focus efforts on this—explicitly accounting for such uncertainties by using more complex models of decisions under risk and uncertainty (Li and Hensher, 2019).

Subjective Waiting Time Uncertainty

Although waiting times are inherently uncertain in public transport networks, like the previous chapter, most research has primarily studied route choice behaviour under (objective) risk. In this chapter we propose a method to assess travellers' route choice behaviour under natural ambiguity. Specifically, we devise a realistic route choice situation whereby travellers' attitudes and perceptions towards waiting time uncertainty as well as the effects of situational contexts thereon can be quantified in terms of a certainty equivalent. The proposed method provides snapshots of travellers' behaviour under uncertainties in real-world public transport systems and can be used to improve transportation models, provide more tailored travel advice, and test the efficacy of different policies.

Two case studies are performed: first, the identified choice situation is contextualised for the Dutch railways within a stated choice experiment; and second, a natural experiment is developed by extracting the situation from smart card data of the urban public transport network of Amsterdam. Using choice observations from the two experiments, we estimate travellers' uncertainty evaluation in the two networks and discuss the impact of explicitly accounting for waiting time uncertainty. The chapter is concluded with an outline of the main contributions, outcomes, and limitations.

This chapter is based on edited versions of the following article and conference presentation:
Shelat, S., Cats, O., van Lint, J.W.C. Quantifying travellers' evaluation of waiting time uncertainty in public transport networks. Travel Behaviour and Society (2021).

Shelat, S., Dixit, M., Cats, O., van Oort, N., van Lint, J.W.C. What does smart card data reveal about subjective beliefs regarding waiting time uncertainty? 8th International Symposium on Transport Network Reliability (2021).

1 Introduction

Analysing route choice behaviour in public transport networks is important for both supply and demand management. It is an essential input for determining network flows which authorities use to manage service levels and prioritize relevant investments. Furthermore, knowing how such decisions are made, travellers can be nudged into choices that are more optimal for them and the system, and can be suggested options that are likely to result in higher traveller satisfaction. Route choice decisions are largely governed by travellers' attitudes and attributes of the public transport system. Increasingly, route choice models have incorporated service attributes beyond travel time components, including, for instance, graphical distortions of transit network representation (Raveau et al., 2011), transfer station layouts (Guo and Wilson, 2011), and on-board crowding (Yap et al., 2018). This study contributes to this line of research by assessing travellers' evaluations of waiting time uncertainty above and beyond nominal values. Given the various sources of stochasticity in public transport networks, its travel time attributes are inherently uncertain; however, as we will show in our literature review, this has not been accounted for properly in existing studies. In order to describe and explain route choice decisions more completely, we develop in this study a route choice model that explicitly accounts for travellers' behaviour under waiting time uncertainties in public transport networks.

First, we clarify what we mean by 'uncertainty'. The Knightian (Knight, 1921) classification of uncertainty is based on whether, for a set of possibly infinite events, objective probabilities exist or not. Decisions under the former regime are said to be made under 'risk' while those under the latter are under 'ambiguity' or 'uncertainty'. Objective probabilities exist either when they are made available to decision-makers (and are trusted by them), there is a consensus amongst decision-makers regarding them, or when they are integrated within the decision problem itself. However, these assumptions are quite stringent and are seldom fulfilled in the real world. Outside of artificial games such as casinos and lotteries, real-world events occur under ambiguity where decisions are based on personal beliefs (Machina and Siniscalchi, 2014). Travellers in public transport networks also do not have access to such objective probabilities for the different attributes involved and make their decisions under uncertainty. Even if information is provided on the various aspects of travel time, it is distorted by travellers' beliefs arising from personal characteristics, habits, experiences, and contemporary contextual variables.

Next, we delineate why uncertainty in waiting time is of special interest. Similar to other industries in the service sector, in public transport systems too, waiting times have been found to play a crucial role in consumers' decision-making and satisfaction (Abenoza et al., 2018). While the cost of waiting can usually be objectively calculated in the manufacturing industry, to describe its manifestation in the service sector, Maister (1985) quotes the copywriters of a parcel delivery service: 'waiting is frustrating, demoralizing, agonizing, aggravating, [and] annoying…'. Arguably, these feelings arise from the uncertainty that is often inherently involved with waiting time as well as the context in which it is experienced. In the service industry, apart from the objective magnitude, the perception of waiting time is critical for customer satisfaction (Maister, 1985) and any disparity between objective and subjective expectations of waiting times may lead to sub-optimal decision-making. Therefore, it is vital to analyse travellers' attitudes and perceptions regarding waiting time uncertainty.

The impact of waiting time on route choice behaviour has been typically studied using either expected values or objective probabilities of risk. Both of these approaches fail to account for travellers' beliefs regarding uncertainties associated with waiting time. To that end, the present study proposes a method to assess travellers' route choice behaviour under natural ambiguity without using objective probabilities or assuming specific learning behaviour—important drawbacks in existing studies. Specifically, a realistic route choice situation is proposed whereby travellers' beliefs towards waiting time uncertainty can be quantified in terms of a certainty equivalent. For any gamble, its certainty equivalent is a risk-less value such that the decision-maker is indifferent between receiving this risk-less value and playing the gamble. For example, if a decision-maker is indifferent between (a) gambling on a fair coin toss winning $0 on heads and $5 on tails; and (b) winning $2 for sure, then $2 is the certainty equivalent of the gamble offered in (a) for this decision-maker. The certainty equivalent in this case indicates the decision-maker's attitude towards risk—risk-aversion in this case. When the gamble is ambiguous/uncertain, such as wining $0 if a train departs within 1 minute of its scheduled time, $5 otherwise, in addition to attitude towards the uncertainty, the certainty equivalent will also indicate how uncertain an outcome is felt to be by the decision-maker. For instance, if the certainty equivalent for the above was $1, we can infer that the decision-maker feels that the train is more likely to be on time than to be late. The identified choice situation also permits the estimation of the effects context variables have on the certainty equivalent for waiting time. The conditions required for the proposed situation are simple enough that it is fairly common for it to take place structurally (i.e., because of service or network design) in real-world public transport networks; also implying that most travellers will be able to identify with the situation. As case studies, the proposed choice situation is (i) contextualized for the Dutch railways and used in a stated preferences experiment and (ii) extracted from the smart card data of the Amsterdam urban public transport network for a (natural) field experiment.

In the next section, studies on travel behaviour under uncertainty are reviewed, classifying them on the type of uncertainty observed. Section 3 lays out a theoretical framework of choice behaviour under uncertainty and section 4 presents the proposed choice situation. This is followed by the design and results of the two case studies in sections 5 and 6. Finally, the main contributions, outcomes, and limitations are outlined in section 7.

2 Literature review

In this section, we briefly review the large body of literature dedicated to analysing the effect of variability in different aspects of travel time on travellers' decisions. While these studies may fulfil their own objectives, here we analyse drawbacks specifically with respect to observing and analysing behaviour under uncertainty. Decisions have been typically observed under risk, simulated uncertainty, or natural ambiguity. Research approaches—stated preference experiments, laboratory experiments, or analysis of actual trips—have been closely associated with the type of uncertainty under which decision-making has been observed and is accounted for in the analysis.

As discussed above, in the real-world, decisions are made under ambiguity—in the absence of objective probabilities. In contrast, however, travel behaviour under uncertainty is most commonly studied by presenting hypothetical route alternatives with objective distributions of travel times. Furthermore, since such probabilities are usually not available to travellers, conveying objective probabilities is notoriously difficult (Bates et al., 2001; Carrion and Levinson, 2012). This is exclusively the type of uncertainty observed in stated preferences (e.g., Small et al. (1999); Swierstra et al. (2017); Tilahun and Levinson (2010)).

A few laboratory experiments have observed choice in traffic networks under partial uncertainty by offering different levels and accuracies of information to respondents within the context created in a 'travel simulator' (e.g., Ben-Elia et al. (2013); Ben-Elia and Shiftan (2010); Bogers et al. (2005); Bogers et al. (2006); Ramos et al. (2011)). Unlike stated preference questionnaires, respondents do not make single-shot decisions but are required to consider a number of choice situations with or without feedback. These experiments typically focus on analysing learning mechanisms (e.g., Avineri and Prashker (2005, 2006)) and the effects of different uncertainty levels (e.g., Ben-Elia et al. (2008)). In an interesting setup, Kemel and Paraschiv (2013) observe choices under artificial ambiguity using Ellsberg's urns (Ellsberg, 1961). Artificial ambiguity is typically created using an unknown mix of differently coloured balls in an urn. This approach is often used in ambiguity studies to control for likelihood beliefs in a laboratory setting. Since participants do not have any information regarding the proportions of different colours, they cannot form any beliefs about this. Kemel and Paraschiv (2013) as well as a number of authors (as summarized in Baillon et al. (2018)) note that the external validity of such studies could be improved by using natural (real-world) events (for example, stock market prices or actual departure times of public transport vehicles).

Studies observing behaviour under natural ambiguity are sparse and typically use revealed preferences from real-world observations in car traffic. While revealed preferences offer high behavioural validity, unlike stated preferences and laboratory experiments, there is little experimental control. A series of papers observed route choice behaviour in two road-pricing demonstrations in California, involving a free but congested route and a (time-varying) tolled route with low congestion levels (and hence an almost certain travel time), just before the beginning of this millennium (see Brownstone and Small (2005) for an overview). While these fairly unique opportunities offered reasonable choice experiment settings, the studies faced significant issues in data collection and preparation.

While in reality travellers do not have access to objective probabilities, studies using observations of choices under risk face an additional problem that is related to the difficulties in conveying probabilities. Empirical findings suggest that for choices under risk, people do not fully distinguish between different levels of probability (Wakker, 2010, section 7.1) as is assumed in the commonly adopted expected utility regime. In recent years, however, a few studies (see Li and Hensher (2011a, 2019); Rasouli and Timmermans (2014) for a review) have used rank-dependent utility (Quiggin, 1982) and cumulative prospect theory (Tversky and Kahneman, 1981) which apply subjective probability weighting that can account for such likelihood insensitivity. However, only a few studies estimate the functional form and parameters of the probability weighting function (Li and Hensher, 2011a, 2019). Finally, an important issue in studies using revealed preferences data is that, although decisions are made under ambiguity, analysis has been commonly carried out using objective probabilities (Carrion and Levinson, 2012) under the assumption that these probabilities are known to the traveller through experience (Ghosh, 2001; Lam and Small, 2001; Small et al., 2005).

An alternative to analysing travellers' attitudes and perceptions regarding uncertainty through choice observations could be to directly ask them about their perceived and expected travel times. The idea is that reported travel time values will incorporate any uncertainties experienced by travellers. This approach has been implemented in a number of studies researching the effect of various aspects of travelling, such as real-time information provision (Dziekan and Vermeulen, 2006; Watkins et al., 2011), on perceived waiting times (see Meng et al. (2018) for a brief overview). This approach is useful to assess a posteriori travel satisfaction. However, Peer et al. (2014) find that reported values do not accurately describe those used for decision making, suggesting that discrepancies between objective and reported values may arise from a number of reasons that do not actually affect travellers' behaviour. Furthermore, even incentivising travellers to report their true beliefs through scoring rules (see Winkler et al. (1996) for an overview) does not seem to reduce discrepancies or improve interpretability (Dixit et al., 2019).

From this review, several drawbacks in existing studies can be identified with respect to analysing route choice behaviour under uncertainty in public transport networks. In most studies, choices observed have been made and/or analysed using objective probability distributions, which are not only missing in the real-world but are also distorted by travellers' prior beliefs that arise from a number of factors such as previous experiences, habits, and contexts, leading to possibly biased outcomes. Studies where choices observed have been made under uncertainty—as in revealed preferences and laboratory studies—were only performed in the context of car traffic networks leaving an important gap for studying behaviour in public transport networks.

To overcome these drawbacks, we identify a choice situation wherein travellers' assessment of waiting time uncertainty in public transport networks can be explicitly quantified directly from observed choices; without external psychometric measurements or collection of reported values. First, however, we present a generic theoretical framework of travel behaviour under uncertainty that outlines the various factors affecting choice and their interactions with the aim of placing the current study in context.

3 Theoretical framework

In order to describe decision-making under uncertainty, we divide the process into three main parts: (i) uncertainty evaluation, (ii) decision-making, and (iii) learning (Figure 12). Evaluation of uncertainty in attributes is a result of the decision-maker's attitudes (e.g., risk aversion) as well as their perception of the system (e.g., feeling that the system is unreliable). Both attitudes and system perceptions, and therefore the uncertainty evaluation, can be affected by the context, which can be situational or affective. The former affects the environment in which the decision is made while the latter relates to the moods and feelings of the decision-maker at the time. These evaluations are then used to assess and compare alternatives leading to a choice. After making a choice, the resolution of some or all of the uncertainty may be observed by the decision-maker, which feeds back to their experience memory. Previous experiences can lead to longer-lasting changes to their attitudes or shorter-term changes to their system perception. This can take place either over several decision outcomes or after a few extreme ones. Experiences also lead to habit formations and the regularity with which the same choices are made can affect perceptions (e.g., regular cyclists might perceive cycling to be safer than occasional cyclists). Finally, the effect the decision-maker's system perception has on their experienced utility (travel satisfaction) and habits closes this short-term learning loop. Note that we do not propose this framework as a validated scheme but use it to highlight and conceptually place the aspects considered in this study.

In this study, we focus on the evaluation of uncertainty which is dependent on personal characteristics developed over a long period of time and system perceptions that are updated more frequently, as well as the effects contexts (we only study situational contexts and not affective ones) have on them. We assume decisions are made under the random utility maximization paradigm. Furthermore, the focus is on capturing snapshots of travellers' attitudes and perceptions; therefore, we do not study the feedback and learning mechanism involved in uncertainty evaluation.

Personal characteristics + system perceptions = uncertainty evaluation

Theoretically, personal characteristics and (subjective) perceptions of risk are distinguished to study which of these are the driving forces behind behaviour under uncertainty (Weber and Milliman, 1997). Anticipation of regret and attitudes towards risk and uncertainty are amongst the most influential personal characteristics for decisions under uncertainty. These personal characteristics are developed over a long period of time and are not susceptible to frequent changes. They have been quantified in literature in a number of ways from Likert scales to various mathematical formalizations in decision models including expected utility, cumulative prospect theory, and regret theory. Unlike attitudes, subjective perceptions are updated frequently based on habits and experiences (gaps between expectations and outcomes). A number of models (e.g., Bayesian updating, weighted average learning) have been proposed for the learning mechanism through which these three aspects—perceptions, habits, and experiences—interact with one another.

Practically, however, it is difficult to disentangle the effects of personal characteristics and perceptions in observed behaviour. For instance: does a person buy theft insurance because she feels theft is likely to occur or because she is generally risk averse in these matters? In single attribute experiments, outcome valuation and subjective probabilities have been successfully disentangled, for instance using the trade-off method (Wakker, 2010) but it is not obvious how this would be done in multi-attribute decisions such as route choice. When using non-expected utility models for decisions under natural ambiguity, only recently have studies explicitly measured ambiguity aversion whilst controlling for likelihood beliefs (Baillon et al., 2018). Indeed some (Nau, 2001) have argued that the separation of preferences arising from personal characteristics and beliefs is neither possible nor required for decision analysis or economic modelling. Therefore, for this study we consider travellers' uncertainty evaluation as a whole which, in fact, is formed by their personal characteristics and perceptions. We will continue using this term in the remainder of the chapter. Note that, we use 'uncertainty evaluation' as an all-encompassing term; referring not only to how likely a decision-maker feels that a particular event will occur but also the impact (or value) thereof.

Situational contexts

Contemporary contexts affect how an attribute (e.g., waiting time) is experienced. For waiting time, Maister (1985) makes a number of propositions that define which contexts make waiting seem longer or shorter than reality; for instance, occupied time feels longer than unoccupied time or that unexplained waits are longer than explained ones. Jones (1996) reviews these propositions in terms of the degree to which service managers can control the related contexts and their impacts on customers. Previous studies in transportation have also explored the differences in value of travel time for different contexts such as free-flow traffic, stop-and-go traffic, and on-ramp delays (Hensher, 2001; Levinson et al., 2004). Ongoing experience is important because it will be taken into account by customers when anticipating the value of uncertain attributes in the upcoming future.

With increasingly prevalent real-time information, seemingly irrelevant information may also affect travellers' evaluation of uncertainty. For instance, delay predictions along the corridor of a traveller or even in other parts of a transportation network might cause increased anxiety and a breakdown of trust in the system, leading to choices that indicate a disproportionately higher degree of pessimism or risk/ambiguity aversion.

As a contextual variable, the amount of waiting time already experienced by the time of decision may have two opposite effects of varying magnitude. On the one hand, greater experienced waiting time translates to increasing stress, frustration (Osuna, 1985), and tiredness (based on waiting conditions); on the other, there may be a sunk-cost effect (Thaler, 1980) wherein having waited for some time is in itself an impetus to wait some more. In an explicit study on the sunk-cost effect for time (rather than money which most authors examine), Soman (2001) finds that because people do not have the ability to account for time in the way they do for money, the effect is not found. However, he does not consider travel time in transportation choices where, often, one time component is traded-off with another in the same trip which could make it easier for people to open and keep mental accounts of time.

4 Choice situation

In this section, we present the choice situation that will be analysed to obtain travellers' evaluations of uncertainty in waiting time and the effects of contexts thereon. Amongst the sequence of choices faced by a traveller, we look at the decision of whether to board a particular vehicle in the following situation.

Consider a traveller who arrives at a public transport stop. From here, either of the next two vehicles can take her to her destination. Both of these vehicles are identical in every way except for their departure and arrival times at the origin and destination stations, respectively. Furthermore, both of these vehicles will take her directly, without any transfers, to the destination station. As is prevalent in many transit systems worldwide, real-time information regarding anticipated departure times and delays is displayed alongside scheduled departure times. Moreover, the traveller is assumed to know the time both vehicles will take to reach her destination station (either from experience or a travel planner). When the first vehicle (VEH1) arrives, she must make a decision, based on the information available to her and her own uncertainty evaluation for the network, whether to board it or to wait for the next one (VEH2). Figure 13 shows the proposed choice situation in a timeline format.

Although the vehicles are identical, the options available to the traveller (unlike route alternatives in most choice situations) are not unlabelled, that is, they have alternative-specific properties—in fact, the traveller is comparing a *certain* (as in risk-less) option against an ambiguous one. The vehicle that has already arrived has a *certain* waiting time which is almost zero due to the, usually, negligible difference between boarding, doors closing and departure. Although the anticipated waiting time for the next vehicle is displayed (either directly or as anticipated departure time of the next vehicle), it is ambiguous for the traveller since no concrete probabilities regarding its accuracy are supplied. Rather, she will draw from her own evaluation of this natural source of ambiguity and make a decision.

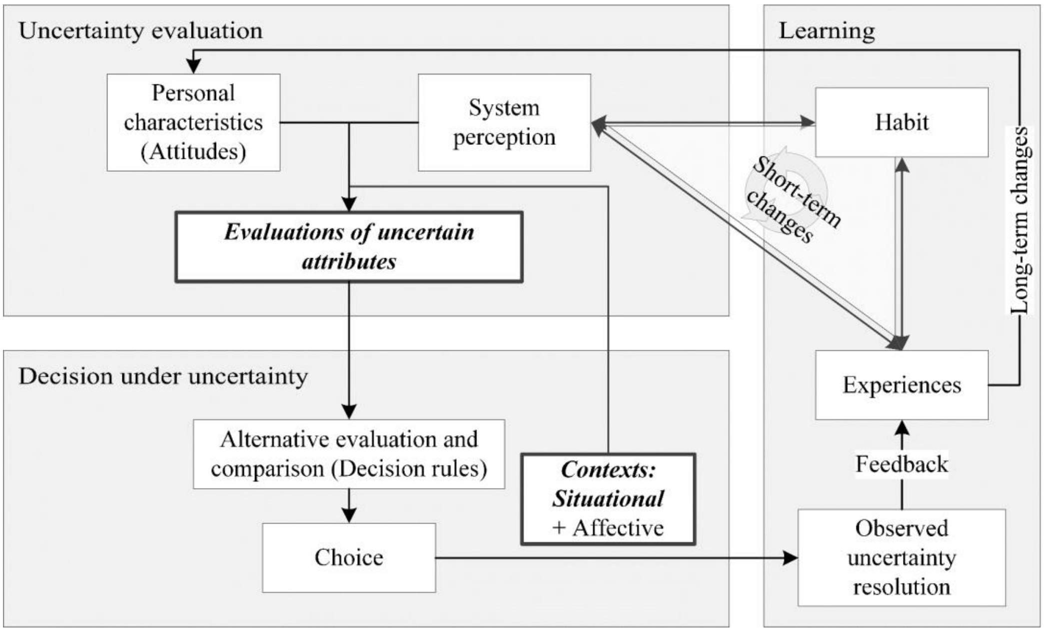


Figure 12: Theoretical framework of decision-making under uncertainty. Components in bold-italics are the focus of this study

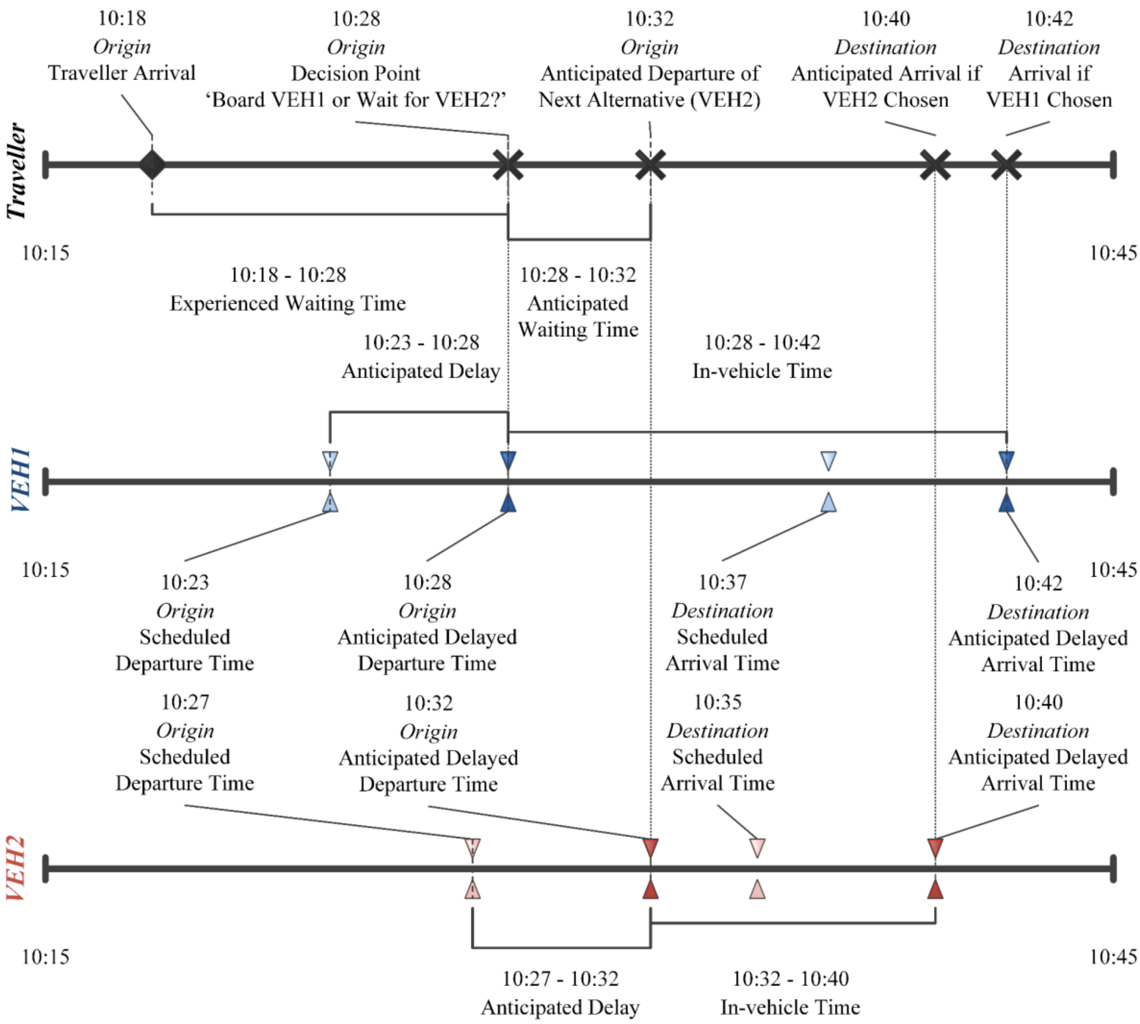


Figure 13: Choice situation presented in a timeline format

Aside from trading-off the difference in in-vehicle times against the anticipated waiting time, the traveller may assign an alternative-specific value to the *certain* option which represents her uncertainty evaluation for the anticipated waiting time. Thus, from this situation, the certainty equivalent of the ambiguous waiting time for the second (uncertain) option can be obtained by estimating the value assigned, holding other things equal, to the *certain* option. The traveller's uncertainty evaluation, and by extension her value of certainty, may also be affected by situational contexts such as delays in the system and the time she already spent waiting before the decision point. Since travellers are not likely to believe that the actual waiting time will be significantly lower than the displayed prediction, it is reasonable to expect that they do not dislike certainty—they are either indifferent or like certainty. This implies that if travellers, in general, believe the shown anticipated waiting time, the value of certainty would be lower than if there is a general perception of poor reliability.

For the proposed situation to take place, there must be a difference in in-vehicle times between the travel options. Moreover, the schedules or real-time delays must be such that the slower and faster vehicles are the certain and uncertain options, respectively (i.e., the slower vehicle arrives first at the origin). To assess the value of certainty in waiting time, choices between non-(strictly)-dominated alternatives must be observed. Assuming that travellers either like or are indifferent to certainty in waiting time, to ensure that the *certain* alternative does not dominate the uncertain one, the former must arrive at the destination later than the latter taking into account any weighting of travel time components. The uncertain vehicle can arrive at the destination before the certain one (i) if it can overtake the latter along a common path or (ii) if they serve two distinct lines.

The conditions outlined for the proposed situation to arise are not stringent and a number of examples can be found in the real world. Using published timetables of real-world public transport networks, specific examples can be found. For instance, the situation arises in the New York City subway and Mumbai commuter railways because express trains can overtake local ones (e.g., local and express lines 1 and 2 between 96 St and Chamber St in New York City; local and express trains between Borivali and Churchgate in Mumbai) (Indian Railways: Western Railway, 2021; MTA New York City Transit, 2020). Examples of the situation arising due to stops being connected by lines with distinct routes can also be found in the New York City subway as well as in the tram network of The Hague (e.g., lines 2 and 4 between 149 St and Franklin Av; lines 9 and 16 between Loevensteinlaan and Station Hollands Spoor) (HTM, 2021; MTA New York City Transit, 2020). Furthermore, even if the situation does not occur in a particular public transport network, given that the setup is fairly common in other networks, it is likely that travellers can identify with the situation. We emphasise that the proposed situation is a probe that permits the measurement of a relevant factor in travel behaviour, that is, uncertainty evaluation. There is little reason to believe that travellers' evaluation of waiting time uncertainty in this situation would be any different in other situations in the networks.

In the following sections, we make use of the proposed choice situation in two case studies: first, within a stated preferences experiment, and second, in a natural field experiment. For each case study, we discuss the experiment design, data collection, choice analysis, and the results.

5 Case study I: Stated choice experiment

In the first case study, we assess the waiting time reliability beliefs of travellers in the Dutch railways by implementing the choice situation presented in the previous section in a stated preferences experiment. In the Netherlands, the railways are used widely, for different trip purposes and over a large range of travel times. Since trains may—and different services such as express (Intercity) and non-express (Sprinter) indeed do—overtake one another by skipping stations, the choice situation would not seem unrealistic to travellers. Furthermore, as required in the proposed choice situation, throughout all railway platforms in the Netherlands, real-time departure information is displayed in a uniform manner.

When the proposed choice situation is presented as a stated preference questionnaire, it has two important advantages over conventional travel time reliability behaviour stated preference experiments. First, since there are no objective probabilities, they do not have to be conveyed to respondents so that everyone can understand them; thus, circumventing a major issue in such experiments. Second, unlike conventional choice experiments where respondents are known to provide protest answers in such experiments to demonstrate (in an exaggerated manner) their dislike towards delays and irregularities in public transport services (Bates et al., 2001), it is less obvious to survey-takers what is being measured and therefore they are likely to indicate their ‘true’ preferences. Next, we discuss the experiment design, data collection, and the choice analyses.

5.1 Experiment design

The choice situation consists of the following variables: (i) time already waited or the experienced waiting time; (ii) the anticipated delays of the two trains; (iii) the in-vehicle times of the two trains; and (iv) the anticipated waiting time for the second train. The first variable, experienced waiting time, is a context variable as it holds true irrespective of the alternative chosen. Since, the objective is to understand how they affect the value of certainty (rather than their marginal disutility), the anticipated delays for the two trains are changed together. Thus, the anticipated delay in the two trains can also be considered to be a context variable.

Attribute values are based on the need to adhere to reality and the ability to obtain the required estimates from choice observations. Since there is no clear indication on the direction or magnitude of the effect context variables have on the value of certainty, it is interesting to test them more closely. To this end, four attribute levels are used for each context variable allowing testing for non-linearity. The selected values (Table 5) are quite realistic as delay information in the Dutch railways is indeed shown in five-minute intervals while experienced waiting time is often rounded as it is difficult to be more precise when thinking about elapsed time.

The selection of attribute values for in-vehicle times and anticipated waiting times is a little trickier. The values of in-vehicle times and anticipated waiting times must be such that, given the expectations of traveller preferences, alternatives presented must not be dominated for a range of trade-off ratios between anticipated waiting time and in-vehicle time. Commonly, studies have found that waiting time is weighed 1.5-2 times compared to in-vehicle time (e.g., Yap et al. (2018)). However, it is also possible that travellers directly compare expected arrival times at the destination, in which case the waiting time and in-vehicle time are weighted equally. Thus, the range of waiting time – in-vehicle time trade-offs considered here is from 1 (arrival time differences) to 2 (higher end amongst most findings). A trial-and-error approach is used to find which attribute values satisfy the set of objectives and constraints described below.

For all three variables—in-vehicle times for the two trains and anticipated waiting time for the second train—only two attribute levels are chosen. This results in 8 (2×2×2) possible utility differences for a given waiting to in-vehicle time coefficient ratio. We would like to select attribute values for these variables such that for both the lowest and highest ratios (i.e., 1 & 2), considering the alternatives to be unlabelled (i.e., without an alternate specific constant), amongst the 8 possible utility differences, there are at least: (i) 4 that are in favour of the second train, (ii) 1 that is neutral, and (iii) 1 that is in favour of the first train. The objectives are tilted in favour of the second train because people are expected to be neutral at the least but in general have a preference for certainty and therefore the alternative-specific utility of a certain waiting time is expected to be positive. The latter two objectives are set to prevent respondents from learning that the first train always arrives second at the destination as well as to allow observations to indicate that our expectation regarding the sign of utility of certainty is incorrect. In addition to these objectives, the following constraints are set on the attribute values: (i) the minimum anticipated waiting time is 4 minutes, (ii) the minimum in-vehicle time is 4 minutes, and (iii) the range of all attributes is at least 4 minutes. The first two constraints ensure realism of attribute values. A minimum attribute value range is set because a larger difference in alternative utilities requires fewer observations to estimate parameters. Note that only even values were used in order to reduce the search space. Table 5 shows the attribute values used in the experiment.

With these attributes and values, a simultaneous orthogonal fractional factorial design is found with NGENE. To limit the number of questions per respondent, the design is blocked into two parts. With this specification, a design with a total of 16 choice situations is found with 8 choice situations per respondent.

5.2 Data collection

The choice experiment was included within a larger survey that consisted of four parts, in this order: (0) screening, (1) socio-demographics, (2) choice experiment, and (3) qualitative measurements. The structure, content, and design of an initial draft of the survey were refined based on comments received from a small pilot of about 20 persons. The final version of the survey was offered in Dutch and had an expected completion time of 10 minutes. It was distributed to a predefined sample size of 700 respondents through an online panel, PanelClix. Given that most people in the Netherlands have access to the internet, this method of data collection does not create any obvious biases. The data collection took place in November–December 2018.

Screening and socio-demographics

Respondents were screened out if they used the trains less than once per month on the basis that if respondents do not meet this criterion, they are likely to not have well-formed evaluations of uncertainties in the railways. Regarding trip purpose, the survey aimed to collect about 80% of responses (550 responses) from those who used the railways for commuting either to work or education, and the rest from those with other purposes. The greater focus on commuters and efforts was, again, to ensure that those travelling more frequently are included since this group is more likely to have more well-formed value systems and uncertainty evaluations. Based on previous experience with the online panel, it was known that unemployed persons and those working part-time were slightly over-represented. Therefore, it was agreed, before the beginning of the distribution, that an additional restriction would be placed in the form of a minimum frequency of travel by commuters, at least twice per week, if too many respondents chose a frequency of once per week or less (enforced after collecting 325 responses).

Table 5: Attribute values used in the choice experiment

Attribute	Attribute values (in minutes)
Experienced waiting time	0, 5, 10, 15
Anticipated delays in both trains	0, 5, 10, 15
In-vehicle time for the first train	14, 28
In-vehicle time for the second train	4, 8
Anticipated waiting time for the second train	4, 10

Desired socio-demographic quotas were obtained from data collected between 2011 and 2015 in a national, one-day, trip diary survey, OViN (*Onderzoek Verplaatsingen in Nederland*) conducted by the Dutch Central Bureau of Statistics (Centraal Bureau voor de Statistiek, 2015). The distribution of age, gender, and household incomes of respondents in that survey who use the railways at least once (during the day of reporting) are used as the desired stratification. It should be noted that these distributions were not weighted by the individual weights given in the survey as the group was reasonably large in itself.

To ensure response validity, those taking less than 4 minutes to complete the survey (40% of the expected time) were eliminated and more responses were added until the predefined target (of 700 responses) was reached. Eventually, a total of 918 responses were collected of which 703 met the completion time threshold.[6] While the survey was expected to take about 10 minutes on average, analysis of completion times after the collection of the required sample size revealed an average of about 6 minutes (after removing 12 respondents taking more than 20 minutes) and a median completion time of a little more than 5 minutes. Table 6 shows the distribution of respondent characteristics for the final set of valid responses.

Choice experiment

The choice experiment section begins with an explanation of the choice situation. Next, the respondent first faces a sample question which is not used in the analysis and then the 8 choice situations that will be used for the analysis. Each choice situation is prefaced by the instruction that there were two trains that could take them to their destination from the platform. To evoke the feeling of actually being at a station, respondents are shown information regarding the waiting times and anticipated delays of the two trains (TRN1, TRN2) in a format similar to the signboards found at platforms of the Dutch railways (Figure 14). Respondents are informed that the images displayed are the state of the signboards at the decision point (as described in section 4). To remind survey-takers of the information shown in different parts of the signboard, an annotated version is also displayed in the example question. Separately from the signboard, information regarding the in-vehicle times and the time already waited is shown as a table and a line of text, respectively. Finally, the respondents are asked to choose whether they would board TRN1 or wait for TRN2. The order of the 8 situations as well as that of the two options in each situation were scrambled to avoid any biases. Figure 15 shows a translated screenshot of a question in the choice experiment.

[6] Analysis of the removed responses revealed very different behaviour from the rest of the sample confirming our suspicion that they were invalid. Using the original sample of 918 respondents, we also did not find any (non-negligible) systematic effects of completion times on attribute weights.

Table 6: Sample characteristics

Total respondents	703		
		Distribution (%)	
Attribute	**Value**	**Actual**	**Required**
Gender	Female	54.8%	50%
	Male	45.2%	50%
Age	<18	0.1%	0%
	18-24	32.7%	36%
	25-34	24.0%	17%
	35-44	15.4%	13%
	45-54	13.2%	16%
	55-64	10.8%	12%
	>64	3.7%	6%
Trip Purpose: Commuting	Work	53.3%	~80%
	Education	27.9%	
Trip Purpose: Others	Errands	0.7%	~20%
	Recreation	18.1%	
	Others	0.0%	
Trips per Week	0	1.8%	
	1	13.2%	
	2	18.8%	
	3	18.9%	
	4	22.0%	
	5	22.0%	
	6	2.4%	
	7	0.7%	

It is likely that respondents' uncertainty evaluations are affected by the time-of-day. Therefore, when not explicitly testing how this belief changes across different time periods in a day, it would be ideal reduce potential bias by not presenting any clock times. However, since the Dutch railways is a schedule-based system, train arrivals are associated with a particular clock-time and travellers are used to seeing this information on the signboards. Therefore, the planned departure time of the first train is fixed at 10:23. This time is somewhat neutral in the sense that it is just outside the morning peak (06:00-09:00) and not too far into the midday off-peak hours. Moreover, respondents may still be able to imagine using this train for different purposes. Finally, a rounded-off time such as 10:00 or 10:15 is intentionally not chosen because it might seem artificial and may induce respondents to act differently than they normally would; for instance, they may become more likely to calculate and focus on the final arrival time as it is easier to do so with round clock-times.

It should be noted that regardless of whether they choose to board the arrived train or wait for the next, respondents are not given any feedback on the outcomes, thus avoiding any learning effects and forcing respondents to continue to depend on evaluations formed in the real-world.

Figure 14: Information displays at a real station (annotated)

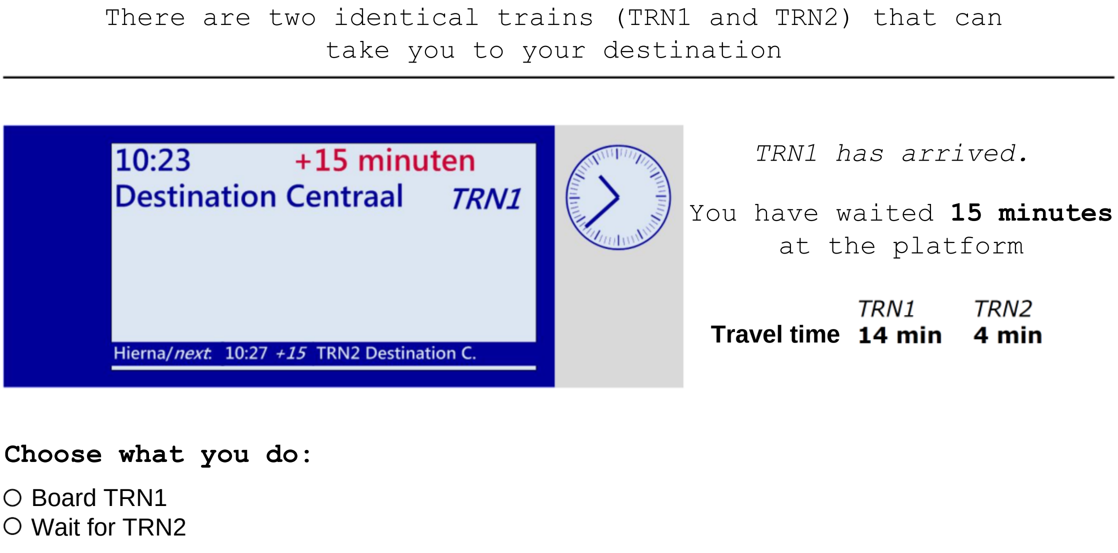


Figure 15: Screenshot of a question in the choice experiment (translated to English)

Qualitative measurements

Finally, the following factors are measured qualitatively on a Likert scale (with 7 levels): (i) regret anticipation, (ii) perception of reliability, and (iii) engagement level while waiting. The intention is not to include them in the modelling of uncertainty evaluation itself but rather analyse potential relationships between these indicators and stated preferences. The first, anticipation of regret, is considered to be one of the main psychological driving forces of risk aversion which leads to a preference for certainty. A standardized regret scale consisting of five items adopted from Schwartz et al. (2002) is used to measure it. This contains statements such as '*Whenever I make a choice, I try to get information about how the other alternatives turned out*' to which respondents indicate their level of agreement. The second factor assesses the perception of reliability of the network in general and in the presence of delays, and the perceived accuracy of displayed real-time information. This is tested using questions such as '*How reliable do you feel is the train arrival information?*' Finally, as discussed in section 3, context can affect how waiting time is experienced. Occupied time has been consistently shown to reduce perceived waiting time (Jones, 1996; Molin et al., 2020) which could in turn affect beliefs regarding anticipated waiting time; therefore, the level of engagement of respondents at train platforms is measured through the following question: '*Usually, how engaged are you with the activity you perform while waiting at a railway platform?*'. The complete set of questions can be found in Table 7.

5.3 Choice analysis

The decisions observed in the stated preference experiment are analysed using discrete choice models under the conventional framework of random utility maximization (RUM). To formulate the utility equations for the two options, we consider the four attributes involved in the choice situation described above: two main variables—in-vehicle times (IVT) and anticipated waiting time (AWT)—and two contextual variables—experienced waiting time (EWT) and anticipated delay (DEL). Furthermore, as the alternatives are labelled, the vehicle that arrives at the origin first (VEH1, the *certain* option) is assigned an alternative-specific constant ($\beta^{\text{certainty}}$) that represents the value of certainty attached to it. Since there are only two alternatives and only differences in utility matter, we set the utility of the second vehicle (VEH2) to zero. The (systematic parts of the) utilities of the two alternatives are then specified as follows:

$$\begin{aligned} V_{\text{VEH1}} &= \beta^{\text{certainty}} + \beta^{\text{IVT}} \cdot (IVT_1 - IVT_2) + \beta^{\text{AWT}} \cdot AWT + \beta^{\text{EWT}} \cdot EWT + \beta^{\text{DEL}} \cdot DEL \\ V_{\text{VEH2}} &= 0 \end{aligned} \quad [8]$$

Table 7: Questions used for psychometric indicators (in English)

Variable Name	Question (less [1] → more [7])
Regret: A	Once I make a decision, I don't look back. (the response order is reversed)
Regret: B	Whenever I make a choice, I'm curious about what would have happened if I had chosen differently.
Regret: C	Whenever I make a choice, I try to get information about how the other alternatives turned out.
Regret: D	If I make a choice and it turns out well, I still feel like something of a failure if I find out that another choice would have turned out better.
Regret: E	When I think about how I'm doing in life, I often assess opportunities I have passed up.
Reliability Perception: A	How reliable do you feel is the train arrival information?
Reliability Perception: B	How reliable do you feel is the Dutch Railways in general?
Engagement while Waiting	Usually, how engaged are you with the activity you perform while waiting at a railway platform?
Effect of Delay	When you are at an NS platform, to what extent is your perception of reliability (for your trip) affected if the next two consecutive trains that you can take to your destination are delayed?

Using the above utility equations, first multinomial logit (MNL) models are estimated to demonstrate the effect of accounting for the value of certainty (or the cost of uncertainty) on other choice parameters and to explore non-linear effects of contextual variables. In the RUM paradigm, the utility of an alternative a, U_a, consists of systematic (V_a) and random (ε) components. The systematic component is the product of the vector of taste preferences (β) and the vector of alternative attributes (x_a). Given that the random component in an MNL model is Gumbel distributed, the probability of choosing alternative i from I alternatives is given by the following:

$$U_a = V_a + \varepsilon;\ V_a = \beta \cdot x_a;\ P_{ni} = \frac{e^{V_i}}{\sum_{a=1}^{I} e^{V_a}} \quad [9]$$

Next, heterogeneity in behaviour is assessed using latent class choice models (LCCM) which capture decision-maker heterogeneity through a discrete mixture of choice models. In LCCM, individuals are probabilistically allocated to latent classes each of which have their own choice models. Depending on the objective, different choice models may be used in each class but in this study, the MNL model, based on the utility equations discussed above, is used as the underlying behaviour model for each class. To represent this mathematically, consider individual n who belongs to class s (amongst S classes) with probability π_{ns}. Then the probability that this individual selects alternative i is the product-sum of the class membership probabilities and the probability of selecting that alternative for each class (given the vector of taste parameters in that class, β_s):

$$P_{ni} = \sum_{s=1}^{S} \pi_{ns} \cdot P_{ni}(\beta_s) \quad [10]$$

If we assume intra-individual homogeneity in sensitivities, that is, account for panel effects, we essentially say that a particular individual is allocated to each class with the same probability for all choices they make. Thus, the likelihood of observing the sequence of choices $i : i_1, \ldots, i_T$ by individual n over T situations is given by the following:

$$L_{ni} = \sum_{s=1}^{S} \pi_{ns} \prod_{t=1}^{T} P_{ni_t}(\beta_s) \quad [11]$$

Apart from accounting for heterogeneity in tastes, an important advantage of LCCM is that individuals' preferences can be explained by using a class membership model to link membership probabilities with individuals' characteristics. The commonly used, logit function is also used here as the class membership model. We use the socio-demographic and qualitative measures collected (see Table 8) as the individual characteristics influencing class membership. For this vector of individual characteristics, z_n, and to-be-estimated, class-specific regression parameters, coefficient vectors, γ_s, and constants, δ_s, the class membership probability is given by:

$$\pi_{ns} = \frac{e^{\delta_s + \mathrm{g}(\gamma_s, z_n)}}{\sum_{a=1}^{S} e^{\delta_a + \mathrm{g}(\gamma_a, z_n)}} \quad [12]$$

The flexibility of the LCCM means that there are a number of ways to specify the model. The researcher needs to decide the number of classes, the parameters to be included in the choice models in each class, and the parameters to be included in the class membership model. Since there are no prescribed methodologies to arrive at the final model, we define here the sequence of steps taken to obtain our models. First, we include all choice parameters and class membership model constants and find the optimal number of classes. The model fit with different number of classes is assessed using the Bayesian information criterion (BIC) which explicitly penalizes the inclusion of extra parameters. Then for the optimal number of classes, the choice models in each class are finalized by removing highly insignificant ($p > 0.2$) parameters one-by-one. Next, all observable individual characteristics (socio-demographics) are added to the class membership function and the model is finalized by removing those that do not have a significant effect. Table 8 shows an overview of all the attributes used in the choice analysis. All choice model estimations are carried out using PythonBiogeme (Bierlaire, 2016).

As noted previously, the collected psychometric indicators were not included in the model itself; instead, the distribution of the unobservable qualitative measures in each class is used for characterising class composition through a posterior analysis of class membership. However, we also estimate a hybrid choice model (HCM) where the class membership model is directly related to the indicators through a measurement model in a framework similar to that employed by Atasoy et al. (2013) and Hurtubia et al. (2014). In this model, the likelihood function given for individual n in the latent class choice model (Equation [11]) is modified. In addition to the likelihood of observing a particular sequence of choices ($\boldsymbol{i} : i_1, \ldots, i_T$), the likelihood of obtaining a particular response pattern ($\boldsymbol{r} : r_1, \ldots, r_K$) for the indicators ($\mathbf{K}$) is also included (Equation [13] below). The probability of obtaining a particular response ($\pi^{k,r}$) is treated as a constant for each class and is estimated directly as a parameter in the model using the indicator responses. Thus, as Atasoy et al. (2011) note, the measurement model for the psychometric indicators helps identifying the latent classes by using responses to these indicators. Since the HCM accounts for these responses, it might lead to different latent classes or newer insights that do not surface in the posterior analysis of the LCCM.

$$L_{nik} = \sum_{s=1}^{S} \pi_{ns} \left\{ \prod_{t=1}^{T} P_{ni_t}(\beta_s) \right\} \left\{ \prod_{k=1}^{K} \pi_s^{k,r} \right\} \quad [13]$$

5.4 Results and discussion

As discussed in section 5.3, we first present results of the multinomial logit models; specifically, the effect of accounting for uncertainty and context variables. Then, heterogeneity in behaviour is presented through distinct behavioural profiles identified by a latent class choice model which also explains membership to these profiles with socio-demographic and other personal factors.

Table 8: Overview of attributes included in the choice analysis

Attributes	Symbol	Explanation	Range
Alternative attribute coefficients			
Certainty constant	$\beta^{certainty}$	–	–
In-vehicle time	β^{IVT}		4-28
Anticipated waiting time	β^{AWT}		4-10
Experienced waiting time	β^{EWT}	All time attributes are in minutes	0-15
Anticipated delays	β^{DEL}		0-15
Personal characteristics			
Socio-demographics			
Age	β^{age}	Ordinal in ascending order: <18, 18-24, 25-34, 35-44, 45-54, 55-64, >64	1-7
Gender	β^{female}	Categorical (effect coded): female, male	
Net personal income	β^{income}	Ordinal in ascending order: unemployed, €0-11K, €11K-19K, €19K-30K, €30-60K, >60K	1-6
Trip purpose	$\beta^{commuting}$	Categorical (effect coded): commuting, non-commuting	
Train use frequency	$\beta^{frequency}$	Average number of days train is used in a week	0-7

Multinomial logit models

To analyse the effect of including a certainty parameter, as opposed to conventional route choice models that consider alternatives to be unlabelled, in addition to the labelled MNL model (MNL^L) that uses the equations presented in section 4, an unlabelled version (MNL^U) that does not include $\beta^{certainty}$ is also estimated (Table 9). The significant and positive alternative-specific constant in the MNL^L model clearly rejects the null hypothesis that there is no effect of uncertainty and shows a preference for certainty. The coefficients for travel time components in both models are also significant and have the expected signs: as the anticipated waiting time increases or the first train is less slow in comparison, the preference for the first train increases. Since the context variable parameters in MNL^L model are small and insignificant ($p > 0.2$), in the model shown in Table 9 they are fixed to zero. The most likely reason for finding these parameters significant in MNL^U but not in MNL^L is that, in the absence of an alternative-specific constant in the former model, these parameters also partially capture respondents' overall preference for certainty. In the MNL^U model, where the contextual variable parameters are significant, the signs of the context variables seem to be reasonable. Regarding delays, one can expect travellers to be increasingly wary of waiting for TRN2 as the delays increase. For experienced waiting time, as discussed in section 3, there is no clear intuition regarding the effect direction since travellers might either experience frustration/increasing tiredness or take into account sunk costs. Moreover, some people may begin to engage in an activity that distracts them from waiting after some threshold of experienced waiting time. An overall positive effect is found and it may be justified—the more time has elapsed, the more travellers just want to take the train that comes first, all other things being equal (Osuna, 1985).

A log-likelihood ratio test between the models shows that, the MNL^{L} model clearly outperforms its unlabelled counterpart ($p < 0.001$). We cross-validate this improvement using a *k*-fold procedure with 14 folds such that all observations from one individual are either in the training or testing data set. The cross-validation reveals similar improvements in likelihood of chosen alternatives in the test dataset: -3335.37 versus -3371.20 with the MNL^{L} and MNL^{U} models, respectively. An important difference between these models is in the β^{AWT}–β^{IVT} ratio. In the unlabelled model this ratio is 1.22, a value close to results in literature which have commonly found that waiting time weighs higher in travellers minds than in-vehicle time (e.g., Yap et al. (2018)). However, once the waiting time uncertainty is accounted for in the MNL^{L} model, the ratio becomes 0.65 indicating the large role of uncertainty in the travellers' assessment of waiting time. Furthermore, the MNL^{L} model also shows that travellers are willing to trade-off 7.70 minutes (0.947÷0.123) of in-vehicle time for certainty in their waiting time.

Although the context variables seem to have no effect in the MNL^{L} model, since four levels were included for each variable, it is possible to check whether they really do not affect decision-making or they have a non-linear nature which averages out. While this is less likely for delays where we have a clear intuition regarding the effect direction, it may very well be true for experienced waiting time where there is an interplay between the effects of frustration and sunk time costs. The variables are effect coded with the level with 0 minutes as the reference. Effect coding allows us to separate the effect of the reference level from the constant. The variables for 10 and 15 minutes of delay, and for 5 minutes of experienced waiting time have high p-values ($p > 0.2$) and are therefore fixed to zero. The final model is shown in Table 9 as $MNL^{L\text{-}nl}$. The results include the coefficient for the reference level which is computed as the sum of the negatives of all the other coefficients for that attribute. Using the log-likelihood ratio test, this model is found to perform better than the MNL^{L} model ($p < 0.001$). The signs for anticipated delays are not as expected and it is difficult to hypothesize why a delay of 5 minutes seems to make it more likely that the second train will be chosen. The signs for experienced waiting time, however, can be explained by a combination of frustration/anxiety/increasing fatigue effects and sunk time/activity engagement effects. The likelihood of choosing TRN1 first increases up to 5 minutes (arguably due to frustration/anxiety/fatigue), then stabilizes between 5 to 10 minutes (more likely to be engaged in an activity), and then falls again (sunk time/activity engagement).

Latent class choice models

Using the steps defined in section 5.3 yields a 4-class model as the one with the best trade-off between efficiency and model fit. However, two classes have a membership of less than 10% which means that the choice parameter estimations within these models would likely have high errors. Therefore, we remove one class and estimate a 3-class model which has a comparable model fit, has reasonable class sizes and offers better interpretability. Table 10 shows the final model. To report results, the class with the smallest size is used as the reference for the class membership model (i.e., for the smallest class, $\delta_s = 0, \gamma_s = 0$).

Table 9: Estimation results of the different multinomial logit models (case study I)

Model	**MNL^{U}**		**MNL^{L}**		**MNL$^{L\text{-}nl}$**	
# parameters	4		3		6	
Initial LL	-3898.260		-3898.260		-3898.260	
Final LL	-3366.782		-3331.959		-3321.126	
Adjusted ρ^2	0.135		0.145		0.147	
BIC	6768.102		6689.822		6694.06	
Parameter	**Coeff.**	**p-val**	**Coeff.**	**p-val**	**Coeff.**	**p-val**
$\beta^{certainty}$	–	–	0.947	0.00	0.94	0.00
β^{IVT}	-0.108	0.00	-0.123	0.00	-0.124	0.00
β^{AWT}	0.132	0.00	0.080	0.00	0.081	0.00
β^{EWT}	0.023	0.00	–	–	–	–
β^{DEL}	0.015	0.00	–	–	–	–
$\beta^{EWT\text{-}0}$	–	–	–	–	-0.206	–
$\beta^{DEL\text{-}0}$	–	–	–	–	0.094	–
$\beta^{EWT\text{-}5}$	–	–	–	–	–	–
$\beta^{DEL\text{-}5}$	–	–	–	–	-0.094	0.03
$\beta^{EWT\text{-}10}$	–	–	–	–	0.091	0.06
$\beta^{DEL\text{-}10}$	–	–	–	–	–	–
$\beta^{EWT\text{-}15}$	–	–	–	–	0.115	0.02
$\beta^{DEL\text{-}15}$	–	–	–	–	–	–

Table 10: Estimation results of the 3-class latent class choice model (case study I)

Model	**LCCM 3-Class**					
# parameters	12					
Initial LL	-4159.808					
Final LL	-3063.483					
Adjusted ρ^2	0.261					
BIC	6230.584					
	Class 1		***Class 2***		***Class 3***	
Class Size	54.74%		28.41%		16.84%	
	Class-specific choice models					
Parameter	**Coeff.**	**p-val**	**Coeff.**	**p-val**	**Coeff.**	**p-val**
$\beta^{certainty}$	1.61	0.00	–	–	0.983	0.01
β^{IVT}	-0.301	0.00	-0.061	0.00	-0.0487	0.01
β^{AWT}	0.258	0.00	–	–	0.126	0.00
β^{EWT}	–	–	–	–	0.0268	0.12
β^{DEL}	0.019	0.18	–	–	–	–
	Class membership model					
	Class 1		***Class 2***		***Class 3 (ref.)***	
	Coeff.	**p-val**	**Coeff.**	**p-val**	**Coeff.**	**p-val**
$\beta^{intercept}$	2.00	0.00	–	–	0	–
β^{age}	-0.236	0.00	0.134	0.00	–	–

In the largest class (55%), behaviour is similar to the MNL^L model with an additional effect wherein the value of certainty increases slightly with delays. Travellers in this class are willing to trade-off about 5.3 minutes of extra in-vehicle time to remove uncertainty in their waiting time. With each minute of delay, travellers are willing to further accept approximately 4 seconds of additional in-vehicle time. Similar to the MNL^L model, once the value of certainty is accounted for, they weigh anticipated waiting time slightly less than in-vehicle time (0.86:1). Membership of this group is more likely for younger travellers. Similar to their preference for certainty here, younger travellers are also found to be more risk averse by de Palma and Picard (2005) in their departure time choice study.

The second group (28%) shows lexicographic preferences (at least in the range of the attribute levels presented in the stated preferences experiment) for faster trains, making decisions only on the basis of in-vehicle time and, apparently, not caring about other factors. In addition to their inherent preferences, it is possible that those who strongly prefer the faster train, may have translated the offered alternatives into real-life services, where the trains are, in fact, different, and thus chosen one train type over another for reasons not measured in the survey. In the Netherlands, the express trains (*Intercity*) offer additional services such as air-conditioning, Wi-Fi internet, and toilets that are not available in some commuter trains (*Sprinter*). Seating configurations are also different with commuter trains typically having more standing space. Older travellers are more likely to be in this class.

Although, the third group (17%) shows some compensatory behaviour, travellers in this group seem to strongly dislike uncertainty and are willing to accept more than 20 minutes of extra in-vehicle time for certainty in their waiting time. Thus, their preferences are nearly lexicographic in favour of the first train to arrive. Furthermore, frustration and/or cumulative waiting fatigue seems to play a substantial role for this group: with every minute spent waiting in the past (which should therefore be irrelevant for the decision at hand), there is a willingness to accept an additional 33 seconds of in-vehicle time for certainty in waiting time. In any case, we note that imagining this frustration/anxiety/fatigue may be somewhat difficult for respondents.

Posterior analysis of the class membership does not reveal substantial differences between classes in terms of distribution of psychometric indicators (Figure 16). Visual inspection of the trends shows that those showing fully compensatory behaviour (Class 1) have a slightly lower trust in the reliability (indicators Reliability Perception A and B) of the system. Moreover, based on regret indicators C, D, and E, respondents in this group are also a little less regret-averse than the sample is on average.

Hybrid choice models

Instead of estimating the LCCM followed by a posterior analysis of the psychometric indicators, we can estimate a hybrid choice model. In the adopted hybrid choice modelling approach, a large number of parameters has to be estimated: if all indicators are used, a total of 162 parameters have to be estimated to obtain the indicator response likelihood (162 = 3 classes × 9 indicators × (7-1) levels). Therefore, we reduce the number of indicators by selecting only one each for regret and reliability perception (from a set of 5 and 2, respectively), and the indicators for engagement while waiting and effect of delays. The indicators for regret and reliability perception are selected based on an exploratory factor analysis and overall model fitness. The full results of the HCM can be found in Table 11[7].

[7] Since the results are not used (as discussed in the next paragraph), we did not refine the model further after the first estimation (e.g., by removing parameters with p-values above our assumed thresholds of insignificance).

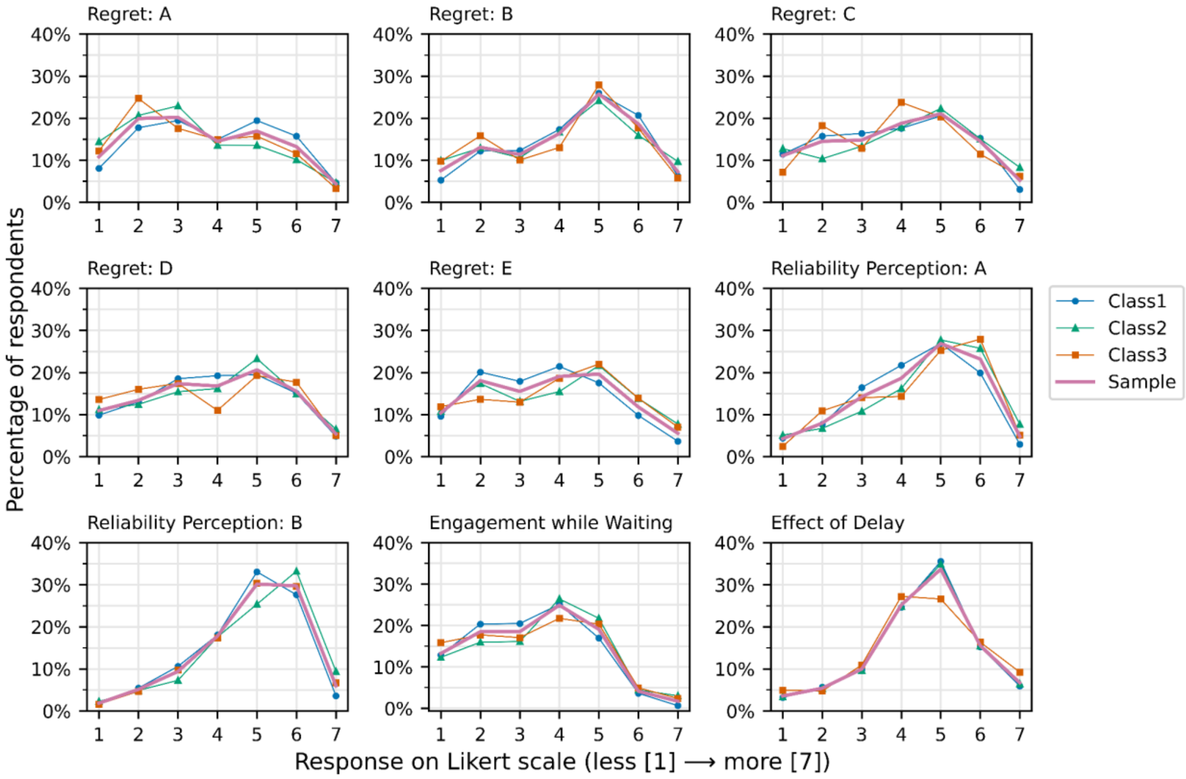


Figure 16: Class profiles of the psychometric indicators (see Table 7)

Table 11: Estimation results of the hybrid choice model (case study I)

Model	**LCCM 3-Class (with indicators)**					
# parameters	90					
Initial LL	-12911.755					
Final LL	-7902.731					
Adjusted ρ^2	0.381					
BIC	16582.593					
	Class 1		***Class 2***		***Class 3***	
Parameter	**Value**	**p-val**	**Value**	**p-val**	**Value**	**p-val**
Choice parameters						
$\beta^{certainty}$	**1.410**	0.000	0.361	0.160	**0.811**	0.030
β^{IVT}	**-0.284**	0.000	**-0.080**	0.000	**-0.033**	0.260
β^{AWT}	**0.251**	0.000	-0.004	0.860	**0.116**	0.000
β^{EWT}	0.010	0.470	-0.003	0.820	**0.025**	0.220
β^{DEL}	**0.026**	0.230	-0.007	0.700	0.002	0.930
Class membership parameters						
$\beta^{intercept}$	**1.99**	0	–	–	**0**	–
β^{age}	**-0.235**	0.01	**0.15**	0.01	–	–

Table 11 (continued)

Model	**LCCM 3-Class (with indicators)**					
	Class 1		***Class 2***		***Class 3***	
Parameter	**Value**	**p-val**	**Value**	**p-val**	**Value**	**p-val**
Indicator probability: Regret B						
$\pi^{\text{Regret\_B},1}$	0.018	0.310	0.136	0.000	0.100	0.080
$\pi^{\text{Regret\_B},2}$	0.114	0.010	0.136	0.010	0.162	0.060
$\pi^{\text{Regret\_B},3}$	0.129	0.000	0.105	0.010	0.094	0.080
$\pi^{\text{Regret\_B},4}$	0.169	0.000	0.177	0.000	0.122	0.070
$\pi^{\text{Regret\_B},5}$	0.288	0.000	0.203	0.000	0.290	0.040
$\pi^{\text{Regret\_B},6}$	0.224	0.000	0.144	0.000	0.178	0.050
$\pi^{\text{Regret\_B},7}$	0.058	–	0.101	–	0.054	–
Indicator probability: Reliability Perception B						
$\pi^{\text{ReliabilityPercep\_B},1}$	0.011	0.120	0.030	0.040	0.013	0.410
$\pi^{\text{ReliabilityPercep\_B},2}$	0.065	0.000	0.036	0.070	0.046	0.240
$\pi^{\text{ReliabilityPercep\_B},3}$	0.127	0.000	0.051	0.060	0.096	0.170
$\pi^{\text{ReliabilityPercep\_B},4}$	0.183	0.000	0.172	0.020	0.175	0.140
$\pi^{\text{ReliabilityPercep\_B},5}$	0.374	1.000	0.208	0.010	0.305	0.130
$\pi^{\text{ReliabilityPercep\_B},6}$	0.236	0.000	0.371	0.000	0.305	0.110
$\pi^{\text{ReliabilityPercep\_B},7}$	0.004	–	0.132	–	0.060	–
Indicator probability: Engagement while Waiting						
$\pi^{\text{WaitEngage},1}$	0.090	0.000	0.170	0.030	0.163	0.150
$\pi^{\text{WaitEngage},2}$	0.214	0.000	0.148	0.040	0.186	0.170
$\pi^{\text{WaitEngage},3}$	0.235	0.000	0.130	0.070	0.168	0.190
$\pi^{\text{WaitEngage},4}$	0.257	1.000	0.258	0.030	0.202	0.180
$\pi^{\text{WaitEngage},5}$	0.173	0.000	0.207	0.040	0.205	0.180
$\pi^{\text{WaitEngage},6}$	0.029	0.020	0.051	0.050	0.054	0.210
$\pi^{\text{WaitEngage},7}$	0.003	–	0.036	–	0.023	–
Indicator probability: Effect of Delay						
$\pi^{\text{DelayEffect},1}$	0.000	1.000	0.074	0.040	0.048	0.190
$\pi^{\text{DelayEffect},2}$	0.049	0.160	0.065	0.070	0.046	0.160
$\pi^{\text{DelayEffect},3}$	0.105	0.090	0.085	0.030	0.116	0.100
$\pi^{\text{DelayEffect},4}$	0.262	0.070	0.223	0.020	0.286	0.070
$\pi^{\text{DelayEffect},5}$	0.382	0.060	0.325	0.030	0.245	0.060
$\pi^{\text{DelayEffect},6}$	0.155	0.060	0.148	0.020	0.170	0.080
$\pi^{\text{DelayEffect},7}$	0.049	–	0.079	–	0.088	–

Indicator descriptions can be found in Table 7. The parameters included in the model without indicators are in bold.

Note: $\pi^{k,7} = 1 - \sum_{r=1}^{r=6} \pi^{k,r}$

Estimates for the parameters common to the hybrid choice and latent class choice models (highlighted in bold in Table 11) were found to be fairly similar. Moreover, the parameters estimated for the indicators in the HCM follow the same trends as their corresponding class profiles in the LCCM posterior analysis. Similarities in the two models may be because, in the HCM, the measurement model does not contribute substantially to the identification of the latent classes in comparison to the class membership model or choice models. Therefore, we choose not to use the HCM results because: first, the more complex HCM offers similar interpretation of the heterogeneity in choice behaviour, hence the parsimonious LCCM is considered superior; and second, the extra information obtained in the HCM as class profiles of indicators can also be obtained through the posterior analysis mentioned above.

6 Case study II: Field experiment

For the second case study, we conduct a natural field experiment wherein real-world instances of the proposed choice situation are extracted from passively collected smart card data in the urban public transport network of Amsterdam, Netherlands. With this analysis of revealed preferences, we make the following contributions: (i) demonstrate the presence of the proposed situation in a real-world network (in addition to the examples cited in section 4); (ii) establish constraints and methodology to study waiting time beliefs from smart card data; and (iii) overcome typical drawbacks of stated choice experiments, in particular, the inability to ‘feel’ the effects of contextual variables.

In the following, we first layout data requirements for field experiments with this choice situation under different smart card systems. Next, we describe our case study and the choice analysis. Finally, we present the results and discussion.

6.1 Data requirements

In general, to analyse route choice from smart card data, information about available and chosen alternatives as well as relevant attributes have to be extracted. Data available from smart cards, depends on the system employed by the operator. These systems vary mainly by how fare is calculated and where the interaction with travellers happens. The fare structure determines whether the traveller interacts with the smart card system at only one end (flat fare) or at both ends (distance- or zone-based) of a trip or journey (sequence of trips without intervening trip-generating activities). Interaction location—that is, where travellers present their smart card—governs where travellers interact with the system and thus what information is available regarding the chosen alternative. Interaction locations may be at stations or on vehicles. Furthermore, for flat fare, station-based systems, interactions may be at the origin or destination. In order to derive feasible[8] route alternatives, the origin and destination station locations must be known. These are directly available from vehicle-based, non-flat fare systems but for flat fare or station-based systems, trip origins or destinations may have to be inferred. Similarly, for data from the latter system type, route assignment is required to obtain the selected route alternatives. Naturally, confidence in the final route choice models is strongly linked to the confidence in each of the required inferences.

The specific situation proposed here, considers the choices to be: (i) board the *first* feasible alternative to depart or (ii) wait for the next one. Thus, when a traveller arrives at the origin stop in relation to the departure of feasible route alternatives is important. Traveller arrival times are known in pay-as-you-enter, station-based systems but not in vehicle-based or pay-as-you-leave, station-based systems. This problem can be circumvented by analysing transfer trips

[8] Given that the feasible choice set will be typically small, it is assumed to be equal to the considered set.

where the end time and location of the previous journey leg is known or can be inferred. Travellers' arrival times at the origin stop for the transfer trip can thus be directly known or inferred using assumed walking speeds (depending on whether the destination of the previous leg and the origin of the transfer trip are the same). As an example, consider a two-leg journey in a flat-fare, vehicle-based system where travellers are required to present their smart card every time they board a vehicle. While the traveller's arrival time at the origin stop for the first leg cannot be known, it might be possible to confidently infer the destination location (based on the origin of the next leg) and thereby the alighting time (from vehicle location data). The arrival time at the origin stop of the transfer trip can then be inferred by adding an assumed walking time to the inferred alighting time of the first leg. Note that this workaround assumes that travellers' evaluation of waiting time uncertainty in direct journeys (i.e., with no transfers) and in the first leg of non-direct journeys is the same as that in latter legs. Our macroscopic route choice analysis using smart card data from The Hague could not conclusively say whether uncertainty was perceived differently in the first and transfer legs.

Using the arrival times of travellers at their origin stop their experienced waiting times (EWT) can be calculated. Anticipated waiting times (AWT) and anticipated delays (DEL) are given by the real-time information displayed for the next relevant vehicle at the decision point (i.e., at the arrival of the first alternative). In-vehicle times (IVT) are assumed to be known to the travellers; thus, we could use planned values or a central tendency of historical realizations. Finally, we reiterate the need for choice situations to be non-(strictly)-dominated alternatives. Given the assumption that travellers do not dislike certainty, these are situations where, based on the displayed AWT and known IVT difference, there would be an *objective* expectation for the alternative with riskless waiting time to arrive at the destination later than the one with uncertainty, potentially taking into account any weighting of travel time components.

6.2 Data collection

For this case study, we use smart card data from Amsterdam's tram and bus networks. We use data from 28 May to 1 July, 2018 when 15 tram lines and 41 bus lines were operational (Figure 17). As nearly all of the trips on these networks are paid for using public transport smart card (*OV-chipkaart*), we can reliably use the data to make inferences regarding behaviour in the population. Both these networks employ a vehicle-based smart card system with a distance-based fare structure, which means that we are able to observe the origin and destination of each trip but not traveller arrival times. For transactions where the travellers fails to check-out at the destination (~4.2% of smart card trips), the destination is inferred based on Trépanier et al. (2007). Trips are combined into journeys by matching them with AVL data and employing existing transfer inference algorithms (Gordon et al., 2013; Yap et al., 2017), resulting in a total of 18.7 million inferred journeys.

6.3 Experiment setup

As discussed in the previous section, to overcome the problem of unobservable arrival times, we use transfer trips. Out of the journeys inferred above, more than 2.5 million have a transfer to the tram or bus network. Travellers' arrival time at the transfer boarding stop, is calculated by their arrival time at the previous alighting stop plus the time required to walk to the current boarding stop. Euclidean distances and a median walking speed of 1.12 m/s (Hänseler et al., 2016) are assumed.

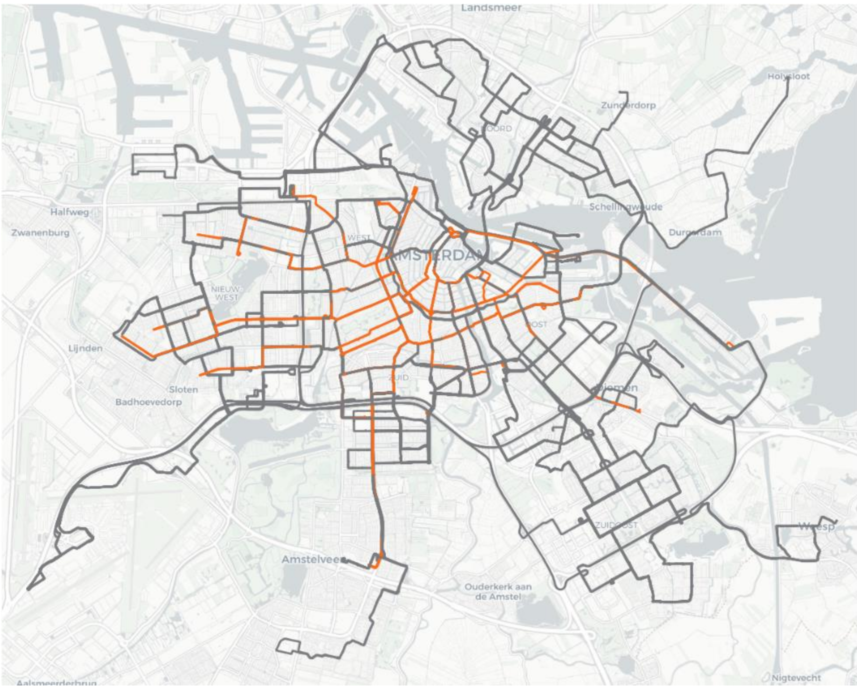


Figure 17: Amsterdam tram (orange) and bus (gray) networks in 2018

Next, a choice set is constructed for each transfer trip. For this, we first add possible alternative destinations within 150m of the original destination (as the crow flies) and then extract which public transport lines the traveller could potentially take. Trips that do not have at least two distinct line options are removed and the options in the remaining trips are assigned the first vehicle of the respective line to arrive after the traveller. Based on this, attributes of the choice alternatives, median planned in-vehicle time, experienced waiting time, anticipated waiting time, and anticipated delay, can be obtained. As the anticipated waiting time and delay information was not available directly, it was derived using the algorithm used for displaying real-time information. The planned arrival time of the vehicle is used to display the anticipated waiting time unless the vehicle is delayed by more than 10% of the planned in-vehicle time, in which case, it is assumed that the driver will cover the distance in 90% of the planned in-vehicle time (and the remaining delay is displayed as the anticipated delay).

To ensure that each choice set is reasonably feasible, further filtering steps are required. Alternatives with an anticipated waiting time of more than 30 minutes are removed, since it is unlikely that anyone would wait so long without an intermediate activity. Furthermore, to ensure that all travellers have the possibility to board the first alternative, choice sets with an experienced waiting time of less than one minute are removed[9]. Also, choice sets with all alternatives arriving at the same time are removed since they do not provide any information on travellers' waiting time beliefs. Two-thirds of the 1.05 million trips remaining consist of only two alternatives. Since analysing trips with more than two alternatives would require further assumptions regarding how travellers' aggregate the value of not boarding the first alternative (e.g., as only the utility of the next alternative or the log-sum of all the remaining alternatives), we choose to discard these, leaving us with 686,000 trips. Most of these trips are, however, dominated (that is, the first alternative reaches the destination first) and thus not useful for the analysis. Based on whether AWT is weighted twice as important as IVT for the dominance check, we obtain unweighted, (4563 trips) and weighted (2128 trips) non-dominated trip sets for the choice analysis.

[9] We note that this did not have a large impact on the final choice model.

6.4 Choice analysis

As in the previous case study, the observed choices are analysed under random utility maximization (RUM). However, since we do not have access to unique smart card identifiers, all observations are treated as independent and modelled using multinomial logit (MNL). Furthermore, given the nature of natural experiments, unlike the previous case study, the two alternatives in each trip have some more differences than just certainty in waiting time. However, we are able to explicitly account for these differences by including them in the utility equations for the two vehicles (Equation [14]). The vehicles can be either a tram or bus (effect coded as TRAM = 1 if tram), the anticipated delays for the two alternatives can be different (DEL_1, DEL_2), or the destination (DEST) can be at most 150m away from the original destination. As the alternative destination will never be chosen, the latter parameter does not offer any interpretation but is merely a correction for the constant. Figure 18 shows the distribution of attributes included in the choice analysis. Separate models are estimated for the unweighted (MNL^{UW}) and weighted (MNL^{W}) non-dominated trip sets.

$$\begin{aligned} V_{\mathrm{VEH1}} &= \beta^{\mathrm{certainty}} + \beta^{\mathrm{IVT}} \cdot (IVT_1 - IVT_2) + \beta^{\mathrm{AWT}} \cdot AWT + \beta^{\mathrm{EWT}} \cdot EWT + \beta^{\mathrm{DEL\_1}} \cdot DEL\_1 \\ &\quad + \beta^{\mathrm{TRAM\_1}} \cdot TRAM_1 + \beta^{\mathrm{DEST}} \cdot (DEST_2 - DEST_1) \\ V_{\mathrm{VEH2}} &= \beta^{\mathrm{DEL\_2}} \cdot DEL\_2 \\ &\quad + \beta^{\mathrm{TRAM\_2}} \cdot TRAM_2 \end{aligned} \qquad [14]$$

6.5 Results and discussion

Similar to the previous case study, to analyse the impact of including the certainty parameter, we estimate both, labelled ($^{-L}$) and unlabelled ($^{-UL}$) versions of the choice model for both trip sets. Table 12 shows the results of the four models after removing insignificant parameters ($p > 0.2$). To allow comparison between the labelled and unlabelled models using likelihood ratio tests, we included all the parameters that were significant in at least one version (except $\beta^{\mathrm{certainty}}$ in the unlabelled versions) to form nested models for each trip set. Note that, as in the previous case study, we tested for non-linear effects of anticipated delay and experienced waiting time but they did not improve the fit of labelled models significantly. Therefore, to reduce model complexity we do not include these in any of the final models or comparison tests. Likelihood ratio tests and 5-fold cross-validations show that the labelled models significantly ($p < 0.001$) improve model fit over their unlabelled counterparts[10].

Anticipated waiting time, in-vehicle time difference, and the certainty parameter all have a significant effect in the expected direction. The positive certainty parameter shows a clear preference for certainty, higher anticipated waiting time results in a preference for the first vehicle, and a higher in-vehicle time difference (in favour of the second vehicle) corresponds to a larger likelihood of choosing to wait. Since anticipated delays in the system would increase the level of associated uncertainty, we expected larger delays on either vehicle to be associated with a stronger preference for the certain alternative. While we generally find this to be true, in the labelled models, we find a small effect in the opposite direction for delays associated with the first vehicle. We do not find any effect of experienced waiting time or mode on the labelled choice models ($p > 0.8$) but similar to the previous case study, these parameters are significant

[10] Labelled models show a significantly ($p < 0.05$) better fit in the likelihood ratio test even if non-linear parameters (up to three degrees) are included. Note that where non-linear parameters were significant (in the unlabelled models), the change in effect direction did not lie within the observed range of values.

in the unlabelled versions: higher experienced waiting time reduces the likelihood of choosing to wait. Again, we hypothesise that the significance of these parameters is because they partially capture respondents' overall preference for certainty. In all cases, there is a stronger preference for trams over buses.

Like the previous case study, the inclusion of the certainty parameter reduces the β^{AWT}–β^{IVT} ratio (from 0.90 to 0.45 and from 1.22 to 0.77, in the unweighted and weighted trip sets respectively). This re-emphasises our earlier hypothesis that uncertainty plays a large role in the value of waiting time. Finally, we find that travellers in the Amsterdam urban public transport network are willing to trade-off about 3.5–3.7 minutes ($\beta^{certainty} \div \beta^{IVT}$) of in-vehicle time for certainty in their waiting time: nearly half the value found for travellers in the Dutch railways.

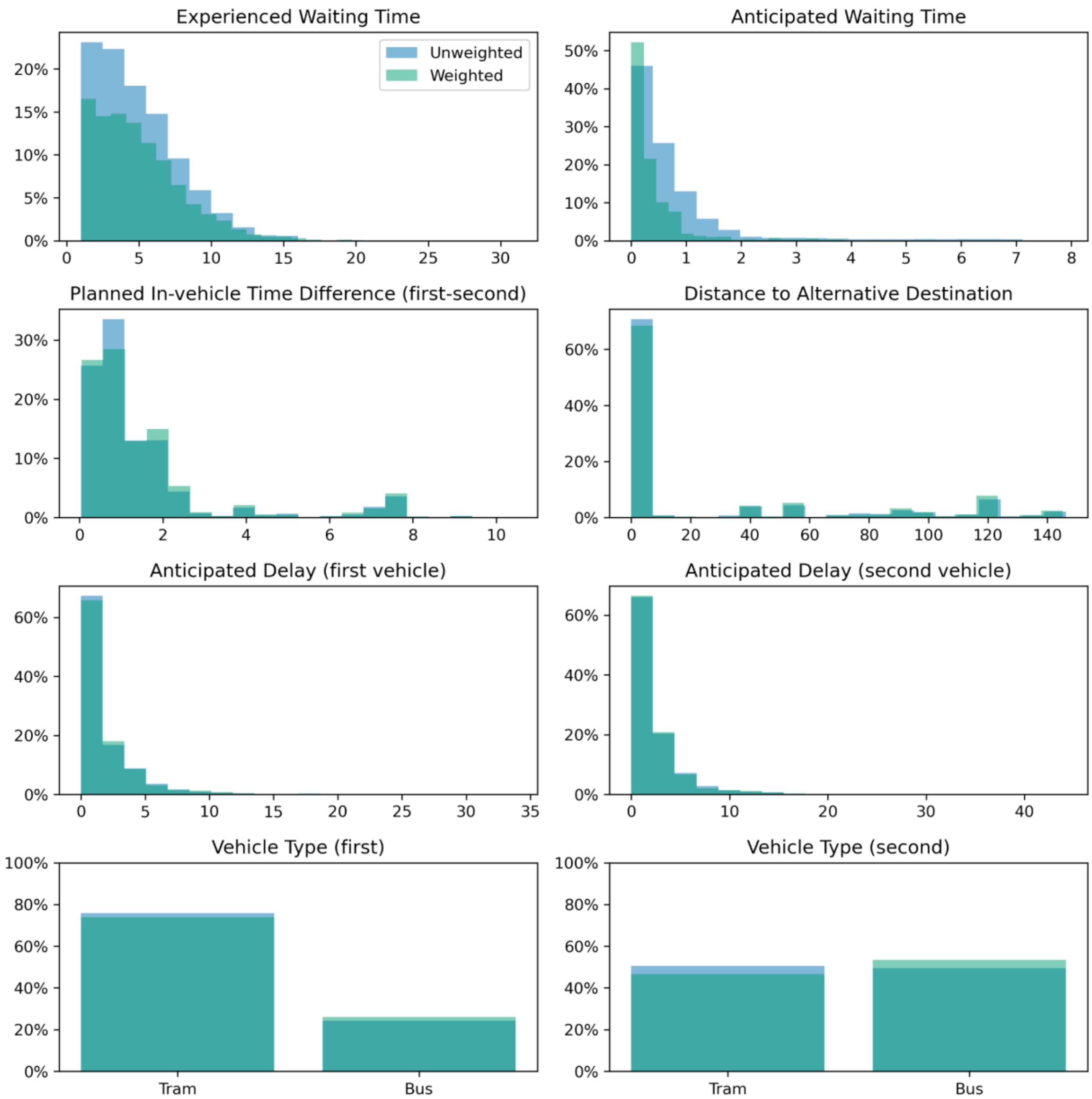


Figure 18: Distribution of choice attributes in the unweighted and weighted non-dominated trip sets

Table 12: Estimation results (case study II)

Model	$\textbf{MNL}^{\textbf{UW-UL}}$		$\textbf{MNL}^{\textbf{UW-L}}$		$\textbf{MNL}^{\textbf{W-UL}}$		$\textbf{MNL}^{\textbf{W-L}}$	
# parameters	7		8		7		8	
Initial LL	-3162.830		-3162.830		-1475.017		-1475.017	
Final LL	-1809.283		-1700.787		-843.362		-801.943	
Adjusted ρ^2	0.426		0.460		0.423		0.451	
BIC	3677.547		3468.980		1740.723		1665.189	
Parameter	**Coeff.**	**p-val**	**Coeff.**	**p-val**	**Coeff.**	**p-val**	**Coeff.**	**p-val**
$\beta^{\text{certainty}}$	–	–	1.482	0.00	–	–	1.340	0.00
β^{IVT}	-0.247	0.00	-0.400	0.00	-0.244	0.00	-0.384	0.00
β^{AWT}	0.223	0.00	0.182	0.00	0.297	0.06	0.296	0.06
β^{EWT}	0.134	0.00	–	–	0.122	0.00	–	–
$\beta^{\text{DEL\_1}}$	–	–	-0.034	0.07	–	–	-0.029	0.23
$\beta^{\text{DEL\_2}}$	-0.085	0.00	-0.034	0.02	-0.091	0.00	-0.040	0.08
β^{DEST}	1.747	0.00	1.771	0.00	1.291	0.00	1.432	0.00
$\beta^{\text{TRAM\_1}}$	0.584	0.00	0.153	0.03	0.536	0.00	0.152	0.13
$\beta^{\text{TRAM\_2}}$	0.286	0.00	0.208	0.00	0.367	0.00	0.270	0.00

7 Conclusion

Although decisions in the real world are almost always taken under uncertainty, that is, in the absence of objective probabilities, most existing studies on the effects of waiting time reliability on travel behaviour observe or analyse travel decisions (as if) made using objective probabilities. Capturing travellers' evaluations, which are a result of complex interactions between their perceptions and attitudes, regarding uncertainty in public transport waiting times is difficult. Therefore, this study identifies a realistic route choice situation where such evaluations can be quantified under natural ambiguity without using objective probabilities or assuming specific learning behaviour. In the slow/fast lines experiment proposed, uncertainty evaluations can be quantified as a certainty equivalent or, as shown, an alternative-specific constant under the random utility maximization regime. Studies in behavioural economics and psychology have indicated that contexts are important in decision making. In addition to quantifying the evaluations in general, we are also able to estimate the effect of contextual attributes on them; for instance, the effect of time spent waiting before making a decision based on anticipated time to be waited.

In the first case study, through a stated preferences experiment with the identified choice situation, we find a strong preference for certainty in travellers of the Dutch railways. Accounting for uncertainty explained away some of the waiting time parameter, reducing the waiting to in-vehicle time ratio to less than one. Contextual attributes do not seem to have an effect on average although small, non-linear effects were found for both experienced waiting time and anticipated delays. A latent class choice model indicated three groups of travellers: the biggest group making fully compensatory choices, weighing uncertainty against travel time attributes, and two others showing lexicographic behaviour, choosing the fastest and the first train, respectively. While age seems to affect association with different behavioural profiles, there are only minor differences between the distribution of psychometric indicators in different classes.

Although the choice situation and stated choice experiment are carefully designed, two limitations potentially affecting estimation and interpretation of results remained. In our experiment, we assume that travellers usually make a conscious choice regarding boarding a train or waiting for the next one. While it is likely that this is a conscious choice, especially for regular travellers who are aware of different lines that can take them to their destination, it is possible that by presenting this choice situation we highlight the uncertainty involved in waiting times thereby making people more averse to it. Another potential limitation of the case study is related to the experiment type itself. We would like to measure the effects of contextual attributes on subjective beliefs but, arguably, it is difficult for respondents to account for such effects separately from their general aversion to uncertainty. For example, respondents may not be able to feel the effect of having waited ten minutes when making a choice in a stated preference questionnaire, yet anecdotal evidence would suggest that this variable indeed has an impact on boarding decision. It is possible that this may be why we do not find strong effects of contextual variables in our case study. Using incentivized laboratory experiments, common in behavioural economics, does not help either because, since these are contextual variables, they cannot be incentivized one way or another.

Since the proposed choice situation is realistic, both of these limitations can be overcome by measuring choice behaviour in a revealed preferences setting, that is, from observations of real-world trips where travellers actually experience the context. We do this in our second case study using smart card data from the urban public transport network of Amsterdam. Results of this natural, field experiment largely follow the first case study. Travellers in this network are willing to trade-off an average of 3.6 minutes of in-vehicle (compared to the average of 7.7 minutes obtained for travellers in the Dutch railways). Similar to the stated choice experiment, accounting for waiting time uncertainty explained away some of the waiting time parameter, reducing the waiting to in-vehicle time ratio. Although, the revealed preferences experiment is intuitively a better method to measure the impact of contextual variables, once we account for the value of uncertainty, we no longer found a significant impact of experienced waiting time or vehicle type. However, this is not necessarily a general result and could be a result of the range of values present in the experiment or the specific nature of urban transportation.

While the field experiment overcomes some limitations of the stated choice experiment it is not without its own limitations. First, while fairly common, the choice situation may not occur in all public transport networks. Moreover, even where the choice situation occurs, the proportion of observations available for analysis may be very small relative to the overall dataset. This problem is particularly compounded in systems such as the one in our case study where we can only utilise observations from transfer trips. Secondly, given the nature of field experiments, we have little control over the range of attribute values available for the choice analysis, limiting the validity of our conclusions for other situations. In the case study here, we note that the distribution of anticipated waiting time and planned in-vehicle time differences is heavily skewed towards zero.

The choice situation proposed in this study offers a relatively simple method to obtain snapshots of evaluations of uncertainties in a real-world public transport network. With respect to planning of services, transportation models can benefit from the added accuracy obtained by explicitly quantifying the effects of uncertainty (as indicated by the improved model fit and predictive value). The proposed situation is used to measure uncertainty evaluation and inferences are not limited to this exact situation—for instance, the finding that associated uncertainty has a large role in travellers' assessment of waiting time holds over all decisions of the type 'whether to board or wait' and could be useful for agent-based models (e.g., Cats and

Gkioulou (2017)) that commonly simulate this choice. Often when biases are pointed out to decision-makers, they choose to correct their choices to more ‘rational’ ones (Gilboa, 2009). Journey planner applications may use choice observations in situations similar to the one used in this study to provide feedback highlighting such potential biases (e.g., loss aversion, overweighting of small probabilities) that travellers might want to correct on reconsideration. Moreover, through association of behaviour under uncertainty with introspective psychological measures, such applications can offer targeted actions to specific groups of travellers to bring their evaluations in line with empirical realities. For instance, applications may work on distracting travellers from the boredom of waiting by engaging them in an activity such as reading. Since the experiment also permits measuring the effects of contextual variables, it may be used to analyse situations which exaggerate feelings of uncertainty and take suitable actions for this. On a related note, the certainty equivalent presented here may also be used as an indicator for A/B type tests when transportation authorities wish to introduce new measures aimed at improving feelings regarding uncertainty. For instance, indicating the cause of delays has been proposed to reduced anxiety associated with uncertainty in waiting time (Maister, 1985). The extent to which this measure is effective may be quantified by comparing certainty equivalents obtained for the identified choice situation in the control and treatment groups.

As discussed in the theoretical framework for this study, we measure the combined outcome of travellers’ perceptions and attitudes on their decisions under uncertainty as subjective beliefs. However, disentangling the effects of these individual determinants on travel behaviour may allow more effective policies and travel advice. In order to analyse attitudes and perceptions separately, we might need to model more complex decision rules and, perhaps, observe different choice situations or sequence of decisions. The challenge will be to do this also directly from observations of real-world trips (i.e., not in a laboratory experimental context), without having to observe risky choices or ask for matching probabilities of uncertain events, both of which require interaction between the researcher and travellers.

Apart from the limitation outlined above, other avenues of research may also be found in the theoretical framework presented in this study. In our analysis, we considered the effects of two situational contexts, namely, experienced waiting time and delays on the travellers’ corridor. Similarly, other situational contexts such as the effects of delays in other parts of the network, or the differences between trip purposes, such as travelling to and from work may be studied. Furthermore, the effects of affective contexts on subjective beliefs can be investigated to assess the indirect effects of various factors affecting moods, such as station lighting. In decisions under ambiguity, such as route choice in public transport networks, where decision-makers can observe the choices of others, herding effects also become important and may be analysed. Finally, while our method provides a snapshot of subjective beliefs towards waiting time uncertainties in real-world networks, it would be interesting to observe the evolution of such snapshots over time for different individuals in various networks.

COVID-19 related Uncertainty

The COVID-19 pandemic (2020–) has led to drastic changes in travel behaviour, particularly a widespread drop in public transport ridership. Not only are travellers more likely to focus on transmission risk determinants in public transport now but, given the sustained the nature of the crisis, also in the post-pandemic phase, leading to a permanent shift to private travel modes. Therefore, in this chapter, we evaluate this change in behaviour and assess traveller responses to future developments in the pandemic. We specifically analyse behaviour related to three criteria affecting the risk of COVID-19 transmission: (i) crowding, (ii) exposure duration, and (iii) prevalent infection rate.

Using stated choice observations from train travellers in the Netherlands at the end of the first infection wave, we model travellers' (heterogeneous) valuation of the aforementioned factors. This is supplemented with an analysis of covariates that include socio-economic characteristics and Likert scale-based risk perceptions and attitudes. Special attention is paid to discussing the heterogeneity in traveller behaviour and comparison of current valuations with those from pre-pandemic analyses. The chapter concludes with a discussion of potential policy implications, limitations of the study, and future avenues of research.

This chapter is an edited version of the following article:
Shelat, S., Cats, O., van Cranenburgh, S. Traveller Behaviour in Public Transport in the Early Stages of the COVID-19 Pandemic in the Netherlands. Transportation Research Part A: Policy and Practice (2022).

1 Introduction

The COVID-19 pandemic has led to unprecedented restrictions on public life globally. Some of the first restrictions in many places were on public transport which, by its very nature of moving people in dense, enclosed spaces, could be a major transmission risk for this highly contagious virus[11]. While some public authorities completely stopped service (e.g., India (Union Home Secretary, 2020)), others restricted or discouraged use other than by essential workers or for urgent needs (e.g., Netherlands, United Kingdom (Department for Transport; Openbaar Vervoer Nederland, 2020a)). Then, at the end of the first infection wave, many authorities, pressed with a need to restart economies and provide essential transportation, eased restrictions and cautiously resumed public transport. Demand levels, however, did not return to pre-pandemic levels (Citymapper, 2021; Google LLC, 2021), at least partly, due to heightened (awareness of the) risk of infection (Beck and Hensher, 2020).

The effect of on-board crowding on travel behaviour has received much attention in literature and has been widely accepted to be a significant influence on various choice dimensions (Hensher et al., 2011; Li and Hensher, 2011b; Tirachini et al., 2013; Wardman and Whelan, 2011). Using choice observations, mainly from stated choice experiments (e.g., Kroes et al. (2013); Sahu et al. (2018)) but also from revealed preferences (e.g., Hörcher et al. (2017); Yap et al. (2018)), a number of studies have estimated the value of crowding in terms of the willingness to pay to reduce it or its impact on the value of travel time. The disutility of crowding in these studies arises primarily from physical and psychological discomfort and exhaustion. However, given the wide and sustained impact of the COVID-19 pandemic, travellers are now likely to want to avoid crowds even more so than under normal circumstances as a measure towards minimizing their exposure to the virus (Tirachini and Cats, 2020).

Travellers may now focus on factors contributing to COVID-19 transmission and for service planners to be able to respond to these changes in behaviour it is essential to have an empirical underpinning of those. The question is then: given the COVID-19 pandemic, how will travellers respond to crowdedness on public transport vehicles and future changes in infection rates? Studies on the COVID-19 pandemic as well as those on previous epidemics resulting from viruses spread through similar means (such as SARS, MERS, swine flu) have shown that people perceive avoiding public transport as a preventive measure (Gerhold, 2020; Kim et al., 2017a; Lau et al., 2003; Rubin et al., 2009). A number of COVID-19 related analyses also indicate a significant mode shift to private modes such as bicycles and automobiles demand (e.g., Bucsky (2020)). While these studies focus on perceptions and aggregate statistics, only a few studies have analysed public transport travellers' choice behaviour.

A Scopus search[12] and other modes of literature collection found only a handful of studies conducting choice analysis in the context of public transport and epidemics. Scorrano and Danielis (2021) conduct a mode choice analysis for before and during COVID-19 in Trieste, Italy. The impact of the pandemic on mode choice is parametrised as mode-specific penalties which they find to be negative (and even more so for COVID-19 risk averse travellers) for

[11] However, there is no conclusive evidence to this end and indeed some suggest that if recommended mitigation measures are implemented, the risk of contracting COVID-19 in public transport could be low (Gkiotsalitis and Cats, 2020; Goldbaum, 2020)

[12] Scopus search term (initial search on 25 March 2021 revealed only two relevant studies; the overview was updated to briefly include studies published during the review based on a search on 25 December 2021):

```
( TITLE-ABS-KEY ( pandemic OR epidemic OR sars OR mers OR "swine flu" OR h1n1 OR ebola OR covid ) AND TITLE-ABS-KEY ( "public transport*" OR transit OR bus OR tram OR train OR metro ) AND TITLE-ABS-KEY ( ( choice OR logit OR probit ) W/2 ( model* OR analys* ) ) )
```

public transport. Cho and Park (2021) and Aghabayk et al. (2021) also conduct a before-after experiment but focus on estimating crowding multipliers using stated choice data from Seoul and Tehran, respectively. They find that crowding multipliers are 1.04–1.44 times higher during the pandemic, confirming expectations that travellers would be more wary of crowds. Awad-Núñez et al. (2021) and Aaditya and Rahul (2021) focus on the impact of COVID-19 safety measures (such as reducing on-board crowding and improving cleaning) on the willingness to use and pay for public transport. They find that a higher safety perception increased willingness to use public transport. Finally, Hensher et al. (2022) present choice models of commuting, work-from-home, or not working using revealed choice data from Sydney and South East Queensland.

We contribute to the growing literature on COVID-19 and public transport by analysing how travellers have adapted their behaviour under these exigent circumstances. A stated choice experiment is conducted to analyse traveller behaviour specifically related to three criteria affecting the risk of COVID-19 transmission (Hu et al., 2020; Prather et al., 2020): (i) distance to other people, (ii) duration of exposure, and (iii) prevalent infection rate. In the context of public transport travel, the first two correspond to on-board crowding and in-vehicle time, respectively. With the choice experiment, we measure travellers' crowding valuation in the backdrop of the ongoing pandemic and how these valuations are affected by factors that might affect the perception of related risk. These model estimates will not only be useful for demand forecasting but could also provide insights that may be valuable for policy designs aimed at managing demand (Gkiotsalitis and Cats, 2020) not only for the ongoing crisis but also the next pandemic. In this study, we report findings from the stated choice experiment conducted with train travellers in the Netherlands towards the end of the first infection wave, just as the first restrictions were being lifted.

In the next section, we describe the survey design, data collection, and choice analysis methodology. This is followed by the results and discussions in section 3. Finally, a summary of the results, potential policy implications, limitations of the study, and future avenues of research are outlined in section 4.

2 Stated choice experiment

To understand traveller behaviour under the new circumstances presented by the pandemic, a stated choice experiment was conducted with Dutch train travellers. The experiment was part of a larger survey that collected, among other things, travellers' socio-demographics, mobility choices, and pandemic-related qualitative measures. Discrete choice analysis is applied on observations from the experiment to measure crowding valuation while the personal characteristics are used to explain heterogeneity in behaviour either *a posteriori* or as part of the choice model.

2.1 Experiment design and presentation

The experiment consists of a series of choice situations in which respondents were asked to assume that they had arrived at a train station from which two trains were available for their destination. They were informed that they were travelling with the same purpose for which they had indicated they most frequently used the train before the pandemic-related restrictions. The train alternatives varied only in terms of on-board crowdedness (distance to other people) and waiting time. We note that this means that crowding valuation will be obtained in terms of waiting time savings instead of the usual money amounts. We did not use different travel costs directly to avoid interactions with respondent income and expectations from a higher travel class. Implying that a more costly train would be less crowded (and therefore safer) could also lead to protest answers. Contextual information about factors potentially affecting the perception of contracting the disease, namely travel time (exposure duration) in either train and prevalent infection rate, was also given. The latter was provided in terms of the proportion of Dutch population that is infectious and capable of transmitting the virus to others. To ensure that only in-vehicle time was considered as the duration of exposure, we noted that it was possible to maintain social distance while waiting. Furthermore, respondents were reminded of the mandatory face mask regulations on-board public transport vehicles (Openbaar Vervoer Nederland, 2020b). While this was not mentioned explicitly in the survey, we note that in majority of the trains, windows cannot be opened, meaning that ventilation conditions cannot be easily changed. Respondents were asked to rank the two train alternatives and the option of not travelling by train for each choice situation. We asked for a ranking rather than a single best choice to enable us to obtain trade-off estimates in the case that the majority of respondents chose to opt-out altogether.

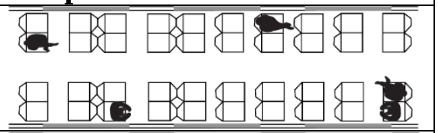
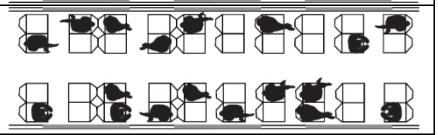
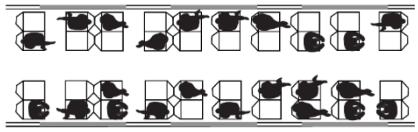
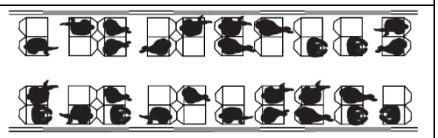
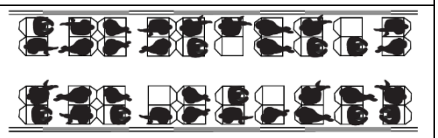
On-board crowding was presented graphically as the seated section in a single coach of a commuter train (known as Sprinter in the Netherlands). Five levels of crowdedness were used: 5, 18, 23, 28, and 36 seats occupied (out of 40), colloquially corresponding to the following labels: ‘almost empty’, ‘able to sit alone’, ‘unable to sit alone but not too crowded’, ‘quite crowded’, and ‘packed’ (Table 13). To avoid confusion, respondents were informed that the indicated crowding level was after everyone else (but the respondent) had boarded. While these trains do have some standing space near the doors (not shown in the graphics), we excluded the option to stand in order to simplify the choice situation. By only offering seating space, we may miss out on capturing the (possibly different) crowding valuation of those who would prefer to stand. However, given the limited and confined standing space, and the relatively long in-vehicle time levels used in the study (see below), it is likely that most travellers would have preferred to find a seat for their trip. As such, we expect the impact of this simplification to be small. To trigger respondents to consider where they would sit, they were asked, in a series of questions prior to the choice experiment, to indicate where they would sit in each of the five crowding levels. Three levels of waiting times were used: 3, 12, and 25 minutes. A wide range was deliberately used to ensure that we would observe trade-offs between on-board crowding and waiting time savings.

It is likely that respondents would find it difficult to respond to infection rate numbers without any real-world references on which to anchor their evaluation of this variable. To help respondents interpret the infection rate numbers, we sought to provide them a best estimate of the infection levels (i) at the time of the survey when restrictions had begun to be lifted (0.1%) and (ii) at the peak of the pandemic (in terms of daily reported cases and hospitalizations) in mid-April (0.43%). The proportion of infectious people in the population is innately unknowable due to the presence of asymptomatic and pre-symptomatic cases, limited testing capacity, and reluctance to get tested. Therefore, in the absence of official estimates (at the time of the survey[13]), we obtained the above numbers from back-of-the-envelope calculations using daily reported infections. In the experiment, five levels around these reference infection rates were used: 0.01% (pre-restriction levels), 0.1% (at the time of the survey), 0.5% (mid-April

[13] Since then, the Dutch government has published these figures (also retroactively) (Rijksoverheid, 2021). Their estimates for April 15 and May 20 (at the time of the survey) are 0.34% and 0.08%, respectively. Although these values are fairly close to ours, they estimate the peak of the infection rate to be around the end of March rather than mid-April.

level), 2% 10% (extremely high). For the other contextual variable, in-vehicle time, three levels were used: 10, 25, and 40 minutes.

Table 13: Graphical presentation of crowding levels

Crowding Level	Graphic
Almost empty	
Able to sit alone	
Unable to sit alone but not too crowded	
Quite crowded	
Packed	

We used a semi-random experiment design: weakly dominated and symmetrical choice situations were removed from the full factorial of the above-described attribute levels; and from these, 4 subsets of 15 choice situations were randomly picked. Respondents then faced one of these subsets at random. Walker et al. (2018) argue that semi-random designs, where dominated choice tasks are eliminated, perform as well as efficient designs, particularly because they are robust against a large range of parameter estimates and model specifications. A screenshot of the experiment is shown in Figure 19.

2.2 Data collection

In March 2020, the Dutch government urged travellers to use public transport only ‘if it is really needed’. By May 2020, having achieved a reduction in the daily reported cases, Dutch authorities announced that certain professions, services, and educational activities could resume by the end of that month (Rijksoverheid, 2020). Furthermore, public transport could be used once again by mid-June 2020 but with new regulations such as mandatory face masks and seat blocking to maintain distance (the latter was stopped in July 2020) (Openbaar Vervoer Nederland, 2020b). Data collection took place from 20 to 25 May, after announcements concerning these measures had been made. A total of 513 valid responses[14] were collected via an online panel. The survey was offered in Dutch and we expected a completion time of 12–15 minutes. In addition to the stated choice experiment described above, three categories of

[14] About 40 responses completed in less than 6 minutes were removed as this was considered to be too fast to have been properly answered. Response time was not a significant indicator in a simple linear-additive multinomial logit model and a model with just these responses returned many insignificant ($p < 0.05$) parameters indicating randomness in the responses given.

personal characteristics were collected to explain potential differences in behaviour: (i) mobility factors, (ii) socio-demographic factors, and (iii) COVID-19-related qualitative measures.

Screening

The survey was distributed to Dutch train travellers who travelled by this mode at least once per month before March 2020 when the first pandemic-related restrictions were imposed. Furthermore, we sought to collect a sample representative of the overall Dutch population in terms of age, gender, and education level (Table 14).

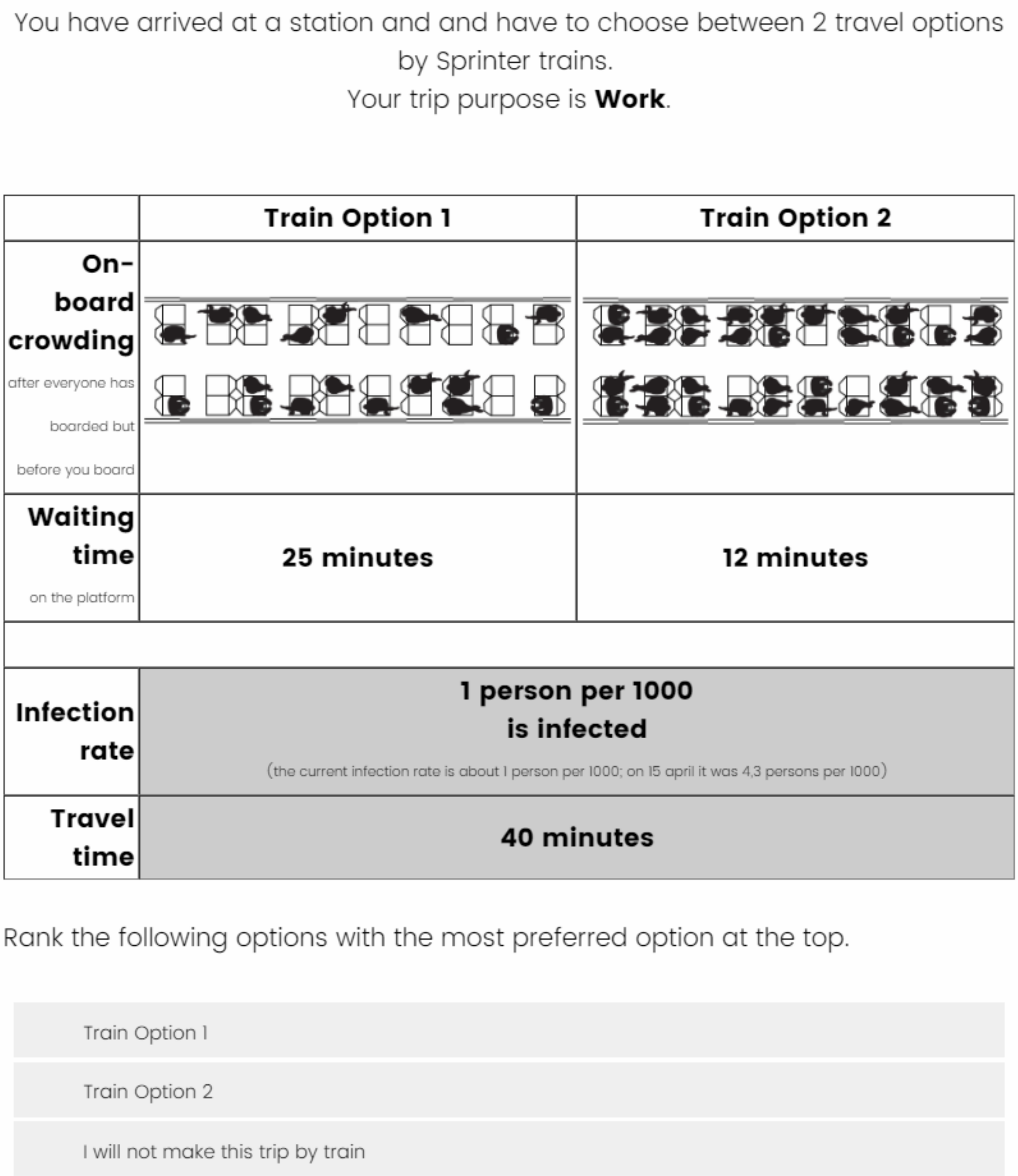


Figure 19: Screenshot of the choice experiment (translated to English)

Table 14: Sample characteristics

Total respondents	**513**		
		Distribution (%)	
Attribute	**Value**	**Actual**	**Required**[15]
Gender	Female	49%	~50%
	Male	50%	~50%
	Other	0%	
Age	18-24	15%	11%
	25-34	18%	16%
	35-44	17%	15%
	45-54	17%	18%
	55-64	19%	17%
	65-74	13%	14%
	>74	2%[16]	10%
Education[17]	Elementary school (*basisonderwijs*)	1%	~29%
	Secondary school (*HAVO/VWO/VMBO*)	27%	
	Vocational diploma (*MBO*)	34%	~37%
	Higher professional education (*HBO*)	25%	~33%
	University education incl. bachelor, master, PhD (*WO*)	13%	

Personal characteristics

In the first category of personal characteristics, we asked travellers how often they travelled with the train before and during the pandemic-related restrictions, the crowding level they usually experienced, their most frequent purpose of travel, and which alternative modes were available for these trips. In the second category, common socio-demographic questions (age, gender, income, employment status, zip code, highest education attained, and household size) were asked. In addition, some variables more specific to the current context were also collected; in particular, ages of household members and past, current, and expected future status of working or studying from home. The final category consists of questions regarding the perceived likelihood of the respondent or someone in their household getting infected and the severity of the disease if they do. Respondents were also asked about the degree to which they think they, themselves, and others follow pandemic related advice and regulations such as frequent hand sanitization and social distancing in public places. Finally, this category also includes questions about institutional trust and frequency of information seeking in relation to the pandemic. Note that all variables in this category except the last one noted here are qualitative Likert scale measures.

[15] Source: Centraal Bureau voor de Statistiek (2020)
[16] The under-sampling here is potentially due to the minimum train trip frequency requirement
[17] Translated to international equivalents

2.3 Choice analysis

Observations are analysed under the conventional random utility maximization framework where the utility of an alternative i for individual n, U_{in}, consists of a systematic (V_{in}) component, capturing the utility associated with factors observed by the analyst, and a random (ε_{in}) component. We assume that the systematic component is linear-additive and is computed by taking the sum of the alternate specific constant (β_i) and the product of taste preferences (β_{ij}) and the values of attributes, j (x_{ijn}) (Equation [15]). The probability of choosing alternative i from I alternatives in a multinomial logit (MNL) model, is given by Equation [16].

$$\begin{aligned} U_{in} &= V_{in} + \varepsilon_{in} \\ V_{in} &= \beta_i + \sum_j \beta_{ij} \cdot x_{ijn} \end{aligned} \qquad [15]$$

$$P_{in} = \frac{e^{V_{in}}}{\sum_{i'=1}^{I} e^{V_{i'n}}} \qquad [16]$$

To assess heterogeneity in traveller behaviour we use a latent class choice model (LCCM) which is a discrete mixture of choice models to which individuals are probabilistically allocated. Although the choice models can have different attributes, structures, or even belong to a completely different framework, we use the same model in each class. The probability of an individual n, belonging to class s (amongst S classes) with probability π_{ns}, choosing alternative i is the product-sum of the class membership probabilities and the probability of selecting that alternative for each class (given the vector of taste parameters in that class, $\boldsymbol{\beta_s}$) (Equation [17]). Panel effects are accounted for by assuming that a particular individual is allocated to each class with the same probability for all their choices. The likelihood of observing an individual's sequence of choices $i{:}i_1,\dots,i_T$ by individual n over T situations is given by Equation [18].

$$P_{in} = \sum_{s=1}^{S} \pi_{ns} \cdot P_{in}(\boldsymbol{\beta_s}) \qquad [17]$$

$$L_{in} = \sum_{s=1}^{S} \pi_{ns} \prod_{t=1}^{T} P_{in_t}(\boldsymbol{\beta_s}) \qquad [18]$$

An important aspect of LCCM is the ability to explain behavioural heterogeneity though class membership probabilities using values of individual characteristics, k (z_{kn}) (Equation [19]). We use socio-demographic and mobility characteristics to explain class membership. For other, generally unobservable variables (such as, worrying about transmitting the infection to someone in the household), we conduct a posterior analysis to find the distributions of these variables in the classes of the estimated model.

$$\pi_{ns} = \frac{e^{\delta_s + \sum_k \gamma_{ks} \cdot z_{kn}}}{\sum_{s'=1}^{S} e^{\delta_{s'} + \sum_{k'} \gamma_{k's'} \cdot z_{k'n}}} \qquad [19]$$

The class-specific taste parameters ($\boldsymbol{\beta_s}$) and membership coefficients (γ_{ks}, δ_s) are simultaneously estimated using PythonBiogeme (Bierlaire, 2016).

3 Results and discussion

Since the experiment was conducted in the context of the COVID-19 pandemic, one fear was that a large number of respondents would simply opt-out of using trains altogether. Ultimately, this was not the case and only about 4% of the respondents always opted-out while 13% never opted-out in the 15 situations they faced. A substantial number of respondents always ranked waiting higher than the crowded train (~13%) or vice versa (~3%). However, only about 1% and 3% always chose waiting or taking the crowded train as their top option, respectively. As expected, the proportion of observations where the respondent takes the crowded train rises with the extra waiting time required to reduce one person on-board (Figure 20).

We tried a number of utility specifications, in particular varying whether attributes were modelled as having a linear or non-linear effect. Since, perceived infection risk (and, therefore, disutility of a train alternative) may be higher when two contributing factors are higher together, we also included interaction effects of crowding with infection rate and in-vehicle times in the utility specification. Ultimately, the specification in Equation [20] within MNL was found to provide the most informative model parameters. Table 15 gives an overview of attributes included in the final model.

The MNL model shown in Table 16 is finalized by removing insignificant (p > 0.10) parameters one-by-one. As shown, all parameters have the expected signs and magnitudes: the likelihood of choosing an option generally decreases with increasing crowding and waiting time while the likelihood of opting out increases with increasing infection rate. A small and large non-linear effect is found for the highest attribute levels of crowding and infection rate, respectively. Since respondents were only shown graphics for on-board crowding, meaningful non-linear effects possibly indicate that respondents may be applying subjective labels (Li and Hensher, 2013); for instance (as will be shown in the LCCM results) assigning a much higher utility to being able to sit alone than to a similar reduction in crowding otherwise. Although the coefficients for in-vehicle time and the interaction between crowding and in-vehicle time were not significant for the MNL model, we keep them in the specification tested for the LCCM.

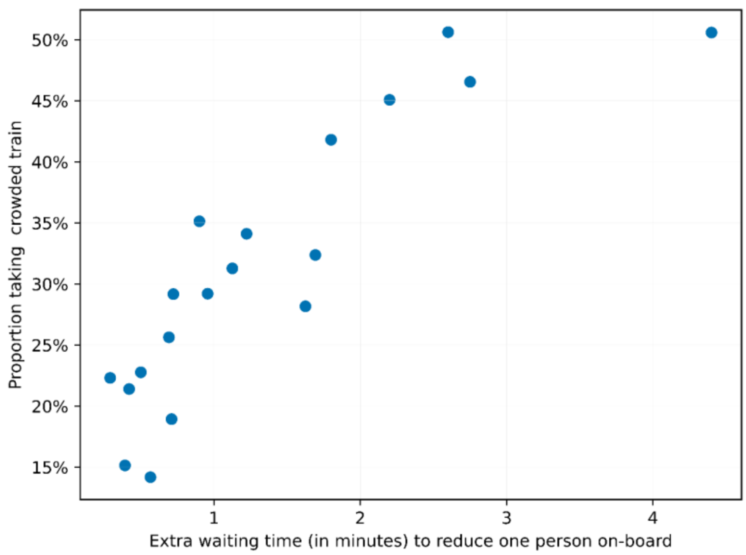


Figure 20: Proportion of observations choosing crowded train versus extra waiting time to reduce one person on-board

$$
\begin{aligned}
V_{\text{train 1}} = {} & \beta^{\text{crowd: can sit alone}} \cdot crowd_{18,\text{ train 1}} + \beta^{\text{crowd: not crowded}} \cdot crowd_{23,\text{ train 1}} \\
& + \beta^{\text{crowd: quite crowded}} \cdot crowd_{28,\text{ train 1}} + \beta^{\text{crowd: almost full}} \cdot crowd_{36,\text{ train 1}} \\
& + \beta^{\text{WT}} \cdot WT_{\text{train 1}} + \beta^{\text{crowd}\times\text{infect}} \cdot crowd_{\text{train 1}} \cdot infect + \beta^{\text{crowd}\times\text{IVT}} \cdot crowd_{\text{train 1}} \cdot IVT \\
V_{\text{train 2}} = {} & \beta^{\text{crowd: can sit alone}} \cdot crowd_{18,\text{ train 2}} + \beta^{\text{crowd: not crowded}} \cdot crowd_{23,\text{ train 2}} \\
& + \beta^{\text{crowd: quite crowded}} \cdot crowd_{28,\text{ train 2}} + \beta^{\text{crowd: almost full}} \cdot crowd_{36,\text{ train 2}} \\
& + \beta^{\text{WT}} \cdot WT_{\text{train 2}} + \beta^{\text{crowd}\times\text{infect}} \cdot crowd_{\text{train 2}} \cdot infect + \beta^{\text{crowd}\times\text{IVT}} \cdot crowd_{\text{train 2}} \cdot IVT \\
V_{\text{opt-out}} = {} & \beta^{\text{opt-out}} + \beta^{\text{IVT}} \cdot IVT \\
& + \beta^{\text{infect: 0.01}} \cdot infect_{0.01} + \beta^{\text{infect: 0.5}} \cdot infect_{0.5} + \beta^{\text{infect: 2}} \cdot infect_{2} + \beta^{\text{infect: 10}} \cdot infect_{10}
\end{aligned}
\tag{20}
$$

We examined whether removing respondents who consistently ranked the crowded train higher or lower than the other train alternative (~16% of the sample) affected the results. The MNL model estimated from this reduced sample shows a slightly lower impact of crowding and infection rates, and a higher (negative) impact of opting out. A possible reason for this is that those who consistently ranked less crowded trains higher often selected opting out as their most preferred option. We also tested consistency in responses to check for respondent fatigue. To do this, we split the choice dataset into two—responses to the first eight choice situations and the last seven choice situations—and estimate the described MNL model for each set. While the value of crowding (relative to waiting time) remains similar, the model for the last seven situation shows larger and smaller impacts for infection rate and opting out, respectively. We suspect that as respondents gain experience with the choice situation, they are able to better discriminate between the contextual variable that is infection rate.

For the LCCM, we first find the optimal number of classes using an intercept-only class membership function. Typically, this is done using the model fit indicators, particularly the Bayesian information criterion (BIC), which explicitly penalizes the number of parameters in the model. In our case, model fit indicators continued to improve as we increased the number of classes (we checked up to 6 classes). Therefore, we chose the 2-class model as it, in our opinion, best described heterogeneity in behaviour. While the 2-class model clearly delineated two behavioural types, adding more classes yielded intermediate groups without adding more insights. Moreover, adding more groups resulted in higher standard errors of estimated parameters and even led to unexpected parameter signs for higher number of classes. Next, the choice models in each class are finalized in the same way as the MNL model: by removing insignificant ($p > 0.10$) parameters one-by-one. Finally, all non-correlated observable individual characteristics are included in the class membership function and eliminated one-by-one if they are insignificant to arrive at the final model shown in Table 16.

The choice parameters in both classes have signs and magnitudes in line with expectations. Results show that, in general, higher levels of crowdedness, waiting times, and infection rates all reduce travellers' willingness to board a particular train alternative and increase the probability of opting out. Surprisingly, for the LCCM too, in-vehicle time—time to be spent in an enclosed train coach—does not affect travellers' decisions indicating that they might be underestimating the importance of duration of exposure on the risk of infection.

Table 15: Overview of attributes included in the final choice model

Attributes	Symbol	Explanation	Range
Choice coefficients			
Crowding (level *i*)	$\beta^{\text{crowd: } i}$	Categorical (effect coded)	
Waiting time	β^{WT}	All time attributes are in minutes	3-25
In-vehicle time	β^{IVT}		10-40
Infection rate (level *i*)	$\beta^{\text{infect: } i}$	Categorical (effect coded)	
Opt-out constant	$\beta^{\text{opt-out}}$		
Personal characteristics			
Age	β^{age}	Ordinal in ascending order: 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84	1-7
Gender	β^{female}	Categorical (effect coded): female, male	
Train use frequency during COVID	$\beta^{\text{train freq. covid}}$	Ordinal in ascending order: never, once per month, 1-3 times per month, more than 4 times per week	1-4

To enable comparisons of coefficients across models, we calculate the ratio of each attribute coefficient to the waiting time coefficient under the scaled values columns in Table 16. As can be seen from the scaled columns, the two estimated classes differ strongly on the relative impact of level of crowdedness and infection rates. Moreover, the general propensity to opt-out has a very large effect in both classes but is in opposite directions. Based on these differences, we call the first class ‘COVID Conscious’ as decisions in this group are more strongly driven by the level of crowdedness, infection rates, and the expected number of infected persons on-board (approximated by the interaction effect of crowdedness and infection rate). In contrast, the second class, which we call ‘Infection Indifferent’, is affected by these factors to a lesser degree and is also less likely to opt-out on average.

Since respondents are assigned to classes probabilistically, econometric indicators calculated for the two classes do not directly apply to individual respondents. Instead, they form the lower and upper boundaries for individual indicators which are obtained as the weighted average of the estimates for the two classes. Averaging the individual-specific indicators gives an aggregate value over the whole sample. The weights—posterior membership probabilities—are calculated by multiplying the prior probabilities (obtained by applying the class membership model) with the likelihood of observing individual respondents’ sequence of choices and then normalizing (Hensher et al., 2015). The sum of the posterior probabilities for each class gives the class sizes shown in percentage in Table 16. To examine the extent to which individuals belong to either class, in Figure 21, we plot the distribution of absolute difference between posterior class membership probabilities (higher values indicate a more deterministic assignment to either class). As can be seen, about 85% of respondents are assigned to one class with a probability of 95% or more. In the following, we discuss traits of the two latent classes—alternatively referring to prototypical (i.e., representative) travellers of either class—rather than focussing on aggregate estimates. Note that, given the posterior probability distribution, for most respondents, individual-specific estimates would be close to those calculated for one of the classes.

Table 16: Estimation results

Model	**MNL**			**LCCM 2-Class**					
# Parameters	10			23					
Initial LL	-8453.822			-8054.030					
Final LL	-7572.643			-6498.217					
Adjusted ρ^2	0.103			0.190					
BIC	15234.77			13202.246					
				Class-specific choice models					
				Class 1: COVID Conscious Travellers			**Class 2: Infection Indifferent Travellers**		
Class Size				53.73%			46.27%		
	Coeff.	p-val	Scaled	Coeff.	p-val	Scaled	Coeff.	p-val	Scaled
$\beta^{\text{crowd: almost empty}}$	1.317	–	-42.76	2.230	–	-159.29	0.690	–	-18.16
$\beta^{\text{crowd: can sit alone}}$	0.245	0.00	-7.95	0.792	0.00	-56.74	0	–	–
$\beta^{\text{crowd: not crowded}}$	-0.188	0.00	6.10	-0.531	0.00	37.93	0.110	0.01	-2.89
$\beta^{\text{crowd: quite crowded}}$	-0.558	0.00	18.12	-0.921	0.00	65.79	-0.262	0.00	6.89
$\beta^{\text{crowd: almost full}}$	-0.816	0.00	26.49	-1.570	0.00	112.14	-0.538	0.00	14.16
β^{WT}	-0.031	0.00	1	-0.014	0.02	1	-0.038	0.00	1
$\beta^{\text{crowd}\times\text{infect}}$	-0.0012	0.00	0.004	-0.0046	0.00	0.33	-0.0026	0.00	0.07
$\beta^{\text{crowd}\times\text{IVT}}$	–	–	–	–	–	–	–	–	–
$\beta^{\text{infect: 0.01}}$	-0.397	–	12.89	-0.720	0.00	51.43	0	–	–
$\beta^{\text{infect:0.1}}$	0.059	–	-1.92	-0.213	–	15.21	-0.488	–	12.84
$\beta^{\text{infect: 0.5}}$	0	–	–	0.133	0.08	-9.5	0	–	–
$\beta^{\text{infect: 2}}$	0.387	0.00	-12.56	0.521	0.00	-37.21	0.254	0.00	-6.68
$\beta^{\text{infect: 10}}$	0.375	0.00	-12.18	0.279	0.05	-19.93	0.234	0.00	-6.16
β^{IVT}	–	–	–	–	–	–	–	–	–
$\beta^{\text{opt-out}}$	-0.424	0.00	13.77	0.927	0.00	-66.21	-2.02	0.00	53.16

Table 16 (continued)

Model	**MNL**			**LCCM 2-Class**					
				Class membership model					
				Class 1: COVID Conscious Travellers			**Class 2: Infection Indifferent Travellers**		
				Coeff.	p-val	Scaled	Coeff.	p-val	Scaled
$\beta^{\text{intercept}}$				0	–	–	-1.160	0.00	1
β^{age}				–	–	–	-0.107	0.08	0.09
β^{female}				–	–	–	-0.275	0.01	0.24
$\beta^{\text{train freq. covid}}$				–	–	–	0.820	0.00	-0.71

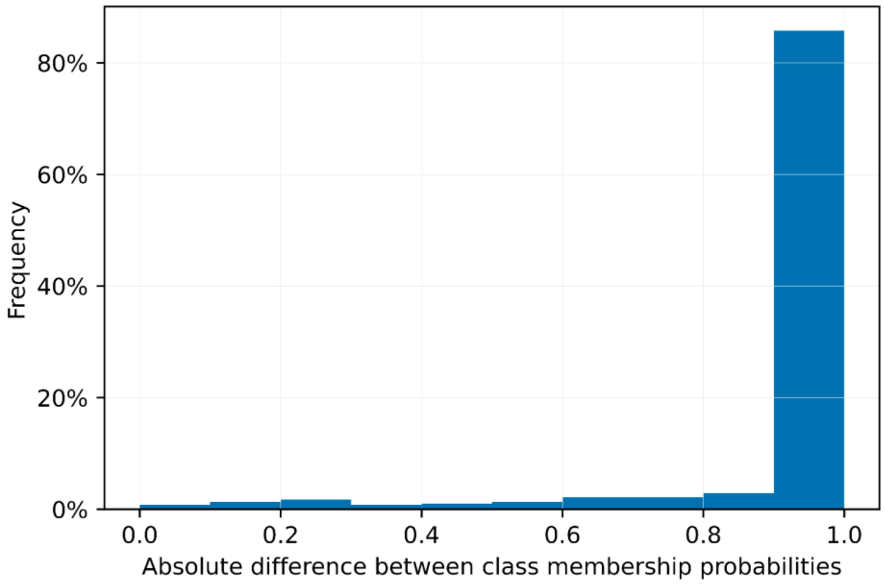


Figure 21: Distribution of absolute difference between posterior class membership probabilities

Typically, the effect of crowding has been modelled as an in-vehicle time multiplier (Li and Hensher, 2011b; Wardman and Whelan, 2011). The idea being that the disutility of crowdedness should be larger for longer trips because passengers have to be in a crowded vehicle for a longer time. We included crowding in our model both as a constant penalty as well as an interaction effect with in-vehicle time (i.e., as a multiplier). As shown in Table 16, the time multiplier parameters were not significant. When constant penalties are excluded, the time multiplier parameters are significant but, similar to Kroes et al. (2014), the model has a significantly lower goodness of fit. We note that constant penalties may only be performing better for the range of in-vehicle times that were tested in this survey or because in-vehicle times were only included as context effects in the experiment.

Equation [21] shows how the values of crowding are calculated. Since we estimate parameters for each crowding level (i), the coefficient for change in crowding between two levels ($\beta^{\text{crowd:}i\rightarrow i+1}$) is given by the difference in the utility contributions divided by the difference in the number of persons on-board (x^i). This crowding coefficient divided by the coefficient for waiting time gives the value of crowding in terms of waiting time between those levels ($\Upsilon^{i\rightarrow i+1}$). The average value of crowding (Υ) is given by the weighted average of these individual values of crowding.

$$\beta^{\text{crowd: }i\rightarrow i+1} = \frac{\beta^{\text{crowd: }i} - \beta^{\text{crowd: }i+1}}{x^i - x^{i+1}}$$

$$\Upsilon^{i\rightarrow i+1} = \frac{\beta^{\text{crowd: }i\rightarrow i+1}}{\beta^{\text{WT}}} \qquad [21]$$

$$\Upsilon = \frac{\Upsilon^{1\rightarrow 2}\cdot(x^2-x^1)+\Upsilon^{2\rightarrow 3}\cdot(x^3-x^2)+\Upsilon^{3\rightarrow 4}\cdot(x^4-x^3)+\Upsilon^{4\rightarrow 5}\cdot(x^5-x^4)}{(x^2-x^1)+(x^3-x^2)+(x^4-x^3)+(x^5-x^4)}$$

Prototypical travellers of the COVID Conscious class, are willing to wait an extra 8.75 minutes, on average, to reduce just one person on-board. As shown in Figure 22, travellers in this class are willing to wait the most when there is a possibility to sit alone. This is indicative of the aversion towards infection risk in this class as well as the general framing of the choice situations in the context of the pandemic. Note that this is in contrast to previous studies which typically report that the impact of crowding increases with the number of persons on-board, especially after 60–80% load factor (Hörcher et al., 2017; Wardman and Whelan, 2011; Yap et al., 2018). The value of crowding for Infection Indifferent travellers seems to be more in line with values from previous studies (albeit on the higher end) with an average willingness to wait of 1.04 minutes to reduce one person on-board. (The average willingness to wait in the MNL model is 2.23 minutes per person.)

Previous studies that evaluate the effect of seating occupancies—either as constant penalties or in-vehicle time multipliers—and waiting times can be compared with our results. To convert coefficients from such studies to units comparable to ours, where required, we assume a total seat capacity of 40 and an in-vehicle time range of 10–40 minutes. Preston et al. (2017) observed stated choices in an experiment where respondents chose between two trains: the first which was due but with no possibility to sit and the second with a given expected waiting time and one of two lower crowding levels. They report constant crowding penalties that lead to values of crowding in the range of 0.52–0.96 minutes per person. From an experiment similar to ours, Kroes et al. (2013) find a willingness to wait of 0.15 and 0.33 minutes per person to reduce crowdedness from 75% to 50% and from 100% to 75%, respectively. Assuming expected waiting times to be half of headways, Tirachini et al. (2013) find similar values of 0.15–0.6 minutes per person from a mode choice experiment. For comparison, note that we found the COVID Conscious and Infection Indifferent classes willing to wait 8.75 and 1.04 minutes per person, respectively. Douglas and Karpouzis (2006) and Sahu et al. (2018) use similar stated choice experiments for Sydney and Mumbai. They find that travellers are willing to wait 1.88–7.52 minutes and 3.58–14.32 minutes, respectively, for an uncrowded seat over a crowded seat alternative. Assuming ‘not crowded’ and ‘almost full’ to be the corresponding categories in our model, we find an extremely high value of 74 minutes and a more moderate 17 minutes in the COVID Conscious and Infection Indifferent classes, respectively. Using revealed preferences from smart card data in The Hague, however, Yap et al. (2018) found significantly lower values between 0.015 and 0.06 minutes per person (depending on in-vehicle times) for trams. Thus, while both the COVID Conscious and Infection Indifferent classes show higher values of crowding, the latter is much closer to pre-pandemic estimates from stated choice experiments.

As shown in Figure 23, for both classes, the tendency to opt-out increases as a concave function of the prevalent infection rate. The effect plateaus at extreme infection rates (2% and 10%) indicating that travellers may be considering a threshold level beyond which the infection rate itself no longer contributes to perceived risk. A t-test could not reject the null hypothesis that the coefficients for these extreme infection rate levels are equal in either class (p-values: 0.17 and 0.91 for each class, respectively). Other consecutive levels are statistically different (p-value < 0.01). The graph for COVID Conscious prototypical travellers demonstrates, again, the strong preference to sit in an almost empty coach or to sit alone. Furthermore, note that the opt-out rates for crowded vehicles in the COVID Conscious class are fairly inelastic in relation to infection rates. This might indicate that travellers in this group would not feel safe travelling in crowded vehicles even when infection rates are back to pre-pandemic levels.

Amongst the individual characteristics collected, three variables contributed to explaining the differences in behaviour between the two classes. Older and female respondents were over-represented amongst COVID Conscious travellers whereas those reporting higher train use during the COVID-19 restrictions were likely to be Infection Indifferent. Presumably, older people, are more risk averse due to a higher vulnerability to the disease. While the disease does not seem to affect women more severely than men, female respondents have often been shown to be more risk averse in their decisions (e.g., de Palma and Picard (2005)). Kluwe-Schiavon et al. (2021) also find that older and female respondents had a lower COVID-19 risk tolerance for economic opportunity. The causal relationship between higher train use during the pandemic and lower risk aversion may be in either direction. Travellers with a higher probability of being Infection Indifferent may have used the train more frequently because they are not particularly averse to the COVID-19 risk. Conversely, lower risk aversion amongst those who use the train more frequently during the pandemic might be explained by the existence of the description-experience gap when evaluating risky choices. When judging the likelihood of contracting COVID-19 on public transport, these travellers may be depending more on their experience rather than the risk described by authorities (Barron and Erev, 2003). When people make decisions based on experience, they do not account for rare events as much as the objective probabilities of such events suggest they should (Hertwig and Erev, 2009). A little surprisingly, having the possibility to conduct the trip with a private mode (e.g., car, bicycle, walk) was not related to the propensity to opt-out of the train alternatives.

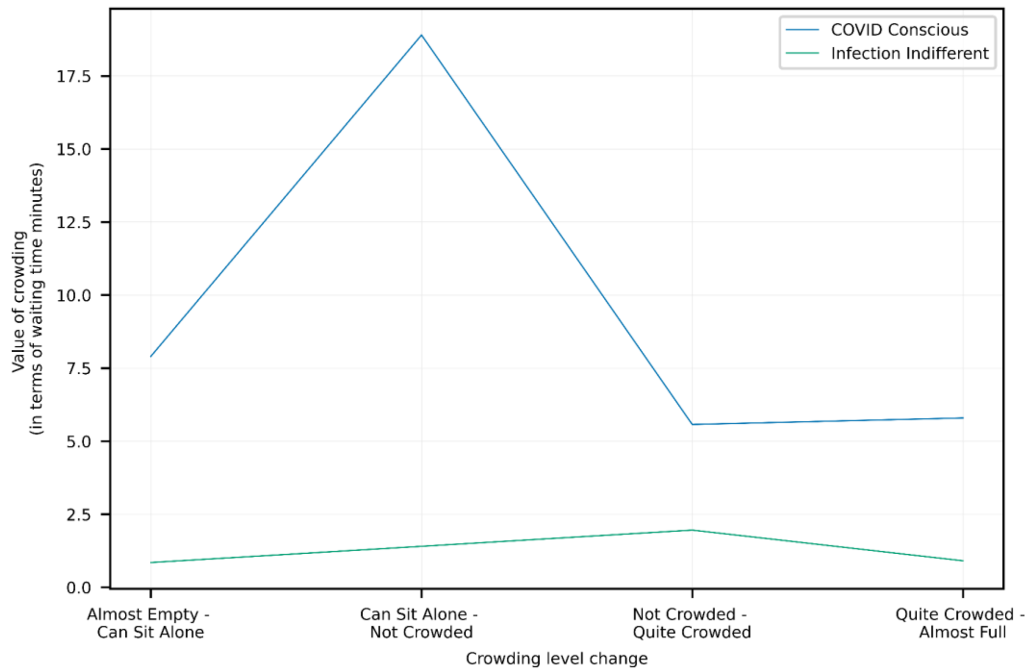


Figure 22: Value of crowding (in terms of waiting time minutes per person on-board) between levels in the two traveller classes[18]

[18] The value of crowding interpolated between can sit alone and not crowded in the Infection Indifferent class

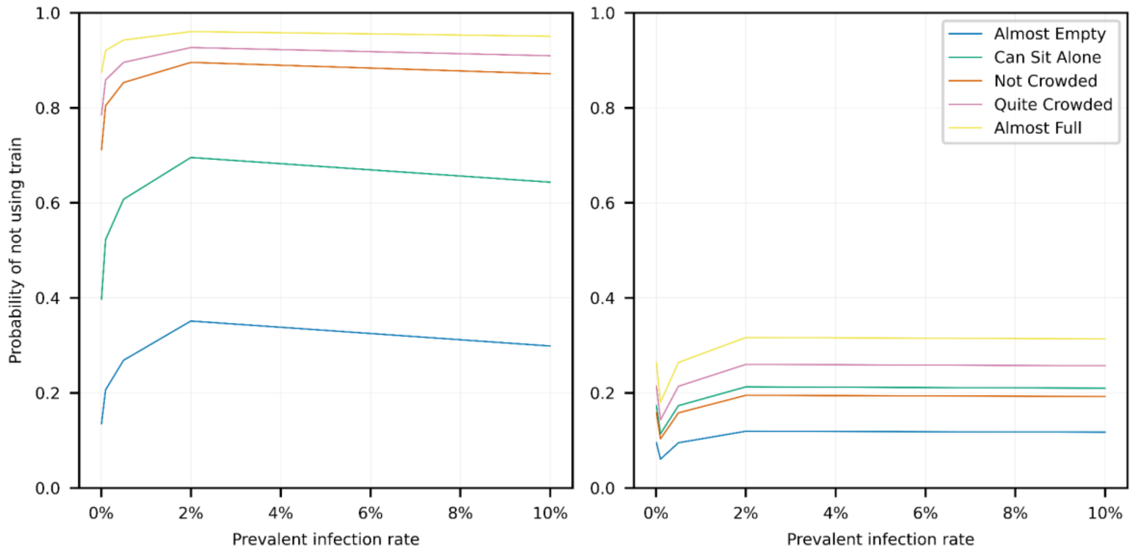


Figure 23: Probability of not using trains versus infection levels for different crowding levels in the two groups (wait time for trains is 12 minutes)

Figure 24 shows the posterior distributions of some unobserved COVID-19-related qualitative measures. To obtain these distributions, the posterior membership probabilities are summed class-wise for all respondents who select a particular level for an indicator. The class-wise sums are then normalized by the respective class-sizes to find the percentage of each class that would select a given level on the indicator being analysed. Since respondents are only probabilistically assigned to either class, we note that the distributions convey the characteristics of the latent class rather than that of the individuals. Although the two classes do not differ too strongly on these factors, small differences can be noted. The COVID Conscious class tends to be more worried about the pandemic, specifically, about being hospitalized and spreading the infection to someone in their household. A moderate correlation exists between age and worrying about being hospitalized but not with other factors (older travellers are over-represented in this class). Additionally, travellers highly likely to be COVID Conscious reported themselves to be more rule-following, indicating that they followed advice such as frequently sanitizing hands and maintaining 1.5 m distances in public places. Moreover, they also had a more negative opinion about the degree to which others followed these rules. Indicative of the long drawn national debate over it (DutchNews.nl, 2020), face mask use is unpopular with both classes and largely uncorrelated with the degree to which other measures are followed.

4 Conclusion

The COVID-19 pandemic has had an extensive impact on public transport. As a result of actions and advisories aimed at containing the disease, public transport ridership has declined sharply, perceptions regarding this mode have become more negative, and there has been a shift to personal transport modes. Consequently, changes in traveller behaviour in order to minimize exposure to the virus are expected. Moreover, these changes may be sustained through different stages of the pandemic and even have a significant effect on public transport demand after the pandemic. While a number of studies examine current ridership patterns and anticipated transport preferences, few have investigated trade-offs in the age of COVID-19 via choice analysis in detail.

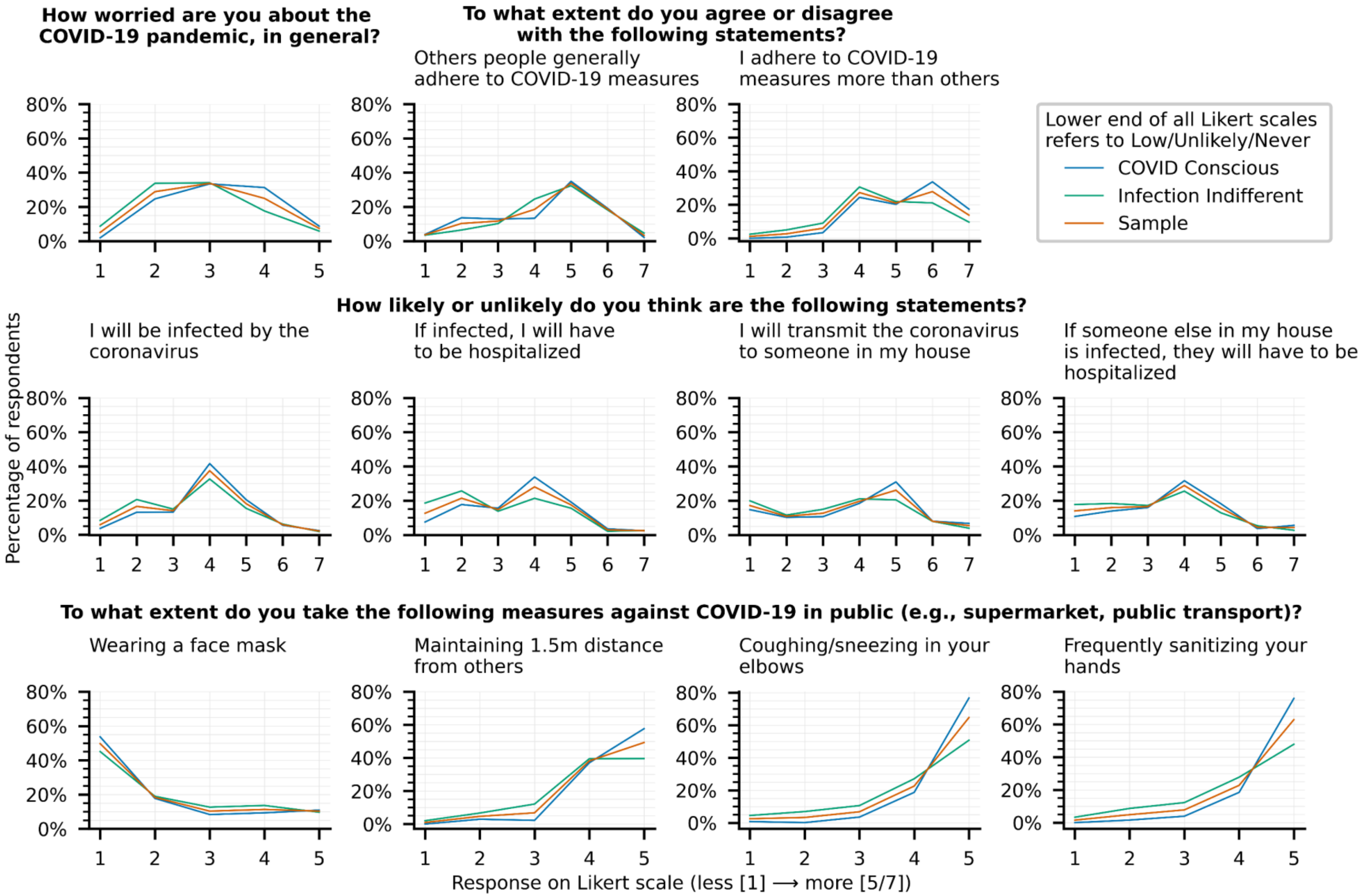


Figure 24: Class profiles of personal, unobserved, qualitative measures related to the COVID-19 pandemic

In this study, we analysed traveller behaviour related to factors affecting the risk of COVID-19 transmission in public transport with a stated choice experiment. Since one of the most important ways to avoid exposure is to reduce contact with other people, we measured travellers' (potentially updated) valuation for on-board crowding. To do this, we obtained respondents' preferences between a crowded-but-low-wait-time and a less-crowded-but-higher-wait-time (and an opt-out) alternative. Choices were presented in the context of exposure duration (operationalised as the in-vehicle time of the alternatives) and infection rate to examine the effects of these risk-contributing factors on choice behaviour.

Responses were collected from train travellers in the Netherlands at the end of the first infection wave (May 20-25, 2020), just as the first restrictions were being lifted and new regulations were setup for travel in public transport. We believe that behavioural insights from this study will contribute to better demand forecasting. In addition to providing insights regarding when travellers opt-out (i.e., choose not to travel), by modelling decisions of the type 'to board or to wait' we can provide an important behavioural input to agent-based models (e.g., Cats and Gkioulou (2017)) that commonly simulate this choice situation. These results will also be valuable in informing public transport policy decisions, not only in the current pandemic but also future ones.

Applying a latent class choice model, we found two, nearly equal-sized traveller segments: COVID Conscious and Infection Indifferent. While higher crowding levels and infection rates reduce the willingness to board a train in both, the effect of these factors is much larger in the COVID Conscious segment. Value of crowding, measured as the number of minutes travellers are willing to wait to reduce one person on-board, is also significantly higher in this class (on average 8.75 minutes/person) and increases sharply with the possibility to sit alone. In contrast, Infection Indifferent travellers' value of crowding (on average 1.04 minutes) is comparable to, although slightly on the higher end of, pre-pandemic evaluations. Moreover, unlike their counterparts, COVID Conscious travellers are highly affected by the prevalent infection rate, particularly at low crowding levels, and are more likely to opt-out in general. Surprisingly, neither group took the exposure duration into account. Older and female respondents are over-represented in the COVID Conscious class while those who report higher train use during the pandemic are more likely to belong to the Infection Indifferent class. Finally, distributions of COVID-19-related indicators showed that COVID Conscious travellers were more worried about the pandemic, considered themselves and household members more likely to be hospitalised if infected, and reported themselves to be following related measures to a higher degree.

Although the crowding valuation for the COVID Conscious class is high, we note that it does not necessarily imply that travellers are willing to accept hour-long waits. Firstly, respondents had the option to opt-out and secondly, applications of the choice model must, technically, be within the range of waiting times included in the experiment (3–25 minutes). Thus, the COVID Conscious class model, in fact, indicates a very strong preference for either selecting the less crowded train or opting out altogether. On the other hand, prototypical travellers of the Infection Indifferent class, show a more nuanced decisions between taking the crowded train or waiting. Furthermore, while we cannot rule out the possibility that the COVID Conscious class has an inherently lower value of time, the results as a whole strongly suggest that the high values of crowding are driven by sensitivity to pandemic risks. For instance, the COVID Conscious class has a higher likelihood of opting out and is impacted more strongly by infection rates. Moreover, it also scores higher on indicators such as worrying about COVID-19, likelihood of being hospitalized, hygiene-related rule following, etc.; all of which imply a strong impact of the pandemic.

A variety of direct and indirect effects of the pandemic have led public transport ridership to plummet. Given the importance of public transport in economic recovery and sustainable mobility, authorities and operators need to work to improve travellers' perception about public transport and slow down the shift to non-sustainable modes. Ridership levels have typically returned to normal at the end of previous localized catastrophic events, such as epidemics and security threats (Gkiotsalitis and Cats, 2020). However, the COVID-19 pandemic is unprecedented in its spatial and temporal scale with regions around the world going in and out of lockdowns over an extended period of time. Thus, authorities cannot depend on ridership to improve by itself but must actively work towards increasing public transport demand while providing this essential service safely.

The apprehension of the COVID Conscious segment seemingly follows calls from authorities to avoid public transport. While such calls are compatible with the intuition that sharing confined spaces may be unsafe, there is little to no hard evidence of outbreaks linked to public transport. This might suggest that public transport travel could be safe if recommended measures are implemented (Gkiotsalitis and Cats, 2020; Goldbaum, 2020; Schive, 2020; UITP, 2020). Yet, the fact that over two-thirds of COVID Conscious travellers in our sample are unwilling to travel if they cannot find an empty row, regardless of infection level, is an indication of how difficult it will be to restore travellers' confidence and foreshadows lingering behavioural adaptations from the pandemic in the future. Where trips cannot be replaced by telecommuting or active modes, providing crowding information for public transport can be the key. For low infection rates (0.1%), when there is a possibility to sit in an almost empty coach or with the adjacent seat empty, 50-80% of COVID Conscious travellers in our sample indicated that they would use the train. By highlighting which trains and coaches will be less crowded, these travellers can adjust their departure times, routes, and even the choice of which coach to board. Assuming that these travellers overestimate the likelihood of contracting COVID-19 on-board public transport, more experience with travelling (even in less crowded vehicles) could bring their assessments in line with reality. Future studies may also look into the effect of other risk mitigating actions, such as mask mandates and increased cleaning, as well as the (perceived) extent to which these are followed on travellers' risk perceptions.

While public transport may be safe with recommended measures, their overall lower concern regarding the virus and absence of substantial behavioural change, indicates that Infection Indifferent travellers may not be motivated to follow them carefully. Poor compliance from these travellers could increase the real as well as perceived risk for everyone and further drive the apprehension of other travellers. Therefore, we must continue to emphasize the need for simple measures such as face masks and recognize that returning to pre-pandemic levels of crowding while the prevalent infection levels are still significant would be reckless.

Although the stated choice experiment provides important behavioural estimates, we note limitations arising from the information provided to respondents and the nature of such experiments. Firstly, the prevalent infection rates given as a contextual attribute can only be estimated, and trusted estimations may not be available everywhere. Even if they were available, one might question whether travellers actually consider this information directly or respond to more abstract cues, such as the intensity of regulations or media coverage. However, since it is difficult to recreate the entire context for the experiment, we used this single indicator (which is correlated to such cues). Furthermore, travellers were helped in anchoring the prevalent infection rates to abstract cues by providing the prevalent infection rates at two different dates. Nevertheless, care must be taken when transferring estimations to other situations. Future studies could focus on analysing the impacts of specific societal, political, and media cues on pandemic-related travel behaviour.

Secondly, crowding information is not commonly available although a growing number of public transport networks and trip planning applications now try to provide predictions in some format (see review in (Drabicki et al., 2020)). In their smartphone application, the Dutch railways show a (qualitative) three-level crowding indicator for most trains and a more precise ‘seat-finder’ on some trains, showing seat availability in different coaches (Nederlandse Spoorwegen). By prominently displaying crowding information, we may have drawn respondents’ attention to this aspect more than usual, leading to higher crowding valuation. Previous studies (e.g., Yap et al. (2018)) have also claimed that travellers tend to demonstrate a higher value of crowding in stated choice experiments than is observed from passively collected data. Moreover, while we presented crowding (and waiting time) information as objectively true, in real life, depending on factors such as trust in the information provided by the operator, travellers may consider attributes of the second train to be uncertain and therefore attach a higher disutility to it. We believe reliable crowding information could be key in regaining travellers’ trust. Since, crowding is likely to be affected by subjective perceptions, studies assessing travellers’ responses to different presentation formats and reliability levels would be critical for such developments.

Thirdly, we note that choices observed here are hypothetical and the situations do not directly reflect the various constraints arising from societal positions of individuals. These constraints have been previously found to play a significant role in travellers actual capacity to change behaviour (Kim et al., 2017a). Thus, although we find an intent to avoid trains amongst COVID Conscious travellers their ability to do so may be limited because of employer constraints. In contrast, some Infection Indifferent travellers may indicate intent to use the trains but do not actually do so as they can work/study from home. We attempted to control for such factors by marking the opt-out option as ‘I will not make this trip by train’ and asking respondents if they had alternative modes for the stated trip purpose. We also asked respondents’ family income range and (for working/studying individuals) how effectively they could work/study at home as well as the frequency of doing so. None of these variables were found to contribute towards explaining choice behaviour. Nevertheless, the precise distribution of travellers between the different classes and their propensity to opt-out must be used with care in applications, accounting for other individual constraints that might affect behaviour. We can, however, confidently interpret crowding valuations and the existence of significant risk-averse and indifferent traveller segments.

Finally, we stress that we have observed a snapshot of behaviour (and intent) for present and future circumstances. Since the global outbreak of the COVID-19 pandemic in March 2020, the situation has developed quickly and unpredictability. Given the widespread and extended impact of this pandemic, people have been rapidly adopting and changing behaviours for the evolving new realities encountered during the course of this crisis. Yet, these acquired behaviours also fluctuate as the level of precaution changes depending on a number of factors, such as local infection rates, vaccination status, personal impact assessment, and ‘pandemic fatigue’. Thus, the trade-offs estimated here may change with new and significant developments; for instance, if a cure for COVID-19 emerges.

However, we emphasise that this does not obviate the utility of our findings and policy recommendations (or those of similar studies) for two reasons. First, given the uncertainty surrounding COVID-19, it has become almost impossible to confidently predict the 'end' of the pandemic and a return to a state of stability. With the emergence of new variants and questions surrounding the duration of vaccine protection and efficacy against new variants, we cannot rule out a return to a situation similar to the one analysed here. Second, and more importantly, our analysis contributes to a larger picture of traveller preferences (and subsequent policy recommendations) in key stages of the pandemic. In time, a meta-analysis charting traveller preferences over the pandemic may also be recommended. What we learn about travel behaviour from this pandemic will be instrumental in supporting policy-makers to act proactively in the next one.

Conclusions and Recommendations

1 Main scientific findings

This thesis studied route choice behaviour under uncertainty in public transport networks with the objective of contributing to the research gaps identified in chapter 1. This objective was broken down into four research questions each of which has been the focus of one of the preceding chapters. Here, we summarise the answers to these questions and briefly discuss empirical and methodological findings.

How to infer route choice sets from passively observed choices using minimal assumptions and producing transferable behavioural insights?
Having identified several drawbacks in existing methods for identifying route consideration sets, a choice set generation method using (deterministic) elimination-by-aspects as its basis is proposed. This decision heuristic combines attribute ranking and setting thresholds (acceptable change over the best possible value), working thus: relevant attributes are listed and ranked, the most important attribute is selected, all alternatives that do not meet the threshold for this attribute are removed, this attribute is removed from the list, this is repeated until all attributes have been checked. An important advantage of this non-compensatory heuristic is that it is likely reflecting the actual underlying process of consideration set formation (as opposed to simply matching the final empirical outcome). Therefore, in addition to route choice sets, insights into their formation are also uncovered from the calibration of the model.

The calibration uses generated feasible and observed routes from GTFS and smart card data, respectively. Generating feasible routes consists of the following steps: (i) representation of the infrastructure and service components of the network as efficient graphs, (ii) constrained enumeration of logical (no loops, no transfers between common lines) routes, and (iii) attribute assignment and removal of infeasible (i.e., unavailable) and state-wise dominated routes. This is followed by the calibration of the elimination-by-aspects model through an optimization procedure with the objective of balancing demand-weighted coverage (of observed routes) and efficiency (of generated routes).

The methodology thus needs as input the list of relevant attributes, the threshold measure (either ratio or absolute), and the required balance between coverage and efficiency. Application to the tram and bus networks in The Hague found that travellers ranked number of transfers as the most important attribute for consideration set formation, followed by waiting and in-vehicle time. Travellers do not seem to include alternatives with extra transfers or with travel times worse by more than 10% in their consideration set. The overall demand-weighted coverage was 63.9% but analysis of the coverage and efficiency trends indicates that this can potentially be improved by modelling closer and farther origin-destination pairs separately.

What is the impact of waiting time uncertainty, expressed as different statistical representations of its historical values, on route choice behaviour?

For this research question, we modelled the impact of waiting time uncertainty as if it were a known risk (i.e., objective uncertainty). The choices analysed were, however, made under natural ambiguity and were obtained from smart card data of the tram and bus networks of The Hague. Travel time characteristics were obtained from GTFS and AVL (automatic vehicle location data). Waiting time uncertainty is included in the choice model by directly introducing statistical measures of its historical values to the utility functions of route alternatives. For this, waiting time is split into (i) scheduled waiting time—the planned value of waiting time, (ii) regular deviations—the difference between the median of realised values and the planned value, and (iii) irregular deviations—the dispersion of the realised values. For the latter, we tested different statistical range and buffer time representations.

We estimated separate multinomial logit models for morning peak (06:00–09:00) and off-peak (09:00–16:00) hours with mode- and origin/transfer station-specific coefficients. Fairly similar fits and coefficient ratios were found for the different statistical measures of irregular deviations, indicating that (at least for this network) these measures can be used interchangeably (our final models use standard deviation). Similar results for total travel time have also been found previously. In the peak hour model, the coefficient for irregular deviations is in the direction opposite to expectations. We proposed that this is likely because we are not able to control for causality in this experiment and the fact that crowding and dispersion of waiting times are positively correlated. Thus, when the effect waiting time unreliability is already low (but still negative), our experiment would find that travellers tend to choose the more unreliable alternative (because their choices and unreliability are not independent). This effect is likely to be more prominent in the peak hours because of the larger demand then.

Our results find reliability ratios (ratios of the coefficients of irregular deviations to the coefficients of scheduled waiting times) in the range of 0.20–1.12. This is on the lower end of reliability ratios reported in literature, implying that our findings concur with recent empirical evidence that revealed preferences analyses find smaller effects of negative attributes (e.g., crowding, unreliability) than their stated preferences counterparts. The overall high reliability of the networks in The Hague could also be a reason why travellers are less averse to waiting time uncertainty. Furthermore, there is also evidence that travellers weight waiting for buses higher than that for trams, potentially indicating that travellers perceive waiting for trams to be less uncertain.

How to evaluate subjective beliefs regarding waiting time for route choices made under the ambiguity that is naturally present in the real-world?

To measure travellers' subjective beliefs regarding waiting time uncertainty, that is, not assuming that waiting time is a known risk, we devise an experiment that takes advantage of a commonly occurring choice situation. Travellers' decisions to either (i) board a public transport vehicle (towards the destination) that departs imminently or (ii) wait for a vehicle that reaches the destination earlier are observed and used to quantify their attitudes and perceptions of waiting time uncertainty in terms of a certainty equivalent. Moreover, the effects of situational context variables, such as elapsed waiting time or time-of-day, thereon can also be measured. The setup is suitable for stated choice experiments as it avoids learning effects or the need to present reliability information. Respondents are thus forced to draw from their own beliefs about the network in question when making decisions. And, as it occurs fairly commonly in the real-world, revealed choices (say, from smart card data) can also be used to perform a more natural experiment. We demonstrate both types of experiments in our case studies with the Dutch railways and the urban public transport networks of Amsterdam, respectively.

From the stated choice experiment with travellers of the Dutch railways, we find that accounting for subjective waiting time uncertainty in the choice model improves both the model fit and prediction performance. Moreover, the waiting to in-vehicle time ratio halves, indicating that uncertainty plays a large role in travellers' assessment of waiting time. On average, travellers are willing to trade-off more than 8 minutes of in-vehicle time to obtain certainty in waiting time. Heterogeneity is analysed through a latent class choice model wherein we find three types of behavioural classes—fully compensatory and in line with the average behaviour (55%), lexicographic preference for faster trains (28%), and very high dislike for uncertainty (17%). Effects of contextual variables (elapsed waiting time, delays) and inter-class differences in qualitative measures of regret propensity and system perception were minor.

We also outlined the data requirements for revealed preference analysis and used smart card data from Amsterdam's tram and bus networks for a case study. Since traveller arrivals at origin stops are not observed (which prevents us from knowing which feasible options were not chosen), transfer trips are used for the analysis (as we know the arrivals from the AVL data). These trips are further filtered to those that meet the requirements for the experiment. Estimation of multinomial logit models on these observations find results very similar to the stated preferences case study. Model fit is improved and including the certainty equivalent causes the waiting to in-vehicle time ratio to shrink. On average, we find that travellers in this system are willing to accept an extra 3.6 minutes of in-vehicle time to remove waiting time uncertainty.

What are the impacts of COVID-19 transmission risk determinants on public transport travellers' route choice behaviour?

In order to analyse public transport travellers' route behaviour during the COVID-19 pandemic we conducted a stated choice experiment with train travellers in the Netherlands at the end of the first infection wave there (20–25 May 2020). The experiment was distributed within an online survey which also collected socio-economic factors and COVID-19 related qualitative measures. The stated choice experiment focussed on three prominent risk criteria in public transport: (i) on-board crowding, (ii) exposure duration, and (iii) prevalent infection rate. In each scenario respondents ranked two train alternatives (described by their waiting time and on-board crowding) and an opt-out option in the context of a given exposure duration (i.e., in-vehicle time), prevalent infection rate, and their typical trip purpose.

We found that a 2-class latent class choice model best described the heterogeneity in behaviour. The two classes, which we call COVID Conscious and Infection Indifferent, indicate distinct behaviour somewhat at the extremes of the spectrum of behavioural response to COVID-19. While the first class has a high propensity of opting-out and its decisions are strongly driven by on-board crowding, infection rates, and the expected number of infected persons on-board; the latter behaves much more as if it is business-as-usual and is less likely to opt-out on average.

The COVID Conscious class is willing to wait on average 8.75 minutes to reduce one person on their carriage and shows the highest preference for sitting alone (i.e., with empty neighbouring seat). Note that this high value does not mean that travellers are willing to accept extremely long waits because the waiting time range tested in the experiment was only 3–25 minutes. On the other hand, the Infection Indifferent class is willing to wait only 1.04 minutes to reduce one person, slightly higher than pre-pandemic (stated preference) evaluations. In both classes, the propensity to opt-out increases with prevalent infection rate up to 2% and then plateaus (for context, the estimated infection rate at the peak of the first infection wave the estimated rate was ~0.95% (Rijksoverheid, 2021)). Although the increase is much more in the COVID Conscious class, at high crowding levels when opt-out rate is already high, it is fairly inelastic to infection rate. This indicates that travellers in this class may not feel safe in crowded trains even when infection rate goes down. The tested range of exposure duration did not have an effect on route choice behaviour.

We find that travellers that are older, female, and who had a lower train use during the pandemic were likely to be COVID Conscious. Posterior analysis of the classes also revealed that travellers likely to be in this group tended to be more worried about the pandemic, about being hospitalized, and spreading the infection. They also reported themselves to be more rule-following than others even though face mask use remained unpopular with both groups.

2 Implications for practice

Our research objective of modelling and analysing the impact of uncertainty on public transport travellers' route choice and our approach to this objective has resulted in a number of practical contributions that are relevant for public transport planners, operators, and policy-makers. First, we improve the performance of route choice models and provide updated estimates of choice behaviour. Second, we propose and demonstrate better experimentation methods that encourage data-driven decision-making and reduce reliance on expert and (post-hoc) traveller judgements. Finally, we provide actionable insights that public transport providers can employ for various objectives, from better system design to tailoring marketing campaigns.

Improved models

Demand estimation and assignment models are critical for planning and designing adequate levels of service in public transport system. Our research improves these models on two fronts. First, we explicitly account for the impact of waiting time uncertainty which is potentially linked with travel satisfaction and negative responses such as anxiety and stress. Both objective and subjective representations of uncertainty led to significantly better model performance (i.e., fit and predictive value). Moreover, the latter was estimated from a 'to board or not' choice situation that could be particularly interesting for agent-based simulations. The models also accounted for various situational contexts (e.g., experienced waiting time) and estimated travel time weights for different modes and origin/transfer stops, that planners may also consider for improving the accuracy of their forecasts.

Second, we estimate choice parameters under the new uncertainties presented by the COVID-19 pandemic. The pandemic has caused significant disruption to public transport travel and it is very likely that travellers now pay much more attention to factors that increase the risk of transmission. Our estimates, which include propensity-to-travel for different infection rates and updated crowding valuations, can help public transport providers accurately forecast demand and the required supply under these new circumstances. Existing models can also be calibrated by comparing similar coefficient ratios to our updated values. Considering travellers' revised trade-offs and planning for sufficient supply will be especially critical in light of the dwindling ridership levels.

Better experimentation

Different studies in this thesis, demonstrate the use of passively collected revealed preferences that have the least hypothetical bias. In particular, we develop a route choice set generation methodology that is necessary because often direct identification of choice sets is either not possible or not suitable. As the methodology requires minimal behavioural assumptions and can be directly calibrated with observed data, it empowers data-driven decision-making and reduces reliance on domain experts. Furthermore, by producing transferable results it enables choice analysis on new (sections of) networks where little data has been observed, thus, improving forecasts and consequently network and supply design. Finally, we note that while the methodology is demonstrated in this thesis on a large dataset of passively observed choices, it can also be used for smaller sets of active observations (e.g., Ton et al. (2020)). Thus, it can also be used to analyse preferences, such as station choices, that are not easily evident from passively collected data but can be known from smaller-scale surveys.

We also propose an experimentation method to quantify subjective beliefs regarding waiting time uncertainty. In addition to improving models as noted above, policy-makers can use this experiment to explicitly calculate the cost of being perceived as unreliable. Here too, the method encourages a shift away from qualitative, post-hoc judgements of travellers to a data-driven paradigm. Along these lines, we also propose that operators leverage the snapshots of uncertainty evaluations captured by this experiment as indicators of satisfaction and anxiety. Using the experiment then, they can analyse which situational or environmental variables cause higher uncertainty perception and assess the efficacy of various measures (from station lighting to provision of extra information) on anxiety reduction via randomized trials. As passively collected data can be used, operators do not have to make any changes other than the treatment (i.e., the applied measure) and can continuously monitor the indicator.

Actionable insights

The models and experimentation methods proposed in this thesis can (or have through the case studies) provide insights on which public transport providers can act. First, from the proposed route choice set generation model, we obtain parameters that give information about how the considered choice set is derived. Policy-makers can use this information to further understand how public transport can be made more attractive. Moreover, results from the non-compensatory model also add to insights from traditional choice models that are fully-compensatory by indicating which trade-offs travellers will not make at all. The derived attribute rankings and thresholds can also be used by journey planner applications (in markets with dense travel options) to provide mode and route recommendations.

Second, the analysis of the impact of waiting time uncertainty gives insights into how it manifests in choice models when it is not accounted for explicitly. When uncertainty is analysed as if it were objective, we find that perceptions that are not captured by the regular and irregular deviations are partially represented in other components such as the tram bonus for waiting time (higher waiting time tolerance for trams over buses). When we account for the subjectivity in uncertainty, we find (in two separate case studies) that the anticipated waiting time to in-vehicle time ratio falls by about half, indicating the large role of uncertainty in the value of waiting time. Policy-makers and planners can use these insights to adjust previously determined values of different travel time components. The quantification of subjective beliefs can also be used to (better) calculate the value of uncertainty for the prioritisation of investments aimed at reducing anxiety associated with public transport travel. Furthermore, as waiting time perceptions that are not aligned with reality can lead to sub-optimal decisions, journey planner applications can use the proposed quantification to highlight biases to travellers and assist them with making better decisions. This could be particularly important in mitigating cascading effects of disruptions.

Finally, we estimate latent classes of travel behaviour during the early stages of the COVID-19 pandemic. In line with anecdotal evidence, we find two distinct segments of traveller behaviour in response to the pandemic, one that is very conscious about the risks and another that is more indifferent. Given the large-scale disruption to public transport systems around the world, policy-makers now have the dual aims of stemming the mode shift trend (i.e., bringing travellers back to public transport) but also doing so safely. The insights into taste heterogeneity and differences in class profiles can be useful to tailor marketing campaigns and policies that balance these dual aims. While we also estimate travellers' intentions under prevalent infection rates, we do note that travel behaviour may evolve with the rapidly changing pandemic. Even so, the estimated behavioural parameters give policy-makers an insight into this behavioural evolution supporting them in becoming more prepared and proactive for the next pandemic.

3 Recommendations for future research

Next, we propose avenues for future research which are designed around research gaps that have arisen from the studies in this thesis. This research agenda serves not only as a natural extension to our research but also as standalone topics to further develop our understanding of route choice behaviour in public transport networks.

Stochastic and heterogeneous route choice set generation

In chapter 2, we develop a route choice set generation methodology that employs the elimination-by-aspects heuristic. The choice sets produced are deterministic (relative to the best possible attribute value amongst the feasible alternatives) and identical for all travellers (with the same origin and destination). However, it is unlikely that the same set of alternatives would have been considered by all travellers and even by the same traveller each time they make a choice. Therefore, we recommend formulating a stochastic version of the elimination-by-aspects heuristic that can account for this. For instance, this could be accomplished by including a random component to the threshold estimation for each attribute. Further heterogeneity can be accounted for by estimating discrete mixtures of ranking and threshold combinations that are used by different travellers. While this might increase computational effort, it could help improve the coverage for individual travellers.

Extending analysis of waiting time uncertainty

We propose extending the waiting time analysis in this thesis in four directions by analysing: (i) the gap between measurements of uncertainty evaluations from stated and revealed preferences; (ii) the development of such subjective beliefs, that is, learning behaviour, (iii) the link between travel satisfaction and waiting time uncertainty evaluations, and (iv) the impact of unplanned and significant disruptions on subjective beliefs.

First, analysing the gap between measured waiting time uncertainty evaluations from stated and revealed preferences. In chapter 4, we analysed stated choices from train travellers in the Netherlands and smart card observations from the public transport networks of Amsterdam. While we found similar results, it would be useful to understand the extent to which stated choice experiment results match those from natural experiments in the same network and under similar conditions. Given these results, operators or policy makers may choose to perform the experiment in a stated choices context if the required situation occurs rarely in their public transport network.

Second, chapter 4 focussed on a limited portion of the theoretical framework proposed therein for decision-making under uncertainty. Previous studies have analysed different components of the framework independently, typically using stated choices. We suggest using longitudinal smart card data in combination with the proposed experimental method for measuring waiting time beliefs in order to study components such as the impact of experienced uncertainty resolution and habit. A sufficiently large (in terms of number of travellers and time span) dataset could help us understand how travellers develop their evaluations of waiting time uncertainty. For example, perceived waiting time distributions may be fuzzier for irregular travellers (or they might be very dependent on displayed information). Using revealed preferences can help overcome some drawbacks of current analyses; specifically, requiring assumptions about initial perceptions and the absence of consequences. However, some aspects of decision-making under uncertainty may not be possible to study without stated choices, such as the effect of affective contexts.

Third, while a link between travel satisfaction and waiting time uncertainty is suggested in the thesis, we do not explicitly analyse this. This relationship could be explored with Likert scale measures of travel satisfaction (e.g., Soza-Parra et al. (2019)) and their impact on the certainty equivalent can be analysed through a structural equation model. The methodology proposed to assess travellers' evaluation of waiting time uncertainty can thus be used to analyse satisfaction. As we show in the case study in Amsterdam, smart card data can be used to continuously monitor subjective beliefs and consequently travel satisfaction.

Finally, we attempted to analyse the impact of contextual effects on waiting time uncertainty evaluations by analysing the relationship between experienced waiting time and (spontaneous and minor) delays. However, more disruptive events are likely to have a more significant impact on perceptions; for instance, a series of delays or a larger-than-normal crowd on the platform. Moreover, if these perceptions are not in line with reality, they may exacerbate the disruption or worsen levels of service (as more travellers attempt to board whichever vehicle arrives first). Therefore, we propose studying the impacts of such disruptions on subjective beliefs and carrying out a simulation analysis of its impact on cascading failures in public transport networks.

A retrospective assessment of travel behaviour in the pandemic

The travel behaviour scientific community has rapidly responded to the pandemic since March 2020. A plethora of studies related to COVID-19 perceptions, aggregate statistics, and to a lesser extent choice behaviour quickly followed. Most studies (including the one in this thesis) analysed a snapshot of travel behaviour during one or two phases of the pandemic and drew comparisons with the pre-pandemic situation. Few exceptions related to mode use and activity location patterns (e.g., Beck et al. (2021); Ton et al. (2022)) aside, longitudinal analysis of traveller behaviour through the crisis is missing. In particular, travellers' (heterogeneous) valuations of on-board crowding and other risk mitigating factors such as mask-mandates or vehicle cleaning are likely to have evolved throughout the pandemic. Therefore, we propose conducting a meta-analysis of choice experiments on this topic. Analysing these valuations may help public transport authorities develop strategies for a robust response to similar crises in the future.

References

Aaditya, B., Rahul, T.M., 2021. Psychological impacts of COVID-19 pandemic on the mode choice behaviour: A hybrid choice modelling approach. *Transport Policy* 108, 47-58.

Abenoza, R.F., Cats, O., Susilo, Y.O., 2018. How does travel satisfaction sum up? An exploratory analysis in decomposing the door-to-door experience for multimodal trips. *Transportation*.

Aghabayk, K., Esmailpour, J., Shiwakoti, N., 2021. Effects of COVID-19 on rail passengers' crowding perceptions. *Transportation Research Part A: Policy and Practice* 154, 186-202.

Alemazkoor, N., Burris, M.W., Danda, S.R., 2015. Using Empirical Data to Find the Best Measure of Travel Time Reliability. *Transp. Res. Record* 2530, 93-100.

Allen, J., Muñoz, J.C., Ortúzar, J.d.D., 2019. Understanding public transport satisfaction: Using Maslow's hierarchy of (transit) needs. *Transport Policy* 81, 75-94.

Anderson, E.C., Carleton, R.N., Diefenbach, M., Han, P.K.J., 2019. The Relationship Between Uncertainty and Affect. 10.

Ansari Esfeh, M., Wirasinghe, S.C., Saidi, S., Kattan, L., 2020. Waiting time and headway modelling for urban transit systems – a critical review and proposed approach. *Transport Reviews*, 1-23.

Atasoy, B., Glerum, A., Bierlaire, M., 2011. Mode choice with attitudinal latent class: a Swiss case-study, *Second International Choice Modeling Conference*.

Atasoy, B., Glerum, A., Bierlaire, M., 2013. Attitudes towards mode choice in Switzerland. *disP-The Planning Review* 49, 101-117.

Avineri, E., Prashker, J.N., 2005. Sensitivity to travel time variability: Travelers' learning perspective. *Transportation Research Part C: Emerging Technologies* 13, 157-183.

Avineri, E., Prashker, J.N., 2006. The Impact of Travel Time Information on Travelers' Learning under Uncertainty. *Transportation* 33, 393-408.

Awad-Núñez, S., Julio, R., Gomez, J., Moya-Gómez, B., González, J.S., 2021. Post-COVID-19 travel behaviour patterns: impact on the willingness to pay of users of public transport and shared mobility services in Spain. *European Transport Research Review* 13, 20.

Axhausen, K.W., Haupt, T., Fell, B., Heidl, U., 2001. Searching for the rail bonus: Results from a panel SP/RP study. *European Journal of Transport and Infrastructure Research* 1, 353-369.

Baillon, A., Huang, Z., Selim, A., Wakker, P.P., 2018. Measuring Ambiguity Attitudes for All (Natural) Events. *Econometrica* 86, 1839-1858.

Barron, G., Erev, I., 2003. Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making* 16, 215-233.

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

Beck, M.J., Hensher, D.A., 2020. Insights into the impact of COVID-19 on household travel and activities in Australia – The early days of easing restrictions. *Transport Policy* 99, 95-119.

Beck, M.J., Hensher, D.A., Nelson, J.D., 2021. Public transport trends in Australia during the COVID-19 pandemic: An investigation of the influence of bio-security concerns on trip behaviour. *Journal of Transport Geography* 96, 103167.
Bekhor, S., Prato, C.G., 2009. Methodological transferability in route choice modeling. *Transportation Research Part B: Methodological* 43, 422-437.
Ben-Akiva, M., Bergman, M., Daly, A.J., Ramaswamy, R., 1984. Modeling inter-urban route choice behaviour, *Proceedings of the 9th international symposium on transportation and traffic theory*. VNU Science Press Utrecht, The Netherlands, pp. 299-330.
Ben-Elia, E., Avineri, E., 2015. Response to Travel Information: A Behavioural Review. *Transport Reviews* 35, 352-377.
Ben-Elia, E., Di Pace, R., Bifulco, G.N., Shiftan, Y., 2013. The impact of travel information's accuracy on route-choice. *Transportation Research Part C: Emerging Technologies* 26, 146-159.
Ben-Elia, E., Erev, I., Shiftan, Y., 2008. The combined effect of information and experience on drivers' route-choice behavior. *Transportation* 35, 165-177.
Ben-Elia, E., Shiftan, Y., 2010. Which road do I take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice* 44, 249-264.
Bierlaire, M., 2016. PythonBiogeme: a short introduction, *Series on Biogeme*. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland.
Bierlaire, M., 2018. PandasBiogeme: a short introduction, *Series on Biogeme*. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland.
Bogers, E.A.I., Van Lint, H.W.C., Van Zuylen, H.J., 2008. Reliability of Travel Time:Effective Measures from a Behavioral Point of View. *Transp. Res. Record* 2082, 27-34.
Bogers, E.A.I., Viti, F., Hoogendoorn, S.P., 2005. Joint Modeling of Advanced Travel Information Service, Habit, and Learning Impacts on Route Choice by Laboratory Simulator Experiments. *Transp. Res. Record* 1926, 189-197.
Bogers, E.A.I., Viti, F., Hoogendoorn, S.P., Van Zuylen, H.J., 2006. Valuation of Different Types of Travel Time Reliability in Route Choice:Large-Scale Laboratory Experiment. *Transp. Res. Record* 1985, 162-170.
Bordagaray, M., dell'Olio, L., Ibeas, A., Cecín, P., 2014. Modelling user perception of bus transit quality considering user and service heterogeneity. *Transportmetrica A: Transport Science* 10, 705-721.
Börjesson, M., Eliasson, J., Franklin, J.P., 2012. Valuations of travel time variability in scheduling versus mean–variance models. *Transportation Research Part B: Methodological* 46, 855-873.
Bovy, P.H.L., 2009. On Modelling Route Choice Sets in Transportation Networks: A Synthesis. *Transport Reviews* 29, 43-68.
Bovy, P.H.L., Fiorenzo-Catalano, S., 2007. Stochastic route choice set generation: behavioral and probabilistic foundations. *Transportmetrica* 3, 173-189.
Bovy, P.H.L., Stern, E., 1994. *Route Choice: Wayfinding in Transport Networks: Wayfinding in Transport Networks*. Springer Science & Business Media.
Bridgman, A., Merkley, E., Loewen, P.J., Owen, T., Ruths, D., Teichmann, L., Zhilin, O., 2020. The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media. *The Harvard Kennedy School Misinformation Review* 1.
Brownstone, D., Small, K.A., 2005. Valuing time and reliability: assessing the evidence from road pricing demonstrations. *Transportation Research Part A: Policy and Practice* 39, 279-293.
Bucsky, P., 2020. Modal share changes due to COVID-19: The case of Budapest. *Transportation Research Interdisciplinary Perspectives* 8, 100141.
Buhr, K., Dugas, M.J., 2002. The intolerance of uncertainty scale: psychometric properties of the English version. *Behaviour Research and Therapy* 40, 931-945.
Bunschoten, T.M., 2012. To Tram or not to Tram. Delft University of Technology.
Cantwell, M., Caulfield, B., O'Mahony, M.J.J.o.p.t., 2009. Examining the factors that impact public transport commuting satisfaction. 12, 1.
Carleton, R.N., 2016. Into the unknown: A review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders* 39, 30-43.
Carrion, C., Levinson, D., 2012. Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice* 46, 720-741.
Carrion, C., Levinson, D., 2013. Valuation of travel time reliability from a GPS-based experimental design. *Transportation Research Part C: Emerging Technologies* 35, 305-323.
Cats, O., Gkioulou, Z., 2017. Modeling the impacts of public transport reliability and travel information on passengers' waiting-time uncertainty. *EURO Journal on Transportation and Logistics* 6, 247-270.

Centraal Bureau voor de Statistiek, 2015. Onderzoek Verplaatsingen in Nederland 2014.
Centraal Bureau voor de Statistiek, 2020. StatLine.
Cheng, Y.-H., Tsai, Y.-C., 2014. Train delay and perceived-wait time: passengers' perspective. *Transport Reviews* 34, 710-729.
Cho, S.-H., Park, H.-C., 2021. Exploring the Behaviour Change of Crowding Impedance on Public Transit due to COVID-19 Pandemic: Before and After Comparison. *Transportation Letters*, 1-8.
Chorus, C.G., Arentze, T.A., Timmermans, H.J.P., 2008. A Random Regret-Minimization model of travel choice. *Transportation Research Part B: Methodological* 42, 1-18.
Citymapper, Citymapper Mobility Index. Citymapper.com/CMI. Accessed on: 20 February 2021
de Haas, M., Faber, R., Hamersma, M., 2020. How COVID-19 and the Dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. *Transportation Research Interdisciplinary Perspectives* 6, 100150.
de Moraes Ramos, G., Daamen, W., Hoogendoorn, S., 2013. Modelling travellers' heterogeneous route choice behaviour as prospect maximizers. *Journal of Choice Modelling* 6, 17-33.
de Palma, A., Picard, N., 2005. Route choice decision under travel time uncertainty. *Transportation Research Part A: Policy and Practice* 39, 295-324.
Department for Transport, U.K., Coronavirus (COVID-19): safer travel guidance for passengers. https://www.gov.uk/guidance/coronavirus-covid-19-safer-travelguidance-for-passengers. Accessed on: 4 December 2020
Dixit, V., Jian, S., Hassan, A., Robson, E., 2019. Eliciting perceptions of travel time risk and exploring its impact on value of time. *Transport Policy* 82, 36-45.
Douglas, N., Karpouzis, G., 2006. Estimating the passenger cost of train overcrowding, *29th Australian Transport Research Forum*, Gold Coast, Queensland, Australia.
Drabicki, A., Kucharski, R., Cats, O., Szarata, A., 2020. Modelling the effects of real-time crowding information in urban public transport systems. *Transportmetrica A: Transport Science*, 1-39.
Dryhurst, S., Schneider, C.R., Kerr, J., Freeman, A.L.J., Recchia, G., van der Bles, A.M., Spiegelhalter, D., van der Linden, S., 2020. Risk perceptions of COVID-19 around the world. *Journal of Risk Research*, 1-13.
DutchNews.nl, Despite the government u-turn, the Dutch are still unwilling to wear masks. https://www.dutchnews.nl/news/2020/10/despite-the-government-u-turn-the-dutch-are-still-unwilling-to-wear-masks/. Accessed on: 20 February 2021
Dziekan, K., Vermeulen, A., 2006. Psychological Effects of and Design Preferences for Real-Time Information Displays. *Journal of Public Transportation* 9, 1-19.
Elias, W., Albert, G., Shiftan, Y., 2013. Travel behavior in the face of surface transportation terror threats. *Transport Policy* 28, 114-122.
Ellsberg, D., 1961. Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics* 75, 643-669.
Engle, S., Stromme, J., Zhou, A., 2020. Staying at Home: Mobility Effects of COVID-19. https://dx.doi.org/10.2139/ssrn.3565703.
Fan, W., Machemehl, R.B., 2009. Do Transit Users Just Wait for Buses or Wait with Strategies?:Some Numerical Results That Transit Planners Should See. 2111, 169-176.
Gerhold, L., 2020. COVID-19: Risk Perception and Coping Strategies. 10.31234/osf.io/xmpk4.
Ghosh, A., 2001. Valuing time and reliability: commuters' mode choice from a real time congestion pricing experiment. University of California Irvine.
Gilboa, I., 2009. *Theory of decision under uncertainty*. Cambridge University Press.
Gilboa, I., 2012. *Rational choice*. MIT press.
Gkiotsalitis, K., Cats, O., 2020. Public transport planning adaption under the COVID-19 pandemic crisis: literature review of research needs and directions. *Transport Reviews*, 1-19.
Goldbaum, C., 2020. Is the Subway Risky? It May Be Safer Than You Think, *The New York Times*. The New York Times.
Google LLC, Google COVID-19 Community Mobility Reports. https://www.google.com/covid19/mobility/. Accessed on: 20 February 2021
Gordon, J.B., Koutsopoulos, H.N., Wilson, N.H.M., Attanucci, J.P., 2013. Automated Inference of Linked Transit Journeys in London Using Fare-Transaction and Vehicle Location Data. *Transp. Res. Record* 2343, 17-24.
Guo, Z., Wilson, N.H.M., 2011. Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. *Transportation Research Part A: Policy and Practice* 45, 91-104.
Haghani, M., Bliemer, M.C.J., Rose, J.M., Oppewal, H., Lancsar, E., 2021a. Hypothetical bias in stated choice experiments: Part I. Macro-scale analysis of literature and integrative synthesis of empirical evidence from applied economics, experimental psychology and neuroimaging. *Journal of Choice Modelling* 41, 100309.

Haghani, M., Bliemer, M.C.J., Rose, J.M., Oppewal, H., Lancsar, E., 2021b. Hypothetical bias in stated choice experiments: Part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods. *Journal of Choice Modelling* 41, 100322.

Han, A.F., 1987. Assessment of transfer penalty to bus riders in Taipei: a disaggregate demand modeling approach. *Transp. Res. Record* 1139, 8-14.

Hänseler, F.S., Bierlaire, M., Scarinci, R., 2016. Assessing the usage and level-of-service of pedestrian facilities in train stations: A Swiss case study. *Transportation Research Part A: Policy and Practice* 89, 106-123.

Hauser, J.R., 2014. Consideration-set heuristics. *Journal of Business Research* 67, 1688-1699.

Hawkins, D.I., Mothersbaugh, D.L., 2010. *Consumer behavior: Building marketing strategy*. Boston: McGraw-Hill Irwin.

Hensher, D.A., 2001. Measurement of the Valuation of Travel Time Savings. *Journal of Transport Economics and Policy (JTEP)* 35, 71-98.

Hensher, D.A., 2010. Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B: Methodological* 44, 735-752.

Hensher, D.A., Balbontin, C., Beck, M.J., Wei, E., 2022. The impact of working from home on modal commuting choice response during COVID-19: Implications for two metropolitan areas in Australia. *Transportation Research Part A: Policy and Practice* 155, 179-201.

Hensher, D.A., Rose, J.M., Collins, A.T., 2011. Identifying commuter preferences for existing modes and a proposed Metro in Sydney, Australia with special reference to crowding. *Public Transport* 3, 109-147.

Hensher, D.A., Rose, J.M., Greene, W.H., 2015. *Applied Choice Analysis*, 2nd ed. Cambridge University Press, Cambridge.

Hertwig, R., Erev, I., 2009. The description–experience gap in risky choice. *Trends in Cognitive Sciences* 13, 517-523.

Holguín-Veras, J., Paaswell, R.E., Yali, A.-M., 2003. Impacts of extreme events on intercity passenger travel behavior: the September 11th Experience. 373-404.

Hoogendoorn-Lanser, S., 2005. Modelling travel behaviour in multi-modal networks. Delft University of Technology, Delft, The Netherlands.

Hoogendoorn-Lanser, S., Bovy, P., van Nes, R., 2007. Application of Constrained Enumeration Approach to Multimodal Choice Set Generation. *Transp. Res. Record* 2014, 50-57.

Hoogendoorn-Lanser, S., van Nes, R., Bovy, P., 2005. Path Size Modeling in Multimodal Route Choice Analysis. *Transp. Res. Record* 1921, 27-34.

Hörcher, D., Graham, D.J., Anderson, R.J., 2017. Crowding cost estimation with large scale smart card and vehicle location data. *Transportation Research Part B: Methodological* 95, 105-125.

Hsu, Y.-T., Kang, L., Wu, Y.-H., 2016. User Behavior of Bikesharing Systems Under Demand–Supply Imbalance. 2587, 117-124.

HTM, Timetable. https://www.htm.nl/en/timetable/. Accessed on: 30 May 2021

Hu, M., Lin, H., Wang, J., Xu, C., Tatem, A.J., Meng, B., Zhang, X., Liu, Y., Wang, P., Wu, G., Xie, H., Lai, S., 2020. Risk of Coronavirus Disease 2019 Transmission in Train Passengers: an Epidemiological and Modeling Study. *Clinical Infectious Diseases*.

Hurtubia, R., Nguyen, M.H., Glerum, A., Bierlaire, M., 2014. Integrating psychometric indicators in latent class choice models. *Transportation Research Part A: Policy and Practice* 64, 135-146.

Indian Railways: Western Railway, Mumbai Suburban Time Table. https://wr.indianrailways.gov.in/view_section.jsp?id=0,6,458. Accessed on: 30 May 2021

Jánošíková, Ľ., Slavík, J., Koháni, M., 2014. Estimation of a route choice model for urban public transport using smart card data. *Transportation Planning and Technology* 37, 638-648.

Jones, P., 1996. Managing perceptions of waiting times in service queues. *International Journal of Service Industry Management* 7, 47-61.

Kemel, E., Paraschiv, C., 2013. Prospect Theory for joint time and money consequences in risk and ambiguity. *Transportation Research Part B: Methodological* 56, 81-95.

Kim, C., Cheon, S.H., Choi, K., Joh, C.-H., Lee, H.-J., 2017a. Exposure to fear: Changes in travel behavior during MERS outbreak in Seoul. *KSCE Journal of Civil Engineering* 21, 2888-2895.

Kim, I., Kim, H.-C., Seo, D.-J., Kim, J.I., 2019. Calibration of a transit route choice model using revealed population data of smartcard in a multimodal transit network. *Transportation*.

Kim, J., Corcoran, J., Papamanolis, M., 2017b. Route choice stickiness of public transport passengers: Measuring habitual bus ridership behaviour using smart card data. *Transportation Research Part C: Emerging Technologies* 83, 146-164.

Kluwe-Schiavon, B., Viola, T.W., Bandinelli, L.P., Castro, S.C.C., Kristensen, C.H., Costa da Costa, J., Grassi-Oliveira, R., 2021. A behavioral economic risk aversion experiment in the context of the COVID-19 pandemic. *PLOS ONE* 16, e0245261.

Knight, F.H., 1921. *Risk, uncertainty and profit*. Houghton Mifflin Company, Boston, New York.

Kroes, E., Kouwenhoven, M., Debrincat, L., Pauget, N., 2013. On the Value of Crowding in Public Transport for Île-de-France, *OECD ITF Meeting on Valuing Convenience in Public Transport*, Paris.

Kroes, E., Kouwenhoven, M., Debrincat, L., Pauget, N., 2014. Value of Crowding on Public Transport in île-de-France, France. 2417, 37-45.

Kugler, T., Connolly, T., Ordóñez, L.D., 2012. Emotion, Decision, and Risk: Betting on Gambles versus Betting on People. 25, 123-134.

Kurauchi, F., Schmöcker, J.-D., Shimamoto, H., Hassan, S.M., 2014. Variability of commuters' bus line choice: an analysis of oyster card data. *Public Transport* 6, 21-34.

Lam, T.C., Small, K.A., 2001. The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review* 37, 231-251.

Lau, J.T.F., Yang, X., Tsui, H., Kim, J.H., 2003. Monitoring community responses to the SARS epidemic in Hong Kong: from day 10 to day 62. 57, 864-870.

Leahy, C., Batley, R., Chen, H., 2016. Toward an automated methodology for the valuation of reliability. *Journal of Intelligent Transportation Systems* 20, 334-344.

Levinson, D., Harder, K., Bloomfield, J., Winiarczyk, K., 2004. Weighting Waiting: Evaluating Perception of In-Vehicle Travel Time Under Moving and Stopped Conditions. *Transp. Res. Record* 1898, 61-68.

Li, Z., Hensher, D., 2011a. Prospect Theoretic Contributions in Understanding Traveller Behaviour: A Review and Some Comments. *Transport Reviews* 31, 97-115.

Li, Z., Hensher, D., 2013. Crowding in public transport: a review of objective and subjective measures. *Journal of Public Transportation* 16, 6.

Li, Z., Hensher, D., 2019. Understanding risky choice behaviour with travel time variability: a review of recent empirical contributions of alternative behavioural theories. *Transportation Letters*, 1-11.

Li, Z., Hensher, D.A., 2011b. Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal. *Transport Policy* 18, 880-887.

Li, Z., Hensher, D.A., Rose, J.M., 2010. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E: Logistics and Transportation Review* 46, 384-403.

Loewenstein, G., Lerner, J.S., 2003. The role of affect in decision making.

Luo, D., Bonnetain, L., Cats, O., van Lint, H., 2018. Constructing Spatiotemporal Load Profiles of Transit Vehicles with Multiple Data Sources. *Transp. Res. Record* 2672, 175-186.

Machina, M.J., Siniscalchi, M., 2014. Chapter 13 - Ambiguity and Ambiguity Aversion, in: Machina, M., Viscusi, K. (Eds.), *Handbook of the Economics of Risk and Uncertainty*. North-Holland, pp. 729-807.

Maister, D., 1985. The psychology of waiting lines, in: Czepiel, J.A., Solomon, M., Surprenant, C.S. (Eds.), *The Service Encounter*. Lexington Books.

Meng, M., Rau, A., Mahardhika, H., 2018. Public transport travel time perception: Effects of socioeconomic characteristics, trip characteristics and facility usage. *Transportation Research Part A: Policy and Practice* 114, 24-37.

Molin, E., Adjenughwure, K., de Bruyn, M., Cats, O., Warffemius, P., 2020. Does conducting activities while traveling reduce the value of time? Evidence from a within-subjects choice experiment. *Transportation Research Part A: Policy and Practice* 132, 18-29.

MTA New York City Transit, Schedules (Subway Timetable, Effective September 13, 2020). https://new.mta.info/schedules. Accessed on: 30 May 2021

Nassir, N., Hickman, M., Ma, Z.-L., 2018. A strategy-based recursive path choice model for public transit smart card data. *Transportation Research Part B: Methodological*.

Nau, R.F., 2001. De Finetti was Right: Probability Does Not Exist. *Theory and Decision* 51, 89-124.

Nederlandse Spoorwegen, Travelling in Crowded Trains (in Dutch). https://www.ns.nl/reisinformatie/service-verbeteren/wat-kunt-je-doen-bij-drukte.html. Accessed on: 17 March 2021

Openbaar Vervoer Nederland, Over het coronavirus (in Dutch). https://www.ov-nl.nl/over-het-coronavirus/. Accessed on: 4 December 2020

Openbaar Vervoer Nederland, Welkom terug; OV-bedrijven staan klaar om reizigers te ontvangen (in Dutch). https://www.ov-nl.nl/welkom-terug-ov-bedrijven-staan-klaar-om-reizigers-te-ontvangen/. Accessed on: 4 December 2020

Osuna, E.E., 1985. The psychological cost of waiting. *Journal of Mathematical Psychology* 29, 82-105.

Peer, S., Knockaert, J., Koster, P., Verhoef, E.T., 2014. Over-reporting versus overreacting: Commuters' perceptions of travel times, *Tinbergen Institute Discussion Paper*. Tinbergen Institute.

Prather, K.A., Wang, C.C., Schooley, R.T., 2020. Reducing transmission of SARS-CoV-2. *Science* 368, 1422-1424.

Prato, C., Bekhor, S., 2007. Modeling Route Choice Behavior: How Relevant Is the Composition of Choice Set? *Transportation Research Record: Journal of the Transportation Research Board* 2003, 64-73.

Prato, C.G., 2009. Route choice modeling: past, present and future research directions. *Journal of Choice Modelling* 2, 65-100.
Prato, C.G., Bekhor, S., 2006. Applying Branch-and-Bound Technique to Route Choice Set Generation. *Transp. Res. Record* 1985, 19-28.
Preston, J., Pritchard, J., Waterson, B., 2017. Train Overcrowding:Investigation of the Provision of Better Information to Mitigate the Issues. 2649, 1-8.
Puto, C., Huber, J., Payne, J.W., 1982. Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research* 9, 90-98.
Quiggin, J., 1982. A theory of anticipated utility. *Journal of Economic Behavior & Organization* 3, 323-343.
Ramos, G.d.M., Daamen, W., Hoogendoorn, S., 2011. Expected Utility Theory, Prospect Theory, and Regret Theory Compared for Prediction of Route Choice Behavior. *Transportation Research Record: Journal of the Transportation Research Board* 2230, 19-28.
Rasouli, S., Timmermans, H., 2014. Applications of theories and models of choice and decision-making under conditions of uncertainty in travel behavior research. *Travel Behaviour and Society* 1, 79-90.
Raveau, S., 2017. Modelling Travel Choices on Public Transport Systems with Smart Card Data, in: Kurauchi, F., Schmocker, J.-D. (Eds.), *Public Transport Planning with Smart Card Data*. CRC Press, Boca Raton.
Raveau, S., Muñoz, J.C., de Grange, L., 2011. A topological route choice model for metro. *Transportation Research Part A: Policy and Practice* 45, 138-147.
Rijksoverheid, Corona-aanpak: de volgende stap. https://www.rijksoverheid.nl/actueel/nieuws/2020/05/19/corona-aanpak-de-volgende-stap. Accessed on: 20 February 2021
Rijksoverheid, Number of Infectious Persons (in Dutch). https://coronadashboard.rijksoverheid.nl/landelijk/besmettelijke-mensen. Accessed on: 20 February 2021
Rubin, G.J., Amlôt, R., Page, L., Wessely, S., 2009. Public perceptions, anxiety, and behaviour change in relation to the swine flu outbreak: cross sectional telephone survey. 339, b2651.
Sahu, P.K., Sharma, G., Guharoy, A., 2018. Commuter travel cost estimation at different levels of crowding in a suburban rail system: a case study of Mumbai. *Public Transport* 10, 379-398.
Schive, K., How safe is public transportation? https://medical.mit.edu/covid-19-updates/2020/09/how-safe-public-transportation. Accessed on: 9 March 2020
Schmöcker, J.-D., Shimamoto, H., Kurauchi, F., 2013. Generation and Calibration of Transit Hyperpaths. *Procedia - Social and Behavioral Sciences* 80, 211-230.
Schwanen, T., Ettema, D., 2009. Coping with unreliable transportation when collecting children: Examining parents' behavior with cumulative prospect theory. *Transportation Research Part A: Policy and Practice* 43, 511-525.
Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., Lehman, D.R., 2002. Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology* 83, 1178-1197.
Scorrano, M., Danielis, R., 2021. Active mobility in an Italian city: Mode choice determinants and attitudes before and during the Covid-19 emergency. *Research in Transportation Economics*, 101031.
Shelat, S., Cats, O., 2017. Measuring spill-over effects of disruptions in public transport networks, *Models and Technologies for Intelligent Transportation Systems (MT-ITS), 5th IEEE International Conference on*. IEEE, pp. 756-761.
Shelat, S., van de Wiel, T., Molin, E., van Lint, J.W.C., Cats, O., 2022. Analysing the impact of COVID-19 risk perceptions on route choice behaviour in train networks. *PLOS ONE* 17, e0264805.
Smail, E., Schneider, K.E., DeLong, S.M., Willis, K., Arrington-Sanders, R., Yang, C., Alexander, K.A., Johnson, R.M., 2021. Health Beliefs and Preventive Behaviors Among Adults During the Early COVID-19 Pandemic in the United States: a Latent Class Analysis. *Prevention Science* 22, 1013-1022.
Small, K.A., Noland, R., Chu, X., Lewis, D., 1999. Valuation of travel-time savings and predictibility in congested conditions for highway user-cost estimation. National Cooperative Highway Research Program.
Small, K.A., Winston, C., Yan, J., 2005. Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability. *Econometrica* 73, 1367-1382.
Soman, D., 2001. The mental accounting of sunk time costs: why time is not like money. *Journal of Behavioral Decision Making* 14, 169-185.
Soza-Parra, J., Raveau, S., Muñoz, J.C., Cats, O., 2019. The underlying effect of public transport reliability on users' satisfaction. *Transportation Research Part A: Policy and Practice* 126, 83-93.
Swierstra, A.B., van Nes, R., Molin, E.J.E., 2017. Modelling travel time reliability in public transport route choice behaviour. *European Journal of Transport and Infrastructure Research* 17.

Teixeira, J.F., Lopes, M., 2020. The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York's Citi Bike. *Transportation Research Interdisciplinary Perspectives* 6, 100166.

Thaler, R., 1980. Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization* 1, 39-60.

Tilahun, N.Y., Levinson, D.M., 2010. A Moment of Time: Reliability in Route Choice Using Stated Preference. *Journal of Intelligent Transportation Systems* 14, 179-187.

Tirachini, A., Cats, O., 2020. COVID-19 and Public Transportation: Current Assessment, Prospects, and Research Needs. *Journal of Public Transportation* 22.

Tirachini, A., Hensher, D.A., Rose, J.M., 2013. Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. *Transportation Research Part A: Policy and Practice* 53, 36-52.

Ton, D., Arendsen, K., de Bruyn, M., Severens, V., van Hagen, M., van Oort, N., Duives, D., 2022. Teleworking during COVID-19 in the Netherlands: Understanding behaviour, attitudes, and future intentions of train travellers. *Transportation Research Part A: Policy and Practice* 159, 55-73.

Ton, D., Duives, D., Cats, O., Hoogendoorn, S., 2018. Evaluating a data-driven approach for choice set identification using GPS bicycle route choice data from Amsterdam. *Travel Behaviour and Society* 13, 105-117.

Ton, D., Shelat, S., Nijënstein, S., Rijsman, L., van Oort, N., Hoogendoorn, S., 2020. Understanding the Role of Cycling to Urban Transit Stations through a Simultaneous Access Mode and Station Choice Model. *Transp. Res. Record* 0, 0361198120925076.

Trépanier, M., Tranchant, N., Chapleau, R., 2007. Individual Trip Destination Estimation in a Transit Smart Card Automated Fare Collection System. *Journal of Intelligent Transportation Systems* 11, 1-14.

Tversky, A., 1972. Elimination by aspects : A theory of choice. *Psychological Review* 79, 281–299.

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

UITP, 2020. Public Transport is COVID-Safe.

Union Home Secretary, 2020. Guidelines on the measures to be taken by Ministries/Departments of Government of India, State/Union Territory Governments and Authorities for containment of COVID-19 Epidemic in the Country (24.03.2020), in: Affairs, M.o.H. (Ed.). Ministry of Home Affairs - Government of India, India.

van Hagen, M., 2011. Waiting experience at train stations. University of Twente.

van Lint, J.W.C., van Zuylen, H.J., Tu, H., 2008. Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transportation Research Part A: Policy and Practice* 42, 258-277.

van Oort, N., 2011. Service reliability and urban public transport design, *Transport and Planning*. Delft University of Technology.

van Oort, N., 2016. Incorporating enhanced service reliability of public transport in cost-benefit analyses. *Public Transport* 8, 143-160.

Van Oort, N., Brands, T., de Romph, E., 2015a. Short term ridership prediction in public transport by processing smart card data. *Transp. Res. Record*.

van Oort, N., Brands, T., de Romph, E., Flores, J.A., 2015b. Unreliability effects in public transport modelling. *International journal of transportation* 3, 113-130.

Wakker, P.P., 2010. *Prospect theory: For risk and ambiguity*. Cambridge university press.

Walker, J.L., Wang, Y., Thorhauge, M., Ben-Akiva, M., 2018. D-efficient or deficient? A robustness analysis of stated choice experimental designs. *Theory and Decision* 84, 215-238.

Wang, K.-Y., 2014. How Change of Public Transportation Usage Reveals Fear of the SARS Virus in a City. *PLoS ONE* 9.

Wardman, M., Whelan, G., 2011. Twenty Years of Rail Crowding Valuation Studies: Evidence and Lessons from British Experience. *Transport Reviews* 31, 379-398.

Watkins, K.E., Ferris, B., Borning, A., Rutherford, G.S., Layton, D., 2011. Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. *Transportation Research Part A: Policy and Practice* 45, 839-848.

Weber, E.U., Milliman, R.A., 1997. Perceived Risk Attitudes: Relating Risk Perception to Risky Choice. *Management Science* 43, 123-144.

Winkler, R.L., Muñoz, J., Cervera, J.L., Bernardo, J.M., Blattenberger, G., Kadane, J.B., Lindley, D.V., Murphy, A.H., Oliver, R.M., Ríos-Insua, D., 1996. Scoring rules and the evaluation of probabilities. *Test* 5, 1-60.

Yap, M., Cats, O., van Arem, B., 2018. Crowding valuation in urban tram and bus transportation based on smart card data. *Transportmetrica A: Transport Science*, 1-20.

Yap, M.D., Cats, O., van Oort, N., Hoogendoorn, S.P., 2017. A robust transfer inference algorithm for public transport journeys during disruptions. *Transportation Research Procedia* 27, 1042-1049.

Summary

Research motivation
Arguably, nearly all real-world decisions are made under uncertainty. Most of these decisions, including travel choices, are inherently associated with subjective uncertainty where decision-makers' personal beliefs regarding the likelihood of events play a significant role. Such uncertainty is largely associated with negative affective responses and has been strongly linked with anxiety and stress. This has also been demonstrated for travel behaviour in public transport systems where several authors have noted that the two main sources of uncertainty—reliability and safety—constitute the most basic of needs. Amongst sources of uncertainty, waiting time is perhaps the single most important aspect of travelling with public transport; it is inherently frustrating and its inevitability in public transport and travellers' apparent lack of control can induce stress and dissatisfaction. More recently, safety-related uncertainty brought about by the outbreak of the COVID-19 pandemic has induced anxiety and public transport avoidance. It is critical that the impact of such pervasive uncertainty is analysed in order to achieve the overarching aim of making public transport a viable and satisfying alternative to motorized individual modes.

Research objective, scope, and questions
Given this ultimate aim, the objective of this thesis is to *model and analyse the impact of pervasive uncertainty on public transport travellers' route choice behaviour*. Because of their significant impact, we study the impact of uncertainty related to waiting time, focussing on research gaps arising from the source of behavioural data and the type of uncertainty analysed. Furthermore, we also assess recent changes in public transport travel behaviour due to the COVID-19 pandemic. To achieve its objective, the thesis aims to answer the following four research questions (RQ) (as also shown in the thesis structure in Figure 1):

RQ1 How to infer route choice sets from passively observed choices using minimal assumptions and producing transferable behavioural insights? (chapter 2)
RQ2 What is the impact of waiting time uncertainty, expressed as different statistical representations of its historical values, on route choice behaviour? (chapter 3)

RQ3 How to evaluate subjective beliefs regarding waiting time for route choices made under the ambiguity that is naturally present in the real-world? (chapter 4)

RQ4 What are the impacts of COVID-19 transmission risk determinants on public transport travellers' route choice behaviour? (chapter 5)

Research approach

In order to address research gaps related to the source of behavioural data, we use both stated and revealed preferences. The former have been collected from (online) stated choice experiments with travellers in the Dutch railways, while the latter have been obtained from passively collected smart card observations from the public transport networks of The Hague and Amsterdam. We contribute to the next set of research gaps by modelling route choice behaviour under different assumptions of waiting time uncertainty. First, we make the conventional assumption that travellers are aware of objective distributions of waiting times and include it as such in our model. Then, we move closer to modelling the real-world by accounting for the subjective nature of uncertainty. Finally, research gaps related to the new uncertainties brought about by the COVID-19 pandemic are addressed by analysing the impact of key transmission risk determinants. While the case studies presented in the thesis use observations from travellers in the Netherlands, the methodologies and, to a limited extent, findings are applicable to developed public transport networks around the world.

In the first chapter, we propose a choice set generation methodology (using smart card data) to answer the first research question. For the remaining questions, choice experiments are carried out; underpinned, as is conventional in transportation modelling, by the random utility maximization paradigm. In particular, multinomial logit models are employed, which assume Gumbel distributed random components, and where panel information is available, taste heterogeneity is analysed with latent class choice models. As shown in Figure 1, the data for choice analyses is obtained from (i) dedicated stated choice experiments for subjective waiting time uncertainty (chapter 4, first case study) and COVID-19 risk determinants (chapter 5) and (ii) smart card data for objective (chapter 3) and subjective waiting time uncertainty (chapter 4, second case study).

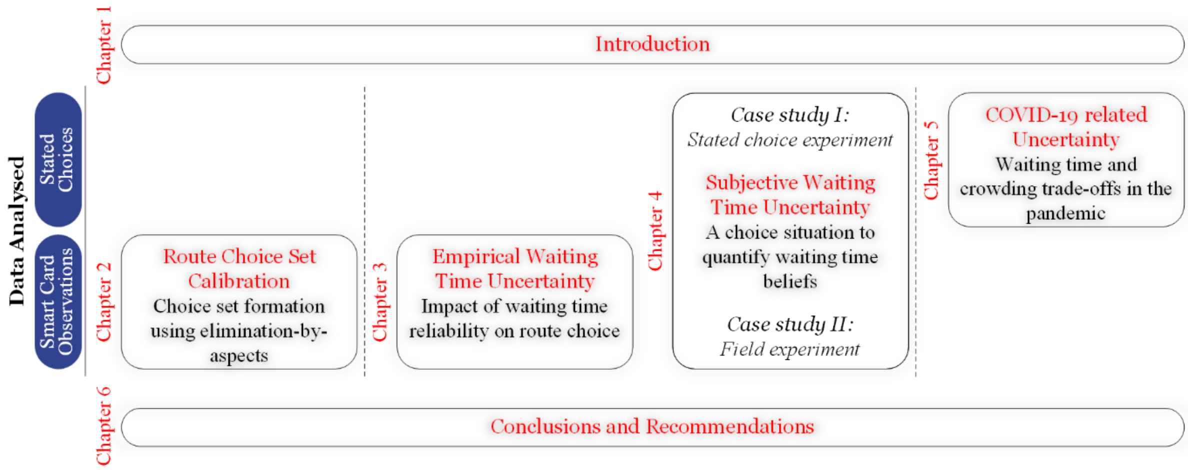


Figure 1: Thesis outline

Scientific contributions

By addressing the research gaps and answering the research questions above, this thesis makes the following scientific contributions:

Development of an assumption-parsimonious and transferable route choice generation methodology for public transport networks (chapter 2): A route choice set generation methodology is developed where consideration set formation is modelled as an elimination-by-aspects process and the parameters of this model are calibrated with passively observed choices from smart card data. Elimination-by-aspects is a non-compensatory heuristic which combines attribute ranking and setting thresholds. It is well-aligned to the actual cognitive process and requires minimal behavioural assumptions while resulting in actionable and transferable insights. The calibration involves an optimization procedure balancing demand-weighted coverage of observed routes and efficiency of generated routes.

Analysis of the impact of waiting time uncertainty on route choices in public transport networks using revealed preferences (chapters 3, 4): In two case studies, we use revealed preferences to analyse route choice behaviour under (objective and subjective) waiting time uncertainty, overcoming hypothetical bias related shortcomings of previous stated preference analyses on the impact of travel time uncertainty. Travellers' choices are obtained from smart card data and combined with vehicle location and schedule data to enable choice analyses.

Suitability comparison of different statistical representations of waiting time uncertainty in route choice models for public transport networks (chapter 3): The impact of waiting time uncertainty is modelled as if it were objective—that is, travellers were aware of historical values and used them for decision-making. Waiting time is included in a choice model by splitting it into scheduled values, regular deviations (median difference between realised and scheduled values), and irregular deviations (dispersion of realised values). For the latter, we tested different statistical range and buffer time representations and their suitability in representing travellers' perception of uncertainty is assessed by comparing model fits and the face validity of other coefficient ratios.

Development of an experimentation method to observe and analyse public transport route choice behaviour under natural waiting time ambiguity (chapter 4): An experimentation method involving a realistic route choice situation is developed to enable quantification of travellers' subjective beliefs regarding waiting time uncertainty and the impact of situational contexts thereon in terms of a certainty equivalent. The choice situation occurs fairly commonly in public transport networks, allowing us to generalize findings beyond the scope of the specific situation. The method lends itself to both controlled and natural route choice experiments as is demonstrated in two case studies.

Analysis of the (heterogeneity in the) impact of COVID-19 transmission risk determinants on public transport route choice behaviour (chapter 5): A choice analysis is performed to examine the trade-offs between on-board crowding, exposure duration, prevalent infection rate, and travel time attributes, with the aim of deriving latent classes of travellers and comparing their crowding valuations and propensities to avoid public transport. Furthermore, class profiles composed of socio-demographics and pandemic-related attitudes and opinions are derived.

Main empirical findings

Next, we discuss the case studies in this thesis and their main empirical findings:

Route choice set calibration (chapter 2): Application of the proposed route choice set generation methodology to the urban public transport network of The Hague found that travellers ranked number of transfers as the most important attribute for consideration set formation, followed by waiting and in-vehicle time. Travellers do not seem to include alternatives with extra transfers or with travel times worse by more than 10% in their consideration set. The overall demand-weighted coverage was 64% but analysis of the coverage and efficiency trends indicates that this can potentially be improved by modelling closer and farther origin-destination pairs separately.

Empirical waiting time uncertainty (chapter 3): We used smart card data from The Hague to model the impact of waiting time uncertainty as an objective risk. Different statistical representations of irregular deviations performed fairly similarly indicating that (at least for this network) they can be used interchangeably. Relatively low reliability ratios (coefficients of irregular deviations divided by the coefficients of scheduled waiting times)—in the range of 0.20–1.12—are found. Furthermore, there is also evidence that travellers weight waiting for buses higher than that for trams, potentially indicating that travellers perceive waiting for trams to be less uncertain.

Subjective waiting time uncertainty (chapter 4): We perform two case studies with our proposed experimentation method. The first uses stated choice data from a survey of travellers in the Dutch railways and the second uses smart card data from the urban public transport network of Amsterdam. In both case studies, we find that explicitly accounting for subjective waiting time uncertainty through a certainty equivalent improves the model fit and predictive performance. Moreover, it causes the waiting to in-vehicle time coefficient ratio to shrink, indicating that uncertainty plays a large role in travellers' assessment of waiting time.

On average, railway travellers are willing to trade-off more than 8 minutes of in-vehicle time to obtain certainty in waiting time. Latent class choice analysis finds three types of behavioural classes in these travellers—fully compensatory and in line with the average behaviour (55%), lexicographic preference for faster trains (28%), and very high dislike for uncertainty (17%). Tram and bus travellers in Amsterdam are willing to accept an extra 3.6 minutes of in-vehicle time to remove waiting time uncertainty.

COVID-19 related uncertainty (chapter 5): A stated choice experiment was distributed to travellers in the Dutch railways at the end of the first infection wave in the Netherlands. We find that the response to COVID-19 risk determinants is best described by two latent classes of behaviour: COVID Conscious and Infection Indifferent. The two classes are willing to wait, on average, 8.75 minutes (with a strong preference to sit alone) and 1.04 minutes (only slightly higher than pre-pandemic stated preference evaluations), respectively, to reduce one person on their carriage. While the COVID Conscious class is more strongly affected by the prevalent infection rate, at high crowding levels when opt-out rate is already high, it is fairly inelastic to infection rate. This indicates that travellers in this class may not feel safe in crowded trains even when infection rates are lower. The class membership model shows that women, older travellers, and those who had a lower train use during the pandemic were more likely to be COVID Conscious.

Practical implications

The thesis makes several contributions relevant for public transport planners, operators, and policy-maker. First, we produce improved models of public transport route choice behaviour by explicitly including waiting time uncertainty (e.g., through irregular deviations, certainty equivalent) in choice models, estimating updated crowded valuations (in light of the pandemic) and the impact of prevalent infection rates, and reducing hypothetical bias by using revealed preferences. These suggested attributes and relevant coefficient ratios can be incorporated into existing models or analyses (e.g., by calibrating against common coefficient ratios) for better forecasting and, consequently, improved system design. Second, our proposals for route choice set generation and capturing subjective beliefs should lead to better experimentation. The former enables experiments with revealed preferences (with minimal behavioural assumptions and when direct identification is not possible), leading to behavioural analysis with smaller hypothetical bias. The latter captures travellers' true perceptions of waiting time uncertainty; and the snapshots of certainty equivalents captured by this method can be used to analyse which situational or environmental variables cause higher uncertainty perception. Third, we provide directly actionable insights, such as: using outputs from the choice set generation methodology (attribute rankings and thresholds) to provide travellers with better route recommendations; making use of descriptive analysis of travellers' waiting time perceptions for prescriptive purposes; or applying knowledge of travellers' response to COVID-19 risk determinants to re-plan supply, tailor marketing campaigns, and support anticipatory preparation for similar future events.

Recommendations for future research

Based on the research gaps arising from the studies in this thesis a research agenda is sketched containing the following topics:

Stochastic and heterogeneous route choice set generation: Formulate a stochastic version of the elimination-by-aspects heuristics and use discrete mixtures of ranking and threshold combinations to account for intra- and inter-traveller heterogeneity in consideration set formation. This will bring the choice set generation methodology further in line with actual behaviour and improve coverage for individual travellers.

Extending analysis of waiting time uncertainty: Analyse (i) the gap between measurements of uncertainty evaluations from stated and revealed preferences; (ii) the development of such subjective beliefs, that is, learning behaviour; (iii) the link between travel satisfaction and waiting time uncertainty evaluations; and (iv) the impact of unplanned and significant disruptions on subjective beliefs. These analyses have the overarching objectives of improving our understanding of travel behaviour under uncertainty and applying that to improve public transport services.

A retrospective assessment of travel behaviour in the pandemic: Conduct a meta-analysis of travellers' (heterogeneous) valuation of (perceived) risk determinants and mitigating factors, such as on-board crowding and vehicle cleaning through the COVID-19 pandemic to develop strategies for a robust response to similar crises in the future.

Samenvatting

Onderzoeksmotivatie

Het is redelijk om te stellen dat bijna alle beslissingen in de echte wereld worden genomen onder onzekerheid. De meeste van deze beslissingen, waaronder reiskeuzes, zijn inherent verbonden met subjectieve onzekerheid waarbij de persoonlijke overtuigingen van besluitvormers over de waarschijnlijkheid van gebeurtenissen een belangrijke rol spelen. Deze onzekerheid gaat grotendeels gepaard met negatieve affectieve reacties en is sterk verbonden met angst en stress. Dit is ook aangetoond voor reisgedrag in het openbaar vervoer, waar verschillende auteurs hebben opgemerkt dat de twee belangrijkste bronnen van onzekerheid - betrouwbaarheid en veiligheid - de meest fundamentele behoeften vormen. Onder de bronnen van onzekerheid is wachttijd misschien wel het belangrijkste aspect van reizen met het openbaar vervoer; het is inherent frustrerend en de onvermijdelijkheid ervan in het openbaar vervoer en het schijnbare gebrek aan controle van reizigers kunnen stress en ontevredenheid veroorzaken. Meer recentelijk heeft de onzekerheid over de veiligheid als gevolg van de uitbraak van de COVID-19-pandemie geleid tot angst en vermijding van het openbaar vervoer. Het is essentieel dat de impact van dergelijke alomtegenwoordige onzekerheid wordt geanalyseerd om het overkoepelende doel te bereiken om het openbaar vervoer een levensvatbaar en bevredigend alternatief te maken voor gemotoriseerde individuele vervoerswijzen.

Onderzoeksdoel, reikwijdte en vragen

Gezien dit uiteindelijke doel, is het doel van dit proefschrift *om de impact van alomtegenwoordige onzekerheid op het routekeuzegedrag van reizigers in het openbaar vervoer te modelleren en te analyseren*. Vanwege hun significante impact bestuderen we de invloed van onzekerheid met betrekking tot wachttijd, waarbij we ons richten op onderzoeksgaten die voortkomen uit de bron van gedragsgegevens en het type onzekerheid dat wordt geanalyseerd. Bovendien beoordelen we ook recente veranderingen in het reisgedrag met het openbaar vervoer als gevolg van de COVID-19-pandemie. Om dit doel te bereiken, heeft de scriptie tot doel de volgende vier onderzoeksvragen te beantwoorden (V) (zoals ook weergegeven in de structuur van het proefschrift in Figure 1):

V1 Hoe kunnen routekeuzesets worden afgeleid uit passief waargenomen keuzes met minimale aannames en het produceren van overdraagbare gedragsinzichten? (hoofdstuk 2)

V2 Wat is de impact van wachttijd onzekerheid, uitgedrukt in verschillende statistische representaties van de historische waarden, op het routekeuzegedrag? (hoofdstuk 3)

V3 Hoe evalueer men subjectieve overtuigingen over wachttijd voor routekeuzes gemaakt onder de ambiguïteit die van nature aanwezig is in de echte wereld? (hoofdstuk 4)

V4 Wat zijn de effecten van COVID-19 transmissierisicodeterminanten op het routekeuzegedrag van reizigers in het openbaar vervoer? (hoofdstuk 5)

Onderzoeksbenadering

Om onderzoeksgaten met betrekking tot de bron van gedragsgegevens aan te pakken, maken we gebruik van zowel gestelde voorkeuren als onthulde voorkeuren. De gestelde voorkeuren zijn verzameld via (online) gestelde keuze-experimenten met reizigers in de Nederlandse spoorwegen, terwijl de onthulde voorkeuren zijn verkregen uit passief verzamelde smartcard-waarnemingen van het openbaar vervoer in Den Haag en Amsterdam. We dragen bij aan het volgende set van onderzoeksgaten door het modelleren van routekeuzegedrag onder verschillende aannames van wachttijd onzekerheid. Allereerst maken we de conventionele aanname dat reizigers op de hoogte zijn van objectieve verdelingen van wachttijden en nemen we dit op die manier op in ons model. Vervolgens komen we dichter bij het modelleren van de echte wereld door rekening te houden met de subjectieve aard van onzekerheid. Ten slotte worden onderzoeksgaten met betrekking tot de nieuwe onzekerheden die voortkomen uit de COVID-19-pandemie aangepakt door de impact van belangrijke determinanten van het transmissierisico te analyseren. Hoewel de casestudy's in het proefschrift gebruikmaken van waarnemingen van reizigers in Nederland, zijn de methodologieën en, in beperkte mate, bevindingen toepasbaar op ontwikkelde openbaarvervoersnetwerken over de hele wereld.

In het eerste hoofdstuk stellen we een methodologie voor het genereren van keuzeset voor (met behulp van smartcard-gegevens) om de eerste onderzoeksvraag te beantwoorden. Voor de overige vragen worden keuze-experimenten uitgevoerd die, zoals gebruikelijk is in de vervoersmodellering, gebaseerd zijn op het paradigma van de willekeurige nutsmaximalisatie. In het bijzonder worden multinomiale logit-modellen toegepast, die uitgaan van willekeurige componenten met een Gumbel-verdeling. Als er panelinformatie beschikbaar is, wordt smaakheterogeniteit geanalyseerd met latente klasse-keuzemodellen. Zoals weergegeven in Figuur 1, worden de gegevens voor de keuzeanalyses verkregen uit (i) specifieke gestelde keuze-experimenten voor subjectieve wachttijd onzekerheid (hoofdstuk 4, eerste casestudy) en COVID-19-risicodeterminanten (hoofdstuk 5) en (ii) smartcard-gegevens voor objectieve (hoofdstuk 3) en subjectieve wachttijd onzekerheid (hoofdstuk 4, tweede casestudy).

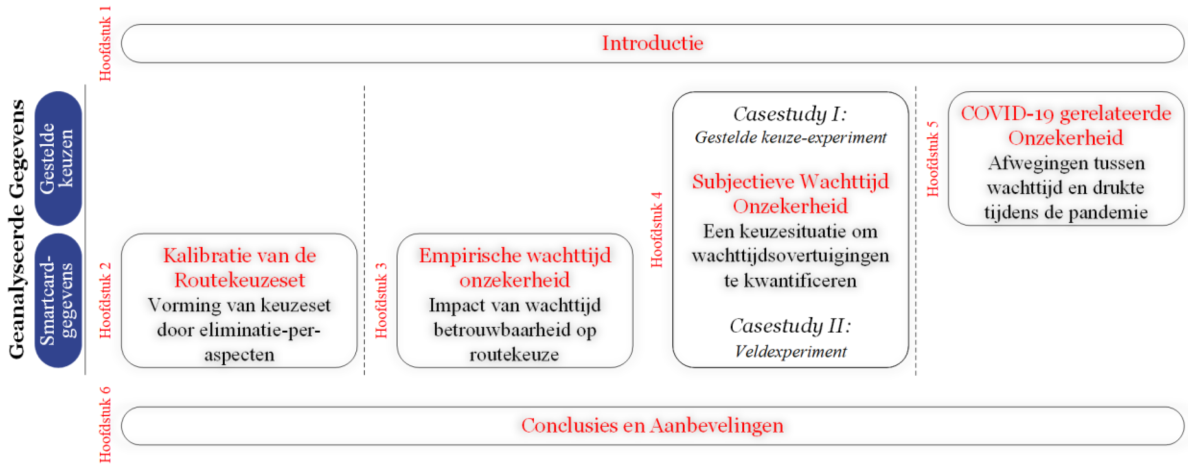


Figure 1: Proefschrift-overzicht

Wetenschappelijke bijdragen

Door de onderzoeksgaten aan te pakken en de bovenstaande onderzoeksvragen te beantwoorden, levert deze proefschrift de volgende wetenschappelijke bijdragen:

Ontwikkeling van een aannamearme en overdraagbare methodologie voor het genereren van routekeuzes in het openbaar vervoer (hoofdstuk 2): Een methodologie voor het genereren van routekeuzesets wordt ontwikkeld, waarbij het vormen van de overwegingsset wordt gemodelleerd als een eliminatie-per-aspecten proces en de parameters van dit model worden gekalibreerd met behulp van passief waargenomen keuzes uit smartcard-gegevens. Eliminatie-per-aspecten is een niet-compensatoire heuristiek die attribuutrangschikking combineert met het stellen van drempels. Het sluit goed aan bij het daadwerkelijke cognitieve proces en vereist minimale gedragsaannames, terwijl het resulteert in bruikbare en overdraagbare inzichten. De kalibratie omvat een optimalisatieprocedure waarbij een balans wordt gezocht tussen de vraaggewogen dekking van waargenomen routes en de efficiëntie van gegenereerde routes.

Analyse van de impact van wachttijd onzekerheid op routekeuzes in het openbaar vervoer met onthulde voorkeuren (hoofdstukken 3, 4): In twee casestudies gebruiken we onthulde voorkeuren om het routekeuzegedrag te analyseren onder (objectieve en subjectieve) wachttijd onzekerheid, waarbij we tekortkomingen van eerdere gestelde voorkeur analyses met betrekking tot hypothetische vertekening over de impact van reistijd onzekerheid overwinnen. De keuzes van reizigers worden verkregen uit smartcard-gegevens en gecombineerd met voertuiglocatiegegevens en dienstregelingsgegevens om keuzeanalyses mogelijk te maken.

Geschiktheidsvergelijking van verschillende statistische representaties van wachttijd onzekerheid in routekeuzemodellen voor het openbaar vervoer (hoofdstuk 3): De impact van wachttijd onzekerheid wordt gemodelleerd alsof het objectief is—dat wil zeggen dat reizigers de historische waarden kennen en gebruiken voor hun besluitvorming. Wachttijd wordt opgenomen in een keuzemodel door deze op te splitsen in geplande waarden, reguliere afwijkingen (mediaan verschil tussen gerealiseerde en geplande waarden) en onregelmatige afwijkingen (dispersie van gerealiseerde waarden). Voor deze laatste hebben wij verschillende statistische bereik- en buffertijdrepresentaties getest en hun geschiktheid om de perceptie van onzekerheid door reizigers weer te geven is beoordeeld door modelpassingen en de gezichtsvaliditeit van andere coëfficiëntverhoudingen te vergelijken.

Ontwikkeling van een experimentele methode voor het observeren en analyseren van routekeuzegedrag in het openbaar vervoer bij natuurlijke wachttijd ambiguïteit (hoofdstuk 4): Een experimentele methode met een realistische routekeuzesituatie wordt ontwikkeld om de subjectieve overtuigingen van reizigers over wachttijd onzekerheid en de impact van situationele contexten daarop in termen van een zekerheidsequivalent te kwantificeren. De keuzesituatie komt vrij vaak voor in openbaar vervoersnetwerken, waardoor we bevindingen kunnen generaliseren buiten de specifieke situatie. De methode is geschikt voor zowel gecontroleerde als natuurlijke routekeuze-experimenten, zoals gedemonstreerd wordt in twee casestudies.

Analyse van de (heterogeniteit in de) impact van COVID-19-transmissierisicodeterminanten op het routekeuzegedrag in het openbaar vervoer (hoofdstuk 5): Er wordt een keuzeanalyse uitgevoerd om de afwegingen tussen drukte aan boord, blootstellingsduur, prevalente besmettingsgraad en reistijdattributen te onderzoeken, met als doel latente klassen van reizigers af te leiden en hun druktewaarderingen en neiging om het openbaar vervoer te vermijden te vergelijken. Daarnaast worden klasseprofielen afgeleid die bestaan uit sociodemografische gegevens en pandemiegerelateerde attitudes en meningen.

Belangrijkste empirische bevindingen

Vervolgens bespreken we de casestudies in dit proefschrift en hun belangrijkste empirische bevindingen:

Kalibratie van de routekeuzeset (hoofdstuk 2): Toepassing van de voorgestelde methodologie voor het genereren van routekeuzesets op het stedelijke openbaar vervoersnetwerk van Den Haag heeft aangetoond dat reizigers het aantal overstappen als het belangrijkste attribuut beschouwen bij het vormen van hun overwegingset, gevolgd door wachttijd en reistijd in het voertuig. Reizigers lijken geen alternatieven op te nemen met extra overstappen of met reistijden die meer dan 10% langer zijn in hun overwegingset. De algehele vraag-gewogen dekking bedroeg 64%, maar de analyse van de trends in de dekking en efficiëntie geeft aan dat dit mogelijk verbeterd kan worden door het modelleren van dichterbij gelegen en verder weg gelegen herkomst-bestemmingsparen apart.

Empirische wachttijd onzekerheid (hoofdstuk 3): We hebben smartcard-gegevens uit Den Haag gebruikt om de impact van wachttijd onzekerheid als een objectief risico te modelleren. Verschillende statistische representaties van onregelmatige afwijkingen presteerden redelijk vergelijkbaar, wat aangeeft dat ze (ten minste voor dit netwerk) onderling uitwisselbaar kunnen worden gebruikt. We hebben relatief lage betrouwbaarheidsratio's (coëfficiënten van onregelmatige afwijkingen gedeeld door de coëfficiënten van geplande wachttijden) gevonden, variërend van 0,20 tot 1,12. Bovendien is er ook bewijs dat reizigers wachten op bussen hoger waarderen dan wachten op trams, wat mogelijk aangeeft dat reizigers het wachten op trams als minder onzeker ervaren.

Subjectieve wachttijd onzekerheid (hoofdstuk 4): We voeren twee casestudies uit met onze voorgestelde experimentele methode. De eerste gebruikt gestelde keuzes uit een enquête onder reizigers in de Nederlandse spoorwegen, en de tweede gebruikt smartcard-gegevens uit het stedelijke openbaar vervoersnetwerk van Amsterdam. In beide casestudies vinden we dat expliciete verwerking van subjectieve wachttijd onzekerheid via een zekerheidsequivalent de model-fit en voorspellende prestaties verbetert. Bovendien zorgt dit ervoor dat de verhouding tussen de coëfficiënten van wachten en voertuig reistijd kleiner wordt, wat aangeeft dat onzekerheid een grote rol speelt in de beoordeling van wachttijd door reizigers.

Gemiddeld zijn treinreizigers bereid meer dan 8 minuten reistijd in het voertuig in te ruilen voor zekerheid over de wachttijd. Een latente klasse-keuzeanalyse identificeert drie soorten gedragsklassen onder deze reizigers—volledig compenserend en in lijn met het gemiddelde gedrag (55%), lexicografische voorkeur voor snellere treinen (28%) en zeer sterke afkeer van onzekerheid (17%). Tram- en busreizigers in Amsterdam zijn bereid om een extra 3,6 minuten reistijd in het voertuig te accepteren om de onzekerheid over de wachttijd weg te nemen.

COVID-19-gerelateerde onzekerheid (hoofdstuk 5): Aan het einde van de eerste besmettingsgolf in Nederland werd een gestelde keuze-experiment verspreid onder reizigers in de Nederlandse. We vinden dat de reactie op COVID-19-risicodeterminanten het best wordt beschreven door twee latente gedragsklassen: COVID Bewust en Infectie Onverschillig. De twee klassen zijn bereid om gemiddeld respectievelijk 8,75 minuten (met een sterke voorkeur om alleen te zitten) en 1,04 minuten (slechts iets hoger dan pre-pandemische gestelde voorkeursevaluaties) te wachten om één persoon in hun rijtuig te verminderen. Hoewel de COVID Bewuste klasse sterker wordt beïnvloed door de prevalente besmettingsgraad, is deze klasse bij hoge drukte, wanneer het afmeldpercentage al hoog is, redelijk inelastisch ten opzichte van de besmettingsgraad. Dit geeft aan dat reizigers in deze klasse zich mogelijk niet veilig voelen in drukke treinen, zelfs wanneer de besmettingsgraad lager is. Het model voor klasse-lidmaatschap toont aan dat vrouwen, oudere reizigers en degenen die minder met de trein hebben gereisd tijdens de pandemie meer geneigd waren om COVID Bewust te zijn.

Praktische implicaties

Dit proefschrift heeft verschillende bijdragen die relevant zijn voor openbaar-vervoersplanners, exploitanten en beleidsmakers. Ten eerste produceren wij verbeterde modellen van openbaar vervoer route keuzegedrag door expliciet wachttijd onzekerheid (bv. door onregelmatige afwijkingen, zekerheid equivalent) in keuzemodellen op te nemen, het schatten van geactualiseerde drukte waarderingen (in het licht van de pandemie) en de impact van prevalente besmettingsgraad, en het verminderen van hypothetische vertekening door het gebruik van onthulde voorkeuren. Deze voorgestelde attributen en relevante coëfficiëntratio's kunnen worden geïntegreerd in bestaande modellen of analyses (bv. door af te stemmen op gemeenschappelijke coëfficiëntratio's) voor een betere prognose en bijgevolg een verbeterd systeemontwerp. Ten tweede zouden onze voorstellen voor het genereren van keuzesets en het vastleggen van subjectieve overtuigingen moeten leiden tot beter experimenteel onderzoek. Het eerste maakt experimenten met onthulde voorkeuren mogelijk (met minimale gedragsaannames en wanneer directe identificatie niet mogelijk is), wat leidt tot gedragsanalyse met minder hypothetische vertekening. Het laatste legt de werkelijke percepties van reizigers over wachttijd onzekerheid vast; de momentopnamen van zekerheidsequivalenten die met deze methode worden vastgelegd, kunnen worden gebruikt om te analyseren welke situationele of omgevingsvariabelen een hogere perceptie van onzekerheid veroorzaken. Ten derde bieden we direct bruikbare inzichten, zoals: gebruik van de resultaten van de methodologie voor het genereren van de keuzeset (attribuutrangschikkingen en drempels) om reizigers betere routeaanbevelingen te geven; gebruik van de beschrijvende analyse van de wachttijdperceptie van reizigers voor prescriptieve doeleinden, of toepassing van de kennis van de reactie van reizigers op COVID-19-risicodeterminanten om het aanbod opnieuw te plannen, marketingcampagnes op maat te maken en de anticiperende voorbereiding op soortgelijke toekomstige gebeurtenissen te ondersteunen.

Aanbevelingen voor toekomstig onderzoek

Op basis van de onderzoeksgaten die voortkomen uit de studies in dit proefschrift, wordt een onderzoeksagenda geschetst met de volgende onderwerpen:

Stochastische en heterogene generatie van keuzesets: Formuleer een stochastische versie van de eliminatie-per-aspecten heuristiek en gebruik discrete mengsels van rangschikkings- en drempelcombinaties om rekening te houden met intra- en inter-reizigers heterogeniteit in het vormen van de overwegingset. Hierdoor komt de methodologie voor het genereren van keuzesets beter overeen met het werkelijke gedrag en verbetert de dekking voor individuele reizigers.

Uitbreiding van de analyse van wachttijd onzekerheid: Analyseer (i) het verschil tussen metingen van onzekerheidsevaluaties uitgesteld en onthulde voorkeuren; (ii) de ontwikkeling van dergelijke subjectieve overtuigingen, oftewel leerprocessen; (iii) de relatie tussen reis tevredenheid en evaluaties van wachttijd onzekerheid; en (iv) de impact van ongeplande en significante verstoringen op subjectieve overtuigingen. Deze analyses hebben als overkoepelende doelstellingen het verbeteren van ons begrip van reisgedrag onder onzekerheid en het toepassen daarvan om het openbaar vervoer te verbeteren.

Een retrospectieve beoordeling van reisgedrag tijdens de pandemie: Voer een meta-analyse uit van de (heterogene) waardering van reizigers voor (waargenomen) risicodeterminanten factoren en mitigatiepunten, zoals drukte in het voertuig en voertuigreiniging, gedurende de COVID-19-pandemie om strategieën te ontwikkelen voor een robuuste respons op vergelijkbare crises in de toekomst.

About the Author

Sanmay Shelat (1993) was born in Vadodara, India. He began his bachelor studies in Civil Engineering in 2011 at Nirma University, Ahmedabad. In 2014, he spent two months as a research fellow at the Indian Institute of Technology Bombay where he researched surrogate measures for traffic safety. After obtaining his bachelor's degree (gold medallist) in 2015, he received an excellence scholarship from the Delft University of Technology to join their Transport and Planning master's degree programme. He graduated from the programme with an honours supplement and *cum laude* in 2017. Published works during his master studies include research on the robustness of public transport networks and on the combination of bicycle and transit. For his graduation project he worked at Philips Lighting, Eindhoven where he developed an activity-based model for pedestrians that leveraged Markov chains.

Eager to continue his academic path, Sanmay began his PhD studies immediately after defending his master thesis in late 2017 at the same department. His research combined discrete choice analysis and decision theory to model traveller behaviour under uncertainty. During his tenure at Delft, he also led a major component of the EU-H2020 project funding his PhD, coordinating as well as contributing to scientific and product deliverables. He also supervised several master theses and projects, organized an international summer school, and served as a reviewer for various journals and conferences. Since early 2022, Sanmay has been working in Hamburg as a Data Scientist for FREENOW's ride-hailing marketplace and pricing team. Here he has developed machine learning products and designed experiments for various aspects of the two-sided marketplace.

In his spare time, Sanmay enjoys urban hiking and cycling, mystery/thriller books, travelling whenever possible, and, more recently, running. He is passionate about the role of mobility in cities and is a strong advocate for public transport and non-motorized modes.

List of Publications

Journal articles

- **Shelat, S**., Huisman, R., van Oort, N. Analysing the trip and user characteristics of the combined bicycle and transit mode. Research in Transportation Economics (2018).
- Mantouka, E. G., Vlahogianni, E. I., Papacharalampous, A. E., Heydenrijk-Ottens, L., **Shelat, S.**, Degeler, V., & van Lint, H. Understanding travel behavior through travel happiness. Transportation Research Record (2019).
- Ton, D., **Shelat, S**., Nijënstein, S., Rijsman, L., van Oort, N., Hoogendoorn, S. Understanding the role of cycling to urban transit stations through a simultaneous access mode and station choice model. Transportation Research Record (2020).
- **Shelat, S.**, Cats, O., van Lint, J.W.C. Quantifying travellers' evaluation of waiting time uncertainty in public transport networks. Travel Behaviour and Society (2021).
- **Shelat, S.**, Cats, O., van Oort, N., van Lint, J.W.C. Evaluating the impact of waiting time reliability on route choice using smart card data. Transportmetrica A: Transport Science (2022).
- **Shelat, S.**, van de Wiel, T., Molin E., van Lint, J.W.C., Cats, O. Analysing the impact of COVID-19 risk perceptions on route choice behaviour in train networks. PLoS ONE (2022).
- **Shelat, S.**, Cats, O., van Cranenburgh, S. Traveller Behaviour in Public Transport in the Early Stages of the COVID-19 Pandemic in the Netherlands. Transportation Research Part A: Policy and Practice (2022).

Conference proceedings

- **Shelat, S.**, Daamen, W., Kaag, B., Duives, D.C., Hoogendoorn, S.P. A Markov-chain activity-based model for pedestrians in office buildings. Collective Dynamics: Proceedings of Pedestrian and Evacuation Dynamics. Lund, Sweden (2018).
- **Shelat, S.**, Cats, O., van Oort, N., van Lint, J.W.C. Calibrating route choice sets for an urban public transport network using smart card data. 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). Krakow, Poland (2019).

Conference presentations

- Mantouka, E. G., Vlahogianni, E. I., Papacharalampous, A. E., Heydenrijk-Ottens, L., **Shelat, S.**, Degeler, V., & van Lint, H. Understanding travel behavior through travel happiness. 99th Transportation Research Board (TRB). Washington DC, USA (2019).
- **Shelat, S.**, Cats, O., van Lint, J.W.C. Quantifying subjective beliefs regarding waiting time uncertainty in public transport networks. 24th International Conference of Hong Kong Society for Transportation Studies (HKSTS). Hong Kong (2019).
- Ton, D., **Shelat, S**., Nijënstein, S., Rijsman, L., van Oort, N., Hoogendoorn, S. The role of cycling towards urban transit stations: Simultaneously modelling the access mode and station choice. 100th Transportation Research Board (TRB). Washington DC, USA (2020).
- **Shelat, S.**, van Oort, N., Mantouka, E., Monteiro, M., Stroumpou, I., Cats, O., van Lint, J.W.C. Subjective beliefs regarding waiting times in public transport networks in the Netherlands, Greece, and Portugal. 48th European Transport Conference (ETC). Milan, Italy (online) (2020).
- **Shelat, S.**, Cats, O., van Oort, N., van Lint, J.W.C. Analysing the role of waiting time reliability in public transport route choice using smart card data. 101st Transportation Research Board (TRB). Washington DC, USA (online) (2021).
- **Shelat, S.**, Dixit, M., Cats, O., van Oort, N., van Lint, J.W.C. What does smart card data reveal about subjective beliefs regarding waiting time uncertainty? 8th International Symposium on Transport Network Reliability (INSTR). Stockholm, Sweden (online) (2021).
- **Shelat, S.**, Cats, O., van Cranenburgh, S. How do COVID-19 related risk criteria affect route choice behaviour in public transport? 102nd Transportation Research Board (TRB). Washington DC, USA (online) (2022).
- **Shelat, S.**, van de Wiel, T., Molin E., van Lint, J.W.C., Cats, O. Perceived COVID-19 risk and route choice behaviour in public transport: A hierarchical information integration analysis. 102nd Transportation Research Board (TRB). Washington DC, USA (online) (2022).
- Spierenburg, L., **Shelat, S.**, van Lint, J.W.C., van Oort, N. Job accessibility improvements with the bicycle-transit combination: Where to implement bike-and-ride facilities in Amsterdam? 102nd Transportation Research Board (TRB). Washington DC, USA (online) (2022).
- Romero, C., Cats, O., **Shelat, S.**, Estimating a seat choice model for Dutch suburban train users. 10th Symposium of the European Association for Research in Transportation (hEART). Leuven, Belgium (2022).
- **Shelat, S.**, Cats, O., van Cranenburgh, S. Crowding valuation in public transportation during the COVID-19 pandemic. 15th International Conference on Advanced Systems in Public Transport (CASPT). Tel-Aviv, Israel (2022).
- Hoobroeckx, T., Cats, O., Molin, E.J., **Shelat, S.** The influence of dynamic pricing on vehicle choice in a shared moped e-scooter system. 103rd Transportation Research Board (TRB). Washington DC, USA (2023).

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 275 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Shelat, S., Route Choice Behaviour under Uncertainty in Public Transport Networks: Stated and revealed preference analyses, T2023/8, June 2023, TRAIL Thesis Series, the Netherlands

Zhang, Y., Flexible, Dynamic, and Collaborative Synchromodal Transport Planning Considering Preferences, T2023/7, June 2023, TRAIL Thesis Series, the Netherlands

Kapetanović, M., Improving Environmental Sustainability of Regional Railway Services, T2023/6, June 2023, TRAIL Thesis Series, the Netherlands

Li, G., Uncertainty Quantification and Predictability Analysis for Traffic Forecasting at Multiple Scales, T2023/5, April 2023, TRAIL Thesis Series, the Netherlands

Harter, C., Vulnerability through Vertical Collaboration in Transportation: A complex networks approach, T2023/4, March 2023, TRAIL Thesis Series, the Netherlands

Razmi Rad, S., Design and Evaluation of Dedicated Lanes for Connected and Automated Vehicles, T2023/3, March 2023, TRAIL Thesis Series, the Netherlands

Eikenbroek, O., Variations in Urban Traffic, T2023/2, February 2023, TRAIL Thesis Series, the Netherlands

Wang, S., Modeling Urban Automated Mobility on-Demand Systems: an Agent-Based Approach, T2023/1, January 2023, TRAIL Thesis Series, the Netherlands

Szép, T., Identifying Moral Antecedents of Decision-Making in Discrete Choice Models, T2022/18, December 2022, TRAIL Thesis Series, the Netherlands

Zhou, Y., Ship Behavior in Ports and Waterways: An empirical perspective, T2022/17, December 2022, TRAIL Thesis Series, the Netherlands

Yan, Y., Wear Behaviour of A Convex Pattern Surface for Bulk Handling Equipment, T2022/16, December 2022, TRAIL Thesis Series, the Netherlands

Giudici, A., Cooperation, Reliability, and Matching in Inland Freight Transport, T2022/15, December 2022, TRAIL Thesis Series, the Netherlands

Nadi Najafabadi, A., Data-Driven Modelling of Routing and Scheduling in Freight Transport, T2022/14, October 2022, TRAIL Thesis Series, the Netherlands

Heuvel, J. van den, Mind Your Passenger! The passenger capacity of platforms at railway stations in the Netherlands, T2022/13, October 2022, TRAIL Thesis Series, the Netherlands

Haas, M. de, Longitudinal Studies in Travel Behaviour Research, T2022/12, October 2022, TRAIL Thesis Series, the Netherlands

Dixit, M., Transit Performance Assessment and Route Choice Modelling Using Smart Card Data, T2022/11, October 2022, TRAIL Thesis Series, the Netherlands

Du, Z., Cooperative Control of Autonomous Multi-Vessel Systems for Floating Object Manipulation, T2022/10, September 2022, TRAIL Thesis Series, the Netherlands

Larsen, R.B., Real-time Co-planning in Synchromodal Transport Networks using Model Predictive Control, T2022/9, September 2022, TRAIL Thesis Series, the Netherlands

Zeinaly, Y., Model-based Control of Large-scale Baggage Handling Systems: Leveraging the theory of linear positive systems for robust scalable control design, T2022/8, June 2022, TRAIL Thesis Series, the Netherlands

Fahim, P.B.M., The Future of Ports in the Physical Internet, T2022/7, May 2022, TRAIL Thesis Series, the Netherlands

Huang, B., Assessing Reference Dependence in Travel Choice Behaviour, T2022/6, May 2022, TRAIL Thesis Series, the Netherlands

Reggiani, G., A Multiscale View on Bikeability of Urban Networks, T2022/5, May 2022, TRAIL Thesis Series, the Netherlands

Paul, J., Online Grocery Operations in Omni-channel Retailing: opportunities and challenges, T2022/4, March 2022, TRAIL Thesis Series, the Netherlands

Liu, M., Cooperative Urban Driving Strategies at Signalized Intersections, T2022/3, January 2022, TRAIL Thesis Series, the Netherlands

Feng, Y., Pedestrian Wayfinding and Evacuation in Virtual Reality, T2022/2, January 2022, TRAIL Thesis Series, the Netherlands

Scheepmaker, G.M., Energy-efficient Train Timetabling, T2022/1, January 2022, TRAIL Thesis Series, the Netherlands

Bhoopalam, A., Truck Platooning: planning and behaviour, T2021/32, December 2021, TRAIL Thesis Series, the Netherlands

Hartleb, J., Public Transport and Passengers: optimization models that consider travel demand, T2021/31, TRAIL Thesis Series, the Netherlands

Invitation

You are cordially invited to attend the public defence of my PhD dissertation entitled:

Route Choice Behaviour under Uncertainty in Public Transport Networks

The defence will be held on Thursday 29 June 2023 at 10.00h in the Senaatszaal of the Aula Congress Centrum, Mekelweg 5, Delft.

Prior to the defence, I will give a brief presentation on my research starting at 09.30h.

Sanmay Shelat