Transit Performance Assessment and Route Choice Modelling Using Smart Card Data

Dixit, M.

DOI
10.4233/uuid:7c2ddb24-da55-4af4-8756-64587de455b7

Publication date
2022

Document Version
Final published version

Citation (APA)
Dixit, M. (2022). *Transit Performance Assessment and Route Choice Modelling Using Smart Card Data*. [Dissertation (TU Delft), Delft University of Technology]. TRAIL Research School. https://doi.org/10.4233/uuid:7c2ddb24-da55-4af4-8756-64587de455b7

Transit Performance Assessment and Route Choice Modelling Using Smart Card Data

Malvika Dixit

Delft University of Technology

Transit Performance Assessment and Route Choice Modelling Using Smart Card Data

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
Monday 10 October 2022 at 10:00 o'clock

by

Malvika DIXIT
Master of Science in Transport, Imperial College, London
born in Indore, India

This dissertation has been approved by the
Promotors: Prof.dr.ir. S.P. Hoogendoorn and Dr. O. Cats
Copromotor: Dr.ir. N. van Oort

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof.dr.ir. S.P. Hoogendoorn	Delft University of Technology, promotor
Dr. O. Cats	Delft University of Technology, promotor
Dr.ir. N. van Oort	Delft University of Technology, copromotor

Independent members:

Prof.dr. Y. Susilo	University of Natural Resources and Life Sciences, Vienna, Austria
Prof.dr. M. Munizaga	Universidad de Chile, Chile
Prof.dr. S. Bekhor	Technion – Israel Institute for Technology, Israel
Prof.dr. M.E. Warnier	Delft University of Technology
Prof.dr.ir. J.W.C. van Lint	Delft University of Technology (reserve member)

TRAIL Thesis Series no. T2022/11, the Netherlands Research School TRAIL

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

ISBN: 978-90-5584-313-8

Printed in the Netherlands

Dedicated to
my grandparents

Acknowledgements

My PhD has been as much of a personal journey for me as it has been a professional one. After working in the industry for three years, it was a chance to discover and engage in things that I love – both at work and outside. The last four and a half years have been deeply fulfilling and enjoyable, and I would like to acknowledge all the people who directly or indirectly contributed to making it that way.

I want to start by expressing my deep appreciation for my supervisory team - Niels, Oded, and Serge – each of you has provided me with much-needed guidance and inspiration in your own different ways. Niels, thanks for being incredibly supportive – I learnt more than just research from you. The way you genuinely care about your students (be it PhD or Masters) is very inspiring. I cannot forget my first TRB presentation where you were sitting in the front row cheering for me. Oded, thank you for asking the right questions and bringing clarity whenever I was stuck along the way. I am in awe of how you are so good at so many things! And Serge, thanks for stepping in with your wisdom whenever needed and for caring so much for all of us at the department, even during the challenging covid times. Lastly, thank you to all of you for giving me the freedom to explore and pursue the research direction I wanted to, and in my own schedule!

The work presented in this thesis benefitted greatly from the many discussions and practical insights from Ties. Ties also played a key role in acquiring the data used in this project and helped provide the local context for the findings. For all of this and more – thank you Ties! I am also grateful to Menno Yap and Ding Luo for graciously sharing with me their knowledge of smart card data analytics during the initial stages of my PhD. This research would not have been possible without the support from the municipality of Amsterdam, Vervoerregio Amsterdam and the AMS Institute, who not only funded this research but took an active interest in the findings – and I am extremely grateful for that. I would also like to acknowledge GVB who shared the smart card and AVL data which was used for this research.

My sincere thanks go to all my committee members Prof. Marcela Munizaga, Prof. Shlomo Bekhor, Prof. Yusak Susilo, Prof. Martijn Warnier and Prof. Hans van Lint for their time in evaluating this dissertation. It’s truly an honour to have you all as part of my PhD journey.

Next, I would like to thank Prof. Ashish Verma for introducing me to the fascinating world of transportation research, and for his invaluable guidance at different points in my career so far. I am also very thankful to Dr. Aruna Sivakumar for being a mentor to me and seeing (research) potential in me when I did not think I had any.

I feel grateful to have had great colleagues at the Transport & Planning department who were always ready to help and made times at the office a lot of fun. Maria, thank you for looking out for me right from my initial days at the department. Giulia and Jishnu, thanks for answering my numerous queries at different stages of the PhD and for being great friends. Sanmay, I really enjoyed our brief collaboration and thereafter the many conversations/rants about our PhD struggles! Danique, Konstanze, Panchamy, Fatemeh, Nejc, Arjan, Peyman, Jamie, Marco, Roy, Alex, Joelle, Nagarjun, Solmaz, Bahman, Alessandro, Lara, Florian, Martijn, Alexandra, and Yan, thanks for the chats and fun times inside and outside the office. Kayhan, Lucas, Mathew and Jesper – thanks for being great office roommates – I wish I could have spent more time with you guys in 4.21. Jesper, thanks also for so kindly translating the summary to Dutch. Edwin, thank you for solving all the issues with the remote desktop setup so swiftly every time. And as I spent much of my work time at the library, I would also like to thank the baristas at coffee star for the best oat milk cappuccinos!

I am also extremely grateful that I got to meet and become friends with so many wonderful people in the Netherlands. Abhishek, Aniket – thank you for all the inspiring conversations about science, spirituality, and more. Subodh and Prajakta, thank you for always being there for anything big or small. You have been like family here - I cannot forget the time you cooked and delivered food to my doorstep when I was down with covid! Alex - thank you for making my life in Delft so enjoyable, for bringing so much positivity during the mundane covid times, and for the many vulnerable conversations. Thanks also for bringing Olaf and Coco into my life, who have been a source of so much love and joy. Thank you to Niharika and Naveen for making Delft feel like home right from my first day there. Niharika, thanks for being my go-to friend for anything and everything, for sharing all the joys and sorrows during the PhD journey, for proofreading over and over, and for being just a number away.

Lastly, I am forever grateful to my family back in India, for giving me the 'wings to fly and roots to come back to', where no matter what I do (or don't do) I know I am always loved! I would have never pursued a PhD if not for the unconditional support from my dad. Thank you Papa for being the north star for me and inspiring me with everything you do. And my mom, for her selfless love and care. My nani – my biggest cheerleader, and my dadaji and dadi for being a constant reminder of what truly matters in life. Chacha, Chachi, Mayank Bhaiya, Niti, Maitreya, and all my cousins for making Indore visits so special. And my parents-in-law and Vriti for showering me with so much love and care, and for making everything more fun! I cannot skip thanking my sister Parul, my best friend, and my constant support system. You inspired me to think about doing a PhD and supported me throughout my journey. And finally, my life partner, Sparsh – thanks for being there at each step through the ups and downs of the PhD journey, for making so many trips to Delft from Luxembourg, and for being so patient, positive, and zen through all of it. Everything is easier (and funnier) with you by my side!

Malvika
September 2022

Content

Acknowledgements i
Chapter 1 - Introduction 1
1.1 Motivation 1
1.2 Research gaps 2
1.3 Research objective and questions 5
1.4 Research approach and theoretical foundation 6
1.5 Research context 7
1.6 Research contributions 8
1.7 Thesis outline 10
Chapter 2 - Passenger Travel Time Reliability for Multi-Modal Public Transport Journeys 13
2.1 Introduction 14
2.2 Methodology 16
2.3 Application 20
2.4 Results and discussion 23
2.5 Conclusions 28
Chapter 3 - Examining Circuity of Urban Transit Networks from an Equity Perspective 31
3.1 Introduction 32
3.2 Literature review 33
3.3 Method 35
3.4 Results 39
3.5 Conclusion 48
Chapter 4 - Perception of Overlap in Multi-modal Urban Transit Route Choice 51
4.1 Introduction 52

4.2 Overlap in transit route choice 53
4.3 Data Preparation 60
4.4 Results and Discussion 64
4.5 Conclusion 70
Chapter 5 - Validation of a Transit Route Choice Model Using Smart Card Data 73
5.1 Introduction 74
5.2 External Validation of Travel Demand Models 75
5.3 Method 77
5.4 Results and Discussion 82
5.5 Conclusion 88
Chapter 6 - Conclusion 91
6.1 Main Findings 91
6.2 Implications for practice 94
6.3 Future research directions 96
References 99
Summary 109
Samenvatting 113
About the author 117
List of Publications 119
TRAIL Thesis Series 121

Chapter 1 - Introduction

1.1 Motivation

More than half of the world's population currently lives in urban areas, and by 2050, this percentage is expected to grow to 68% (United Nations, 2018). Most cities today struggle to improve mobility for their residents while minimizing congestion, accidents, and pollution (European Court of Auditors, 2014). Public transport[1] serves as a possible solution to these problems, and forms the backbone of modern urban infrastructure. Besides contributing to economic benefits (Hensher et al., 2012), an efficient public transport system (along with walking and cycling infrastructure) can make a city more sustainable and livable by being more environmentally friendly, and potentially reducing inequality and social segregation, by providing accessibility to all (UITP, 2020; van Oort and Yap, 2021). However, as our cities continue to grow, transit agencies worldwide strive to make public transport more attractive for travelers by providing optimal services, often with limited budgets.

Transit performance assessment is essential for identifying where improvements are most needed, and for measuring the impact of improvements once they are implemented. Service quality, measured in terms of frequency, travel times, transfers, crowding, reliability, comfort, etc., influences the way passengers perceive public transport, which is reflected in their travel behavior, and ultimately ridership (van Lierop and El-Geneidy, 2016). However, providing a higher service quality also implies a higher cost for the operator. The Transit Capacity and Quality of Service Manual (TCQSM) (Kittelson & Associates, 2013) recommends transit operators to "strike a balance between the quality of service that passengers would ideally like and the quality of service that a transit agency (a) can afford to provide or (b) would reasonably provide, given the demand for transit service." Transit route choice research helps to understand the relationship between various transport network supply variables and the routes chosen by individual travelers. By providing a relative valuation of journey attributes by the travelers

[1]Public transport and transit have been used interchangeably throughout this document, and for the purpose of this dissertation refers to all modes of transportation available for use by the general public such as bus, tram and metro.

(including service quality and network characteristics), transit route choice models can support informed decision-making by identifying policies with the highest impact and predicting the changes in network outcomes because of changes to the supply variables.

Traditionally, public transport performance assessment, as well as route choice modeling, relied primarily on data from surveys, travel diaries, and transit schedules (Bertini and El-Geneidy, 2003; Kim et al., 2019; Zhao et al., 2020). However, for more than a decade now, automated systems have been introduced for public transport fare collection, providing access to a massive amount of passively collected data, as opposed to a small sample from the traditional data collection methods. Public transport smart cards are one of the most popular media for implementing Automated Fare Collection (AFC) systems. Smart card data can provide information on network-wide travel patterns at the most disaggregate spatial and temporal scales (Pelletier et al., 2011). When combined with Automatic Vehicle Location (AVL) data, it can enable a precise assessment of public transport performance including ridership and service quality measurement. In most cases, smart card data does not enable the identification of individual travelers or their demographic characteristics owing to privacy regulations. In such cases, it can potentially be combined with external sources of socio-demographic data to measure the distributive impacts of transport, enabling equity analysis at a network-wide scale. Moreover, the calculated service quality measures along with individual choices can be used to understand transit route choice behavior and reveal relative valuation of transit journey attributes, which can be further used to estimate route-shares in response to network changes.

However, being relatively new, automated sources of transit data have still not been explored to their full potential, and the methods for performance assessment and route choice modelling used with this data source are often borrowed from past research based on traditional data sources. This dissertation aims to develop methods that leverage the key strengths of automated transit data to advance transit performance assessment and route choice modelling, while accounting for its limitations. The research is undertaken in the context of urban multi-modal transit networks. In doing so, it addresses multiple scientific gaps which are described in the next section.

1.2 Research gaps

To improve the performance of a transit system, one first needs to measure its current performance in terms of the quality of service being offered as well as how it is being used. Once that is measured, this information can be used to study the impact of transit network supply on travel behavior. This study looks at these two steps sequentially and the subsequent sections describe the knowledge gaps in each of these subject areas, which this thesis attempts to address.

1.2.1 Transit performance assessment from a passenger perspective

Transit performance has several dimensions and can be measured from (at least) two perspectives - the service providers’ and the passengers’. Transit Capacity and Quality of Service Manual (Kittelson & Associates, 2013), first published in 1999, defines a transit performance measure as “a quantitative or qualitative factor used to evaluate a particular aspect of transit service”. From the service providers’ perspective, automated data sources like AVL and Automatic Passenger Count (APC) can enable calculation of supply-oriented performance measures, such as average speeds, distance travelled by vehicles, number of trips, observed travel times etc. (see for examples Bertini and El-Geneidy (2003), Erhardt et al. (2017)). In addition, smart card data can be used to provide information on the usage of the network relevant for the service provider, such as passenger-km, passenger hours and average trip length (such as in Trépanier et al. (2009)).

When considering the impact on ridership or travel behavior, the performance from a passenger perspective becomes more relevant. Service quality is defined as “the overall measured or perceived performance of transit service from the passenger's point of view”, and ideally includes everything that could impact passengers’ decisions, including travel times, availability, service, delivery, safety and security, and maintenance and construction (Kittelson & Associates, 2013). However, since many of these cannot be quantified easily, service quality measurement is often focused on the ‘availability’ (including frequency, service span and access) and ‘comfort and convenience’ (including travel times, reliability, and crowding) dimensions. In this research, we focus on two specific gaps regarding transit performance assessment, primarily from a passenger perspective. We look at each of these subsequently in the following paragraphs.

Travel time reliability forms an essential component of transit service quality, and its importance to customer satisfaction has been repeatedly highlighted in the literature (Gittens and Shalaby (2015); Jenelius (2015); Van Oort (2011)). For urban multimodal transit networks, passengers’ experience of reliability is based on the whole journey experience, including transfers. Although there are multiple passenger-oriented reliability indicators available (Chan (2007); Gittens and Shalaby (2015); Jenelius (2018); Uniman et al. (2010); Van Oort (2011)), the majority are restricted to single-leg journeys (without transfers) and do not consider different modes and their interactions. Some of the relatively recent work has looked at journeys with transfers but focuses primarily on the reliability of transfer time (Lee et al. (2014)) or travel time from the time the passenger boarded the vehicle (Bagherian et al. (2016)), ignoring the waiting time at the origin. Hence, there is a need for a methodology that can be used for travel time reliability measurement for multi-modal public transport journeys with transfers, which is sensitive to all travel time components. This leads to the following research gap:

- **Research Gap 1: Reliability quantification for multi-modal public transport journeys**

Along with operational efficiency, providing equitable access is often considered one of the primary goals of transit planning (Wei et al., 2017). Equity in this case is defined as the ‘fairness in distribution of resources’ (Litman, 2002). The importance of equity in transit planning is being increasingly recognized, and equity consideration during transport planning is, in many countries, required by legislation (Delbosc and Currie, 2011). These two goals of efficiency and equity are sometimes conflicting, and often there is no one perfect solution but many Pareto-optimal ones. The best solution depends on the specific needs of the urban area and usually involves trade-offs between different goals. For example, during public transport network design, patronage and coverage goals are often conflicting (Walker, 2008). A transit authority looking to maximize financial returns may focus on having higher patronage while compromising on coverage. To make the optimal choice, decision-makers need information on what factors impact each of these goals and what trade-offs exist between them.

There is a plethora of literature looking at the diverse perspectives and dimensions of equity in transit planning. Equity can be measured in terms of several different transport outcomes; some of the common ones include travel times, accessibility and fare paid. While there are plenty of studies analyzing equity in terms of these outcomes (such as Delbosc and Currie (2011), El-

Geneidy et al. (2016), Guzman et al. (2017), Neutens et al. (2010)) limited attention has been paid to the equity of transit network design, and its role in determining these outcomes. The few studies that do include it do so by measuring equity in terms of transit coverage (for example in Camporeale et al. (2017) or Wei et al. (2017)). However, if we consider the eventual outcomes of travel times and fare paid, these are impacted not just by the coverage but also by the design of routes.

Circuity is defined as the ratio of the network and Euclidean distances between an origin-destination (OD) pair (Barthélemy, 2011) and is a measure of directness of transit networks. Circuity has been found to impact travel behavior in multiple ways, including mode and route choice decisions of travelers (Huang and Levinson, 2015; Kim et al., 2019; Raveau et al., 2014). Transit routes with higher circuity imply a longer network distance (i.e. a longer detour) for the same Euclidean distance covered. This impacts the passengers in two ways. Firstly, longer network distance implies longer travel times for the passengers, all else being equal. Secondly, for transit networks where the fare is calculated based on the network distance travelled, higher circuity directly impacts the fare paid. Essentially, travelers using highly circuitous routes will pay higher for a worse-off connection. Hence, an uneven distribution of circuity can contribute to the disparity in travel times, and fare paid, making circuity relevant from an equity perspective. However, there is limited research on the distribution of circuity observed within a transit network and its impact on travelers from different population groups. In sum, in this dissertation, we consider the following research gap:

- **Research Gap 2: Evaluation of transit circuity from an equity perspective**

1.2.2 Transit route choice modelling

Compared to traditional data sources like stated preference surveys and travel dairies, automated data sources have a significant advantage of providing network-wide travel choices and precise measurement of travel attributes experienced by the passengers, making them an attractive data source for route choice analysis. Another major advantage, generic to all revealed preference data sources is avoidance of response bias as in the case of stated preference data. Notwithstanding, the literature on transit route choice modelling using smart card data is scarce. Many of these studies are limited to one transit mode only, typically metro (like Guo and Wilson (2011), Hörcher et al. (2017), Kusakabe et al. (2010), Raveau et al. (2014), Tirachini et al. (2016), Zhao et al. (2017)). Only a few (namely Arriagada et al. (2022), Jánošíkova et al. (2014), Kim et al. (2019), Tan et al. (2015), Yap et al. (2020)) have analyzed a large-scale multi-modal transit network.

In route choice modelling, capturing unobserved correlations between overlapping routes is a non-trivial problem. For road networks, this has been a well-researched topic with a consensus that overlapping routes are perceived to be similar by the travelers, leading to a reduction in their utility compared to completely independent routes (Bovy et al., 2008). However, for transit networks, research so far has been inconclusive on how this overlap should be defined and is perceived. Specifically, there are three gaps in knowledge on this topic. Firstly, there is no consensus on how travelers perceive overlap between routes – with some studies reporting a positive valuation (for example, Hoogendoorn-Lanser and Bovy (2007); Anderson et al. (2017)) while some others a negative valuation (for example Yap et al. (2020); Tan et al. (2015)). Secondly, when considering path overlap, it is not clear whether it should be defined in terms of overlapping links (i.e. part of the route between consecutive stops) (like in Tan et al. (2015)) or entire journey legs (like in Hoogendoorn-Lanser et al. (2005)). So far, none of the studies compare these two different ways of defining path overlap, and how each of them is perceived by travelers. Lastly, for multi-modal transit networks, correlation between routes is not just restricted to path overlap but could also be because of common nodes, service, runs or modes (Hoogendoorn-Lanser et al., 2005). However, the studies for urban transit networks so far have only considered overlap in terms of path. This leads us to the following research gap:

- **Research Gap 3: Definition and valuation of different types of overlap for route choice models in case of urban transit networks**

Generally, the methods and models used for travel behavior research should be such that they highlight the strengths of the data source being used, while managing its limitations (Zhao et al., 2020). The route choice models currently in use were originally developed based on data actively collected for model estimation purposes (Zhao et al., 2020). Although smart card data provides several advantages over traditional data sources as discussed earlier, there are also some major shortcomings with regards to their application for transit route choice modelling. Firstly, no information is available on the trip purpose, journey origin location and often the time of arrival at the boarding stop, which requires assumptions to be made regarding consideration choice set and perceived travel attributes for route choice modelling. Secondly, due to privacy regulations, socio-demographic characteristics of card holders are not available in most cases. Hence, models estimated using smart card data cannot incorporate differences in travel behavior resulting from variation in individual characteristics. Lastly, privacy regulations often do not allow for a cardholder to be tracked across days, making it challenging to incorporate the panel structure of the data. Despite these differences between passively collected smart card data and other data sources, no research so far has assessed how stable are the estimates from models estimated based on such data and how applicable the models are for demand forecasting.

When the selected model is close to reality, the estimated parameters are expected to be stable for a reasonable range of temporal and spatial conditions, and the model predictions are expected to resemble observed demand. Before the availability of automated data sources, getting detailed time-varying demand patterns was a challenge (Poon et al., 2003). This made it difficult to validate the transit route choice models once they had been estimated. With the availability of smart card data, the observed flows on each transit link are available and have been utilized to validate the flows obtained from schedule or frequency-based transit assignment models for the rail/metro modes (see Fung et al. (2005) or Poon et al. (2003)). On the other hand, many studies have used smart card data to estimate the route choice models (Hörcher et al., 2017; Kim et al., 2019; van Oort and Yap, 2021, amongst others), but, to the best of our knowledge, no study so far has undertaken such an external validation of these models. Hence, this dissertation seeks to address the following research gap:

- **Research Gap 4: External validation of transit route choice models estimated using smart card data**

1.3 Research objective and questions

Based on the research gaps identified in the previous section, the overarching research objective considered in this thesis is as follows:

"*To improve performance assessment and route choice modelling for urban multi-modal transit networks using smart card data*".

To fulfill this objective, we formulate four research questions that we intend to answer through this research:

1. How can travel time reliability for multi-modal transit journeys from a passenger perspective be quantified?
2. How can the effects of network design on distributional aspects of travel times and fare paid in the network be characterized?
3. How can travelers' perception of overlap between alternative routes be incorporated in models of transit route choice?
4. How valid are models of transit route choice estimated using smart card data?

1.4 Research approach and theoretical foundation

To answer the research questions, we divide our research into two steps, as shown in **Figure 1.1**. The first part seeks to improve transit performance measurement and the second route choice modelling. The research is performed using the urban transit network of Amsterdam, the Netherlands, as a case study. However, the methods developed are applicable to other urban settings where a comparable data set is available.

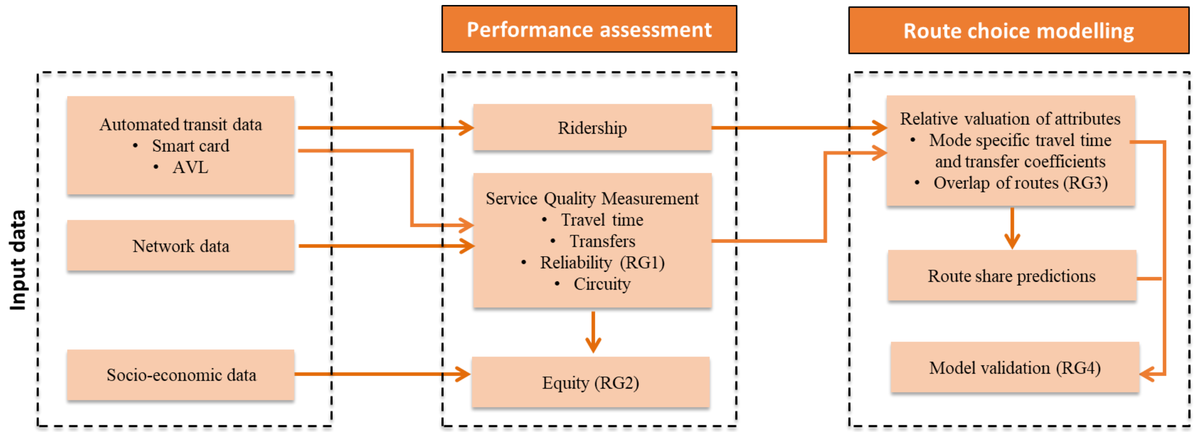


Figure 1.1 Overall research approach, including research gaps (RGs).

As a first step, passenger journeys are inferred by fusing smart card and AVL data using existing destination and transfer inference algorithms (Gordon et al., 2013; Trépanier et al., 2007; M. D. Yap et al., 2017). This results in a network-wide passenger journeys database that can be aggregated at any spatial or temporal level for performance assessment. Using this database, a reliability metric is proposed, which is applicable for transit journeys with transfers, addressing Research Gap 1. Along with the reliability metrics, other service quality measures are also calculated for each journey including different components of travel times, number and type of transfers and circuity. The circuity and travel times, combined with the zonal level income distribution of the population is used to undertake an equity analysis for the network (Research Gap 2). A spatial regression analysis is undertaken to understand the relationship between income, circuity and distance covered by the travelers.

In the second step, we use the service quality measures calculated in the first part of the research for understanding the route choice behavior of transit travelers. For this, discrete choice models based on the Random Utility Framework (McFadden, 1974) are used to obtain the valuation of mode-specific travel time and transfer attributes. Further, the definition and perception of three different types of overlap in the networks is explored – at the link, leg, and transfer node levels

(Research Gap 3). The underlying assumption here is that an overlap between alternative routes results in correlations between unobserved components of routes' utilities, unless explicitly accounted for.

Lastly, we examine how transferable are the developed models of route choice by undertaking an external validation using a real world case of introduction of a new metro line in the network (described in the subsequent section), addressing Research Gap 4. The model is estimated using data from before the network change, and is used to predict the impact of the network change on passenger shares. The parameters estimated for the travel attributes are also compared by estimating a similar model using the data from 'after' the network change.

The research assumes the availability of smart card and AVL data, and focuses specifically on the challenges and opportunities when using these data sources for transit performance assessment and route choice modelling. Although the smart card data we use is from a closed AFC system (where both check-in and check-out information is available), the methods can also be applied to open AFC systems after inferring destinations using existing methods (Munizaga and Palma, 2012; Trépanier et al., 2007).

1.5 Research context

The research presented in this dissertation was conducted in the context of the North South metro line research project initiated by the Municipality of Amsterdam, the Netherlands. The new north-south metro line was introduced in Amsterdam, the Netherlands on 22nd July 2018 adding significant capacity to the existing metro, bus and tram network in the city. Along with the introduction of the new line, changes were made to the rest of the transit network including the removal or re-routing of existing transit lines, as shown in **Figure 1.2**.

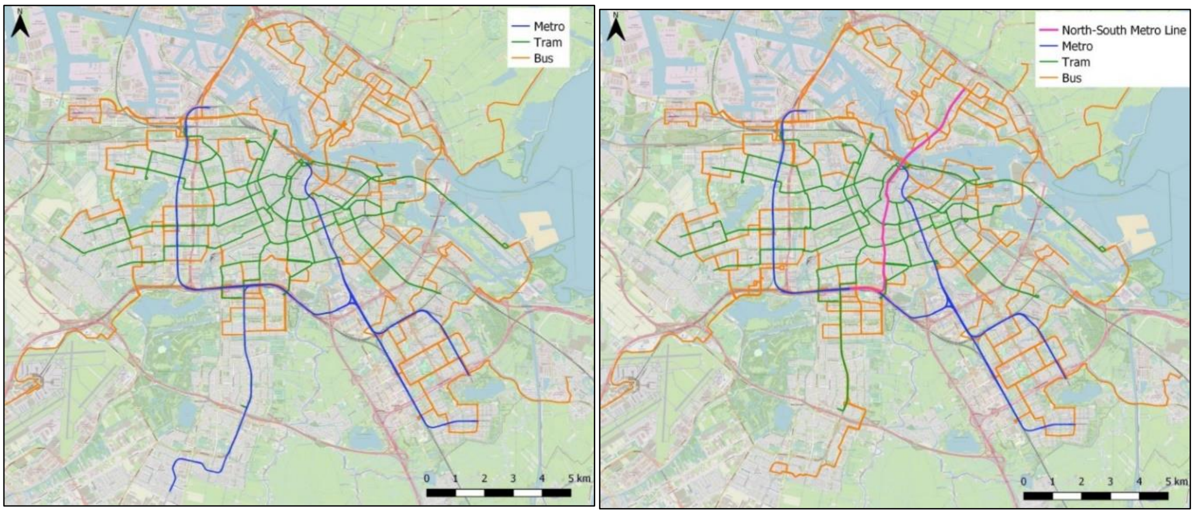


Figure 1.2 Public transport network of Amsterdam before (left) and after (right) opening of the north-south metro line.

The North South metro line project was aimed at studying the impact of the new metro line on mobility patterns and behavior, socio-economic and spatial aspects, as well as quality of life. The project partners include two industry partners (Municipality of Amsterdam and Vervoerregio Amsterdam) and five academic partners (Delft University of Technology, University of Amsterdam (UvA), Vrije Universiteit Amsterdam (VU), Centrum Wiskunde &

Informatica (CWI) and The Amsterdam Institute for Advanced Metropolitan Solutions (AMS)). In addition, the Gemeentelijk Vervoerbedrijf (GVB), the public transport operator of Amsterdam, provided the public transport data used in this research and contributed to the discussion on network impacts.

This PhD research is one of the main contributions of Delft University of Technology to the broader north-south metro line project (the final report of which is available at AMS (2021)), and aims to developing methods to assess the impact of the new metro line on the quality of public transport network and passenger ridership. The methods developed in this research were applied to undertake empirical analyses for the multi-modal urban transit network of Amsterdam. It also utilizes the network change to assess the applicability of transit route choice model estimations to predict the impact of changes in demand in response to major infrastructural changes.

The addition of the new metro line has several impacts on the quality of public transport - not just on the affected corridor(s) but also on the rest of the network. **Figure 1.3** shows one way of looking at these impacts, which are relevant for the North-South metro line project. The introduction of the new line impacts the travel times, travel time reliability, number and type of transfers, in-vehicle crowding, and the detours experienced by passengers between different OD pairs. The changes in such service quality attributes are expected to lead to a change in travel behavior in terms of public transport route choice, mode choice (including that between public transport and private modes), destination choice, departure time choice, or addition of new trips (induced demand). These changes result in a change in ridership for public transport, which in turn impacts some of the service quality attributes. This project focuses on the measurement of change in service quality attributes (or performance assessment) and their impact on public transport route choice. The project scope is shown with highlighted (orange) boxes in **Figure 1.3**.

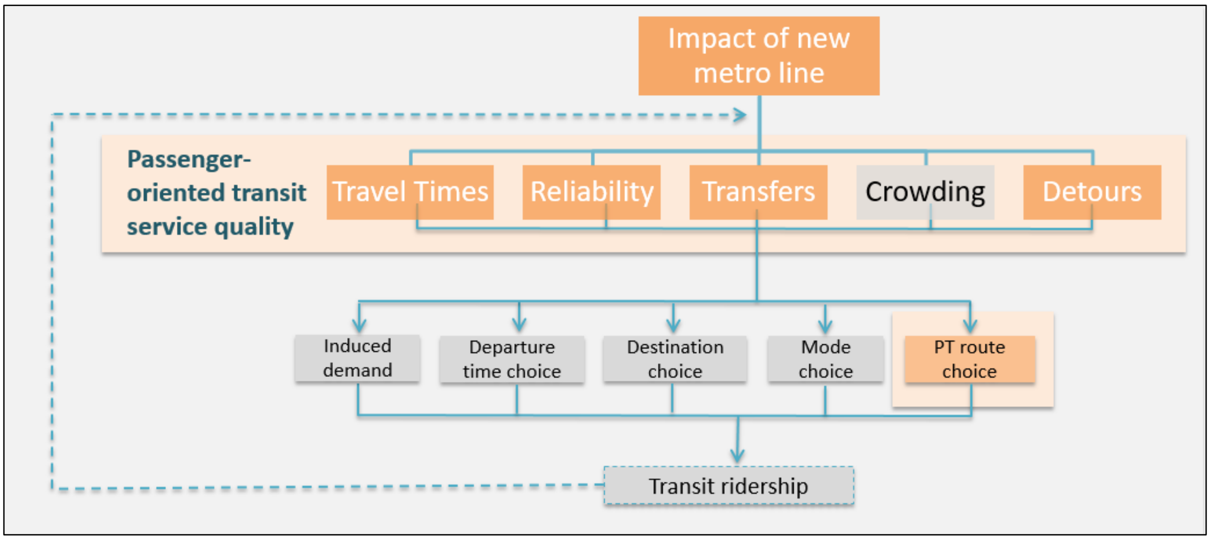


Figure 1.3. Project scope within the North-South metro line research project.

1.6 Research contributions

This section summarizes the main contribution of the thesis, distinguishing scientific (theoretical and methodological) contributions, and societal contributions.

1.6.1 Scientific contributions

This dissertation attempts to advance the ways in which automated sources of transit data are currently used for public transport perforamance assessment and route choice modelling. A key theoretical contribution pertains to stitching together the existing methods offered in literature to establish their suitability for working with these new transit data sources. The specific scientific contributions of each chappter are described below:

1. *Developing a methodology for travel time reliability measurement from a passenger perspective for multi-modal transit journeys with transfers (Chapter 2)*
   We extend the existing Reliability Buffer Time (RBT) metric (Chan, 2007; Uniman et al., 2010) to multimodal transit journeys with transfers and develop a methodology to calculate it using a combination of smart card and AVL data. The resulting metric enables comparison of reliability between different modes and routes, and is sensitive to the variability in all components of travel time experienced by a traveler including initial waiting time, in-vehicle times and transfer times for all journey legs. We demonstrate the methodology and some potential applications of the metric by applying it to the urban transit network of Amsterdam. The metric can also be used as an input to route choice models, particularly since it is calculated from a passenger perspective.

2. *Inestigating the relationship between circuity of public transport networks and its equity outcomes (Chapter 3)*
   This study provides a new perspective on the role of public transport network design in determining its equity outcomes, and how it can be used to reduce existing inequities. Specifically, we disentangle the impact of land-use distribution and transit network design on the disparity in distance travelled by different income groups in case of urban transit networks. We further explore its implications on travel times and fare paid for networks with distance-based fare. We leverage the travel patterns obtained from smart card data to undertake a network-wide investigation. Smart card data is linked to aggregate-level income data to undertake the equity analysis.

3. *Evaluating the perception of different types of overlap during route choice modelling for urban transit networks (Chapter 4)*
   The main contribution of this study is a better understanding how travelers perceive different types of overlap between alternate routes when making transit route choice decisions in case of urban transit networks. In addition to the traditional definition of overlap in terms of path (in the form overlapping links or legs), we also define the overlap in terms of transfer nodes. This study also adds to the limited literature using large-scale revealed preference data for understanding route choice behavior, and provides valuations of mode-specific travel time & transfer attributes using smart card data from Amsterdam, the Netherlands.

4. *Establishing the temporal transferability of transit route choice models estimated using smart card data (Chapter 5)*
   This study adds to the scarce literature on the validation of travel demand models and, to the best of our knowledge, is the first to undertake an external validation of a transit route choice model. We use smart card data from before and after the opening of a new metro line in Amsterdam, the Netherlands for model estimation and validation, respectively. We examine the transferability of transit route choice models developed using smart card data, and how well they perform for forecasting changes in demand because of a network change. We also compare alternate specifications of the model to

empirically study the impact of adding/omitting relevant variables on model transferability.

1.6.2 Societal contributions

This research proposes novel ways in which automated data sources can be used for improving transit planning which can be useful for transit service providers and policy makers. With the eventual aim of making transit more attractive as well as equitable, practitioners can apply methods developed in this research to improve transit performance monitoring as well as enhance their understanding of transit user behavior, as described in the subsequent sections.

Improvements in transit performance assessment and monitoring

Firstly, the reliability measure developed can be used by practitioners to evaluate the reliability of current services from a passenger perspective. Passively collected data sources can be used to calculate this metric on a regular basis, which can then be aggregated to identify trends in the reliability from the passenger's perspective. The metric can also be used as an input to existing travel demand models to incorporate the impact of reliability on travel demand. Secondly, this research demonstrated how smart card data can be combined with aggregate-level socio-demographic data to undertake an equity analysis of travel times and fare paid in the network. The proposed method can be used by practitioners to assess the contribution of transit network design toward the observed equity outcomes. The methodology for calculating travel times, transfers and reliability developed in this study has already been applied for undertaking an ex-post impact evaluation of the new north-south metro line in Amsterdam on transit ridership and service quality (see Brands et al. (2020), Dixit et al. (2019a)).

Improvements in route choice modelling

Our study presents the application of a transit route choice model on the network-wide smart card data for Amsterdam. First, the research provides behavioral insights into how overlap between alternate routes is perceived by travelers when making route choice decisions, which has implications for transit network design. Moreover, findings regarding the definition of different types of overlap can be used to enhance existing transit assignment and route choice models. Additionally, the relative valuation of different mode-specific attributes from the route choice models provides valuable input for improving the outcomes of transport models and evaluation methods such as cost benefit analyses. The research also presents a systematic validation methodology that can be used by practitioners for validating route choice models. Lastly, the results of the model validation analysis undertaken has implications for practitioners looking to apply these models for policy analysis.

1.7 Thesis outline

The outline of this dissertation is shown in **Figure 1.4**. The dissertation is divided into two main sections. Smart card and AVL data are first used to calculate transit performance metrics. Thereafter, some of the calculated metrics are used as an input for route choice modelling, and the models are validated by using them for predicting the outcome of a network change.

Chapter 2 presents a method to use raw smart card and AVL data to develop a passenger-oriented reliability measure which considers multi-modal journeys with transfers (Research gap 1). The application of the metric is demonstrated on a small sample of smart card data. In this chapter we also discuss the steps followed for pre-processing and fusion or smart card and AVL data to form a passenger journey database, which is used a base for all the further analysis in this research. Next, in *Chapter 3* we analyse the circuity experienced by travelers of different income groups, and its impact on the disparity in travel times, distance travelled and fare paid by them (Research gap 2). In doing so, we explore the role of network design on the equity outcomes for the network. Circuity measured in this study is also found to be a significant factor in explaining the transit route choice of passengers which is explored in *Chapter 4*. In addition to circuity, the outcomes of the route choice models include mode-specific travel time and transfer valuations, as well as how the overlap between alternate routes is perceived by travelers (Research gaps 3). Three types of overlap are considered and compared – overlap of links, journey legs and transfer nodes, and the implications of the results for network design are discussed. Finally, in *Chapter 5*, we explore how transferable are the models of transit route choice developed using smart card data (Research gap 4) using the data from before and after opening of the new metro line, thereby undertaking a validation assessment of the models estimated in *Chapter 4*. Finally, in *Chapter 6* we summarize the main conclusions from our research and its implications for practice and future research.

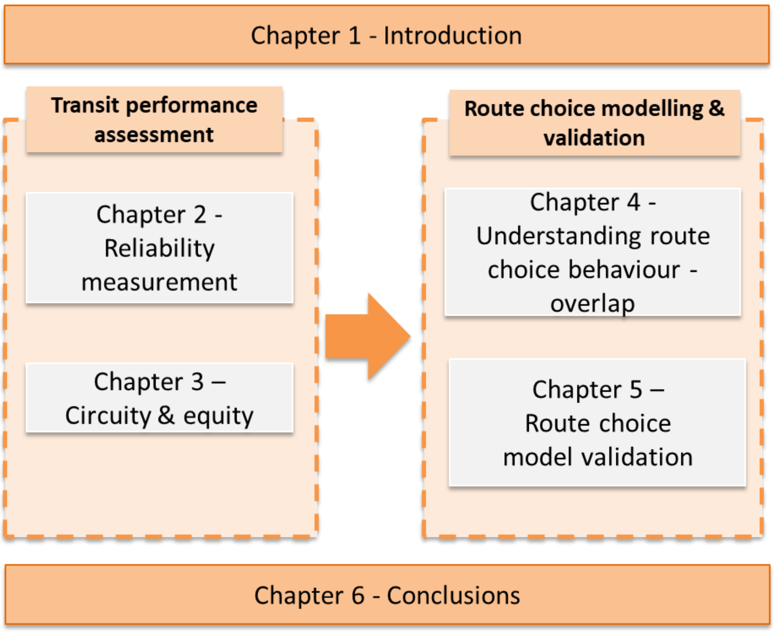


Figure 1.4. Thesis outline.

Chapter 2 - Passenger Travel Time Reliability for Multi-Modal Public Transport Journeys

Travel time reliability is an important determinant of the service quality of transit networks. This chapter contributes to improving transit performance assessment by focusing on the quantification of travel time reliability for urban transit networks from a passenger perspective. Urban transit networks typically consist of multiple modes and the journeys may involve a transfer within or across modes. Hence, the passenger experience of travel time reliability is based on the whole journey experience including the transfers. Although the impact of transfers on reliability has been highlighted in the literature, the existing indicators either focus on uni-modal transfers only or fail to include all components of travel time in reliability measurement.

This chapter extends the existing 'Reliability Buffer Time' metric to journeys with multi-modal transfers and develops a methodology to calculate it using a combination of smart card and automatic vehicle location data. The developed methodology is applied to the data from the Amsterdam transit network and potential applications are demonstrated. The measure is calculated for each OD-route pair and can be aggregated to compare reliability between different routes, modes, transit stops, or time periods. The measure can also be used as an input for travel behavior models of mode, route, or departure-time choice.

This chapter is based on the following article:

Dixit, M., Brands, T., van Oort, N., Cats, O., Hoogendoorn, S. (2019) Passenger Travel Time Reliability for Multimodal Public Transport Journeys. *Transportation Research Record,* 2673(2), 149–160, doi:10.1177/0361198118825459.

2.1 Introduction

2.1.1 Multimodal service reliability

Transit service reliability forms an important component of service quality and its importance to customer satisfaction has been repeatedly highlighted in the literature (Gittens and Shalaby, 2015; Jenelius, 2018; Van Oort, 2011). From the operator's perspective, improved reliability can reduce operational costs and increase revenue by potentially increasing the ridership and retention (Diab et al., 2015).

Urban transit networks typically consist of multiple modes and passenger journeys may involve a transfer within or across modes. Hence, the passenger experience of reliability on such networks is based on the whole journey experience including the transfers. Although there are multiple passenger-oriented reliability indicators available (Chan, 2007; Gittens and Shalaby, 2015; Jenelius, 2018; Uniman et al., 2010; Van Oort, 2011), the majority are restricted to single leg journeys (without transfers), and do not consider different modes and their interactions. Some of the recent work has looked at journeys with transfers, but focus primarily on the reliability of transfer time (Lee et al., 2014) or travel time from the time passenger boarded the vehicle (Bagherian et al., 2016), ignoring the waiting time at the origin.

This study uses an existing indicator – Reliability Buffer Time (RBT, described by Chan (2007), Uniman et al. (2010)) – as a point of departure, and extends it to journeys with multiple legs and modes for urban, high-frequency transit networks using smart card and automatic vehicle location (AVL) data. The developed metric aims to

- measure reliability for multi-modal public transport journeys;
- enable the comparison of different transit modes and routes;
- be sensitive to the variability in waiting time, in-vehicle time and transfer time for all legs of the journey

The method developed is applied to a real-life case study of the Amsterdam transit network to demonstrate its implementation in practice. The methodology however is independent of the data system(s) in use and could be applied to any transit network where smart card and AVL data sources are available.

By using a consistent method for all journeys using all available transit modes, reliability can be compared between any route for any OD pair in a multi-modal transit network. The developed metric can be used to study the reliability impacts of policies affecting multiple transit modes. Additionally, it could also be used as an input to behavioral models such as mode, route or departure time choice models.

The rest of the chapter is structured as follows: first a background on reliability and the application of automated data for reliability measurement is presented. *Section 2.2* then introduces the new metric and the methodology to calculate it using smart card and AVL data. The developed methodology is applied to the case study in *Section 2.3,* for which the results are discussed in *Section 2.4*. Lastly, *Section 2.5* presents the conclusions and limitations of the study.

2.1.2 Travel time reliability

Reliability in this context is defined as 'certainty of service aspects compared to the schedule (such as travel time (including waiting), arrival time and seat availability) as perceived by the user' (van Oort, 2016). Traditionally measured in terms of service-oriented indicators (such as on-time performance, headway regularity (Kittleson & Associates et al., 2003)), lately there has been a shift towards passenger oriented measures, as they can better capture the effectiveness of reliability improvement strategies by including the end-user perspective (Bagherian et al., 2016). A review of the existing passenger-oriented reliability measures can be found in Gittens and Shalaby (2015) and in Currie et al. (2012).

Reliability may be measured in terms of travel time regularity (consistency of experienced travel times) or punctuality (deviation from the scheduled arrival time/travel time). Cats (2014) notes that in case of urban high-frequency services, passengers arrive randomly without consulting the schedule, making travel time regularity more relevant than punctuality. Our study is based on dense high-frequency urban transit networks, hence reliability has been measured in terms of regularity of travel time.

2.1.3 Reliability and automated data sources

The smart card data source has been utilized repeatedly in the recent past for a range of applications in transport planning. Pelletier et al. (2011) provide a review of these applications of smart card data for strategic, tactical and operational levels of transport planning. For service reliability measurement also, much of the recent research utilizes smart card and AVL data sources (Chan, 2007; Uniman et al., 2010). In practice also, many transit agencies are moving towards such data sources due to lower data collection costs and better quality of data (Kittleson & Associates et al., 2003).

The AVL data provides spatio-temporal information on vehicle movement, which can be used directly to calculate vehicle-oriented passenger reliability metrics (van Oort et al., 2015b). For estimating passenger-oriented metrics, AVL data can be used in combination with Automated Passenger Counts (APC) data to weigh the calculated metrics based on demand. Furth and Muller (Furth and Muller, 2006) used the observed headways from AVL data to obtain the waiting time distribution at origin stop for buses. Similar approach was also employed by Ehrlich (2010) to estimate the travel time (waiting time + in-vehicle time) distribution for bus journeys in London.

Lee et al. (2014) highlighted the importance of including the impact of transfers on reliability assessment. Using AVL data, they estimate the additional delay due to transfer synchronization. Jenelius (2018) also used the AVL data to estimate the transfer times by tracing a virtual "probe traveler" undertaking the journeys between different OD pairs. Since AVL data does not directly provide any information on transfers, assumptions need to be made to estimate transfer time(s) experienced by the passengers. However, with AFC data, the total journey time including the experienced transfer time can be inferred precisely for each passenger. In addition, the number of transferring passengers for each OD pair can also be derived.

Bagherian et al. (2016) used the AFC data for measuring regularity and punctuality for journeys with transfers. However, the travel time is measured from tap in at first stop to tap out of last stop of the journey, ignoring the waiting time at the origin stop.

Table 2.1 provides a summary of some of the notable studies using automated data sources for reliability measurement, in terms of the components of travel time reliability included and the type of journeys they are applicable to.

Table 2.1. Existing travel time reliability measures using automated data sources.

Study	Measure(s) developed	Data used	Travel time component included			Modes applicable to
			In-vehicle time	Waiting time (origin)	Transfer time	
Furth and Muller (2006)	Waiting cost	AVL	No	Yes	No	All
Uniman et al. (2010)	Reliability Buffer Time, Excess Reliability Buffer Time	AFC	Yes	Yes	Only for metro	Metro
Van Oort (2011)	Additional travel time, Reliability buffer time	AVL	Yes	Yes	No	All
Lee et al. (2014)	Additional Travel Time, Reliability Buffer Time	AVL, APC	No	No	Yes	Train and tram
Gittens and Shalaby (2015)	Journey Time Buffer Index	AVL, APC	Yes	Yes	No	Bus
Bagherian et al. (2016)	Passenger Journey Time Variability, Passenger Schedule Deviation Reliability	AVL, AFC	Yes	No	Yes	Bus/ tram

From a passenger perspective, the reliability of a journey should incorporate the variation in all the travel time components – the waiting time, in-vehicle time and transfer time. Although the various components of reliability have been addressed individually or combined in the existing literature, none of the existing indicators incorporate sensitivity to all components of travel time for multi-modal public transport journeys. Our research aims to fill that gap.

2.2 Methodology

2.2.1 Definitions

In this study, the word ‘journey’ refers to the travel made by a passenger from an origin transit stop to a destination transit stop including transfers, if any. A ‘leg’ of a journey consists of the travel made using a single transit line, without any transfers. A journey may include multiple legs by same or different transit modes. The term Origin-Destination (O-D) pair is used to denote a combination of transit stops (a stop-stop pair). An O-D pair may be connected by more than one transit ‘route’ which refers to the combination of transit lines a passenger may choose, where each route may or may not include a transfer.

2.2.2 Reliability Buffer Time

The Reliability Buffer Time (RBT), first introduced by Chan (2007) and later updated by Uniman et al. (2010), measures the variability in travel time as the absolute difference between an extreme N^{th} percentile and the 50th percentile travel times. The value of N is typically set to 95 (Wood, 2015), where RBT can be interpreted as the additional time passengers have to budget for their travel to ensure on-time arrival one out of twenty times, a value considered acceptable in the literature. RBT and its variations (such as individual RBT and platform to platform RBT (Wood, 2015)) are one of the popular ways of measuring travel time reliability. Besides being easy to compute, some of its key advantages include its representation of passenger perspective, ease of interpretation for non-experts and flexibility of calculation across time and space (Wood, 2015). Although originally developed for metro, RBT has also been applied for reliability evaluation of bus networks (Ehrlich, 2010). However, so far, the majority of applications of RBT have focused on unimodal journeys. Our study extends the calculation of RBT to multi-modal transit journeys. The RBT is calculated for each transit route alternative per OD pair, which may or may not include transfers.

For a journey with multiple legs, the total travel time (of the transit part of a journey) includes the waiting time at the origin stop, the in-vehicle time for the first leg, the waiting times and in-vehicle times for the subsequent legs of the journey, and the transfer (walking) times between each leg – shown in **Figure 2.1** for a journey with two transit modes. Since the passenger experience of reliability is defined by the variation in all of these components of travel time, the ideal reliability metric should be sensitive to each of these components, where each component may be perceived differently by travellers. However, our research assigns equal value to the variability in each individual component of travel time, since, in the end, the travel time variability of the entire journeys is relevant for the traveler. Furthermore, these components may be correlated, i.e. the variation in in-vehicle time for the first leg of the journey could be absorbed by the waiting time at the transfer. Therefore, in our study we measure variability of the total travel time over the entire journey.

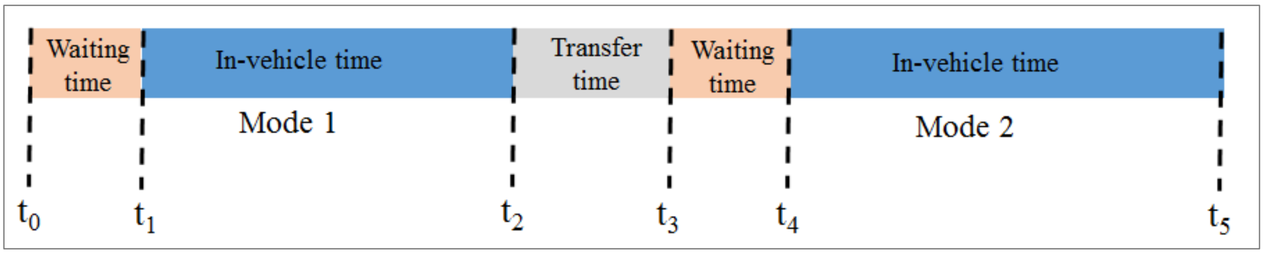


Figure 2.1. Components of passenger experienced travel time for a transit journey with two legs.

For each multi-leg journey, the following generic parameters are defined:

$wt_{o,d,r,l,i}$ Waiting time for leg '*l*' of journey '*i*' using route '*r*' between origin–destination pair *o,d*

$ivt_{o,d,r,l,i}$ In-vehicle time for leg '*l*' of journey '*i*' using route '*r*' between *o,d*

$x_{o,d,r,l,i}$ Transfer time between leg '*l*' and '*l+1*' for journey '*i*' using route '*r*' between *o,d*

$n_{o,d,r}$ Number of legs in journey between *o,d* using route '*r*'

$tt_{o,d,r,i}$ Total travel time for journey '*i*' using route '*r*' between *o,d* over all legs

For each journey 'i', the total travel time is given as

$$tt_{o,d,r,i} = \sum_{l=1}^{n_{o,d,r}} \left(wt_{o,d,r,l,i} + ivt_{o,d,r,l,i}\right) + \sum_{l=1}^{n_{o,d,r}-1} x_{o,d,r,l,i} \quad \forall o,d,r,i \tag{2.1}$$

The individual travel times are aggregated over all journeys that belong to a specific o-d pair and route combination, by calculating the median value and the 95th percentile value, given as

$tt_{o,d,r}^{55}$ 50th percentile travel time over all journeys between origin–destination pair o,d
$tt_{o,d,r}^{95}$ 95th percentile travel time over all journeys between origin–destination pair o,d

These values are used to calculate the RBT for each o-d pair for each route:

$$RBT_{o,d,r} = tt_{95}^{o,d,r} - tt_{50}^{o,d,r} \tag{2.2}$$

The RBT measures the absolute difference (in minutes) in travel times, as opposed to the relative values. This has been consciously chosen for this study, because of three reasons:

1. Different modes have different speeds, i.e. metro routes are expected to have a shorter travel time than trams for the same OD pair. Since one of the aims is to be able to compare reliability between modes, the relative values may underestimate the reliability of faster modes.
2. For OD pairs very close to each other (for example next stop on metro), the ratio of 95th to 50th percentile travel times may be very high, since the waiting time component is large compared to the in-vehicle travel time. If the reliability is measured as a percentage of median travel times, this will lead to RBT values exceeding 100% which are difficult to compare.
3. When using travel time variability in policy evaluation (using a value of travel time reliability) absolute values are preferred.

The RBT can be measured for any selected time interval such as the peak hour or three hours, provided that enough data points are available. It is recommended to choose a time period where the frequency of services is consistent, since variation in frequencies can contribute to higher variation in waiting and transfer times, leading to a higher RBT.

2.2.3 RBT calculation using smart card data

Since the aim is to measure variability in travel times, the large amount of observations provided by the smart card data allow for a realistic measurement (Uniman et al., 2010). For this study, data from the Dutch smart card system (see Van Oort et al. (2015a) for details) and AVL data (see Van Oort et al. (2015b) for details) has been utilized. The Dutch smart card requires tap-in and tap-out for all modes, implying both boarding and alighting locations and times are available. However, for systems where tap-out is not recorded (London buses, New York metro etc.), destinations can be inferred using a combination of smart card and AVL data (Trépanier et al., 2007; Zhao et al., 2007). Once the destination is available for all transactions, transfer inference is undertaken to combine individual transactions to journeys (Gordon et al., 2013; M. D. Yap et al., 2017).

Depending on the system, the tap-in is either *at the stop* (e.g. most current metro systems, including Amsterdam) or *on-board the vehicle* (most current bus and tram systems, e.g. London bus and Amsterdam bus and trams). When tap-in is in the vehicle, the difference between tap-in and tap-out times for each transaction correspond to the in-vehicle times only (t_2-t_1 and t_5-t_4 in **Figure 2.1**), whereas in the station-based tap-in, this time also includes access/egress time to the vehicle, waiting time at the platform and the transfer time within the same mode (if applicable) (t_2-t_0 and t_5-t_3 in **Figure 2.1**). For the sake of simplicity, and assuming they form a small component of the overall travel time, the access and egress time to the vehicle from the fare gates is not explicitly included in the specification of the measure used in this study. It is assumed that this time is constant across passengers and hence does not contribute to reliability.

For Amsterdam, for journeys with metro as the first mode, the total travel time can be calculated directly from the smart card data as the difference between the last tap-out and first tap-in (t_5-t_0 in **Figure 2.1**). Whereas for journeys with buses and trams as the first mode, the waiting time at the origin stop needs to be measured/estimated separately (only t_5-t_1 in **Figure 2.1** is measured). This is represented as

$$tt_{o,d,r,i} = \begin{cases} \left(\tau^{out}_{i,l=n_{o,d,r}} - \tau^{in}_{i,l=1}\right) & if\ mode\ is\ metro\ for\ l = 1 \\ \left(\tau^{out}_{i,l=n_{o,d,r}} - \tau^{in}_{i,l=1}\right) + wt_{i,l=1} & if\ mode\ is\ bus\ or\ tram\ for\ l = 1 \end{cases} \tag{2.3}$$

Where,

$\tau^{in}_{i,l}$ Tap-in time for the leg *'l'* of journey '*i*'

$\tau^{out}_{i,l}$ Tap-out time for the leg *'l'* of journey '*i*'

In this study, the waiting time at origin $wt_{i,l=1}$ is estimated for each individual (journey) and then added to the time measured by smart card data for journeys with bus or tram as the first mode.

For short headway services, it has been known that passengers arrive at the transit stops without consulting the schedules (Furth and Muller, 2006; Kittelson & Associates, 2013). Hence, within this short interval of time between the arrivals of consecutive buses/trams, a uniform distribution of passenger arrivals can be assumed. Continuous random variables are then generated and sampled over the uniform distribution of [0, observed headway] to obtain waiting time for each individual journey. The observed headway is obtained from the AVL data. Since the waiting time is sampled over all the passengers arriving during an observed headway, this method captures the ramifications of uneven headways on passenger waiting times as more waiting times are sampled for the longer headways.

Once the waiting time is assigned for journeys where the first mode is bus/tram, the total travel time is calculated for each journey (using Equation (2.1)) and aggregated to percentile values for each route for each stop-stop pair. Subsequently, RBT is calculated based on Equation (2.2). RBT can be compared between multiple routes for the same OD pair. It can also be aggregated for each mode or mode combination by using a demand weighted average, given by

$$RBT_m = \frac{\sum_{o,d,r \in R_{m,o,d}} (N_{o,d,r} * RBT_{o,d,r})}{\sum_{o,d,r \in R_{m,o,d}} N_{o,d,r}} \quad \forall\, m \tag{2.4}$$

Where,

$N_{o,d,r}$ Total passengers travelling on route 'r' between origin destination pair *o,d*

$R_{m,o,d}$ Set of all routes on origin destination pair *o,d* using mode '*m*'

Using a similar approach, it can also be aggregated or segmented to other dimensions such as for the number of transfers involved, or for the whole population or groups of users within the population. **Figure 2.2** summarizes the steps that can be followed to derive RBT for multi-modal journeys using smart card and AVL data.

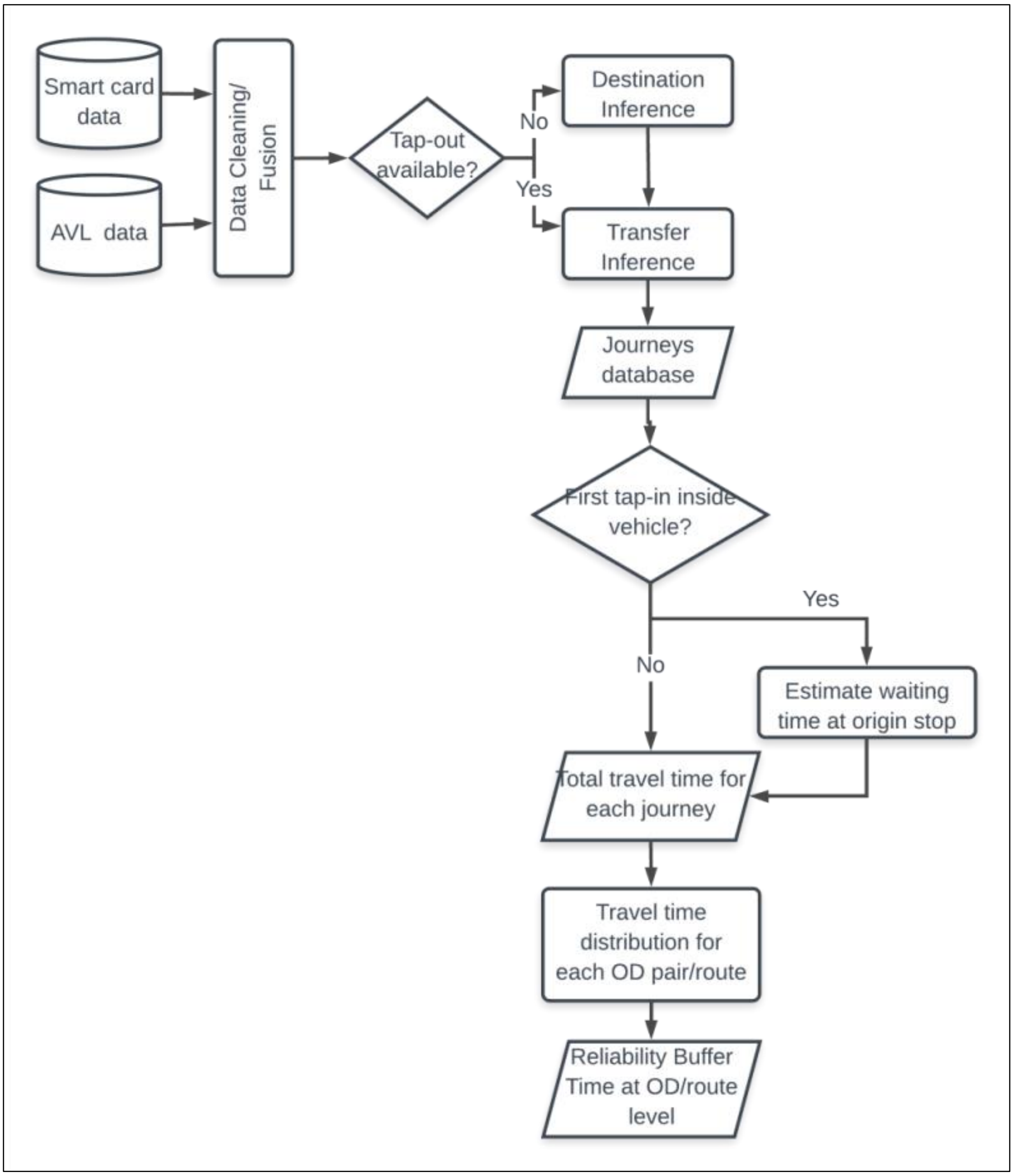


Figure 2.2 Analysis steps to derive RBT from raw smart card and AVL data.

2.3 Application

2.3.1 Case study description

The proposed method is applied to a real-world case study consisting of bus, tram and metro lines operated by GVB – the urban transit operator of Amsterdam. The study area consists of 4 metro lines, 15 tram lines and 25 bus lines spread over 1,282 stops, shown in **Figure 2.3**.

Figure 2.3 Amsterdam transit network (Source: GVB, 2018).

The smart card and AVL data set used for the analysis are for two weekdays (1st and 2nd March 2018), and consists of more than 750,000 transactions per day for each day spread over more than 80,000 OD pairs. The average frequency for metro in the dataset is 6-8 trains per hour per direction during 7am to 7pm. For bus, this number ranges between 4-10 and for tram between 5-12 vehicles per hour per direction. Four bus lines (29, 231, 240, and 248) with a frequency of less than four vehicles per hour have been removed from the dataset used for the analysis.

The planned headways for the lines are found to be homogenous throughout the day, and the RBT analysis for this study has been performed for the entire day (7am – 7 pm). From the realized headways it can be observed that in regular situations during transit operation in the study area vehicle bunching is not common. The OD/route combinations with less than twenty observations during this period have been excluded from the analysis.

2.3.2 Implementation

Each record/transaction of the smart card data received from the operator consists of a combination of tap-in and tap-out, which in case of buses and trams includes one leg of the journey only (without transfers). For metro it may include transfers within the metro, because passengers transfer without using their smart card. Each record consists of boarding and alighting times, locations and the mode used. For bus and tram trips, the line number and vehicle number are also provided. The data however does not provide a smart card id due to privacy restrictions, implying that the transactions cannot be tracked within the day. Instead, a journey id is provided, which combines individual transactions based on the transfer inference criteria applied by the operator which identifies any transaction by the same user within 35 minutes of the previous transaction as a transfer.

Data Cleaning

The raw smart card and AVL data are first cleaned to remove erroneous or incomplete transactions such as unrealistic travel times, departure time before arrival time, missing origin and destination information etc. (7.4% in the dataset). For metro trips, some extreme values of travel times were observed in the data, possibly due to passengers taking the wrong train or waiting for a friend at the platform. To ensure that such passenger behavior does not lead to unrealistic reliability measurement, the following procedure was applied to identify and remove records with odd passenger behavior for metro, ensuring that large disturbances are retained:

1. For each OD pair, select the records for which the travel time was more than double of the median travel time and which exceeded more than 15 minutes of the median travel time for that OD pair. This value was decided by observing the outliers in the data, taking into consideration both very short and very long metro journeys.
2. For each selected record, check if there was another record in the smart card data that started after the tap-in and ended before the tap out of the selected record. If the difference between tap-outs of these two records is more than one headway (10 minutes) - the selected record is considered an outlier and removed from the data set. This resulted in the removal of 0.25% of metro trips.

Data Fusion

Next, the individual transactions in the AFC data are matched with the AVL data to obtain the actual vehicle arrival and departure times at stops for bus and tram trips. Since the smart card data for Amsterdam does not provide a vehicle trip number which can be matched directly to the AVL data, the matching is undertaken based on the vehicle number and boarding time and location. If a tap-in time lies between the arrival of a vehicle at the boarding stop and the arrival of the vehicle at the next stop on that line, the passenger trip is assigned to that vehicle trip id. For the first stop in a vehicle trip, a buffer time of 5 minutes before the departure of the vehicle is considered for assigning the passenger trip to that vehicle trip. With this algorithm, 92% of the bus and tram trips could be matched to a corresponding vehicle trip; 88% of which could be matched based on destination also. It is noted that there are other more rigorous methods available to match the remaining smart card data to AVL data (Luo et al., 2018), but since that is not the focus of this work, this was not undertaken. Instead, for the passenger trips where a corresponding vehicle trip could not be found in the AVL data, the tap-in and tap-out times have been considered as the trip start and end times.

Although the Dutch smart card system requires tap-in and tap-out for all modes, ~3.5% of transactions have missing tap-outs for buses and trams. In the absence of the smart card id, it is not possible to infer destinations for these records, and they are hence removed from the data.

Transfer Inference

Since the purpose of this study is reliability measurement, it is crucial that this step is carried out accurately as an incorrectly classified transfer may add extreme values of travel times, increasing thus the measured reliability. The transfer criterion of 35 minutes applied by the operator is very generous and may include some activities wrongly classified as transfers. Hence, this study applies four additional transfer inference criteria derived from (Gordon et al., 2013; M. D. Yap et al., 2017):

1. Two consecutive journey legs on the same line in either direction are not classified as a transfer.
2. Only those legs are considered as a transfer where the passenger boarded the first vehicle which arrived after passenger reached the next boarding stop. This has been calculated by estimating the walking distance as $\sqrt{2}$ times the Euclidean distance between the two stops. The 2.5th percentile of walking speed (Hänseler et al., 2016) has been assumed to ensure that this criteria does not eliminate passengers with walking speed on the lower side. Additionally, a minimum transfer buffer time of 5 minutes is applied.
3. Transfers occurring with an Euclidean distance of more than 750m (Gordon et al., 2013) between the two stops are not considered as transfers; and
4. A circuity ratio of more than 2.5 is classified as an activity. This has been applied to prevent back and forth trips on different but parallel lines from being classified as a transfer.

Waiting Time Distribution

The waiting time for journeys with bus/tram as the first mode is estimated by assuming uniform arrivals over the observed headway for each vehicle trip for each stop. For the vehicle trips where the observed headway was not available (such as the first vehicle trip in the day), the headway was assumed as the average of the observed headways during the hour. Additionally, for the stops with headway longer than fifteen minutes for a line, such as when a stop was skipped during certain runs, the waiting times have been distributed over fifteen minutes (maximum planned headway) only.

RBT calculation

With all the components of travel time available for each journey, RBT is calculated for each stop-stop pair and route combination where a minimum of 20 observations (journeys) has been recorded. For Amsterdam, this represents 673,767 journeys spread over a total of 7,531 OD/route combinations.

2.4 Results and discussion

2.4.1 Reliability per mode

The RBT is calculated for each mode combination available in the data (**Table 2.2**) as the demand weighted average of RBT for each OD pair/route, as shown in Equation (2.4). Based on the observed data, journeys with only metro are found to be the most reliable, followed by the single leg journeys using bus or tram modes. Due to separate right-of-way and no disruptions observed during selected days, there is negligible variation in in-vehicle time component for metro journeys and the RBT is primarily contributed by the variation in waiting time component of the journey (including at transfers).

The tram network in Amsterdam serves the congested city center with mixed traffic, including bikes, and faces high passenger demand potentially causing lower reliability levels. Buses on the other hand, tend to run on less central streets with lower demand levels and hence less crowding variation, making them relatively more reliable for single leg journeys. However, the frequency of buses are typically lower – making the journeys with intra-modal transfers less reliable for buses than for trams.

Journeys with inter-modal transfers do not show major variations between different mode combinations. The fact a transfer is involved increases RBT, with a lesser importance to the specific combination of bus, tram or metro between which the transfer is made. It should be noted here that since RBT is measured in absolute terms, it is expected to be higher for journeys with transfers due to typically longer journey lengths.

Table 2.2. RBT for different mode combinations.

Mode(s) used	Number of journeys	Number of OD pair/route combinations	Median travel time (mins)	RBT (mins)
Unimodal				
Metro (+ Metro-Metro)	235,287	1,189	14.7	5.9
Tram	315,410	4,094	15.4	6.6
Bus	104,495	1,703	14.8	6.2
Tram-Tram	1,755	60	23.2	7.2
Bus-Bus	130	5	20.5	9.1
Multimodal				
Metro-Tram	7,588	213	25.0	7.6
Metro-Bus	747	26	28.8	7.8
Tram-Metro	6,665	179	26.3	8.3
Tram-Bus	115	5	21.6	8.3
Bus-Metro	1,336	48	28.7	8.5
Bus-Tram	239	9	24.8	7.9

Differences are also found in the distribution of RBT across OD pairs by mode(s) used (**Figure 2.4**). Only the mode(s) with a minimum of 40 OD/route combinations have been presented here. It is observed that the spread of RBT values across OD pairs is much wider for journeys with transfers compared to the ones without a transfer, possibly due to longer journey lengths for transfer journeys. Routes connected by single leg bus journeys are found to have not only the lowest average RBT value but also the lowest variation of RBT across different routes. The largest spread of RBT is seen for the tram-tram transfer journeys. It is noted here that **Table 2.2** shows the RBT as the demand weighted average whereas **Figure 2.4** shows the spread of RBT over OD pairs. Hence, although RBT for metro seems to be higher when looking at the distribution across OD pairs, more passengers seem to use the lower RBT metro routes, bringing down the overall RBT (**Table 2.2**).

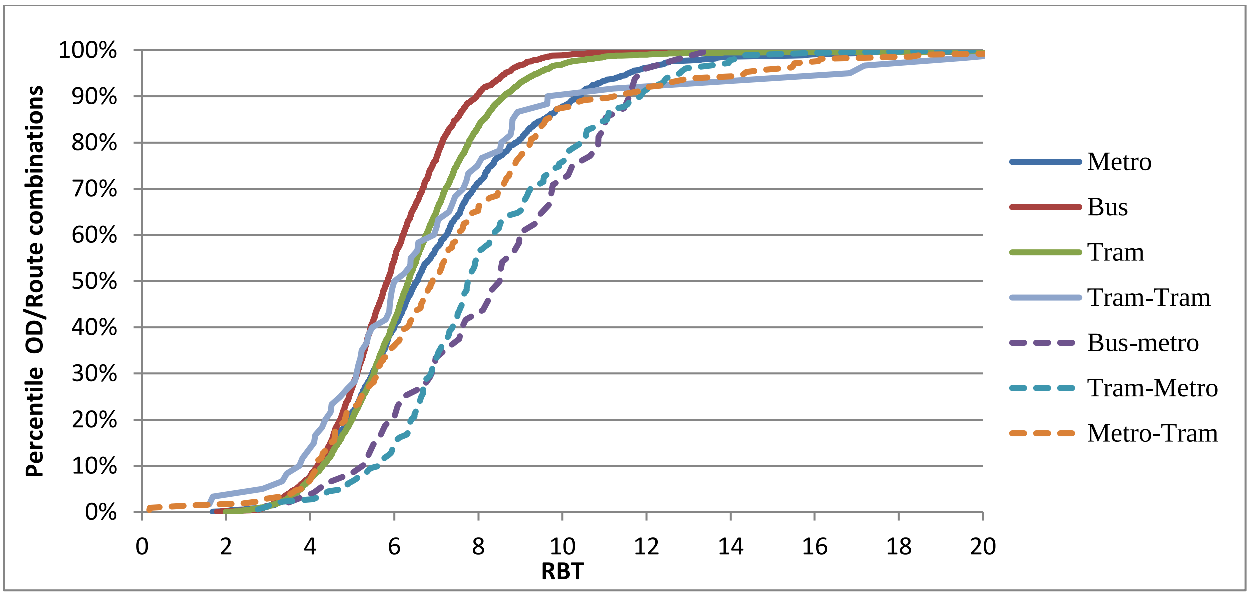


Figure 2.4 Distribution of RBT across OD pair/route combination by mode(s) used.

2.4.2 Reliability of accessing transit hubs

The developed RBT can be used to analyze the reliability of a transit stop to/from all other origins/destination transit stops, as obtained using Equation (2.2). **Figure 2.5** shows the spatial distribution of RBT values for journeys from various transit stops to two major train stations (Amsterdam Central and Sloterdijk) by various modes. The size of the circles represents the RBT from that origin to the selected train station.

The Sloterdijk station is situated outside the city center. Consequently, it can be reached in a relatively reliable way, from all directions and with all modes. In the South and in the Southeast of the city two metro branches can be observed that are less reliable, which makes sense since a transfer is needed to reach Sloterdijk from these branches. Also, the combined mode of tram and metro seems relatively less reliable for Sloterdijk.

To and from the Central Station most destinations are reached without a transfer. The metro branches to the Southeast of the city as well as the buses to the North are found to be relatively reliable. Reliability by trams seems relatively lower except for tram line 26 to the East of the city which avoids the crowded city center.

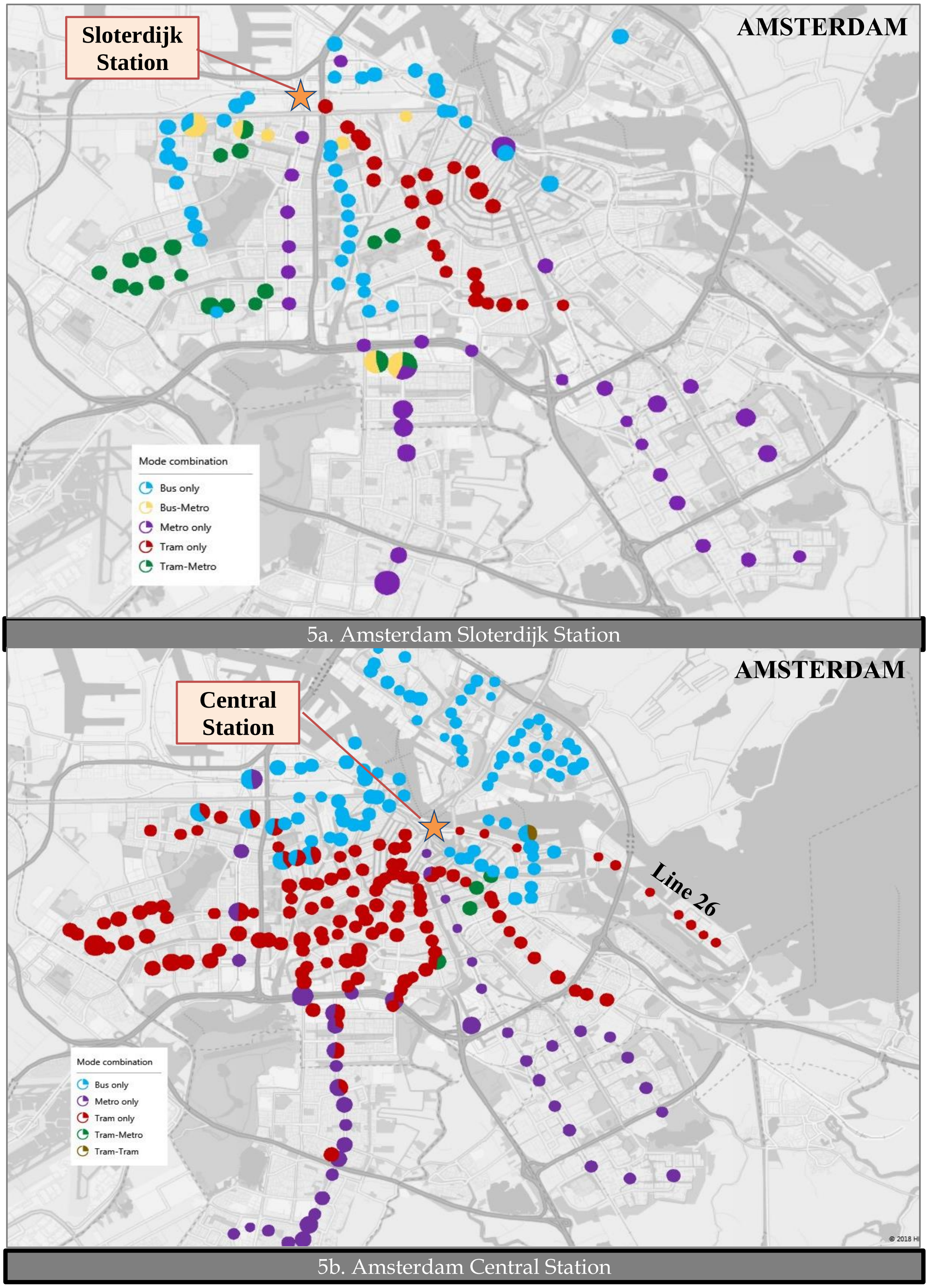


Figure 2.5 Reliability to transit stops using different modes.

2.4.3 Reliability per route

Next, the RBT for different routes for the same OD pair is investigated. Four route alternatives are available between the origin-destination pair of Station Sloterdijk to Boelelaan (**Figure 2.6**).

One could either take a metro (with a transfer at Amsterdam Zuid station), or take the metro till Amstelveenseveg station and from there bus 62 or trams 16 or 24 could be taken.

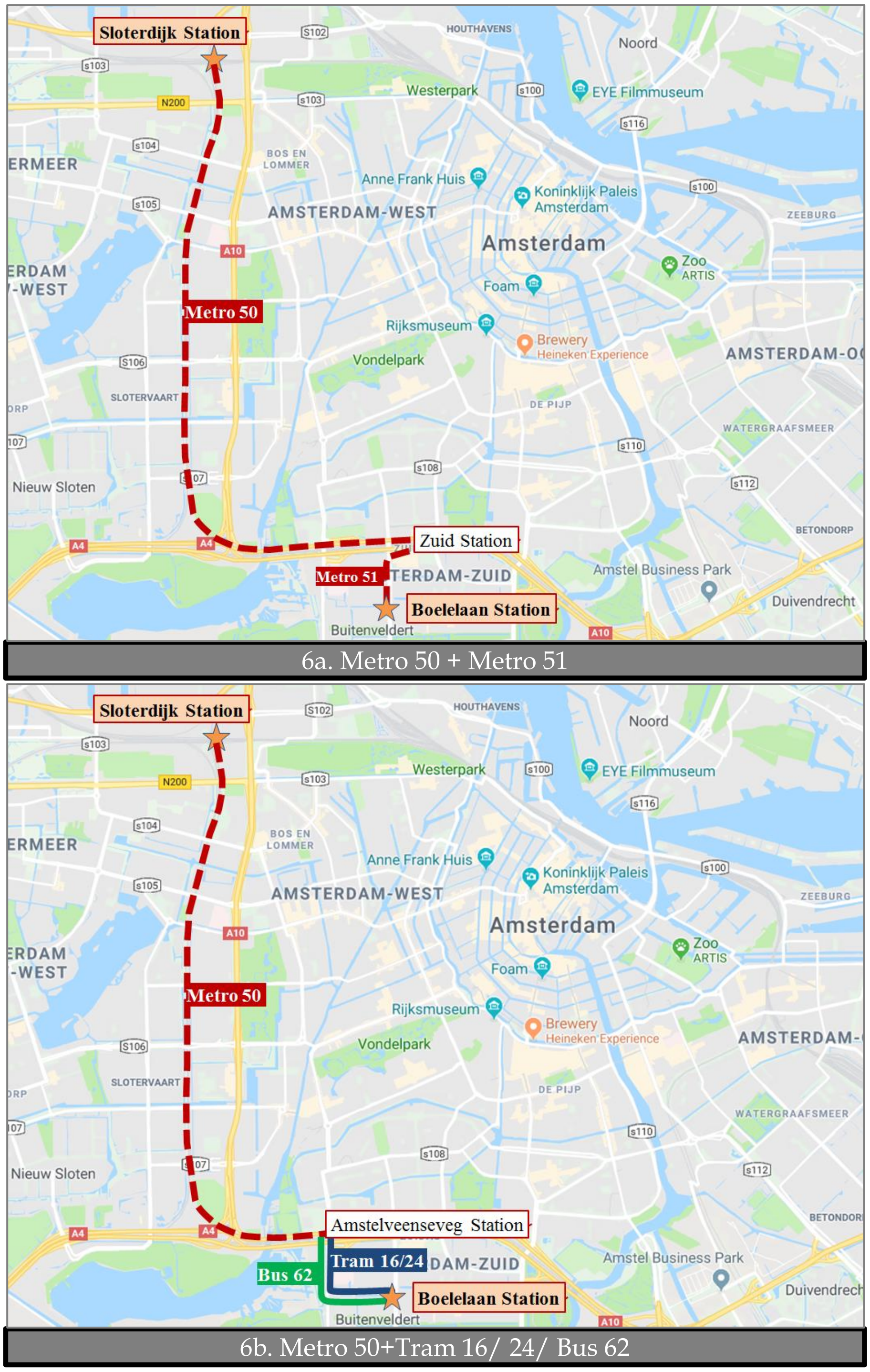


Figure 2.6 Observed passenger routes from Sloterdijk to Boelelaan Station.

Based on the data, it is seen that RBT is in this case in fact the highest for metro and lowest for the route with metro and bus together (**Table 2.3**). This example highlights that aggregating RBT across routes/modes/OD pairs ignores the variations observed between different areas and routes. Measuring RBT at a route level gives more detailed and actionable results of reliability.

It is also noted that the number of journeys are least for the metro-metro route. This is expected as not only the RBT but the median travel time is also the highest for this route. Tram routes 16 and 24 overlap, which is also reflected in the similar travel times and RBT for these routes.

Table 2.3. RBT per route for station Sloterdijk to Boelelaan.

Journey Origin	Journey Destination	Route	Number of journeys	Median (min)	95th Percentile (min)	RBT
Sloterdijk	Boelelaan	Metro-Metro	25	26.0	36.8	10.8
		Metro-Bus 62	66	22.4	27.8	5.4
		Metro-Tram 16	217	23.1	32.2	9.1
		Metro-Tram 24	105	22.9	31.4	8.6

Looking at the journey time distribution for different routes (**Figure 2.7**), it is noted that the route with metro and bus services has a steep slope of travel time indicating a more reliable service. The metro route on the other hand has a jump in travel time just before the 90th percentile value –potentially due to the passengers missing the metro at the transfer station and having to wait another headway (10 minutes).

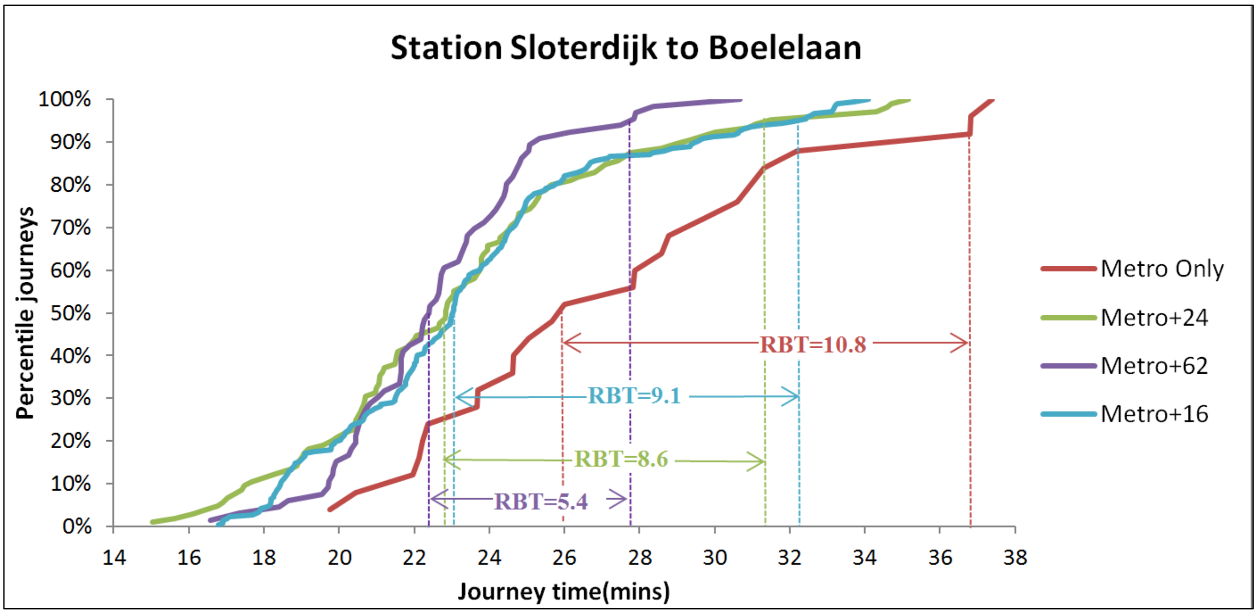


Figure 2.7. Journey time distributions and RBT for different routes from station Sloterdijk to Boelelaan.

2.5 Conclusions

This research proposes a new metric for travel time reliability measurement, considering multimodal transit journeys, including waiting and transfer times for all legs of the journey. The developed metric can be calculated using a combination of smart card and AVL data, which was demonstrated by applying it to the Amsterdam transit network. Since the chosen smart card data does not measure the waiting time for bus and tram journeys, a method to estimate the waiting time for each journey starting with these modes was proposed, based on the observed headway from AVL data. Based on the semantics of the smart card system applicable, the method can be modified and applied to other networks.

Since the reliability metric is developed at a very disaggregate level (i.e. for each OD pair/route combination), it provides flexibility of aggregation across various dimensions depending on the goal. The case study demonstrated its application by aggregating across three dimensions – mode, transit stop and routes.

Aggregating the RBT at a larger scale such as at the mode level provides with an overall picture of reliability. This can for example be used for reliability impact analysis of policies affecting one or more transit modes. In the Amsterdam case study, it was observed that for single leg journeys, trams in the city centre are the least reliable. For multiple legs however, a bus to bus transfer was found to be the least reliable of all, possibly due to the longer headways and consequently longer transfer times due to missed connections.

However, aggregating at such a large scale ignores the variations in reliability between different OD pairs and routes. For example, in the case study it was seen that for the journeys from Sloterdijk to Boelelaan station, the RBT was highest for the route with metro only compared to the other three routes consisting of metro in combination with either bus or trams. This type of information can be used to address specific unreliability issues for a particular route/OD pair. Route level RBT can also be used as a direct input to behavioral models such as route choice models.

The RBT can also be compared for different origin/destinations from a selected transit stop/hub. This can for example be used to analyse from which locations and using which modes is the unreliability largest, providing policy makers with information on urgency of reliability issues across the city.

The case study demonstrated some of the potential applications of the developed method. However, due to the low sample size (only two days), the analysis was undertaken for a longer time period within the day (7am to 7pm). As future work, this method could be applied to a larger dataset enabling analysis at an hourly level. The RBT can then be used to compare progress in performance over time or between different time periods during the day.

Based on the available data, simplifications were made leading to some limitations of the work. Firstly, it was assumed that the passengers boarded the first vehicle that arrived in case of bus and tram modes. In case of overcrowded vehicles (for example due to vehicle bunching) in reality passengers may prefer to take the next arriving vehicle. However, for metro the time measured by smart card already includes the potential delay due to denied boardings. Additionally, this study did not consider the impacts of availability of real-time information on passenger arrivals and their waiting time distribution. Although only short headway services were considered, it is common for passengers to consult the real time arrival information before arriving at the transit stop. Further research could focus on addressing these limitations of the analysis.

Chapter 3 - Examining Circuity of Urban Transit Networks from an Equity Perspective

In chapter 2, we focused on travel time reliability and developed a measure using smart card data that can be applied to multi-modal transit networks. In this chapter, we explore another service quality aspect – the circuity of transit networks. Defined as the ratio of the network to Euclidean distance traveled, circuity has been known to influence travel behavior. In addition to the longer time spent in travel, for networks where the fare is based on distance traveled, higher circuity also means higher fare for the same Euclidean distance. This makes circuity relevant from an equity perspective.

Using a case study of the urban transit network of Amsterdam in the Netherlands, this chapter explores the role of transit circuity on the disparity in distance traveled by travelers' income profiles and its implications on travel times and costs for networks with distance-based fares. This enables us to characterize the contribution of transit network design in determining the equity outcomes in a network, and how it could exacerbate or reduce the existing disparities in distance traveled in the network.

This chapter is based on the following article:

Dixit, M., Chowdhury, S., Cats, O., Brands, T., van Oort, N., Hoogendoorn, S. (2021) Examining circuity of urban transit networks from an equity perspective. *Journal of Transport Geography*, 91.

3.1 Introduction

Transit networks are often optimized to maximize directness and minimize transfers (Zhao and Ubaka, 2004) while minimizing costs and travel times. Circuity[2] is defined as the ratio of the network and Euclidean distances between an origin-destination (OD) pair (Barthélemy, 2011), and is a popular measure to quantify the directness of road and transit networks. Circuity of transit networks has been found to influence travel behavior at various decision-making levels. Lee et al. (2015) studied five Korean cities and found evidence of a strong relationship between circuity and transit ridership. At a long-term decision level, Levinson and El-Geneidy (2009) found in their study of twenty US cities that people tend to locate themselves in areas with smaller circuity for home-work trips – with the circuity of used routes being smaller than randomly selected routes in the network. At a short-term level, Huang and Levinson (2015) found that circuity can explain the mode choice of commuters in Minneapolis-St. Paul, Minnesota – a low transit mode share was found to be associated with higher circuity. Transit circuity was also found to explain transit route/path choice in some studies (Kim et al., 2019; Raveau et al., 2014). Such a direct relationship with travel demand makes circuity an important transit performance measure.

In the case of transit routes that follow the road network, circuity is a function of the street network layout. In addition to the circuity of individual transit lines, service network structure and transfer locations also impact the circuity of journeys experienced by passengers. For example, radial networks are expected to have a higher circuity for journeys between two suburbs that require transferring in the core compared to tangential or ring networks that may provide a direct connection. Sometimes transit agencies intentionally design routes with high circuity to maximize coverage, even though it may discourage ridership on those routes (Huang and Levinson, 2015).

It is common for transit networks to be designed based on efficiency and demand, without explicitly focusing on the equity aspect (Soltani and Ivaki, 2011). Such a design may end up favoring a particular section of the population (high income) over others. This is particularly true for mono-centric European cities where low-income residents typically live away from the city center in areas with lower population density, leading to more circuitous routes. Transit routes with higher circuity imply a longer (network) distance traveled for the same Euclidean distance covered. The impact of this on passengers is two-fold. Firstly, longer network distance results in longer travel times for passengers, all else being equal. Secondly, for transit networks where the fare is calculated based on network distance traveled (such as in Amsterdam, and Beijing metro), circuity directly impacts the fare paid by travelers. Essentially, travelers using highly circuitous routes end up paying more for a worse-off connection. Hence, in such networks, an uneven distribution of circuity can result in an uneven distribution of both travel times and fare paid per Euclidean distance covered. Both aspects make circuity relevant from an equity perspective. However, to the authors' knowledge, there is limited research on the distribution of circuity observed within a transit network, and its impact on travelers from different population groups.

This study investigates the role of transit circuity on the disparity in distance traveled, and its implications in terms of travel times and costs for three income levels. This is done by

[2] Circuity is similar to the road 'detour factor' as introduced by Cole and King (1968). However, in transit literature, Circuity is a more prevalent term. Hence we have used the term 'Circuity' in this dissertation.

undertaking an empirical assessment of circuity for the urban transit network of Amsterdam using a combination of anonymized smart card data (which contains automatic fare collection (AFC) data), and automatic vehicle location (AVL) data. Based on the information on circuity for all transit journeys made within the network (by metro, tram, and bus), the study addresses the following questions for the case study system:

- Do travelers from lower income areas have more circuitous transit journeys?
- What is the contribution of circuity to the distribution of distance traveled by different income demographics?
- What implications does this have on the travel times and fare paid by them?

3.2 Literature review

Transport equity is a complex topic with multiple definitions and interpretations. For this study, one of the commonly used definition in transportation studies, the 'fairness in distribution of impacts', is adopted (Litman, 2002). Martens et al. (2019) highlight three key components of a transport equity analysis: defining what impacts (burdens or benefits) are considered, which population or social groups are they distributed over, and what constitutes as being fair. The literature so far has included a wide range of impacts associated with transport provision: road and transit network supply (Ahmed et al., 2008; Delbosc and Currie, 2011), environment and health externalities (Feitelson, 2002), travel costs, taxes and subsidies (El-Geneidy et al., 2016; Eliasson and Mattsson, 2006; Pucher, 1981), and access to jobs and other opportunities (Guzman et al., 2017; Neutens et al., 2010a). Further, there are a range of groups emphasized in equity analyses, including but not limited to genders, income classes, and spatially, mentally or physically disabled groups. Litman (2002) classifies equity in two types - horizontal and vertical. Horizontal equity refers to fairness between individuals of the same ability, income and social class. Vertical equity includes fairness between individuals across different abilities, income and social classes.

Accessibility has been one of the most common impacts (benefit) of transport that is subject to a transport equity analysis. This is because any change in policy or intervention has an impact on accessibility, both short and long term. While most of the research on accessibility focuses on travel times or distances, travel cost has also been recognized as an important barrier to transport access (Foth et al., 2013; Kaplan et al., 2014; Pritchard et al., 2019). Hence, it is typically included in equity evaluations, either exclusively (El-Geneidy et al., 2016), or along with other factors (Currie, 2004).

Several studies (Bandegani and Akbarzadeh, 2016; Brown, 2018; Farber et al., 2014; Rubensson et al., 2020) have investigated how fare is distributed across population groups, and evaluated the impact of alternate fare policies on equity. In Utah, Farber et al. (2014) found that lower socio-economic groups tend to travel shorter distances with high ridership – making distance-based fare policy more vertically equitable than zonal fares. Similarly, in Toronto, Foth et al. (2013) found that residents in lower socio-economic areas had shorter travel times due to proximity to city center. In contrast, Rubensson et al. (2020) noted that for Stockholm, lower income travelers made a higher proportion of longer journeys, for which distance-based fare was vertically inequitable. As highlighted by them, the equity outcome of a fare policy is dependent on the geographical distribution of income levels, land-use and travel patterns.

In many European cities, the city center typically has better access to amenities, which increases land value in close proximity to it. This results in a decline in income with increasing distance from the center (Brueckner et al., 1999). With this pattern of income distribution, low income residents need to travel longer (Euclidian) distances to reach the city center, where most opportunities are located. In addition, the disparity in (network) distance traveled could be either alleviated or exacerbated by differences in circuity of transit routes serving different areas. The variation in distance traveled is expected to be a combination of these two effects.

Figure 3.1 shows the relationship between the land-use patterns, transit network design and the outcomes of fare paid and travel times observed in the network. The socio-demographic characteristics of a person impact the need for travel. Observed travel behavior in the network is a function of both land-use and transport network. Examining the factors separately can help to provide tailored solutions for addressing each of these issues based on their respective contributions. However, the literature to date has primarily focused on the contribution of land-use patterns to distance traveled, and not enough attention has been given to the contribution of transit network design. Our study aims to address this gap by examining the contribution of circuity in the distribution of distance traveled and in turn the fare paid and travel times.

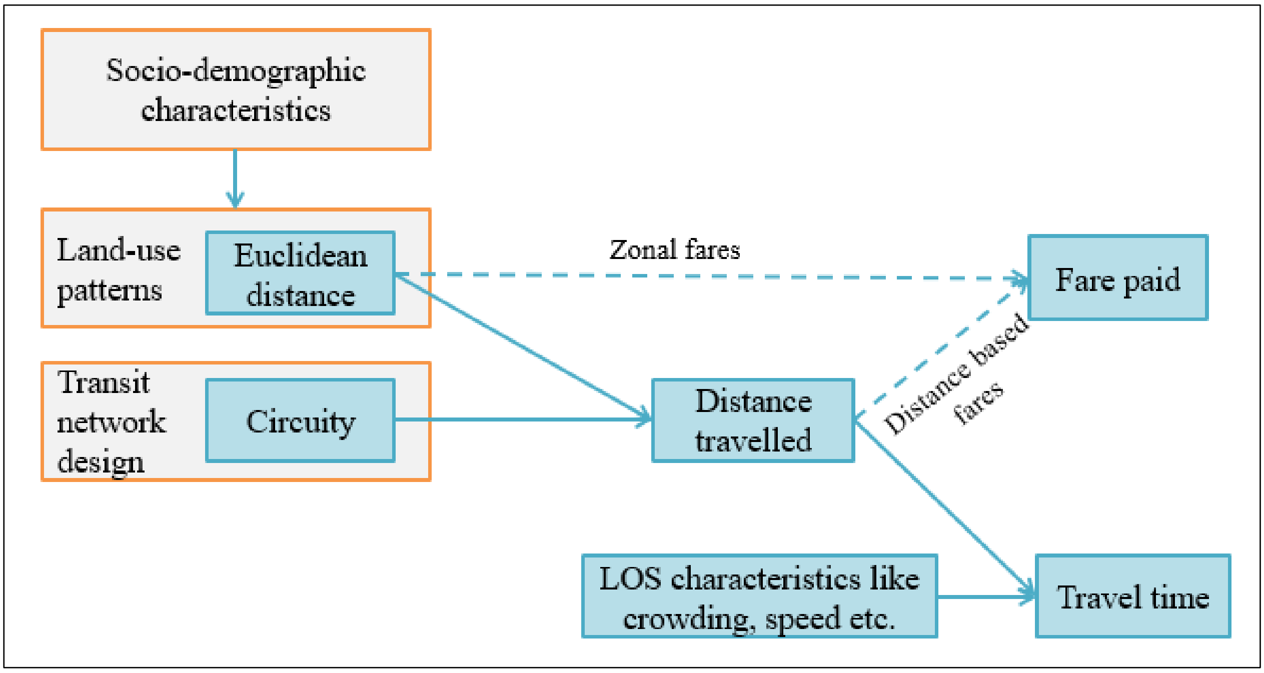


Figure 3.1. Relationship between circuity of a network, travel times and fare paid.

A key question underlying all equity analysis is how fairness is defined. Carleton and Porter (2018) emphasize that most transport equity studies measure the level of equality. To move from equality to equity, it is paramount to define what is considered fair, for which several, often conflicting theories of justice exist. Pereira et al. (2017) provide a detailed review of these theories in the context of transport. We start by measuring the levels of equality in the current distribution of circuity in the network, and its contribution to the (in)equality of distance traveled in the network. We specifically focus on measuring vertical equity by investigating whether the distribution favors an income group. Next, by the means of Gini index, horizontal equity of distribution of circuity in the network is analyzed. However, we refrain from giving absolute judgements on equity, with respect to suggesting appropriate corrections for mitigating inequity concerns, which will depend on the specific theory of justice chosen to be followed.

3.3 Method

3.3.1 Transit Circuity

Transit circuity of a (passenger) journey is calculated as the ratio between the network distance traveled and the Euclidean distance between the origin and destination of the journey. A journey may or may not include transfer(s) within or across transit modes. A route is defined as the combination of transit lines and transfer stops used by a passenger in his/her journey. Mathematically, it can be expressed as,

$$C_{o,d,r} = \frac{\sum_{l=1}^{L_{o,d,r}} x_l^n + \sum_{l=1}^{L-1} x_{l,l+1}^t}{x_{o,d}^e} \quad \forall o,d,r \tag{3.1}$$

Where,

$C_{o,d,r}$ is the circuity for a journey between origin-destination transit stops *o,d* using route *r*;
x_l^n is the network distance traveled on leg *l* of route *r* between *o,d*;
$x_{l,l+1}^t$ is the transfer distance between leg *l* and *l+1* of route *r* between *o,d*;
$x_{o,d}^e$ is the Euclidean distance between *o,d*; and
$L_{o,d,r}$ is the number of legs in the journey between *o,d* using route *r*.

Figure 3.2 shows a schematic representation of it.

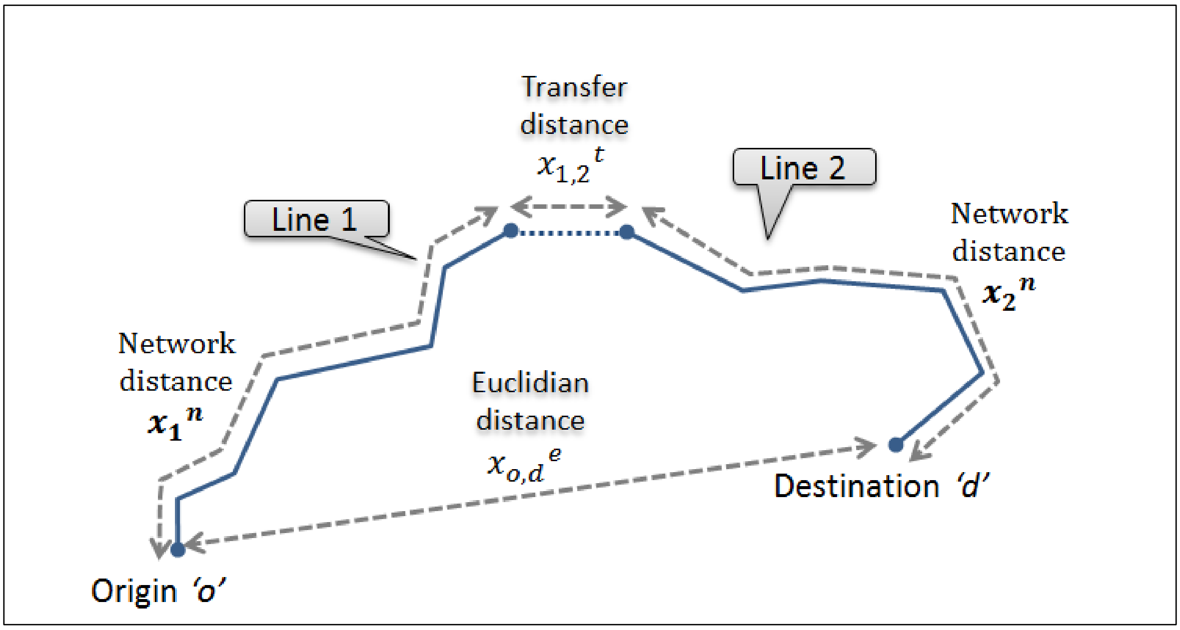


Figure 3.2. Schematic representation of circuity measurement for a two-leg transit journey.

The term '*realized transit circuity*' is used to indicate the circuity that is obtained based on the actual routes used by the journeys made in the network (as opposed to potential ones based on shortest path between an O-D pair). The information on journeys made in the network is obtained from smart card data. The subsequent sections first describe the case study network, followed by how the smart card data is processed to obtain the realized transit circuity, and how it is linked to the income data to facilitate an equity analysis.

3.3.2 Introduction to case study network and data sources

The analysis is performed for the urban transit network of Amsterdam (**Figure 3.3**), and includes all bus, metro and tram lines operated by GVB, the transit network operator of Amsterdam. The time period for analysis is spring 2018 (28th May – 1st July). During this time, 41 bus lines, 15 tram lines and 4 metro lines were operational. The city center of Amsterdam is served by a dense network of tram lines, mainly connecting the center with large residential areas. The metro provides connections between the south-eastern suburbs and the city center, and a ring line to the west of the city. The bus completes the network, mainly to and from the northern part of the city, as feeder links to the metro, and some tangential and some radial services where tram and metro services are missing.

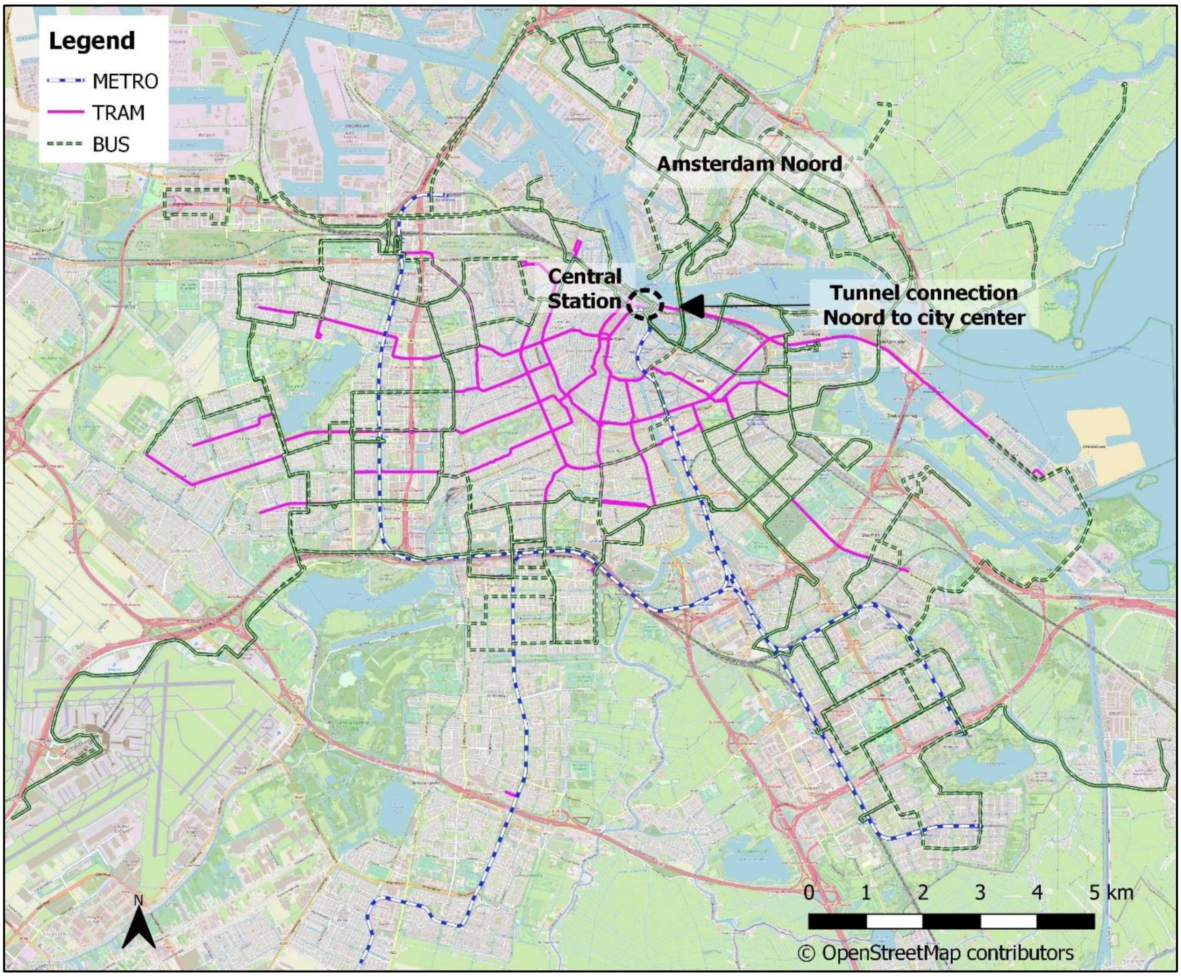


Figure 3.3. Urban transit network in Amsterdam.

This study uses a combination of anonymized smart card data and automatic vehicle location (AVL) data to obtain information on the routes used for all transit journeys made in the network. The Dutch smart card (called OV-chipkaart) records information on both check-in and check-out for all modes (for more information see van Oort et al. (2015a)). For the urban transit network of Amsterdam, it provides approximately 675,000 transactions per day on average for the study period. The AVL data is publicly available for all transit modes in the Netherlands (see van Oort et al. (2015b) for more details).

The smart card data used in this study does not provide any information on the socio-demographic characteristics of the traveler or the type of fare paid. Hence, we use the income data from Central Bureau of Statistics (CBS) Netherlands (2020a), where this information is available at a neighborhood level (with 470 neighborhoods in Amsterdam). Two relative income indicators per neighborhood have been used for this study: the share of persons belonging to the top 20% or the bottom 40% of the national personal income distribution (Bresters, 2019).

3.3.3 Data processing steps

The first step in data processing is to convert raw (anonymized) smart card transactions to linked trips (or passenger journeys). This process is described as below:

1. Data cleaning: The smart card and AVL data were first cleaned to remove incomplete, invalid or unrealistic records (~3.3%).
2. Destination Inference: This was carried out for records with missing check-outs (4.2% in the data) using the method detailed in Zhao et al. (2007).
3. Assigning journey length: For buses and trams in Amsterdam, the check-in and check-out happen inside the vehicle, and the information of the transit line used is recorded in the smart card data. Based on the origin, destination and transit line used, the network distance traveled is added for each bus and tram trip. For metro, the check-in and check-out happen at the station entrance/exit, and the information on transit line(s) used is not directly available from the smart card data. For the purpose of this research, the network distance corresponding to the shortest path is used for these trips.
4. Transfer inference: Individual smart card transactions (trips) are matched with the corresponding AVL data to identify transfers using existing algorithms (for more details see Dixit et al. (2019b)). For each journey, the network distance and transfer distance for each leg of the journey is recorded.

After processing the data and accounting for transfer inference, over 500,000 journeys per day were obtained. Once the origin, destination and route are known, circuity of each journey is calculated as the ratio of the sum of traveled (network) distance and transfer distance, and the Euclidean distance between the origin and destination stops of the journey, as expressed in Equation (3.1). Journey level circuity values are then aggregated by mode(s) used, distance traveled and origin transit stop by taking an average across all journeys.

Some journeys in the network might include unnecessary detours which are made by choice. Such detours are more likely to happen for leisure trips than for commute trips. Restricting the time period to the weekday morning peak period is expected to minimize the proportion of leisure trips. In addition, only origin-destination and route pairs with a minimum of 20 journeys over the study period have been considered, to ensure only reasonable routes are included. With this threshold, we retain 87% of the journeys made in the network.

3.3.4 Linking income with travel data

To study the relationship between income and circuity, the next step is to link the neighborhood level income data to the observed circuity data from smart card. Since the residential location is not directly available from the smart card data, it is assumed that the travelers reside in the catchment area of the transit stop from which they start their transit journey during the weekday morning peak period. Accordingly, the income characteristics of the catchment area of the transit stop have been used as a proxy for income profiles of the travelers using the transit stop during the morning peak period. An (area) weighted average of all neighborhoods within the catchment area (400 m radius) of a transit stop has been used for this. For this reason, only the journeys starting in the morning peak period (7AM to 10AM) on weekdays are considered for this study, which constitute approximately 16% of the total journeys in the processed data.

Figure 3.4 shows the resulting spatial distribution of the share of low-income persons by transit stop. The areas in the north, and south-east and west peripheries of the city have a higher than average share of low-income residents. The city center of Amsterdam has a relatively lower concentration of low-income residents. However, unlike the typical pattern of a European mono-centric city, some higher-income areas are also located further away from the city center in southern and eastern peripheries of the city.

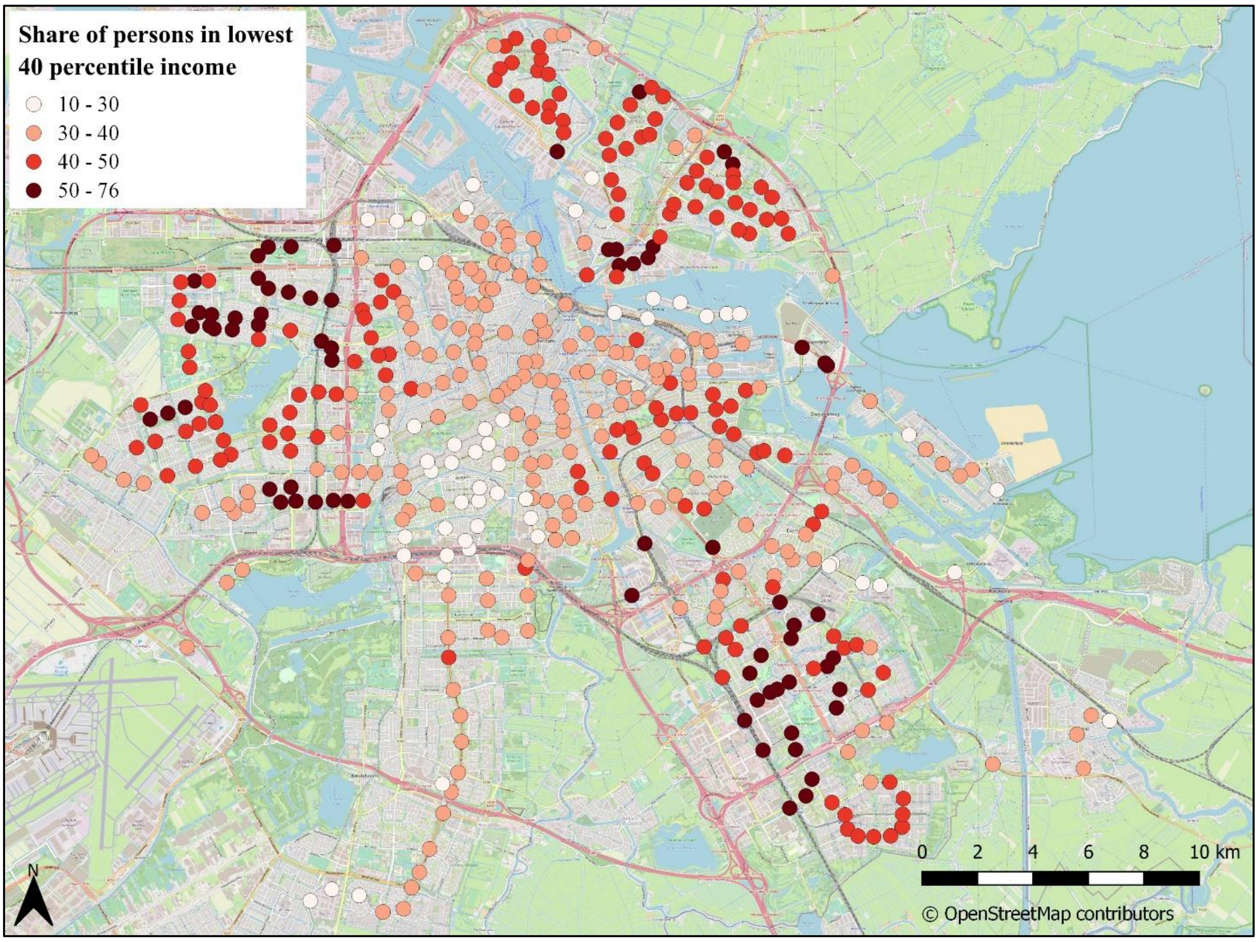


Figure 3.4. Share of people in the bottom 40% of the national personal income distribution.

3.3.5 Equity analysis

Once the income profile is assigned to each transit stop, the distribution of Euclidean distance, circuity and network distance by income is analyzed to identify patterns. Next, a multiple regression is conducted to disentangle the impact of income on circuity, while controlling for the Euclidean distance covered in the journey. First, an Ordinary Least Squares (OLS) regression was conducted, and the residual errors were tested for spatial autocorrelation. For defining neighbors, a distance based weights matrix with inverse distance weighting was used. After testing different options of distance, a threshold of 600 m was identified as providing the best results. Using the resulting weights matrix, Moran's I statistic was applied to detect the presence of spatial autocorrelation in the data. On identifying the presence of spatial autocorrelation, the Lagrange Multiplier (LM) tests were conducted to identify the appropriate spatial model. Based on the test results, Spatial Error Model (SEM) was chosen for this analysis. For more details on spatial autocorrelation and spatial models, the readers are referred to Anselin (1988) and LeSage (2008).

Finally, the equity of fare paid and travel times is evaluated using Gini coefficient (Gini, 1912). Gini coefficient quantifies the horizontal (in)equity of an outcome (an equity indicator such as accessibility), and has been a popular measure for horizontal equity analysis in transport (Delbosc and Currie, 2011; Rubensson et al., 2020). It varies between 0 and 1, with 0 signifying perfect equality, and 1 the perfect inequality where the entire outcome is concentrated with one individual.

3.4 Results

3.4.1 Transit Circuity in Amsterdam

The majority of transit trips in Amsterdam in the morning peak have a circuity of 1.4 or lower (**Figure 3.5**), with an average circuity of 1.28 for the entire network. A large share of trips (38%) in the study period are made exclusively by metro, where the network distance between subsequent stops is close to the Euclidean distance, resulting in circuity values close to 1 for shorter distances. Large circuity values also occur for metro (for longer distances) which makes the average circuity value for metro as 1.21. The circuity of bus trips is found to be 1.54, which is the highest amongst the three modes. A reason for this is that they typically run in low density areas of the city which often have indirect routes to maximize coverage. Trams on the other hand have a dense network in the city center with relatively less detours (average circuity of 1.18).

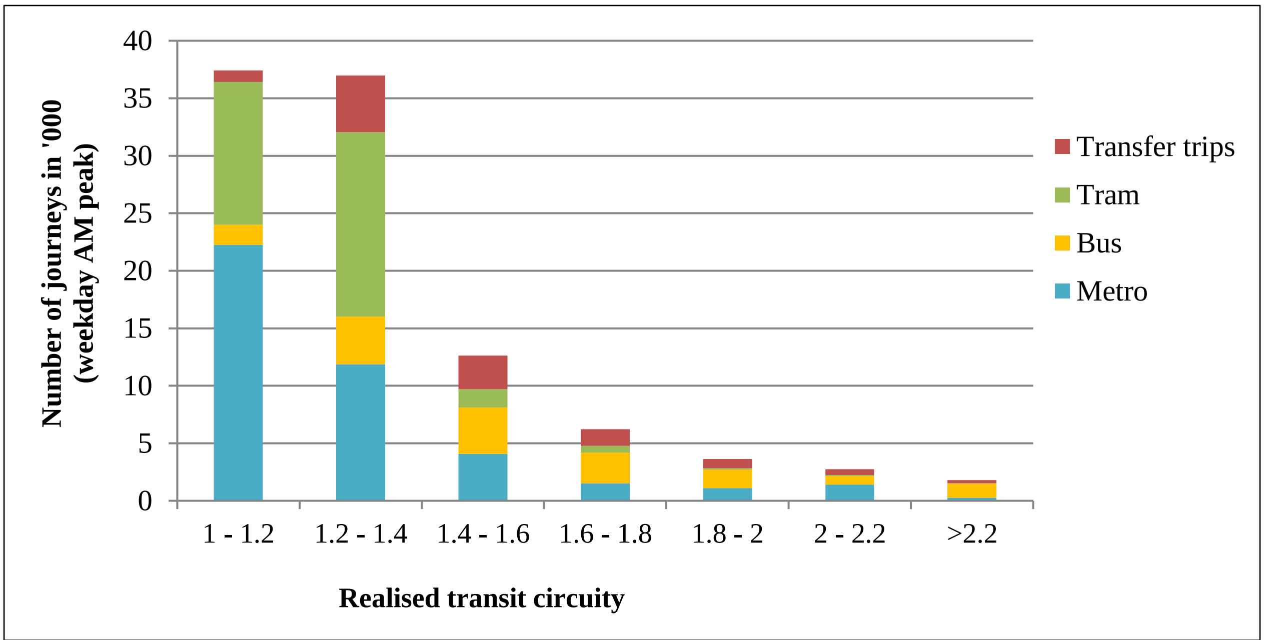


Figure 3.5. Realized transit circuity in Amsterdam.
Note: Circuity for metro includes metro-to-metro transfers

Figure 3.6 shows the spatial distribution of circuity, measured as the average circuity of all transit trips originating from a certain stop. The size of the bubbles indicates the relative number of trips originating from the respective stop. The areas to the north of the river (known as Amsterdam Noord) show distinctly higher values of circuity, with the majority of stops having a circuity of 1.6 and above. This is expected as the only transit connection from the Noord to the city center in the study period was via buses that used a single tunnel to cross the river (**Figure 3.3**). In addition to Noord, all higher circuity stops are found in the peripheral areas of the city, whereas most stops in the city center have an average circuity of 1.4 or lower. However, it is worth noting that many of the peripheral areas of the city also have a low circuity, for example those in the south-east parts of the city, due to the presence of direct metro and tram lines.

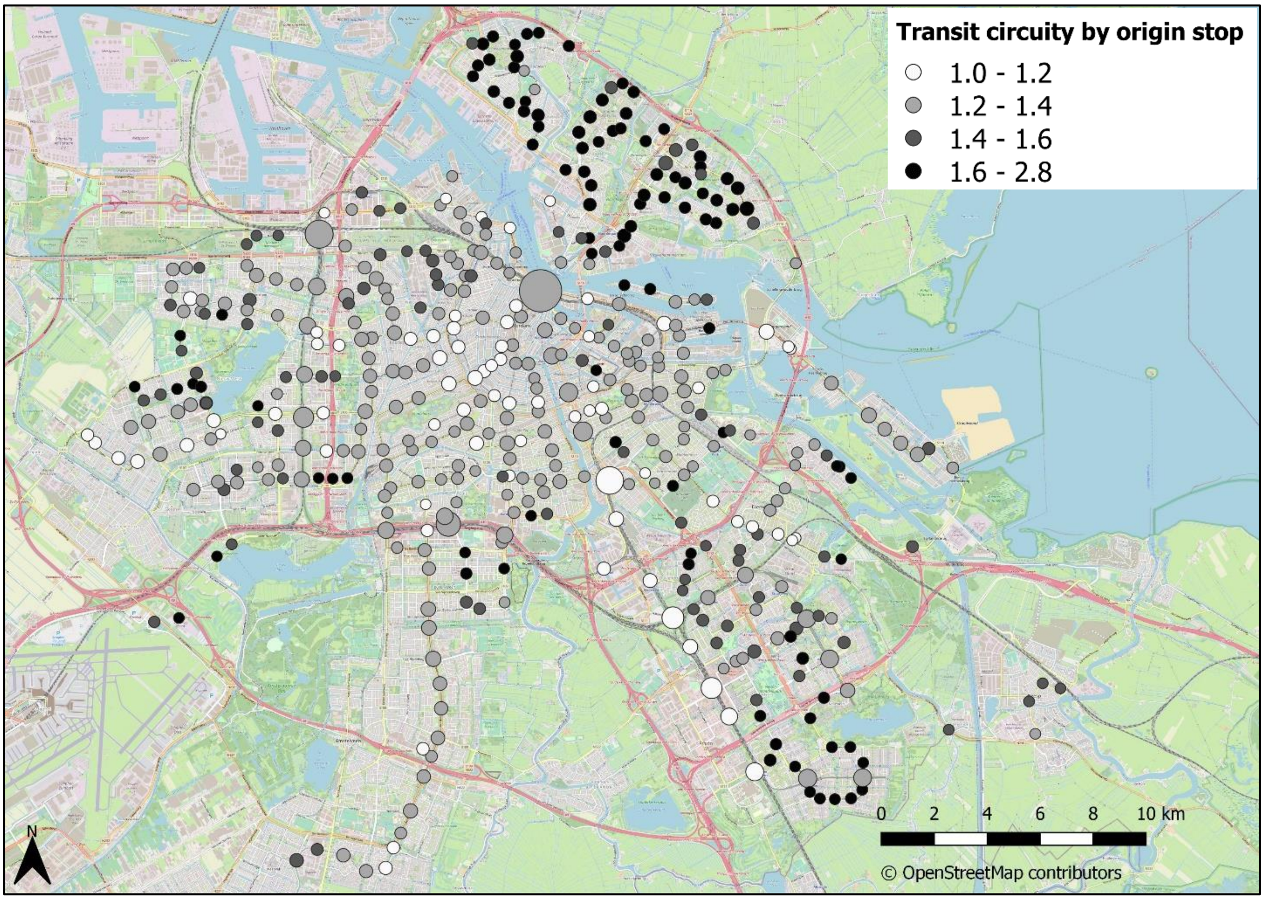
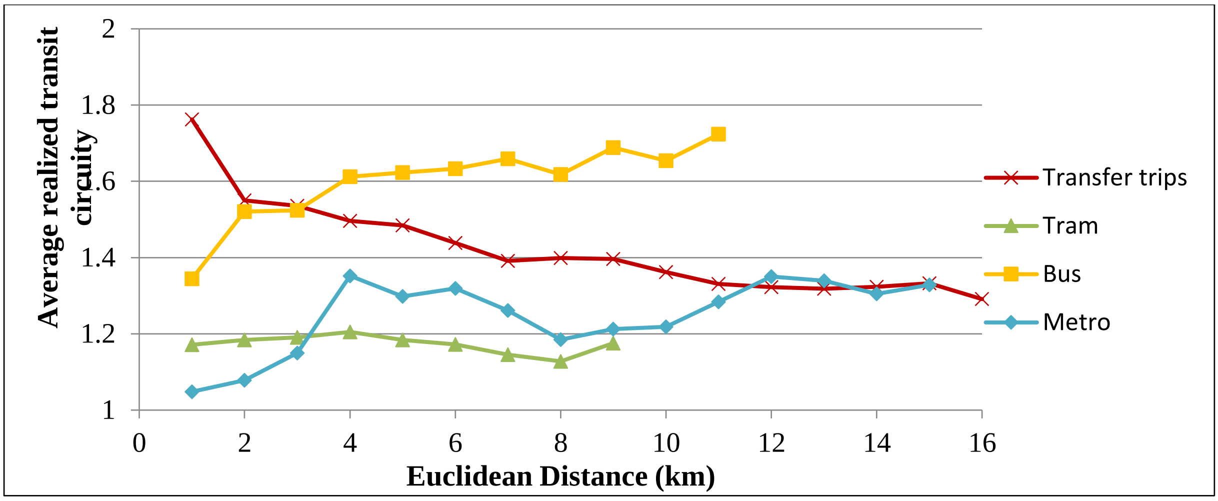


Figure 3.6. Spatial distribution of transit circuity in Amsterdam by origin transit stop.

The realized circuity increases marginally with Euclidean distance traveled for journeys without transfers, especially for metro and tram (**Figure 3.7**). On the other hand, it decreases with increasing Euclidean distance for journeys with transfers. The circuity for tram journeys is largely unaffected by journey length. As expected, metro journeys have the lowest circuity for shorter distances. The steep increase in circuity after 3 km and the drop around 8 km could be due to the circumferential nature of the metro lines in the (relatively small) network. In line with **Figure 3.5**, bus is found to be the most circuitous of the modes (including transfer trips), regardless of the distance covered, with an overall increasing trend for longer distances. Most of these larger distances traveled are to and from Amsterdam Noord. The trends for bus and metro modes are in contrast with those reported by Huang and Levinson (2015) for Minneapolis–St. Paul region, where circuity was found to decrease with increasing Euclidean distance. As discussed, this contrast could be attributed to the geometry of metro and bus routes in Amsterdam.



Note: Circuity for metro includes metro-to-metro transfers

Figure 3.7. Circuity by Euclidean distance covered and mode used.

As discussed earlier, we measure the 'realized' instead of 'potential' circuity in this study. However, this could lead to a selection bias. For example, a traveler could have chosen a route with sub-optimal circuity due to other desirable characteristics such as lower travel times or less crowding. We further investigate this by comparing the observed circuity values with the shortest-path circuity for each observation in our data. The results show that for 96% of journeys, the difference between circuity of observed routes and the shortest path circuity is less than 0.01 units. This means that the observed circuity distribution is close to the potential circuity distribution in our case, and we therefore conclude that our data contains minimal selection bias.

3.4.2 Circuity, income and distance traveled

Next, the relationship between the income, circuity and distance traveled is explored. As described in section 3.3.4, transit journeys have been assigned the income profile of their origin transit stops. For this analysis, the transit stops have been divided into three categories based on their share of low income residents, roughly corresponding to the mean ± standard deviation in the study area:

- *Group 1* with a share of low-income residents of less than 30%
- *Group 2* with a share of low-income residents between 30 and 50%
- *Group 3* with a share of low-income residents of more than 50%

We first establish how far the travelers from each of these three groups travel by transit, as measured by the Euclidean distance of their journeys (**Figure 3.8**). This gives an indication of the proximity of travelers to various opportunities they need to access. It is noted that travelers from predominantly low-income areas (group 3) have a much higher proportion (52%) of longer journeys (>4km) compared to the rest of the travelers, for whom this proportion is only ~34%. The results support the amenity-based theory (Brueckner et al., 1999) that higher income persons locate themselves in places with greater proximity to amenities, with travelers from Group 3 traveling longer Euclidean distances (median distance of 4.1 km), compared to the rest (median distance between 2.8 and 3 km). The difference in distribution between areas of low to medium share of low-income people (group 1 and 2) is less pronounced.

Figure 3.8. Distribution of Euclidean distance covered by income profile of transit stops.

If the circuity is the same across all journeys made in the network, the distribution of network distance will follow the distribution of Euclidean distance. However, the uneven distribution of circuity could either reduce or exacerbate the differences in journey length distribution in the network. To investigate this, we plot next the circuity distribution for the three income groups (**Figure 3.9**). Circuity is found to have the highest median value (1.24) for predominantly low-income areas (group 3). However, the spread of circuity distribution is also found to be the widest for this income group, with 25% travelers having circuity values of less than 1.05 – the lowest between the three groups. The least amount of detours are experienced by travelers from Group 2 with a median circuity value of 1.19 for this group. Even though the distribution of Euclidean distance is comparable for group 1 and 2, the relatively favorable circuity distribution of group 2 is expected to reduce the network distance travelled by this group. Concurrently, the higher circuity values for low-income areas may worsen the disparity in distance traveled for this group compared to the rest of the population.

Arguably, people with higher income are likely to locate themselves in areas with higher proximity to opportunities due to which they need to travel shorter (Euclidean) distances (**Figure 3.8**). In addition, these areas may also be served by a denser transit network with direct routes to most destinations, because of which they benefit from smaller detours, leading to lower circuity values. To isolate the relation between high-income areas and circuity, a regression analysis is conducted with Euclidean distance as a control variable to represent the proximity to opportunities for different income groups. The analysis is carried out on data aggregated for each origin stop with the natural logarithm of average circuity as the dependent variable and share of high-income residents as one of the independent variables. Additionally, all stops in Amsterdam Noord have systematically higher circuity values (**Figure 3.6**). To control for these differences due to the structure of the city, a dummy variable for transit stops located in Amsterdam Noord is added. First an OLS regression was undertaken and based on the tests for spatial autocorrelation as described in section 3.5, a Spatial Error Model was implemented to incorporate the spatial dependence in the data. **Table 3.1** shows the results of the Spatial Error model.

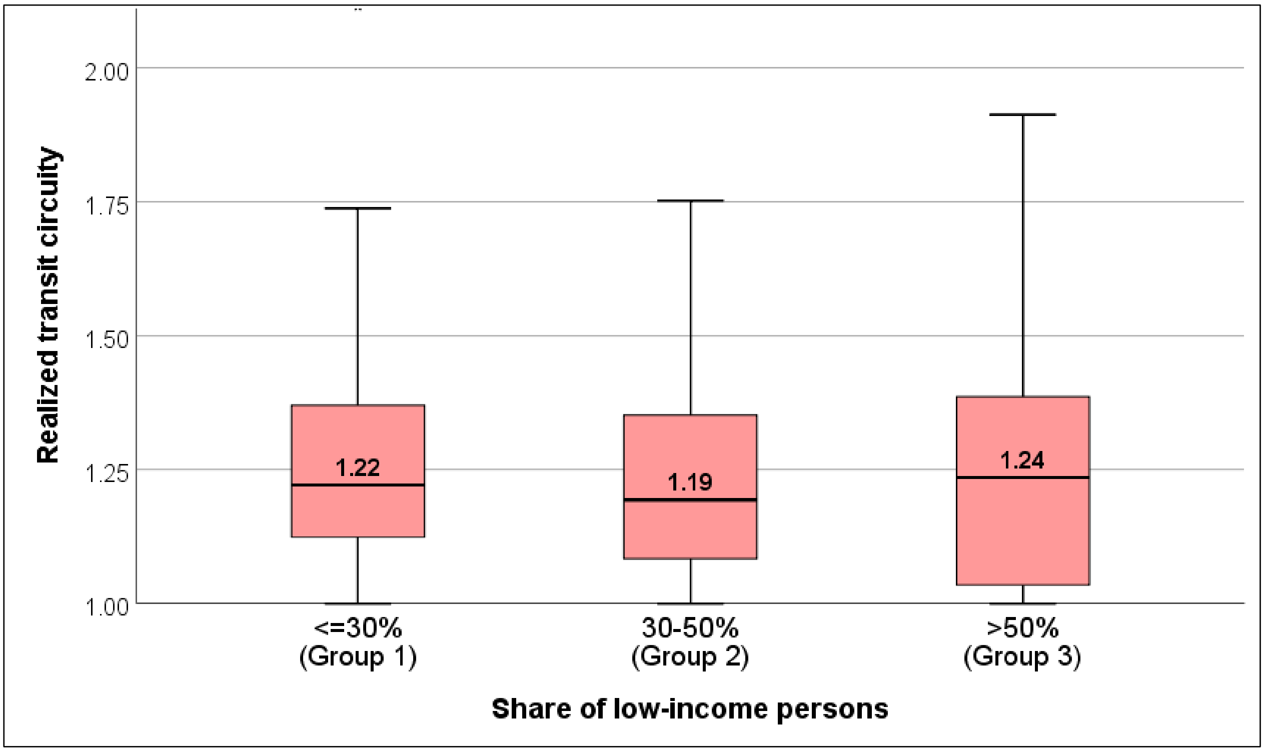


Figure 3.9. Distribution of circuity by income profile of transit stops.

Table 3.1. Spatial error model estimation results.

Variable	Coefficient	Std. error	Probability
Dependent variable = Log (Circuity)			
Constant	0.474	0.029	0.000
Percentage share of high-income persons	-0.003	0.001	0.001
Dummy for Amsterdam North	0.185	0.029	0.000
Average Euclidean distance (km)	-0.032	0.004	0.000
Spatial coefficient on errors (Lambda)	0.536	0.040	0.000
Number of observations = 472			
Log likelihood = 347.77			
AIC = -687.55 (AIC for OLS = -587.87)			

All dependent variables are found to be statistically significant. As expected, transit stops in Amsterdam Noord have 20.3% (=exp (0.185)-1) higher circuity on average compared to the rest of the city, all else being equal. The average Euclidean distance traveled for a transit stop represents the proximity to opportunities of the transit stop. For every km increase in Euclidean distance, the average transit circuity of a stop decreases by 3.2% – implying the longer transit routes tend to be more direct. However, even when controlling for the Euclidean distance, stops in higher income areas are associated with lower circuity values. For every percent increase in share of high-income residents, the circuity decreases by 0.3%, all else being equal. The share of high-income residents within the study area ranges between 3% and 54%, implying a maximum circuity difference of up to 14% between neighborhoods due to income effect.

The regression analysis confirms that the transit routes being used by travelers from high-income areas indeed have lower circuity for the same Euclidean distance covered, even when controlling for Amsterdam Noord and spatial dependence. This could be a result of two contributing factors. Firstly, the circuity of routes serving high-income areas could be low. But it could also be that the destinations of travelers from high-income areas have more direct routes. Although both scenarios highlight the underlying inequity, different solutions are needed for each. To confirm if there are differences in the types of destinations visited, we analysed the distribution of destinations for each of the three income groups. However, no substantial differences were found between travelers from the three groups, suggesting that the differences in circuity by income observed in the data are primarily due to the routes serving these areas as opposed to the differences in destinations.

3.4.3 Impact on travel times and fare paid

As a combined effect of the distribution of circuity and Euclidean distance, travelers from predominantly low-income areas in Amsterdam do indeed have longer transit journeys on average compared to the rest of travelers (**Figure 3.10**). The share of longer journeys (>8.5 km) is found to increase with the share of low-income residents. Moreover, substantial difference is found in the median journey length for group 3 (4.9 km), compared to group 1 (3.9 km) and group 2 (3.6 km). Overall, the differences between group 1 and 2 are found to be less pronounced than those of either of them with group 3, as in case of the distribution of Euclidean distance.

Transit fare in Amsterdam is calculated based on the network distance traveled, with the fare increasing linearly with distance. The journey length distribution in **Figure 3.10** hence implies that travelers from lower income areas travel longer on average, and in turn pay a higher fare per trip, before accounting for redistribution measures such as special subscriptions and concessions. The circuity of transit networks is a function of the network design. It can be argued that for a horizontally equitable distribution of transit services, every traveler in the network should pay the same fare per Euclidean distance covered, which means equal distribution of circuity over the network. Here we evaluate the horizontal equity of the network in terms of circuity using Gini coefficient. **Figure 3.11** shows the Lorenz curve with the share of accumulated circuity by share of population, and the Gini coefficient. The Gini coefficient of 0.11 indicates that the fare paid per Euclidean distance traveled is slightly unevenly distributed in the network. In relative terms, it is not possible to comment on how fair this distribution is, as such an analysis of circuity has not been undertaken for any other network yet.

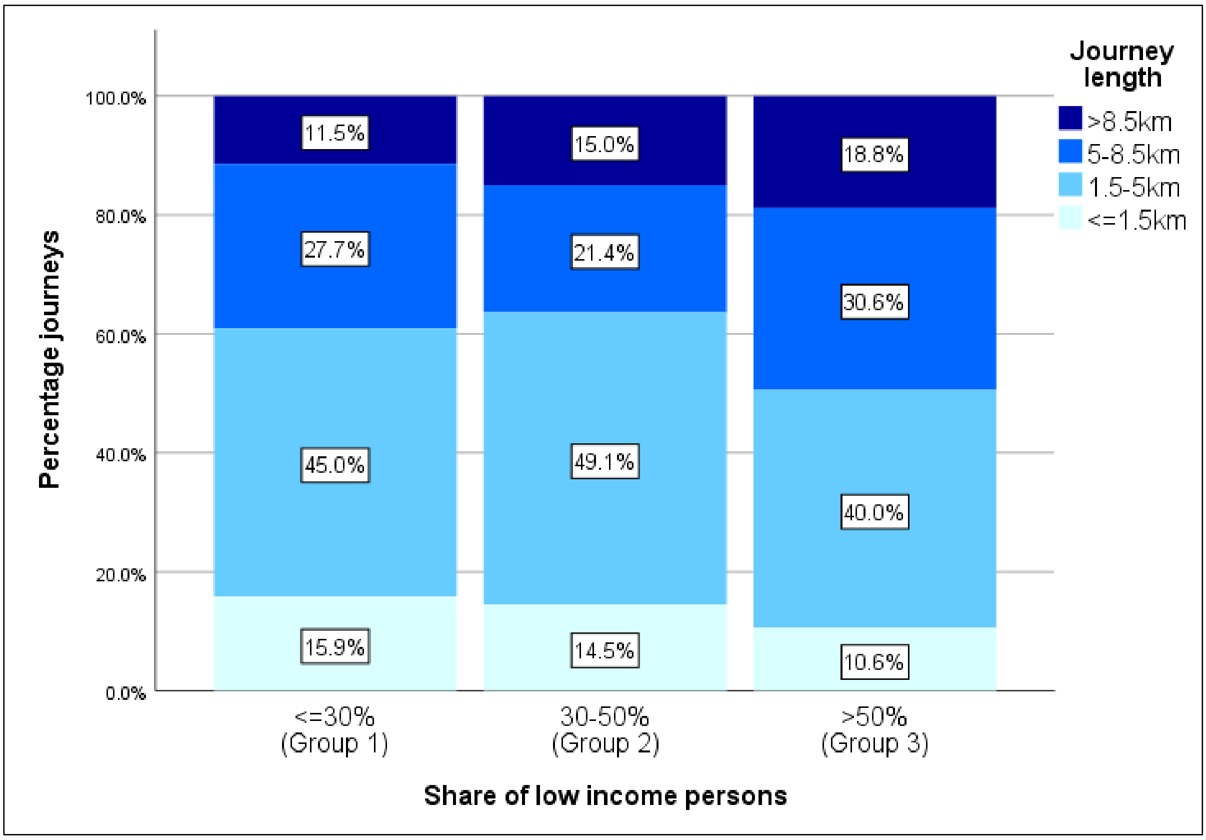


Figure 3.10. Distribution of journey lengths by income category.

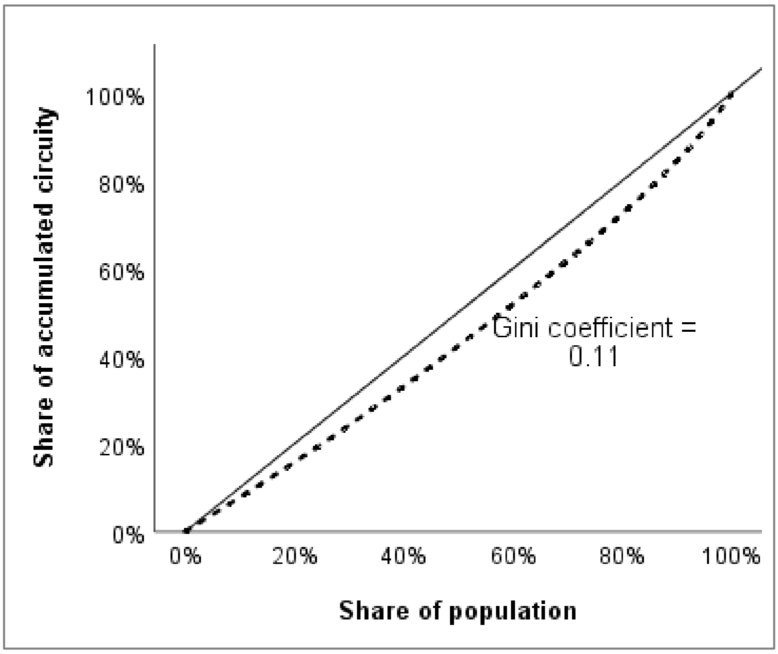


Figure 3.11. Lorenz curve and Gini coefficient for distribution of transit circuity in the network.

Transit circuity is expected to also impact the observed travel times. To analyze this relationship, we normalize the travel time by the Euclidean distance covered. **Figure 3.12** shows the distribution of this metric with the realized transit circuity across the network. In the absence of congestion effects, as the circuity of journey increases, longer time is spent on average to cover the same Euclidean distance, which is found to be the case for Amsterdam network.

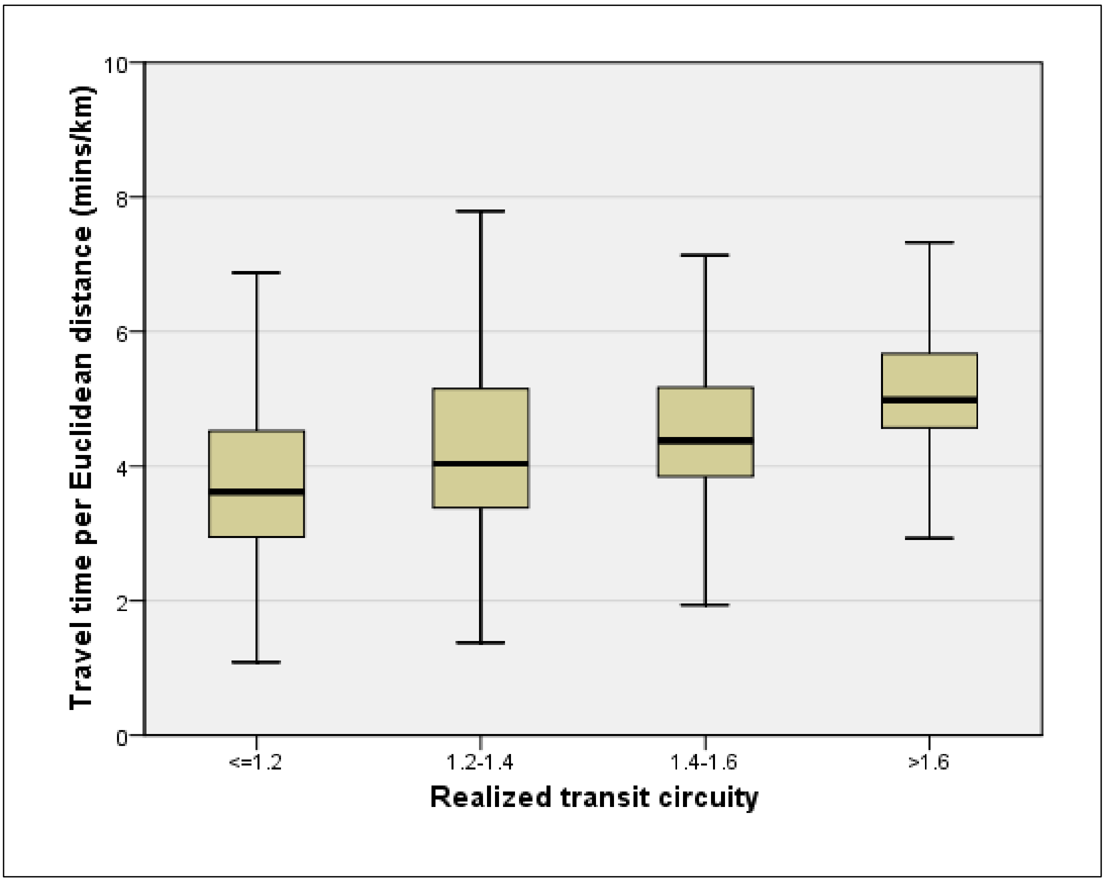


Figure 3.12. Travel time per Euclidean distance and realized transit circuity.

However, when we examine the distribution by income categories (**Figure 3.13**), the travel time per Euclidean distance covered does not follow the trend of circuity distribution, with group 2 having the highest median value (4.2 min/km), followed by group 1 and group 3 (4.0 and 3.8 min/km, respectively). Perhaps for group 1 and 2, although the circuity of routes is lower, other network characteristics such as shared right-of-way, on-road congestion and crowding compensate for the reduced travel time. Similarly, on routes with higher circuity serving lower income areas, the vehicle speeds may be higher. In the case of Amsterdam, tram services have lower speed in the historical core which is characterized by higher income levels. In contrast, low income areas are often located in proximity to tram corridors with a designated right-of-way or metro lines - especially the high circuity categories (>2) have a large share of metro (see **Figure 3.4**), and therefore high speeds.

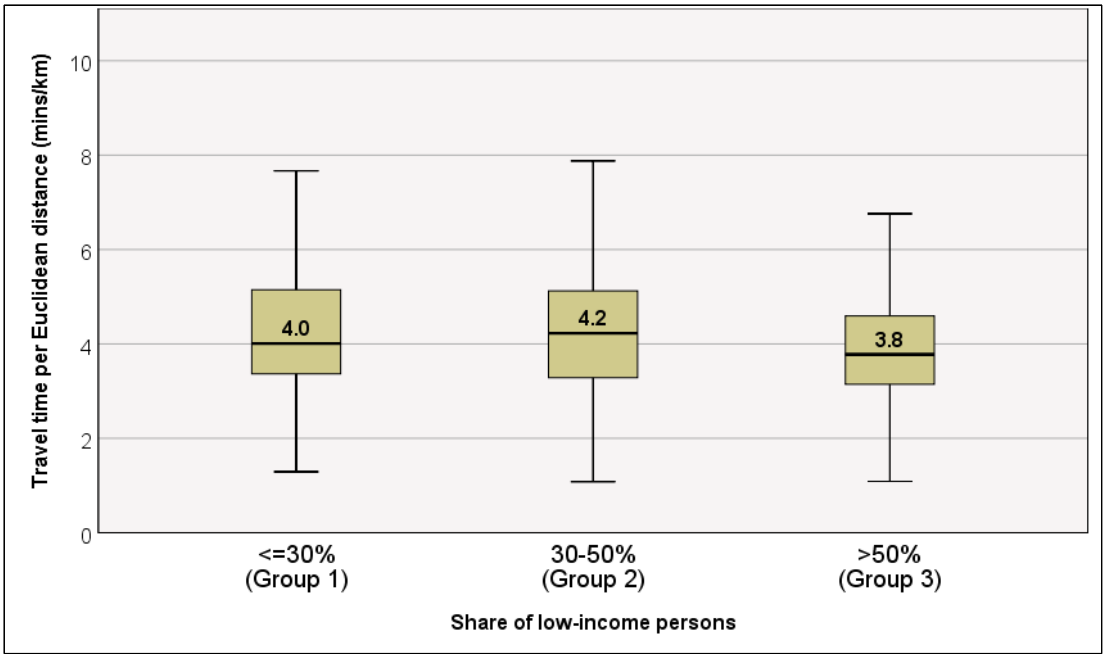


Figure 3.13. Travel time per Euclidean distance by income category.

3.4.4 Discussion

The circuity of transit networks has an impact on the distance traveled in the network. For Amsterdam network, the distribution of circuity favors the higher income groups, exacerbating the differences in distance traveled by income groups. This directly impacts the fare paid by the travelers. The Gini coefficient quantifies the equity of distribution of fare paid for every km of Euclidean distance covered under the current distance based fare structure. The circuity patterns observed in a network are a function of network design, which is often derived from a city's natural terrain and evolution of urban form. By improving the circuity of transit routes serving low income areas, the distribution of fare paid per Euclidean distance covered can be made more vertically equitable. This may however come at the cost of compromising other network planning considerations. Alternatively, with an egalitarian perspective, fares could be charged based on the Euclidean distance covered instead of network distance to address equity concerns.

This study highlighted the contribution of network design to the inequity of fare paid in a network, and how it can be used to address equity concerns. Camporeale et al. (2017) highlight that equity concerns have traditionally been ignored during network planning, and have been "in the best cases an afterthought during service provision". The process and analysis used for this study can be adapted for different network configurations (in combination with the fare structures) to assess the equity of a system. Where it is not possible to reduce circuity, other mitigation measures could be applied to compensate for the disparity caused by circuity of routes, such as different fare structures (based on Euclidean distance or flat fare). A key advantage of measuring equity using circuity is that such an analysis not only highlights the problems, but also provides insight on possible solutions. Incorporating equity concerns at the network design stage can lead to improved equity of outcomes such as fare paid and travel times.

Ridership and coverage are considered two of the primary goals of public transport, which are often opposing (Walker, 2008). Lower circuity is typically associated with shorter travel times (as also in the case of our study) leading to higher ridership, but lower coverage. Conversely, higher circuity can provide more coverage but it comes at the cost of longer travel times which can negatively impact ridership. Coverage goals are often social ones relating to serving the needs of disadvantaged population, or providing geographic equity of transit provision (Walker, 2008). However, as shown in our study, the higher circuity required to fulfil these goals can result in inequity of distance travelled and fare paid. Eventually, these trade-offs need to be weighed against each other based on the planning goals for the network under consideration. To that end, it will be useful to have more empirical research looking at these trade-offs in greater detail in the future.

Although smart card data enabled an extensive analysis of circuity by providing information on all journeys undertaken in the urban network, some limitations cannot be ignored. Firstly, the smart card data used in this study does not distinguish between tourists and residents. This may impact some results of the study as tourists are more likely to travel in the higher income areas in the city center, and tend to make shorter journeys. This might have overestimated the number of shorter trips associated with high income residents in our analysis. However, the proportion of such journeys is expected to be small, especially in the AM peak period (Central Bureau of Statistics (CBS) Netherlands, 2020b). Secondly, since our data is restricted to only the urban transit network of Amsterdam (excluding regional buses/trains), we cannot differentiate between travelers coming into Amsterdam from neighboring regions and trips originating within the case study network. People travelling to Amsterdam by train or regional bus services are now assigned to the income levels associated with the station where the traveler transfers to the urban network. Since people who need trains or regional bus services to reach Amsterdam have larger travel distances (and therefore higher fares), this assumption may have underestimated journey lengths. Thirdly, we have used a commonly used catchment area radius of 400m (El-Geneidy et al., 2014) for assigning income and our results are subject to this assumption. The analysis could be improved with a more precise value of this catchment area obtained from additional data sources. Lastly, our analysis was restricted to morning peak period due to the unavailability of income information for evening/off-peak journeys. However, considering that low-income persons often travel during off-peak periods, including such time periods can provide a more comprehensive picture of equity and could be undertaken as further research. This would however require additional data sources to estimate the income levels for off-peak journeys.

3.5 Conclusion

This study examined the contribution of transit circuity to the disparity in distance traveled between different income groups in a network. Furthermore, its implications on the travel times and the fare paid in a distance-based fare system were discussed. This was done for the case study of the multi-modal urban transit network of Amsterdam, using the demand data from smart card paired with the neighborhood level income data.

Travelers from predominantly lower income areas in Amsterdam were found to have more circuitous journeys compared to the rest of the travelers. For the same Euclidean distance covered and residential location with respect to the river (north/south), circuity was found to decrease with an increasing share of high-income residents, when controlled for spatial-autocorrelation. This exacerbated the already existing differences in Euclidean distance traveled between the income groups. As a result, travelers from lower income areas need to travel longer distances and pay a higher share of the fares in the network. The Gini coefficient also indicates a horizontal inequity in the distribution of fare paid per Euclidean distance.

However, the differences in travel time (per Euclidean distance) were in favor of lower income areas (3.7 min/km as opposed to 4-4.2 min/km for other areas). These are presumably compensated in the Amsterdam case by the various network characteristics experienced by the respective groups.

Overall, this study highlighted the role of transit network design in determining the equity outcomes of travel time and fare paid in a network. The equity outcomes in a network depend on the specific interaction between the land-use distribution, transit network design, and the fare policy employed. As further research, it will be valuable to compare the results obtained in this study with those for other cities, and could be utilized to compare different network structures or fare policies in terms of equity.

Chapter 4 - Perception of Overlap in Multi-modal Urban Transit Route Choice

The previous two chapters focused on service quality measurement for urban transit networks. In this chapter, we aim to understand the impact of such service quality and network design attributes on the transit route choice decisions of travelers. In doing so, we contribute to the scarce literature on modeling transit route choice using network-wide data for large-scale multi-modal transit networks, and use mode-specific travel attributes to incorporate the differences in perceptions of different modes. In addition, it also particularly focuses on exploring how travelers perceive overlap between alternate transit routes that lead to unobserved correlations between them. Capturing this unobserved correlation between overlapping routes is a non-trivial problem in route choice modeling. For urban transit networks, research so far has been inconclusive on how this overlap is perceived by travelers. We estimate a series of path size correction logit (PSCL) models to account for alternative specifications of route overlap in the context of multi-modal urban transit networks. In addition to the conventional path-based overlap of links or complete journey legs, an alternative definition of overlap in terms of transfer nodes is proposed for multi-leg journeys. A better understanding of how the overlap is perceived and should be incorporated into transit route choice models is expected to improve the accuracy of transit route choice models as well as provide more realistic behavioral insights.

This chapter is based on the following article:
Dixit, M., Cats, O., Brands, T., van Oort, N., Hoogendoorn, S. (2021) Perception of overlap in multi-modal urban transit route choice. *Transportmetrica A: Transport Science*, DOI: 10.1080/23249935.2021.2005180

4.1 Introduction

Public transport plays an important role in making cities more sustainable and liveable. To that end, policy makers and transit agencies are always striving to make transit more attractive to its users. Understanding how travellers choose between alternate transit routes is useful when planning and designing systems. It can help improve transport models and their predictions, and better assess interventions and network improvements, eventually leading to an increased transit usage.

Route choice models were traditionally developed for road networks, but the last decades have seen a rise in its applications to transit networks. Until recently, these models were mainly based on stated preference data sources, which although valuable in its own right, suffers from a common drawback of discrepancy between stated and actual behaviour (Yap, Cats, and van Arem 2018). Smart card provides a rich data source for analysing route choice by providing information on the actual choices made in the network, as well as the observed travel times at a high spatio-temporal resolution. Yet, only a handful studies have used it on a large scale multimodal transit network – namely Jánošíkova *et al.*(2014), Kim *et al.*(2019), Tan *et al.*(2015), and Yap, Cats, and van Arem (2020). The aim of this study is to leverage the large revealed preference dataset provided by the smart card to improve transit route choice modelling, by investigating in-particular the perception of travellers regarding *overlap* between routes.

Unless explicitly accounted for, the overlap between alternative routes results in correlations between the unobserved components of routes' utilities For road networks, it is widely accepted, and has been shown empirically, that this overlap is valued negatively by travellers (Bovy et al., 2008). However, this is not necessarily true for transit networks, where the negative perception of overlapping routes may be masked by the positive utility of having an improved level-of-service (e.g. shorter waiting time on a shared corridor) or more alternatives available in case of disruptions. Overlap between transit routes has been argued to add to the robustness of the trip (Anderson et al., 2017), which is further improved when complemented with coordinated schedules (van Oort and van Nes, 2009). The research on how overlap is perceived by travellers during transit route choice is inconclusive so far, with some studies reporting a positive valuation (Anderson et al., 2017; Hoogendoorn-Lanser and Bovy, 2007), while others reporting a negative valuation (Yap, Cats, and van Arem 2020; Tan et al. 2015). In this chapter, we investigate this issue further by analysing the *different specifications of consideration of overlap* between transit alternatives, to identify the circumstances and underlying reasons for its impact on passengers' route choice.

Similar to road networks, transit routes can have a partial overlap with one or more links being shared by multiple routes. Moreover, for routes that involve a transfer, there could be an overlap of entire journey leg(s). Hoogendoorn-Lanser *et al.*(2005) defined the overlap in terms of number of legs overlapped, as opposed to the links overlapped for road networks. In Tan *et al.*(2015), the authors used link-level overlap, but proposed additional definitions in terms of travel time of overlapped links, also incorporating the frequency of overlapped routes. In case of urban transit networks, it is not yet clear how each of these types of path overlap (link and leg) is perceived by travellers, as, to the best of our knowledge, these have not been compared in the literature so far.

Furthermore, literature so far has defined and considered overlap exclusively in terms of paths (or path-based overlap). In this study, we propose an additional definition of overlap between routes in terms of common transfer nodes (or transfer stops). From a traveller's perspective, each transfer node is a decision point, where he/she can choose between alternative transit lines. The expectation is that there are utility benefits associated with routes that share a transfer node, because of the multiple transit options it provides to the travellers, making the overlapped routes more robust compared to independent routes. This alternative definition of overlap is compared against the usual definition based on overlapped links and legs. Further, we distinguish between the valuation of overlap of paths versus nodes, by considering them separately as well as together.

The main contributions of this study are twofold. Firstly, it adds to the handful of empirical studies using large-scale revealed preference (smart card) data for estimating multi-modal transit route choice models. In doing so, it provides RP-based valuations of *mode-specific* travel time and transfer attributes, which to our knowledge are not available at such granularity in the literature so far. Secondly, it undertakes a comprehensive investigation of overlap between transit routes by defining overlap in terms of both the path (links and legs) and transfer nodes. We report results from our application of route choice models using the smart card data for the urban transit network of Amsterdam, the Netherlands.

We start with a Multi Nomial Logit (MNL) model of route choice that includes mode-specific in-vehicle and waiting times; number and type of transfers; transfer time; circuity of routes; and mode-specific constants. The base MNL model is then compared against the alternate formulations of Path Size Correction Logit (PSCL) models defining overlap in terms of links, legs, and transfer nodes. Lastly, the path-based and node-based PSCL formulations are considered together to investigate the relative contribution of each of these to the utility of overlapping routes.

The rest of the chapter is organized as follows: *Section 4.2* describes the approach used for quantifying and incorporating overlap in route choice models. In *Section 4.3*, the steps followed for processing and preparing smart card data are described including the model specifications. *Section 4.4* presents and discusses the results of model estimation and validation, followed by the conclusions in *Section 4.5*.

4.2 Overlap in transit route choice

4.2.1 Definitions

In this study, a transit *journey* refers to the travel made by an individual from an origin transit stop to their destination transit stop, using a *route* that may involve transfer(s) within or across different transit modes, such as bus, tram or metro. A journey may contain multiple *legs*. A leg represents a part of the journey undertaken using a single transit vehicle. A *transfer node* is defined as the transit stop where the traveller transfers between multiple legs of a journey. Each leg may consist of multiple *links* which refer to the physical path connecting two consecutive transit stops on the route.

4.2.2 Background

Both traffic and transit route choice typically consist of overlapping route alternatives, resulting in correlation between unobserved characteristics of the overlapping routes. In the case of road networks, overlap between routes results in the utility of the overlapping routes being overestimated. This is because the routes with an overlap may not be perceived as being distinct from the perspective of a traveller, and are hence less likely to be chosen compared to comparable independent routes. The basic MNL formulation assumes the unobserved characteristics of alternatives to be independent, i.e. Independence of Irrelevant Alternatives (IIA) property. To incorporate the overlap between alternatives, there are two common approaches – either explicitly modelling it by making assumptions on the correlation between error terms (such as in error component logit model), or adding a deterministic term in the utility function to approximate the correlation (such as in C-logit (Cascetta et al., 1996) or path-size logit (Ben-Akiva and Bierlaire, 1999)). This study follows the latter approach, which can be more directly specified and interpreted, and is commonly adopted in practice (Frejinger and Bierlaire, 2007).

Approaches such as C-logit, path-size logit (PSL) and path-size correction logit (PSCL) aim to reduce the utility assigned to overlapping routes, thus resulting in a lower probability compared to completely independent routes (Prato, 2009). The reduction in utility in these models is often proportional to the length (Cascetta et al., 1996) or cost/time (Ramming, 2002) of overlapping links. This is intuitive in case of road networks – the higher the proportion of the route overlapped, the more they are expected to be considered alike by travellers. C-logit model has been found to be generally outperformed by the PSL, because of which most recent studies use path size based models (Prato, 2009). In terms of performance PSL and PSCL have been found to yield similar results (Bovy et al., 2008). We choose to use PSCL in this study owing to its stronger theoretical foundation(Bovy et al., 2008; Tan et al., 2015).

Under the PSCL model, as defined by Bovy, Bekhor, and Prato (2008), the expression for probability of a route alternative *'i'* is given by

$$P_i = \frac{\exp(V_i + \beta_{PSC} PSC_i)}{\sum_{j \in C} \exp(V_j + \beta_{PSC} PSC_j)} \tag{4.1}$$

Where
V_i = deterministic utility of route alternative *i*,
PSC_i = path size correction term of route alternative *i*,
β_{PSC} = parameter for the PSC term to be estimated, and
C = choice set of all alternative routes.

The path size correction (PSC) factor in its original form is given by,

$$PSC_i = -\sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \, ln \sum_{j \epsilon C} \delta_{aj}\right) \tag{4.2}$$

Where
l_a = length of link *a* within alternative route *i*,
L_i = total length of alternative route *i*,
Γ_i = set of all links for route *i*,
C = set of all routes between the chosen origin-destination pair, and
δ_{aj} = link-route incidence between link *a* belonging to alternative route *j*.

The PSC term has a maximum value of 0 for completely independent routes and decreases as the overlap between routes increases, with a theoretical lower bound of -∞. For road networks, β_{PSC} associated with the PSC term is typically positive, resulting in a reduction of utility for overlapped routes (since PSC itself is negative for such routes).

While for road networks, there is a consensus on how the route overlap is defined and perceived by travellers, in case of transit networks the answer is not as clear. Hoogendoorn-Lanser, van Nes, and Bovy (2005) were the first to incorporate overlap in case of transit route choice. They defined overlap in terms of number of journey legs, travel time, and distance on those legs, and found that the overlap is valued negatively for all of these definitions (i.e. overlapped routes are less likely to be chosen). Contrastingly, Hoogendoorn-Lanser and Bovy (2007) found that the overlap in the train-leg of the multi-modal inter-urban journey was valued positively by the travellers, unlike the access and egress parts which were valued negatively. For urban transit networks also, there is evidence of a positive valuation of overlap (Anderson et al., 2017). As argued by Hoogendoorn-Lanser, van Nes, and Bovy (2005), the negative perception of overlapping routes in case of transit networks may be compensated by their contribution to robustness of the routes in case of disruptions.

One of the important questions in case of transit networks is how the overlap should be defined and formalized mathematically. The formulation in Equation (4.2) was developed for road networks, and is often directly adopted for transit networks by defining overlap in terms of common physical links and their properties (for example in (Anderson, Nielsen, and Prato (2017)). Tan *et al.*(2015) proposed a formulation based on travel time on the overlapping links and frequency of services, rather than the link length. Proposing a different approach, Hoogendoorn-Lanser, van Nes, and Bovy (2005) defined overlap in terms of common trip legs (as opposed to links) for inter-urban multi-modal transit routes. They found that the overlap defined in terms of number of trip legs explained the observed choices better, as opposed to travel times or distances on those legs. None of the studies so far have compared the alternative ways for defining and quantifying path overlap (link and leg) in the context of urban transit route choice.

Further, as per Hoogendoorn-Lanser, van Nes, and Bovy (2005), apart from physical path, overlap between routes can also be defined in terms of nodes, services, runs or modes. However, applications so far have been limited to path overlap only. We argue that in case of transit route choice, decision points are pertaining to transfer nodes where you may interchange, as opposed to links in case of road networks, where each intersection is a decision point. Hence, in this study, we include both path (link & leg) overlap and transfer node (decision point) overlap.

Based on the literature reviewed, we conclude that the following two questions remain unanswered regarding the perception of overlap in transit route choice:

- Is the overlap between alternate transit routes perceived positively or negatively by the travellers?
- Which way of defining overlap - link, leg or transfer node - best captures the perception of travellers for urban multimodal transit route choice?

In the next sub-section we describe the approach adopted in this study for addressing the abovementioned questions.

4.2.3 Research approach

We now discuss the different possibilities of such overlaps in case of urban multi-modal route choice, and our approach for addressing overlap in the form of common links, journey legs, and/or transfer nodes. We start with journeys without transfers (i.e. single leg journeys), and subsequently extend the approach to journeys with transfers. **Figure 4.1** shows the possible overlaps for transit route alternatives without a transfer (single leg journeys), which could either be a partial or complete overlap of physical paths of the transit lines. Routes A and B/C have a partial overlap of physical paths with only two of the links overlapping, whereas Routes B & C have a complete overlap.

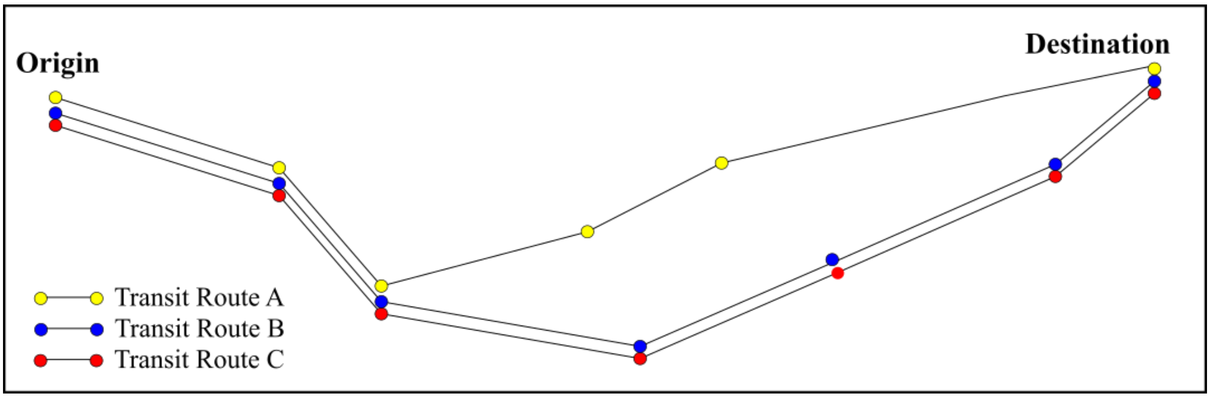


Figure 4.1. Overlap between transit routes without a transfer.

In Amsterdam, where we perform our case study, a map showing physical path of transit lines is displayed at the transit stops, providing travellers with information to choose an alternate overlapping transit line. Moreover, real-time passenger information systems are provided at the majority of stops, showing the next arriving transit vehicle(s). Hence, in this study, we assume that completely overlapping transit lines using the same mode are perceived as being the same by travellers. Accordingly, for such lines, the effective waiting time at the origin stop is calculated based on the combined observed headway of overlapped lines, as derived from the Automatic Vehicle Location (AVL) data. For the other case when there is a partial overlap of transit lines (Routes A and B/C), a link-based overlap is considered. In some cases, there may be a complete overlap of physical path but different modes are used – bus and tram in our case. For such cases, a leg-based overlap (and no link-based overlap) is considered.

Next, we consider routes with multiple legs. For such routes, in addition to link and leg based overlap, there could be a common (or overlapping) transfer node amongst the alternatives. **Figure 4.2** shows an example of leg and transfer node overlap for routes with one transfer. Alternative E & F share overlapped leg 1, and all alternatives share an overlapped transfer node. Note that it is possible for routes to have no overlap of path (link or legs), but still have an overlap of transfer node(s) (alternatives D and E/F in **Figure 4.2**).

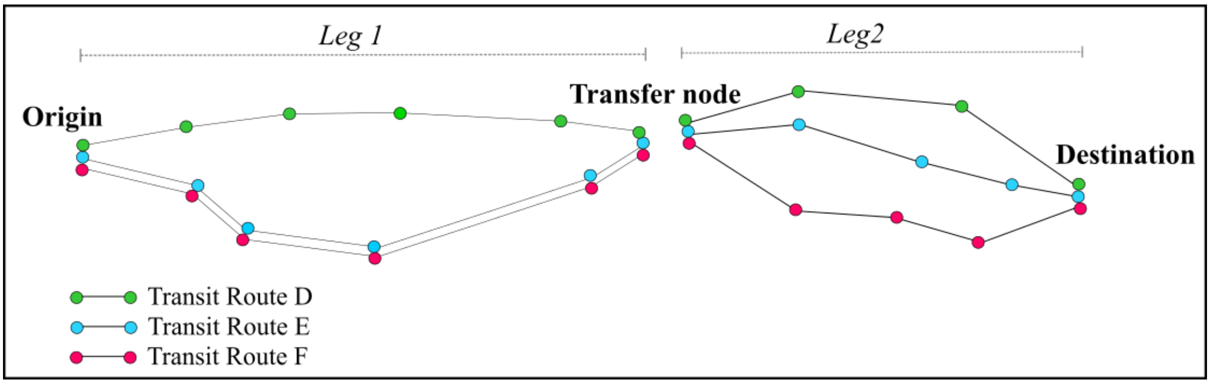


Figure 4.2. Overlap between transit routes with one transfer.

For all such possibilities of overlap, the unobserved characteristics of overlapping routes may be correlated. Hence, the utility is modified to take into consideration such overlap. We specify and analyse four PSC formulations for such routes - one each for link and node based overlap, and two for leg-based overlap:

1. **Link-based PSC:** This follows the traditional definition of PSC, as presented in Equation (4.2), and is based on the length of overlapping links as proportion of the total route length.

2. **Leg-based PSC - number of overlapped legs:** The hypothesis here is that travellers perceive the overlap in terms of number of overlapped legs, rather than the travel times or distance on those legs. In Hoogendoorn-Lanser, van Nes, and Bovy (2005), a similar definition was used for calculating the path size logit term for inter-urban transit travel. The path-size correction term in this case is given by:
$$PSC_i^L = -\sum_{l \in \Gamma_i} \left(\frac{1}{N_i} \, ln \sum_{j \epsilon C} \delta_{lj}\right) \tag{4.3}$$
Where
N_i = Number of journey legs in route i,
Γ_i = set of all legs for route i,
C = set of all routes between the chosen origin-destination pair, and
δ_{lj}= leg-route incidence between leg l belonging to alternative route j.

3. **Leg-based PSC - travel times on overlapped legs:** The PSC term in this case is calculated based on the travel time on the overlapping leg as a proportion of total travel time of the route, as proposed by Tan *et al.* (2015):
$$PSC_i^T = -\sum_{l \in \Gamma_i} \left(\frac{t_l}{T_i} \, ln \sum_{j \epsilon C} \delta_{lj}\right) \tag{4.4}$$
Where
t_l = travel time for journey leg *l* in route *i*,
T_i = total travel time for route *i*,
Γ_i = set of all legs for route *i*,
C = set of all routes between the chosen origin-destination pair, and
δ_{lj} = leg-route incidence between leg *l* belonging to alternative route *j*.

4. **Transfer Node-based PSC:** This factor captures the overlap in terms of number of decision points for multi-leg journeys, and is given by:

$$PSC_i^X = -\sum_{n \in K_i} \left(\frac{1}{X_i} \, ln \sum_{j \epsilon C} \delta_{nj}\right) \tag{4.5}$$

Where

X_i = Number of transfer nodes in route *i*,

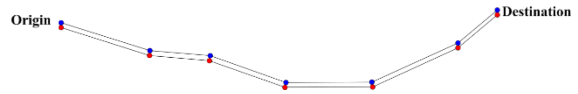
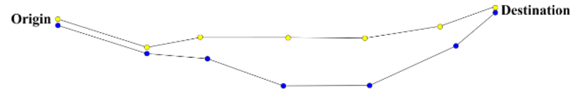
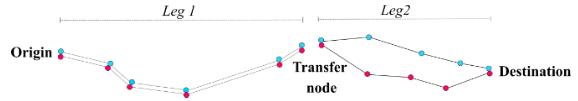
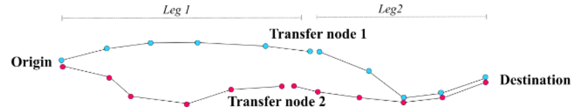
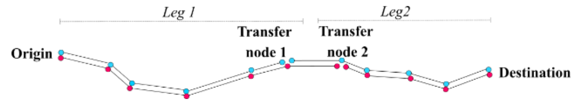
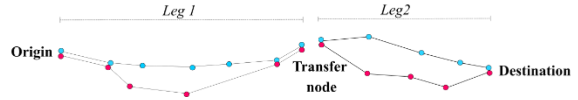
K_i = set of all nodes for route *i*,

C = set of all routes between the chosen origin-destination pair, and

δ_{nj} = node-route incidence between node *n* belonging to alternative route *j*.

The models incorporating above PSC factors are tested against the MNL model to establish how the addition of each of the PSCs impact the model fit. **Table 4.1** summarizes the different possibilities of path and node overlap in our data, and the PSC formulations applied in each case.

Table 4.1 Types of overlap considered.

<table>
<tr><th>Journey type</th><th>Type of overlap</th><th>Example</th><th>Link-based PSC</th><th>Leg-based PSC</th><th>Transfer node-based PSC</th></tr>
<tr><td rowspan="3">No transfer (single leg)</td><td>Complete overlap of path using the same mode</td><td rowspan="2">
</td><td colspan="3">Assumed to be perceived as the same alternative</td></tr>
<tr><td>Complete overlap of path using different modes</td><td>-</td><td>✓</td><td>-</td></tr>
<tr><td>Partial overlap of path using same or different modes</td><td>
</td><td>✓</td><td>-</td><td>-</td></tr>
<tr><td rowspan="4">With transfer (multi-leg)</td><td>Complete overlap of one or more legs using same or different modes</td><td>
</td><td>-</td><td>✓</td><td>✓</td></tr>
<tr><td>Partial overlap of one or more legs using same or different modes</td><td>
</td><td>✓</td><td>-</td><td>-</td></tr>
<tr><td>Different transfer nodes but using the same route</td><td>
</td><td>✓</td><td>-</td><td>-</td></tr>
<tr><td>Same transfer node but different/partially overlapped routes</td><td>
</td><td>if applicable</td><td>-</td><td>✓</td></tr>
</table>

4.3 Data Preparation

We perform our analysis on the urban transit network on Amsterdam, consisting of bus, tram and metro modes. The time period of analysis is 28th May to 29th June 2018, during which 41 bus lines, 15 tram lines and 4 metro lines were operational in the network. **Figure 4.3** shows a map of the transit network during our analysis period. The metro network forms a part-ring structure around the city centre, with two of the lines providing direct connections from the south-eastern and southern peripheries of the city to the city centre. The tram lines have a dense network in the city centre, while also serving as feeders to the metro network. The bus network mainly serves the outskirts of the city with relatively lower density areas, but provide some important connections, especially from the areas in the North to the city centre.

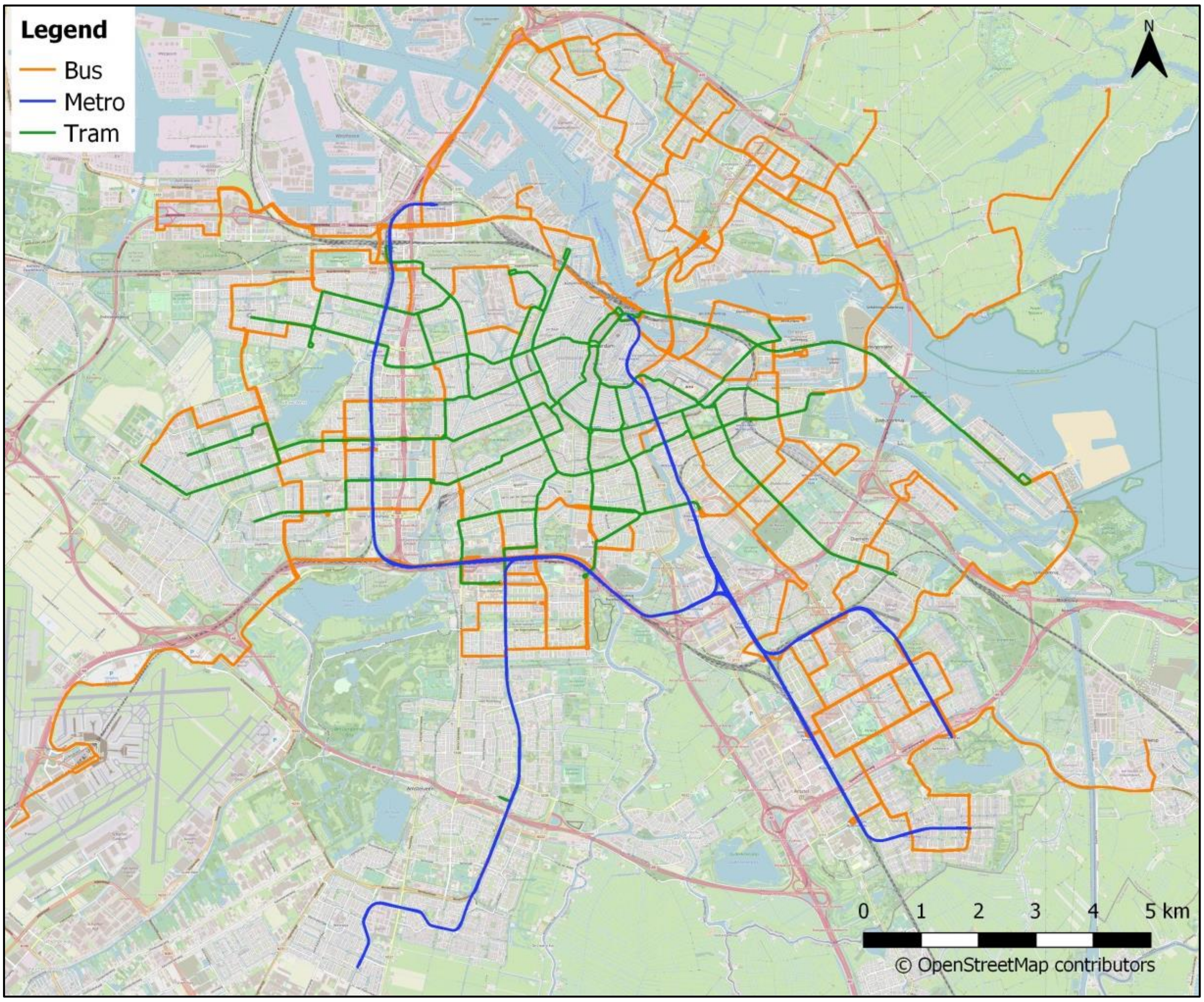


Figure 4.3. Amsterdam transit network.

We use (anonymized) smart card data for our analysis, which captures (nearly) all journeys made in the network. On an average day during our study period, over 675,000 smart card transactions were recorded in the network. We restrict our analysis to weekday AM peak period (7 am to 10 am), in order to maximize the proportion of commuters and regular travellers in the data, who are expected to be more familiar with the route options, thereby making an informed route choice decision. The following subsections describe the steps undertaken to convert the raw smart card data to the required format for route choice analysis.

4.3.1 Trips to journeys

The Dutch smart card system provides information on both boarding and alighting transit stops and times (for an overview of the Dutch smart card system see van Oort, Brands, and de Romph (2015a)). Each transaction in the raw smart card data represents a check-in and check-out made by a passenger. For the route choice analysis, it is required to trace the entire journey of the travellers from their origin transit stop until their destination transit stop. For this, we combine individual smart card transactions to form passenger journeys by identifying transfers. For Amsterdam, a maximum time threshold of 35 minutes is applied by the operator to classify consecutive trips by an individual as transfers. However, we apply additional criteria to ensure that trip generating activities conducted within the 35-minute criteria are not wrongly classified as transfers. We do this by fusing the smart card data with AVL data, and applying the transfer inference algorithm as proposed in Gordon *et al.* (2013) and Yap *et al.* (2017). For more details on how the algorithm is applied to our case study network, the readers are referred to Dixit *et al.* (2019b). The transfers made within the metro network are not directly available from the smart card data, since the travellers do not need to check-in and out within the metro system. Hence, we have inferred the number of transfers using a shortest path (time-wise) approach for such journeys. The Amsterdam metro network consists of only 4 lines with no loops, and all the origin-destination pairs within the network can be reached with a maximum of one transfer. In case of origin-destination pairs where more than one transfer stop is possible (due to parallel lines), we have assigned the transfer stop corresponding to the shortest travel time (calculated as the sum of in-vehicle and expected transfer waiting times).

After performing transfer inference, the journeys were filtered to remove those with any missing route attribute (such as observed headway from the AVL data). Further, in Amsterdam's network, almost all origin-destination pairs can be reached with a maximum of two transfers. Hence, for our analysis, we excluded journeys with more than 2 transfers (<0.01% of all journeys), to minimize irrational traveller behaviour in our data. After applying all the filters, this resulted in a dataset of 2.9 million journeys for the whole study period (weekday AM peak).

4.3.2 Aggregating transit stops

Smart card data does not provide any information on where passengers actually began or ended their journey - only the origin and destination transit stops are known. Further, the smart card data used for this study is anonymized, and the users cannot be tracked across multiple days. Hence, no information is available from the smart card data on which transit stops were considered by the traveller while making their route choice decision. A simplistic assumption is to restrict a traveller's route choice set to all the route alternatives available at his/her boarding stop and alighting stop only. However, such an assumption greatly reduces the number, and diversity of alternatives considered in the choice set. More importantly, it is an unrealistic assumption for a city like Amsterdam, where the median walking feeder distance for bus stops is more than 300m (Brand et al., 2017), while the distance between neighbouring transit stops may be as low as 100m in transit dense areas. Therefore, assumptions are needed regarding passenger's access/egress stop choice set. For example, Kim *et al.* (2019) aggregated all transit stops at an intersection into a node. We follow a similar approach of clustering neighbouring stops together, by means of agglomerative hierarchical clustering. In this method, starting with each transit stop forming its own cluster, the closest clusters are merged until a maximum distance threshold of 500m between any two stops within a cluster is achieved. The 500m distance threshold is chosen to achieve compact clusters, as measured by the silhouette score. 651 transit stops resulted in 279 stop clusters. The resulting clusters are shown in **Figure 4.4** where each dot corresponds to a transit stop and the transit stops belonging to the same cluster are shown in the same color. Once the transit stops are aggregated, all the routes connecting the traveller's origin and destination stop clusters can be included in his/her choice set.

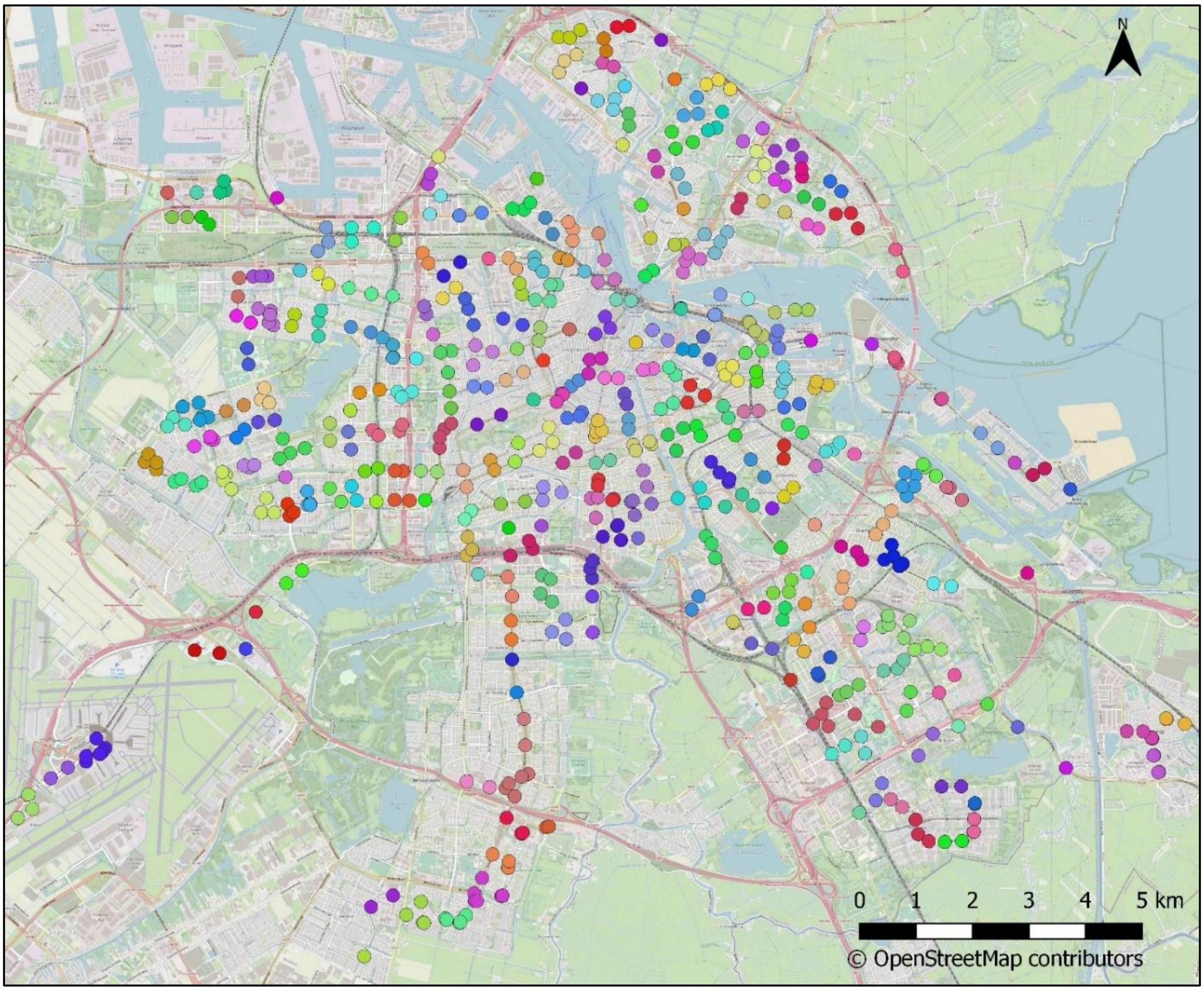


Figure 4.4. Clustering of transit stops in Amsterdam.

4.3.3 Route definition and choice set preparation

A route between an origin and destination stop cluster (further on referred to as O-D) is defined as a combination of the physical path(s) followed by the transit line(s), and the transfer stop cluster(s) used for each leg of the journey. The choice set in this study is based on the observed routes taken by the travellers in the data. This eliminates the need for assumption on the feasibility of non-observed routes. Further, it allows us to calculate the route attributes (such as in-vehicle and transfer times) purely based on the observations from the smart card data.
To make a realistic representation of the available route choices, we split the study period (weekday AM peak) into 6 half hour time slices and calculate the expected value of route attributes during each time slice. For example, a traveller starting their journey anytime between 7:00 to 7:30 am is assumed to have in his/her choice set the route alternatives that were used between that time window only. Accordingly, the expected value of route attributes is calculated by taking the median over all observations using the O-D route during the chosen time slice. Although seemingly large, the half hour time slice is intended to maximize the number of observed journeys to have reliable estimates of route attributes. For that, only those routes with a minimum of 20 journeys in the time slice (over all days) were used for the

analysis. Further, only those O-D pairs with a minimum of two route alternatives were included. Majority (91%) of O-D pairs had only two route alternatives, with a maximum of four alternatives for any O-D pair. Overall, the busiest time period within the morning peak was between 08:00 to 08:30 am, capturing 25% of all journeys. **Figure 4.5** shows the number of distinct OD pairs used in the data for each time slice, and of those how many include an overlap in at least of the alternative routes. The proportion of OD pairs with at least some type of overlap varies between 56% and 67% for each time-slice. Of these, majority have an overlap of link (only), which ranges between 42% and 49% of all O-D pairs for each time-slice.

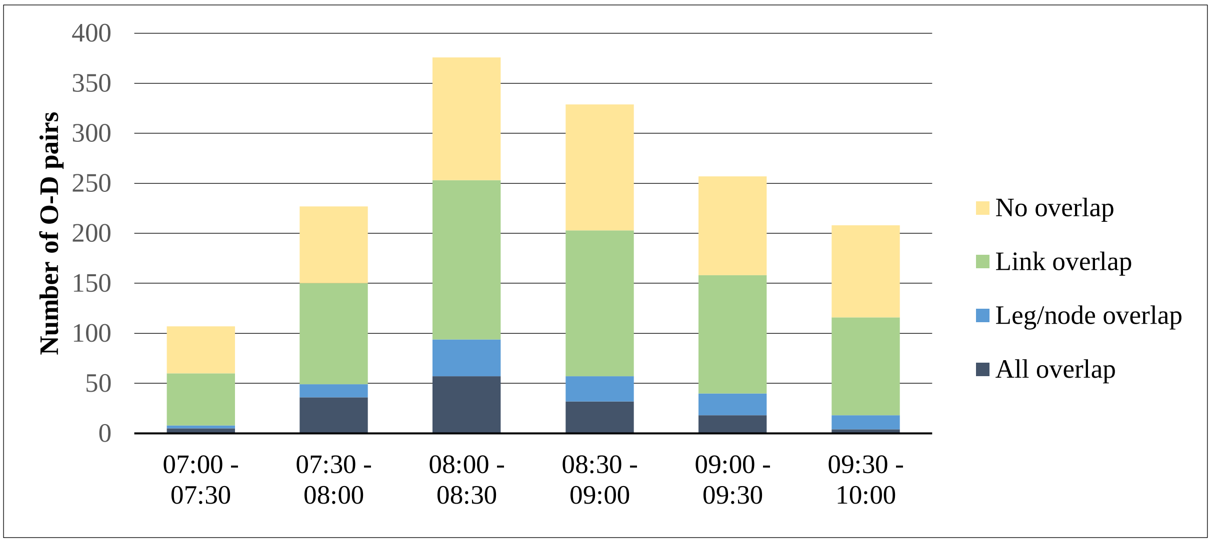


Figure 4.5. Number of unique O-D pairs in the data by type of overlap.

4.3.4 Route attributes and model specification

The way travel time is measured is different for bus and tram versus metro in Amsterdam. For buses and trams, the smart card is tapped in/out inside the vehicle, whereas for metro, this happens at the station. Because of that, the in-vehicle travel time for metro journeys is not directly available from smart card data. Hence, for this study, we derive the in-vehicle, waiting and transfer times (if applicable) for metro journeys based on the AVL data. While the average in-vehicle time between all O-D pairs is directly available from the AVL data, the effective waiting and transfer times are derived based on the observed headway at each origin and transfer stations.

The following route attributes are populated for each O-D route and time-slice combination, which are subsequently used in our choice model specification:

1. **In-vehicle time by bus and tram (IVT_{bus} and IVT_{tram}):** These correspond to the total in-vehicle time by bus and tram modes summed over all legs in a journey.
2. **Expected waiting time for bus and trams (WT_{bt}):** This is calculated for the first leg of the journeys starting with bus and trams. For common lines, this is calculated based on their combined headway.
3. **Metro time (TT_{metro}):** This includes in-vehicle time, and effective waiting time at the origin metro station.
4. **Number and type of transfers ($Trans_{bt}$, $Trans_{btm}$ and $Trans_{m}$):** Transfers made are distinguished as being within bus/tram network (which includes bus-bus, tram-tram and bus-tram transfers); transfers between metro and bus/tram; and transfers within the metro network.

5. **Transfer time (TrT):** This includes the transfer time for all types of transfers.
6. **Circuity ($Circ$):** This quantifies the detours made in the route, and is calculated as the ratio of network to Euclidean distance of the route. It ranges from a minimum of 1 (for very short routes), to approximately 4 in our dataset.
7. **Mode-specific constants for bus, tram and metro (MSC_{bus}, MSC_{tram} and MSC_{metro}):** These incorporate the preferences for a particular mode that is not captured by any of the above attributes.

A total of six route choice models are estimated. We start with an MNL model, the systematic utility of which is specified in Equation (4.6). All β represent the coefficients of the attributes which are estimated.

$$\begin{aligned} V^{MNL} = \beta_{ivt_{bus}} * IVT_{bus} + \beta_{ivt_{tram}} * IVT_{tram} + \beta_{wait_{bt}} * WT_{bt} + \beta_{tt_{metro}} \\ * TT_{metro} + \beta_{trans_{bt}} * Trans_{bt} + \beta_{trans_{bt}} * Trans_{btm} \\ + \beta_{trans_{m}} * Trans_{m} + \beta_{TrT} * TrT + \beta_{Circ} * Circ + MSC_{Bus} \\ + MSC_{Tram} + MSC_{Metro} \end{aligned} \tag{4.6}$$

Taking the above utility function as a basis, we add the path size correction term to the utility function (as shown in Equation (4.1) earlier) and estimate five PSCL models. PSCL Models 1 to 4 define the overlap based on link (Equation (4.2)), number of legs (Equation (4.3)), travel time on legs (Equation (4.4)) and number of transfer nodes (Equation (4.5)), respectively. Lastly, we combine the leg-based and node-based models to account for the contribution of each of those elements to the perception of route overlap. Different combinations of such model specifications were tried, and the best one is presented in PSCL Model 5.

4.4 Results and Discussion

4.4.1 MNL Model without overlap

The five PSCL models along with the MNL model as described in the previous section were estimated using the BIOGEME estimation package (Bierlaire, 2020). **Table 4.2** shows the results of the estimation. For all models, we present the final log-likelihood, rho-square-bar and likelihood ratio statistic (LRS) with respect to MNL for comparison.

All parameters are found to be significant at $p<0.01$ level, and the signs are as expected. All observed travel attributes being the same, there is a preference amongst travellers for using routes with metro over tram and bus. This is expected owing to the higher reliability the metro lines provide. Additionally, the simplicity of the metro network, weather protection at the stations, and a more comfortable waiting environment could be some other factors contributing to this preference. After metro, tram is found to be preferred by travellers over bus. Moreover, the in-vehicle time of bus is valued more negatively compared to tram, in line with findings from other studies reporting a 'tram bonus' in the Netherlands (Bunschoten et al., 2013). For buses and trams, one minute of waiting time at the origin stop is valued as much as 1.8 minutes of bus in-vehicle time, which is comparable with the values reported by Yap, Cats, and van Arem (2020) for The Hague, the Netherlands (1.5 -1.6 minutes).

Table 4.2. Model estimation results

		Link-based	**Leg-based**		**Node-based**	**Combined**
Description	**MNL Model**	**PSCL Model 1 - links**	**PSCL Model 2 - number of legs**	**PSCL Model 3 - travel time**	**PSCL Model 4 - transfer node**	**PSCL Model 5 – travel time +transfer node**
Number of observations	382,295	382,295	382,295	382,295	382,295	382,295
Estimated parameters	11	12	12	12	12	13
Final log likelihood	-233,892	-233,767	-233,714	-233,764	-233,513	-233,473
Rho-square-bar	0.178	0.178	0.178	0.178	0.179	0.179
Likelihood Ratio Statistic (compared to MNL)	-	249.6	355.6	256.2	759.0	838.6
Parameter estimates*						
Mode-specific constant for bus <fixed>	0.00	0.00	0.00	0.00	0.00	0.00
Mode-specific constant for tram	0.49	0.48	0.50	0.50	0.51	0.50
Mode-specific constant for metro	0.84	0.84	0.86	0.88	0.90	0.90
Bus in-vehicle time (mins)	-0.11	-0.11	-0.10	-0.11	-0.11	-0.11
Tram in-vehicle time (mins)	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09
Effective wait time bus/trams (mins)	-0.19	-0.19	-0.19	-0.19	-0.19	-0.19
Metro time[1] (mins)	-0.09	-0.09	-0.09	-0.09	-0.10	-0.10
Number of transfers between bus & tram[2]	-1.24	-1.24	-1.18	-1.26	-1.37	-1.41
Number of transfers between metro and bus/tram	-2.38	-2.40	-2.42	-2.47	-2.62	-2.65
Number of transfers within metro	-1.50	-1.50	-1.42	-1.44	-1.47	-1.51
Transfer time (mins)	-0.25	-0.25	-0.25	-0.24	-0.23	-0.23
Circuity	-0.43	-0.41	-0.43	-0.42	-0.39	-0.39
Path size correction – link	-	-0.53	-	-	-	-
Path size correction – leg	-	-	-0.90	-0.63	-	0.52
Path size correction – transfer nodes	-	-	-	-	-1.02	-1.39

*p<0.01 for all estimates
[1]includes in-vehicle time and origin waiting times.
[2]includes bus-bus, tram-tram and bus-tram transfers

The pure transfer penalty between bus/tram modes, which is valued at 11.5 (= -1.24/-0.11) minutes of bus in-vehicle time in the MNL model, is found to be much higher than the one reported by Yap, Cats, and van Arem (2020) for the buses and trams for The Hague (3.8-5.2 minutes), perhaps owing to Amsterdam's larger network with relatively longer transfer distances. Amsterdam also has a relatively higher share of tourists that may prefer direct routes. Although on the higher end, our transfer penalties are found to be comparable to other values reported in the literature, such as those observed by Garcia-Martinez *et al.* (2018) for the multi-modal transit network of Madrid, Spain (15.2-17.7 minutes), or by Anderson, Nielsen, and Prato (2017) for the regional multi-modal transit network of Greater Copenhagen Area (14.1-17.9 minutes). Between the different types of transfers made in the network, the ones between metro and bus/tram are the least preferred. This is expected, since in case of Amsterdam, transferring to metro from bus/tram typically involves walking a longer distance and climbing (deep) stairs. In comparison, the transfers within the metro network are more convenient as metro stations are relatively small and transfers are often cross-platform and do not involve long passageways or multiple level changes. The transfer waiting times are also more reliable in case of metro, because many of these transfers are synchronized during operation. However, the most preferred transfers are within the bus/tram network which are usually at the same level.

In addition to the pure transfer penalty, travellers also have a strong dis-preference for transfer time compared to the corresponding in-vehicle time. Detour of a route, as measured by the circuity, is also found to play a role in explaining the route choice of travellers, with a trade-off of 4 minutes of bus in-vehicle travel time for 1 unit change of circuity. The part-ring structure of the metro network results in some O-D pairs having high circuity. Moreover, having a distance-based fare system implies that higher circuity also results in higher fares being paid. Despite that, our parameter value for circuity is found to be much lower than that reported by Kim *et al.* (2019) for the transit network of Seoul, which to our knowledge is the only other study that includes circuity in multi-modal transit route choice. They report a value of about 22 minutes of IVT for 1 unit change of circuity, possibly due to the larger scale of the network.

4.4.2 Incorporating overlap

Next, we examine how travellers perceive different aspects of overlap between alternate transit routes, looking at each type of overlap individually (PSCL Models 1 to 4). Firstly, all PSCL models offer an improvement of model fit compared to the MNL model, as demonstrated by the LRS=249.6 for the worst performing PSCL model (PSCL Model 1), exceeding the critical χ^2 value of 6.6 at 1% significance level (df=1).

Secondly, the sign of the PSC parameter for all models considering overlap individually (PSCL models 1 to 4) is significant (at $p<0.01$ level) and negative, implying that the overlap between transit routes is perceived positively by the travellers in general, and the utility of overlapping routes is underestimated by the MNL model. This means that the travellers prefer routes having an overlap between alternatives, over completely distinct transit routes – be it an overlap of links, legs or transfer nodes. This result contradicts with the findings from analysis of route overlap in road networks (such as in Bovy, Bekhor, and Prato (2008)), as mentioned earlier in section 2.2. However, it is in line with some studies for transit route choice (Anderson et al., 2017; Hoogendoorn-Lanser and Bovy, 2007), and is perhaps explained by the claim of Anderson, Nielsen, and Prato (2017) that the PSC "can be seen as a measure of robustness of the trip by the traveller".

From the three path based PSCL models (PSCL Models 1 to 3), using a PSCL based on number of journey legs (PSCL Model 2) explains the travellers' route choice better than the link-based or travel time-based ones, as demonstrated by the final log-likelihood and LRS (with the same number of parameters in the three models). However, out of the four proposed PSC formulations, the one based on transfer nodes (PSCL Model 4) is found to have the highest final log-likelihood, meaning that it best explains the observed data, significantly better than any of the path-based PSCL models. As with the path-based PSC parameters, travellers are more likely to choose a route that includes a transfer stop that is shared by other routes for the OD pairs, implying more options of travelling to their destination stop, which could also be considered to be more robust. It is also noted that once the node-based overlap is added, the transfer penalty is found to increase steeply for transfers within bus/tram and from bus/tram to metro. This implies that when such overlap is ignored, the positive utility derived for routes with a transfer node overlap is captured by the transfer penalty instead, leading to an underestimation of the disutility induced by the number of transfers. Simultaneously, the transfer time parameter is found to decrease marginally, further implying that travellers dislike transferring irrespective of the transfer time.

4.4.3 Combining path and node-based overlaps

As shown in **Table 4.1,** the notions of overlap in terms of transfer nodes and overlap in terms of journey legs are distinctive, i.e. there can be routes where there is an overlap of transfer nodes, but no overlap of journey legs and vice-versa. Notwithstanding, often the overlap of transfer nodes and legs occurs simultaneously, leading to a high correlation between them (correlation ~0.7). In that case, omitting one of these factors may result in endogeniety. To understand and isolate the valuation of overlap of transfer nodes versus overlap of journey legs in more depth, we include them simultaneously. Different combinations of model formulations were estimated, and the best combination of parameters is presented (PSCL Model 5).

Estimating a model that includes both the leg-based and node-based PSC results in a slightly better model fit compared to the model with node-based PSC alone. In the combined model, the PSC parameters for overlap of transfer node is found to follow similar signs as in the individual model. However, surprisingly, the results indicate that once the overlap of transfer nodes is accounted for, subsequent overlap of journey legs is valued negatively. This means that travellers ideally prefer routes which have an overlap of transfer nodes, i.e. decision points, rather than of journey legs per se. The latter is perceived negatively once overlap in transfer nodes is accounted for. The main argument explaining the positive perception of route overlap for transit networks is the availability of alternate travel options in case of disruptions. From that perspective, having a common transfer point but distinct journey legs between routes meets the objective. On the other hand, having the overlap in journey legs, i.e. the same travel option before or after transfer for the two overlapping routes, does not help in case of disruptions, and is hence found to reduce the attractiveness of overlapped routes, compared to completely distinct routes. This could also be interpreted as the routes with a complete overlap of legs being considered similar (rather than as two distinct routes).

Conversely, the partial overlap of journey legs (in the form of links), is found to be valued positively implying that routes with partial overlap, i.e. some but not all common links, are preferred over completely independent journey legs. This is expected as routes with some common links could provide more options for at least a part of the journey-leg(s) in case of disruptions. Moreover, the routes with partial overlap are more likely to be situated in transit-dense city centre areas as opposed to outskirts. This could contribute to the preference towards them, everything else being the same, due to feelings of familiarity to the route, safety, preferred surroundings (especially for tourists in the case of Amsterdam), and perhaps a better level of service. The overall impact of route overlap depends thus on the relative values of link, leg and node overlap.

4.4.4 Cross-validation and sensitivity analysis

We undertook out-of-sample validation for each of our models using a cross-validation approach. Each model was re-estimated on (randomly selected) 80% of the data (305,836 observations), and the remaining 20% of the data (76,459 observations) was used as a validation dataset. This process was repeated five times (with replacement) for each model, and the average probability of chosen routes across all validation runs is presented in **Table 4.3**. In terms of prediction performance also, the node-based overlap model is found to perform better than the link or leg-based models for our data. The best performance of all is found to be for the combined node and leg-based model, marginally better than the node-based model.

Table 4.3. Cross-validation results.

Description	MNL Model	Link-based	Leg-based		Node-based	Combined
		PSCL Model 1 - links	PSCL Model 2 - number of legs	PSCL Model 3 - travel time	PSCL Model 4 - transfer node	PSCL Model 5 - travel time +transfer node
Average log-likelihood of validation data	-46,778	-46,753	-46,743	-46,753	-46,703	-46,695
Average probability of chosen route for validation data	58.89%	58.91%	58.92%	58.91%	58.96%	58.97%

Next, we perform a sensitivity analysis of our results to the choice set size. In this study, we use the ‘observed’ choice set as opposed to a synthetic choice set generation method. To ensure reliable estimates of travel attributes, only those routes with a minimum of 20 journeys in the half hour time slice (over all days) were used for the analysis. However, this results in many less frequently used routes being excluded from our analysis, reducing our choice set size. In case of choice models, it is well known that choice set size and composition may greatly impact results (Prato and Bekhor, 2007). For models that include route overlap (such as the PSCL), including irrelevant alternatives has been found to bias the results, hence it is advised to include the attractive routes only in the choice sets (Bliemer and Bovy, 2008; Bovy et al., 2008). However, one could argue that routes with less than 20 observed journeys in the time period of our analysis (specially in the outskirts of the city) could still be considered ‘relevant’. Hence, we test the sensitivity of our conclusions to choice set size by reducing the threshold of minimum journeys needed to include a route in our analysis. As we lower the threshold on the number of journeys, more ‘less preferred’ routes are included in the data set. In general, these ‘less preferred’ routes have more overlap between them – with 74% of OD pairs having some type of overlap with a threshold of 5 journeys, as compared to 63% for a threshold of 20 journeys.

Table 4.4 shows the model estimation statistics and path size factors for our best performing model (PSCL Model 5) for different journey thresholds. As expected, the model fit statistics improve as the number of observations increase. The PSC parameters are still found to be significant and with the same signs irrespective of the sample size, although their magnitudes decrease with reducing threshold. Amongst other parameters, the transfer related parameters and circuity were found to be sensitive to the composition of choice set, whereas the in-vehicle and waiting times were observed to be relatively stable. Similar to Ton et al. (2018), this study has undertaken a data-driven choice-set generation approach. The impact of choice-set size when using this approach on model estimates remains a topic for further research.

Table 4.4. Model sensitivity to choice set size

Description	**Minimum 20 journeys**	**Minimum 10 journeys**	**Minimum 5 journeys**
Number of observations	382,295	538,696	756,467
Maximum number of alternatives	4	6	7
Estimated parameters	13	13	13
Likelihood Ratio Test (compared to respective MNL)	838.6	987.2	988.6
Rho-square-bar	0.179	0.227	0.304
Parameter estimates*			
PSC – travel time	0.52	0.48	0.16
PSC – transfer nodes	-1.39	-1.10	-0.61

*p<0.01 for all estimates

4.4.5 Discussion

Our findings highlight the importance of transfer hubs for passenger route choice decisions. The perception of overlap is found to refer to decision points such as interchange locations. Having routes with common transfer locations that offer distinctive travel options to and from transfer locations is ideal from the perspective of travellers. Network topology analysis has demonstrated that having multiple (back-up) links in transit networks increases the robustness of a transit network in case of disruptions (Jenelius and Cats, 2015). Our findings imply that this also translates into increased attractiveness of overlapping routes compared to independent routes due to their contribution to journey-level robustness.

The analysis performed in this study can be extended to access stop choice. The models used in this study attempt to capture the value of robustness of routes with overlapping links or transfer nodes. More generally, this preference for more robust routes may also be reflected in case of the transit travellers' access stop choice. However, in most existing models of route or access stop choice, this impact is not considered. The dataset used in this study does not provide information on door-to-door journeys, hence it is not possible to observe the preferences of travellers on their choice of origin transit stop. However, it is an interesting research direction to check for this using a data set (such as travel diary) which will allow for such an examination. Lastly, our results show that depending on the types of modes between which the transfers are occurring, some transfers are preferred over others, in line with the findings from Garcia-Martinez *et al.* (2018) for intermodal transfers. Transfer penalty is expected to be a function of the transfer environment, such as level difference, number of crossings, shelters, availability of information, etc. In the case of Amsterdam, as with many other transit networks, many bus/tram lines are intended to serve as access/egress modes for the metro which is limited to major corridors. However, the higher transfer penalty for such transfers (as opposed to transfers within the bus/tram network) indicates that more attention should be given to making such intermodal transfers more seamless, thereby reducing the associated transfer penalty and making such journeys more attractive.

Finally, our study is subject to three main limitations. Firstly, crowding was not included as an attribute in any of our models. Even though the Amsterdam transit network is not very crowded, in some contexts crowding may have an impact on other attributes. Secondly, while the PSCL models used in this study allowed for capturing the correlation between alternatives, they do not capture the heterogeneity amongst travellers or the correlation due to the panel characteristic of the data. While the former could be explored using other model structures such as the mixed-logit, the latter is not possible given the characteristics of the dataset which does not contain panel information. Lastly, this study compared the alternate specifications of route overlap, including the new node-based overlap, using PSCL model only. It might be interesting to explore how these alternate definitions of overlap perform under other model structures used to incorporate overlap, such as the C-Logit, the PSL or more complex ones like the Paired Combinatorial Logit or Cross Nested Logit.

4.5 Conclusion

The main contribution of this study is that for the first time, we provide insight into how travellers perceive different types of overlap between routes while making urban transit route choice decisions. An empirical analysis by means of choice modelling was conducted for the urban multi-modal transit network of Amsterdam using smart card data. We defined route overlap in terms of overlapping links, journey legs, and transfer nodes. Overall, incorporating route overlap resulted in a significant improvement in model fit compared to the basic MNL model.

Our findings support the argument of Anderson, Nielsen, and Prato (2017) and Hoogendoorn-Lanser and Bovy (2007) that having multiple options of travel enhances the attractiveness of routes that have an overlap. On the one hand, our results show that the partial overlap of routes with some links overlapping is found to be preferred by travellers, presumably because it provides more options for travel in case of disruptions. Further, travellers ideally prefer routes that have common transfer locations, but not completely overlapping journey legs. This is intuitive, as having multiple (distinct) travel options at a transfer location adds to the robustness of their route choice decision. On the other hand, completely overlapping journey legs does not add any value in terms of robustness, and is hence found to reduce the attractiveness of overlapped routes, compared to distinct ones.

The majority of studies in the literature that consider transit route overlap measure it in terms of path only. In this study, not considering the overlap in terms of transfer nodes led to the contrasting conclusion of a positive valuation of overlapped legs by travellers. Hence, a key take-away from our results is that for transit route choice, it is important to define the overlap in terms of both path and nodes.

Overall, this study contributed to advancing the understanding of travellers' perception of overlap during transit route choice. It also added to the limited studies that empirically analyse route choice behaviour for large-scale multi-modal transit networks using smart card data. The results show differences in perceptions of travel times and transfer penalties by mode(s) used. The trade-off values between different route attributes obtained in this study also provides behavioural insights to transit planners and policy makers. Moreover, the methodology proposed to incorporate route overlap could be adopted for other transit networks to improve the performance and accuracy of route choice models, leading to better predictions.

Chapter 5 - Validation of a Transit Route Choice Model Using Smart Card Data

Validation of travel demand models, although recognized as important, is seldom undertaken. This chapter contributes to advancing transit route choice models by undertaking an out-of-sample validation of the models estimated in Chapter 4. Specifically, the estimated models are used to predict changes in travel demand in response to a major network restructure that included the introduction of a new metro line, and the predicted travel behavior is compared with the observed behavior. The smart card data from before and after the network change is used for estimation and validation, respectively. First, the MNL model with mode-specific travel attributes is estimated using data from each of the two time periods, and the estimated parameters are compared for stability. Next, the predictive ability of the model estimated on the data before the network change is evaluated by comparing it against the observed demand at different aggregation levels. Lastly, two alternate specifications of the model including the best performing model from Chapter 4 are compared in terms of their predictive ability. By conducting a posterior validation analysis we derive insights for improving future transit route choice models, specifically the ones estimated based on smart card data.

This chapter is based on the following article:

Dixit, M., Cats, O., van Oort, N., Brands, T., Hoogendoorn, S. (Under review) Validation of a Multi-Modal Transit Route Choice Model Using Smart Card Data.

5.1 Introduction

The last few decades have seen substantial research into discrete choice models of transit route choice (Bovy and Hoogendoorn-Lanser, 2005; Guo and Wilson, 2011; Liu et al., 2010). These models aid in understanding transit riders' preferences by revealing the relative valuation of various travel attributes, often specifically focusing on service quality characteristics such as those related to transfers (Garcia-Martinez et al., 2018; Guo and Wilson, 2011; Nielsen et al., 2021), crowding (Hörcher et al., 2017; Kim et al., 2015; Yap et al., 2020) or reliability (Swierstra et al., 2017). The relative valuations obtained from these models can be used for predicting passenger flows in response to changes in policy, enabling the comparison of alternative policy scenarios. When the selected model is close to the true representation of reality, the estimated parameters are expected to be stable for a reasonable range of temporal and spatial conditions, and the model forecasts are expected to resemble the observed demand. However, the process of model validation is only seldom undertaken, and model selection is typically made based on goodness-of-fit statistics such as log-likelihood and rho-squared (Parady et al., 2021). Although useful in their own right, models with high goodness-of-fit may not necessarily be well-specified and hence may not be transferable (Koppelman and Wilmot, 1982). Issues like overfitting, endogeniety, omission of variables, measurement errors, incorrect model structure or incorrect theoretical assumptions about the travel behaviour could lead to a misspecified model, which may still have an acceptable goodness-of-fit statistic.

Model validation can be defined as "the evaluation of generalizability of a statistical model" (Parady et al., 2021) and includes both internal validation or reproducibility and external validation or transferability. External validation can be further divided into spatial transferability, temporal transferability, and methodological transferability (Parady et al., 2021). Although recognised as important, external validation is rarely undertaken in the case of travel demand models, probably due to the lack of suitable data. Parady et al. (2021) highlight that only 4% of transport academic literature published between 2014 and 2018 conducted an external validation.

In recent years, revealed preference data in general, and smart card data, in particular, has become increasingly available for inferring route choices of transit travellers (see for example Hörcher et al. 2017; Jánošíkova et al. 2014; Kim et al. 2019; Yap et al. 2020). Depending on the penetration rate amongst transit riders, smart card data can provide information on almost all journeys made in the network at a highly disaggregated level. However, no information is available on the intention of the travellers, their origin location, and in many cases the time of arrival at the origin stop. Due to these limitations, several assumptions need to be made along the modelling process, specifically regarding the travellers' consideration choice set and the perceived level of service values. However, to the best of our knowledge, none of the studies that elicits route choice preferences from smart card data has attempted to validate their performance. This study aims to address this gap in the literature by undertaking an external validation of a transit route choice model using smart card data and thereby provide valuable insights on how transferable such models are and how we can facilitate their transferability.

A model of transit mode-route choice was developed for the urban transit network of Amsterdam, where a new North-South metro line was added to the existing bus, tram, and metro network in July 2018. Along with the addition of the new line, significant changes were made to the rest of the network (see Brands et al. (2020) for details). This major network change provides an opportunity to perform an ex-post evaluation of the route choice model developed based on data before the network change. Two types of validation tests are undertaken. First, we compare the model parameters estimated for the ‘before’ model with the locally estimated model developed based on the data ‘after’ the network change. The two data sets used are ~3 months apart. Second, the demand changes estimated using the ‘before’ model are compared with the observed demand after the network change. The results aim to establish the validity of route choice models estimated using smart card data for predicting the change in travel behaviour because of a major network change.

The rest of the chapter is structured as follows: we start with reviewing the literature on external validation of travel demand models in *Section 5.2*. *Section 5.3* describes the study setting and the various statistical tests used for model validation, along with the model specifications. *Section 5.4* presents the results of the validation tests undertaken on our data, and *Section 5.5* discusses the main conclusions.

5.2 External Validation of Travel Demand Models

External validation of models, or transferability, implies the ability of a model developed in one context to be useful in another context. Transferability is implicitly assumed when models are used to predict change in demand in response to a policy change. Some of the earliest literature on (external) model validation dates back to Atherton and Ben-Akiva (1976) and Train (1978). Since then, most work in this area has focused on the temporal transferability of models over long time horizons (often more than 10 years) and/or their spatial transferability. The primary motivation for such studies was to reduce costs of data collection and model development by using an existing model for a comparable region or during a different time period for the same region. Parady et al. (2021) provide a comprehensive review of the recent literature on the validation of discrete choice models in transportation. Here we summarise the main issues considered in external validation studies and the corresponding learnings from them.

A fundamental theoretical assumption behind any model transferability is the consistency of underlying behavioural theory in both contexts. Koppelman and Wilmot (1982) highlight that model transferability is a “property of the estimation and application contexts, as well as the specification of the model”. Naturally, a model with highly context-specific variables will not be transferable to a new context. Sometimes the ASCs are updated based on the application context to account for average changes in unobserved variables between the two contexts (Atherton and Ben-Akiva, 1976; Badoe and Miller, 1995; Sanko and Morikawa, 2010). While the updated ASCs capture the mean contribution of the unobserved terms, there could also be differences in the variance of these unobserved terms. Hence, before transferring a model, the scale for the transferred model needs to be updated to match the scaling differences between the two contexts (Swait and Louviere, 1993). In cases where the estimation and application contexts are widely different, one could implement a partial model transfer with varying transfer scales for different sub-groups of variables (Gunn et al., 1985). This is especially applicable when some parameters are more transferable than others. For example, Fox et al. (2014) found the level of service parameters to be more transferable than cost parameters in their study of mode-destination choice models.

Multiple studies have noted that, generally, an improved model specification improves transferability (Badoe and Miller, 1995; Fox et al., 2014; Rossi and Bhat, 2014). However, some others also highlight the risk of overfitting which may reduce the transferability of models. For example, Fox (2015) found that although incorporating taste heterogeneity in time and cost parameters improved model fit for their base data, it did not necessarily result in enhanced transferability. Badoe and Miller (1995) also report a similar finding where over-specification led to reduced transferability. Overall, it is noted that a good fit in the estimation context may not be sufficient.

Another issue of concern is the ability of a model to capture causal relationships. As clearly highlighted by Atherton and Ben-Akiva (1976): “To be transferable, then, it is not enough that the model merely fit existing data; it must also explain why travel behaviour changes as conditions change. Rather than simply correlating existing travel behaviour with socioeconomic characteristics and transportation level of service, the model specification must represent the causal relationships between these variables. Thus, the causal specification of a model is a precondition to its consideration for transferability.” For example, Chorus and Kroesen (2014) argue against the transferability of hybrid choice models for predicting policy outcomes, as these models (theoretically) cannot capture the causal relationship between the latent variable and the travel choice.

The only way to empirically establish whether a model is under/over specified or if it captures the causal relations required for transferability is to undertake a posterior analysis of transferability. Nonetheless, as Koppelman and Wilmot (1982) note, such posterior analyses of transferability are undertaken with the intent to provide insights that can be helpful for (future) prior transferability studies. This study aims to get such insights for the case of transit route choice models, specifically the ones estimated based on smart card data.

So far, most validation studies in the literature have been for mode or mode-destination choice models. In the case of route choice models, some studies undertake an internal validation (see for examples Lai and Bierlaire (2015); Mai (2016)), but very few undertake an external validation. Bekhor and Prato (2009) were the first to consider the issue of transferability of route choice models. They undertook a spatial transferability assessment of traffic route choice models based on two independent revealed preference survey data sets, one each for Boston and Turin networks. In addition to assessing the transferability of the route choice models, they also evaluated the transferability of path generation techniques. In their case, the transferability of route choice model parameters could not be verified, partly due to the dissimilarity in characteristics between the two networks.

To the best of our knowledge, none of the studies so far have undertaken an external validation of a transit route choice model. This study addresses this gap by undertaking a transferability analysis across two closely spaced time periods for the same urban area, which allows for many exogenous factors to be controlled for, including any major changes in the underlying population. Specifically, the following issues are investigated using the smart card data from before and after a major network change:

1. How transferable are models of transit route choice estimated using smart card data, and can they be used for forecasting the changes in demand because of network changes?
2. How does omitting/adding relevant variables (determined based on improved goodness of fit measures in the base context) impact models’ prediction performance?

5.3 Method

5.3.1 Case study context

In July 2018, a new metro line (the north-south line) was introduced in the urban transit network of Amsterdam, the Netherlands, adding significant capacity to the existing network of metro, bus and tram lines. The new metro line runs through the dense historical city centre, and connects the northern part of the city with the centre – a connection which was made earlier via buses with highly circuitous routes. The opening of the new metro line was accompanied by changes to the existing bus and tram network, including the addition of new feeder routes and re-routing or removal of duplicate routes. The new metro line differs from the existing ones in a few aspects – some of the stations (especially the ones in the city centre) are deeper than the existing metro stations implying a longer access time to the metro. In addition, the frequency for the new line is higher than the frequencies offered on the other metro lines (see Brands et al. (2020) for details).

This significant change in public transport supply provides an opportunity to undertake a transferability analysis for the transit route choice models developed for the network using the two time periods corresponding to 'before' and 'after' the opening of the new north-south metro, as shown in **Figure 5.1**. Although the time periods used in this study are very close apart, the major changes to the transit network supply cause significant changes to the flow patterns (as shown in Brands et al. (2020)), making this case study ideal for undertaking a model transferability analysis.

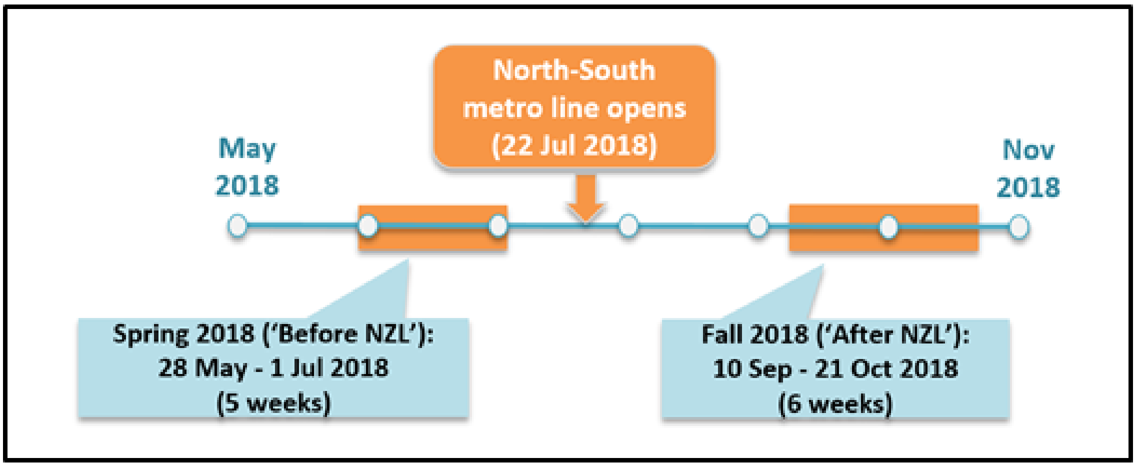


Figure 5.1. The time period for validation analysis.

5.3.2 Validation assessment

We divide the validation metrics into three categories, similar to those defined in Koppelman and Wilmot (1982). The first category relates to the stability of estimated parameters, while the next two assess the predictive ability of the model. **Figure 5.2** shows this classification, and the sets of metrics used for each category. Subsequent sections elaborate on each of the metrics used.

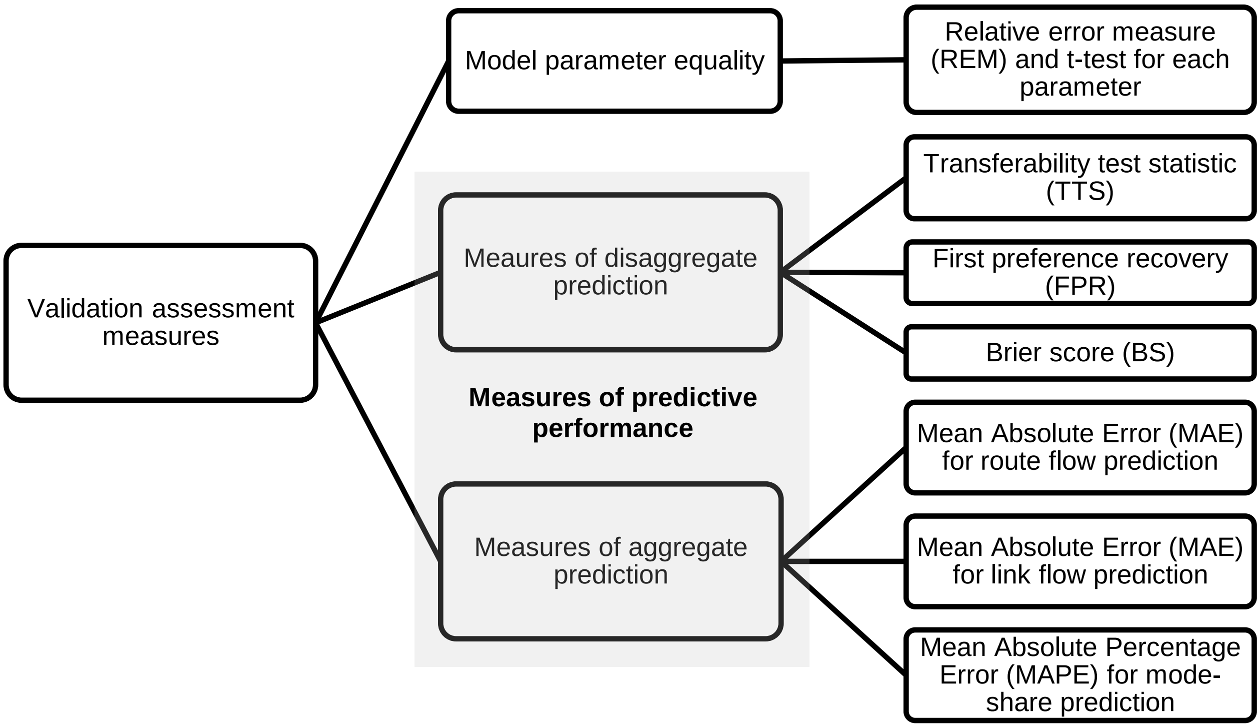


Figure 5.2. Validation assessment metrics used

Model parameter equality
The first test of transferability consists of comparing the parameters estimated for the base and transfer contexts (the before and after situations in our case). This helps establish whether some parameters are more transferable than others. Since the two datasets are from different time points, we first check for differences in scale parameters between the two. For this, the two datasets are pooled together and the scale parameter is estimated relative to the 'before' dataset.

After adjusting for scale differences, the parameters estimated for the two cases are compared. For each estimated parameter, the relative error measure (REM) is calculated as:

$$REM_{\beta_k} = \frac{\mu\beta_k^{after} - \beta_k^{before}}{\beta_k^{before}} \quad (5.1)$$

Where,
μ is the scale parameter to account for differences in error variance between the two cases,
β_k^{after} is the parameter for attribute 'k' estimated for the after case, and
β_k^{before} is the parameter for attribute 'k' estimated for the before case.

Next, we check for statistical significance of the differences in each of the model parameters by means of a t-test, as described in Fox (2015). The t-statistic, in this case, is given by,

$$t\left(\mu\beta_k^{after} - \beta_k^{before}\right) = \frac{\mu\beta_k^{after} - \beta_k^{before}}{\sigma\left(\mu\beta_k^{after} - \beta_k^{before}\right)} \quad (5.2)$$

The denominator σ corresponds to the standard error of the difference in parameters. In our study, although the two datasets were collected a few months apart, we do not link individual observations collected in different periods. When the covariance is assumed to be zero, the standard error of difference is given by,

$$\sigma\left(\mu\beta_k^{after} - \beta_k^{before}\right) = \sqrt{(\mu\sigma\left[\beta_k^{after}\right])^2 + (\sigma\left[\beta_k^{before}\right])^2} \tag{5.3}$$

Where,
$(\sigma\left|\beta_k^{after}\right|)^2$ is the standard error of β_k^{after}, and
$(\sigma\left|\beta_k^{before}\right|)^2$ is the standard error of β_k^{before}.

Disaggregate measures of predictive ability
Next, we assess how well the 'before' model can predict the outcome of the network change. For this, the model parameters estimated based on the 'before' data are used to estimate the probabilities for the 'after' situation. The outcomes obtained using the transferred model are then compared to those from the locally estimated model (i.e. model estimated with same the specification but using the 'after' data). In this section, we discuss the methods for comparing the performance for individual-level predictions. As there is no agreement in the literature on the best metric for this, we use multiple metrics, each providing a different perspective on it, as described below:

- **Transferability Test Statistics (TTS):** The TTS statistic is similar to a likelihood ratio test undertaken between transferred and the locally estimated model, both applied to the 'after' data. It is a strict pass/fail test and is chi-squared distributed with degrees of freedom equal to the number of model parameters. This has been used by Atherton and Ben-Akiva (1976) and Koppelman and Wilmot (1982) among others to test model transferability.

  $$TTS_{after}(\beta_{before}) = -2 * (LL_{after}(\beta_{before}) - LL_{before}(\beta_{before})) \tag{5.4}$$

  Although commonly noted, it has been observed that almost all models fail this strict test of transferability (Badoe and Miller, 1995; Fox, 2015).

- **First preference recovery (FPR)**: Also referred to as 'percentage of correct predictions', this shows the percentage of choices correctly estimated by the model, given by**:**

  $$FPR = \frac{100}{N}\sum_{i=1}^{N} y_i^p = y_i^o \tag{5.5}$$

  Where,
  y_i^p is the predicted choice for individual 'i', and
  y_i^o is the observed choice for individual 'i', and
  N is the number of individuals (observations) in the data.

  As opposed to TTS, the FPR provides an indication on the degree of transferability and can be used to compare alternative models in terms of how well they can predict individual choices. It can also be useful when comparing the results with similar studies in the literature. However, a major limitation of this measure is its inability to differentiate between the range of probabilities assigned to the chosen alternatives (de Luca and Cantarella, 2016). Hence, we look at another measure – Brier score - of disaggregate predictive performance that considers the probabilities assigned to the chosen and non-chosen alternatives.

- **Brier score (BS):** Brier score (Brier, 1950) is an absolute measure used to quantify the accuracy of probabilistic predictions. For each alternative in each observation, the predicted probability of choosing it is subtracted by the actual outcome. The square of this value is summed across all alternatives for each observation, and averaged across all observations. Mathematically, it is given by:

$$BS = \frac{1}{N}\sum_{i=1}^{N}\sum_{r=1}^{R_i}(P_{ir} - y_{ir})^2 \tag{5.6}$$

  Where,
  P_{ir} is the predicted probability an individual 'i' chooses alternative route 'r'
  y_{ir} is equal to 1 is alternative route 'r' is chosen by individual 'i' and 0 otherwise,
  R_i is the number of alternative routes available to individual 'i', and
  N is the number of individuals (observations) in the data.

  The Brier Score has a minimum value of 0 for perfect predictions, and a maximum value of 2 for the worst possible prediction.

Aggregate measures of predictive ability

Next, we compare the aggregate shares estimated by the models. To do this, individual probabilities are summed to calculate the market shares for each alternative route for each OD pair. The aggregate predictions are assessed using three metrics addressing different levels of aggregation, as discussed below:

- **Predictions per route - Mean Absolute Error (MAE):** The predicted shares (passenger flows) are then compared to the observed ones, and the Mean Absolute Error (MAE) for each origin-destination (OD) pair is calculated as:

$$MAE_{od} = \frac{1}{R}\sum_{\mathrm{r}=1}^{R}\left|S_r^p - S_r^o\right| \tag{5.7}$$

  Where,
  S_r^p is the predicted flows for alternative route 'r' for origin-destination pair 'o-d',
  S_r^o is the observed flows for alternative route 'r' for origin-destination pair 'o-d', and
  R is the number of available routes for origin-destination pair 'o-d'.
  The MAE_{od} is then averaged across all ODs to get an average MAE per route.

- **Predictions per link - Mean Absolute Error (MAE):** The passenger flows per route are aggregated to calculate flows for each link. A link here refers to the path connecting two consecutive transit stops, which may be used by multiple transit routes. Similar to the route level, MAE is calculated for each link, and a mean MAE over all links is reported. The percentage error in predicting the flow on each link is also visualized to identify patterns.

- **Predicted modal shares – Mean Absolute Percentage Error (MAPE):** The predicted passenger flows on each route are further aggregated to calculate the market share for each of the mode combinations. This is specifically relevant in our case as we would like to know how well the model performs when estimating the impact of the network changes on the public transport mode-shares. The observed and predicted mode shares are compared, and the Mean Absolute Percentage Error (MAPE) is calculated for each mode as,

$$MAPE_{modes} = \frac{1}{M}\sum_{\mathrm{m}=1}^{M}\left|P_m^p - P_m^o\right| \tag{5.8}$$

Where,
P_m^p is the predicted mode-share for mode (combination) *'m'*,
P_m^o is the observed mode-share for mode (combination) *'m'*, and
M is the number of mode (combinations).

5.3.3 Data preparation and model specification

We use a combination of smart card and Automated Vehicle Location (AVL) data for the route choice model estimation and validation analysis (see van Oort et al. (2015a) for an overview of the Dutch smart card system). The raw data is processed by undertaking cleaning, destination inference and transfer inference to form a journey database (see Dixit et al. (2019b)) for more details on these steps). For undertaking route choice analysis, we use only the morning peak period for our model estimation, as it is expected to have a higher share of commuters during this time, which are typically more regular travellers making their travel choices more conscious (Fox and Hess, 2010). The choice set is derived based on observed routes in the data set. Transit stops in close proximity are clustered together to form a more realistic consideration choice set, and a threshold of minimum 20 journeys for each route in the before period is applied to ensure only reasonable routes are included (see Dixit et al. (2021) for more details on this). After applying all filters, a dataset of 382,295 observations for the before period and 563,210 for the after period is obtained which is used for estimation and validation of the model, respectively.

We specify and test three models of transit route choice. We start with an MNL model with mode-specific travel attributes, with the deterministic component of utilities as specified in Equation (5.9).

$$\begin{aligned} V^{MNL\_specific} = {} & \beta_{ivt_{bus}} * IVT_{bus} + \beta_{ivt_{tram}} * IVT_{tram} + \beta_{wait_{bt}} * WT_{bt} + \beta_{tt_{metro}} \\ & * TT_{metro} + \beta_{trans_{bt}} * Trans_{bt} + \beta_{trans_{bt}} * Trans_{btm} + \beta_{trans_{m}} \\ & * Trans_{m} + \beta_{TrT} * TrT + \beta_{Circ} * Circ + MSC_{Bus} + MSC_{Tram} \\ & + MSC_{Metro} \end{aligned} \tag{5.9}$$

Where,
IVT_{bus} and IVT_{tram} are the in-vehicle times by bus and tram in minutes, respectively,
WT_{bt} is the initial waiting time for bus and tram modes,
IVT_{metro} is the travel time by metro including the initial waiting time at the platform,
$Trans_{bt}$, $Trans_{btm}$ and $Trans_{m}$ are the numbers of transfers made within the bus/tram network (which includes bus-bus, tram-tram and bus-tram transfers); transfers between metro and bus/tram; and transfers within the metro network, respectively,
TrT is the transfer time in minutes,
$Circ$ is the circuity of the route measured as the ratio of network to Euclidean distance, and
MSC_{bus}, MSC_{tram} andMSC_{metro} are the mode-specific constants for bus, tram and metro.

Next, to analyse the impact of omitting variables, instead of mode-specific in-vehicle time and transfer penalties, we use generic ones. The deterministic utility function in this case is shown in Equation (5.10).

$$V^{MNL\_generic} = \beta_{ivt} * IVT + \beta_{wait_{bt}} * WT_{bt} + \beta_{trans} * Trans + \beta_{TrT} * TrT + \beta_{Circ} * Circ + MSC_{Bus} + MSC_{Tram} + MSC_{Metro} \quad (5.10)$$

Where,
IVT corresponds to the in-vehicle time in minutes, and
$Trans$ is the number of transfers made within or across modes.

Lastly, we test the model which incorporates the overlap between alternative routes. For this, we use a Path Size Correction Logit model which includes overlap of path and transfer nodes. The path size correction terms for journey legs (PSC_i^T) and transfer nodes (PSC_i^X) are as defined in Dixit et al. (2021), given by

$$PSC_i^T = -\sum_{l\in\Gamma_i}\left(\frac{t_l}{T_i}\, ln \sum_{j\epsilon C}\delta_{lj}\right) \quad \text{and} \quad PSC_i^X = -\sum_{n\in K_i}\left(\frac{1}{X_i}\, ln \sum_{j\epsilon C}\delta_{nj}\right) \quad (5.11)$$

Where,
t_l = travel time for journey leg l in route i,
T_i = total travel time for route i,
Γ_i = set of all legs for route i,
C = set of all routes between the chosen origin-destination pair,
δ_{lj} = leg-route incidence between leg l belonging to alternative route j,
X_i = Number of transfer nodes in route i,
K_i = set of all nodes for route i, and
δ_{nj} = node-route incidence between node n belonging to alternative route j.

The path size correction terms defined above are added to the deterministic utility function with mode-specific travel attributes as defined in Equation (5.9).

5.4 Results and Discussion

We first evaluate the validity of the MNL model with mode-specific travel attributes as described in Equation (5.9) using the measures discussed in **Section 5.3.2**. Then, the impact of variable omission is examined by testing the validity of the two alternate model specifications.

5.4.1 Model parameter equality

We start with examining the stability of the estimated model parameters across the before and after time periods. Before comparing the parameters, the two models were checked for differences in scale parameters. The scale difference was found to be significant with a value of 0.92 for the ‘after’ model relative to the ‘before’ model, implying a lower variance in the unobserved parameters for the ‘after’ case as compared to the ‘before’ case.

Table 5.1 shows the parameters estimated from the two models after scaling and the corresponding REM and t-test statistic for each. The relative error measure is the highest for the mode-specific constants, implying a significant difference in the average effect of unobserved (excluded) variables specific to each mode between the two contexts. This could include attributes like comfort, safety, cleanliness, reliability, weather protection at stations, availability of information, ease-of-navigation or any other inherent preference/dispreference for any particular (public transport) mode. Some of these attributes are expected to change after the introduction of the new line. For example, deeper stations of the new metro line may reduce the attractiveness of the mode. On the other hand, the higher frequency and more options for travel may lead to it being more attractive. Amongst the rest of the parameters, circuity is found to have the highest change (an increase of 45% in magnitude), followed by the number of transfers within the metro which is found to decrease in magnitude by 19%. Although the REM values for all other parameters are approximately 10% or less, the null hypothesis of the parameters being identical across the two cases is rejected for most of them (with a 95% confidence interval). Only the travel time by metro, number of transfers between bus and tram, and the transfer time are found to be stable across the two time periods as per the t-statistic.

Table 5.1. Model parameter comparison between models estimated on 'before' and 'after' datasets.

Parameter*	Before	After**	REM	t-statistic	Significantly different?
Mode-specific constant for bus <fixed>	0.00	0.00	-	-	-
Mode-specific constant for tram	0.49	0.25	-48.5%	-13.90	Yes
Mode-specific constant for metro	0.84	0.37	-56.0%	15.04	Yes
Bus in-vehicle time (mins)	-0.11	-0.12	10.4%	4.97	Yes
Tram in-vehicle time (mins)	-0.09	-0.10	11.1%	6.54	Yes
Effective wait time bus/trams (mins)	-0.19	-0.20	5.6%	4.48	Yes
Metro time[a] (mins)	-0.09	-0.10	3.6%	1.22	No
Number of transfers between bus & tram[b]	-1.24	-1.21	-3.1%	0.80	No
Number of transfers between metro and bus/tram	-2.38	-2.19	-9.0%	6.83	Yes
Number of transfers within metro	-1.50	-1.22	-19.3%	6.30	Yes
Transfer time (mins)	-0.25	-0.25	0.2%	0.07	No
Circuity	-0.43	-0.63	45.1%	9.17	Yes

*p<0.01 for all estimates
**the reported estimates are after adjusting for scale differences
[a]includes in-vehicle time and origin waiting times.
[b]includes bus-bus, tram-tram and bus-tram transfers.

There could be several reasons for the differences in estimated parameters between the two contexts. Firstly, there could be contextual factors that are not captured by the observed variables that could differ between the two contexts. Secondly, the underlying population may have changed – some travellers may have stopped using transit after the network change while other new travellers may have been added. Also, some existing travellers may have reduced/increased their travel frequencies. A related point is the possible presence of endogeniety, especially since our study is based on observational data. The models assume the explanatory variables to be exogenous, which may not be true. This could be due to multiple reasons, including omitted variables. For example, the new metro line has newer, cleaner and more aesthetically pleasing trains and stations – contributing toward comfort, which is not included in our model(s). Concurrently, the new metro line provides direct routes with lower circuity values compared to the rest, making the circuity correlated with the unobserved attribute of comfort. When using a model to predict the demand in response to changes in policy, it is important to have a model capturing the causal relationship between them. The smart card data used for this study does not provide the origin (home) location of the travellers. The missing attributes (such as access/egress distance or time, comfort levels, reliability, and accessibility of modes among others) and/or endogeniety may hamper the establishment of a causal relationship. Lastly, one cannot theoretically rule out that our model may have been misspecified (wrong model structure or non-linear relationships between variables), or that the behavioural theory is altogether inconsistent with the observed behaviour. Irrespective of the reasons behind the instability of model parameter values, our results imply that one should be cautious when making inferences on the relative valuation of travel time or service quality attributes from such models, specifically if they have not been thoroughly validated.

5.4.2 Predictive performance

Disaggregate measures

Next, we test the predictive ability of the model by forecasting the impact of the network change at an individual, route, link and mode level. The predictions are compared with the predictions from a locally estimated model (i.e. model estimated with same the specification but using the 'after' data) to benchmark the performance. We start with the measures of predictive performance at a disaggregate level, which are shown in **Table 5.2**. The TTS, compared against the chi-squared distribution for 11 degrees of freedom, strongly rejects the hypothesis that the two sets of parameters are equal. However, as many other studies note, most models fail this test of model transferability, but may still be good in their predictive abilities (see for example Badoe and Miller (1995); Forsey et al. (2014); Fox (2015)).

Table 5.2. Disaggregate measures of predictive ability for locally estimated and transferred models.

Statistic	Locally estimated model	Transferred model	Difference
Log-likelihood	-328,646	-330,299	0.5%
TTS	-	3,306	
First preference recovery	71.7%	71.5%	0.7%
Brier score	0.370	0.372	-0.3%

The FPR of over 70% is found to be rather high compared to values reported for other route choice models in the literature, where this percentage ranges between 51% and 73% for some of the recent studies (Parady et al., 2021). Moreover, both FPR and the Brier score of the transferred models are found to be very close to the locally estimated models (<1% difference), with the FPR being slightly higher and the Brier score marginally lower in the case of transferred models. Hence, although many of the parameter estimates differ for the two periods, the predicted choice probabilities of the transferred model at an individual level are found to be close to the locally estimated model.

Aggregate measures

Disaggregate measures like FPR and Brier score are often used for assessing the models in terms of their ability to predict individual-level choices in the new context. However, in most applications, one is more interested in the predictions at the mode, route, or link levels. Hence, we analyse the performance of the ‘before’ model to predict the market shares at each of these levels.

First, we use the MAE to compare the local and transferred models in terms of their predictions at the route level (**Table 5.3**). The MAE for the transferred model shows an average error of 45 journeys per route for the transferred model as compared to 43 for the locally estimated model. Examining the predicted flows at a link level, we observe an MAE of 328 passengers per link over the entire morning peak period when predicted using the transferred model, ~8.6% higher than that those obtained by the locally estimated model.

Table 5.3. Aggregate measures of predictive ability for locally estimated and transferred models.

Statistic	Locally estimated model	Transferred model	Difference
MAE per route	43.2	45.4	5.0%
MAE per link (AM peak)	302.2	328.2	8.6%

Figures 5.3 and 5.4 show the error in flow prediction at the link level by the local and transferred models, respectively. The width of the lines corresponds to the observed flow on the link. A positive error implies that the model overestimated the flow on the link, while a negative error means an underestimate of observed flow. The maps show that the link-level flow predictions using the two models are overall similar, implying that using the ‘before’ data set for estimation of the model is not a problem per-se, compared to the inherent estimation errors when using such a model. The maps can give an indication of the possible causes of such errors. For example, close to the central station, there are two parallel tram routes with one showing an underestimation while the other an overestimation of flows (highlighted with a red circle in the maps). This could be attributed to the assumptions made regarding the consideration choice set for the model. The smart card data does not provide information on the origin (home) location of the travellers. Hence, stops within a maximum distance of 500m were clustered together to form the consideration choice set for travellers (for more details on the clustering process see Dixit et al. (2021)). In the absence of the actual origin location, all routes between the origin-destination stop-clusters are assumed to be equally accessible for the travellers. However, for the origin-destination pairs such as the one highlighted where the distance travelled is very short, travellers are likely to choose the transit stop closest to them as opposed to the one with the shortest generalized cost that is predicted by the model. Hence, in such cases the link-level predictions can be erroneous, and should be used with caution.

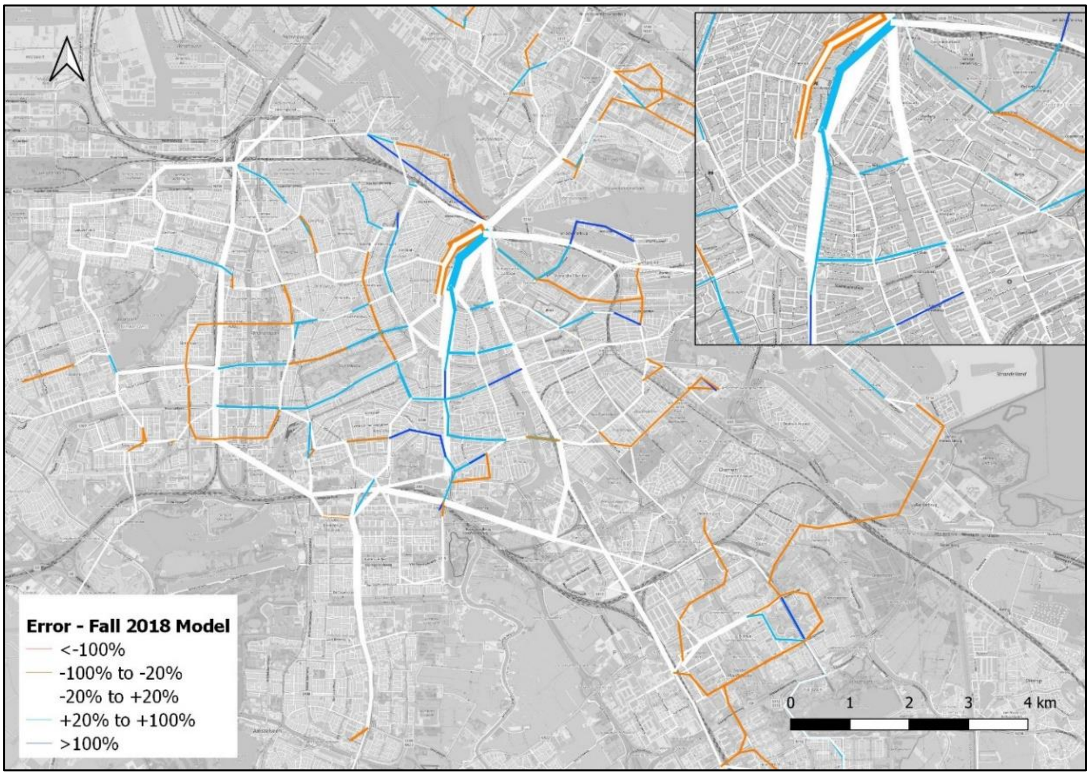


Figure 5.3. Percentage error in prediction per link for the locally estimated model.

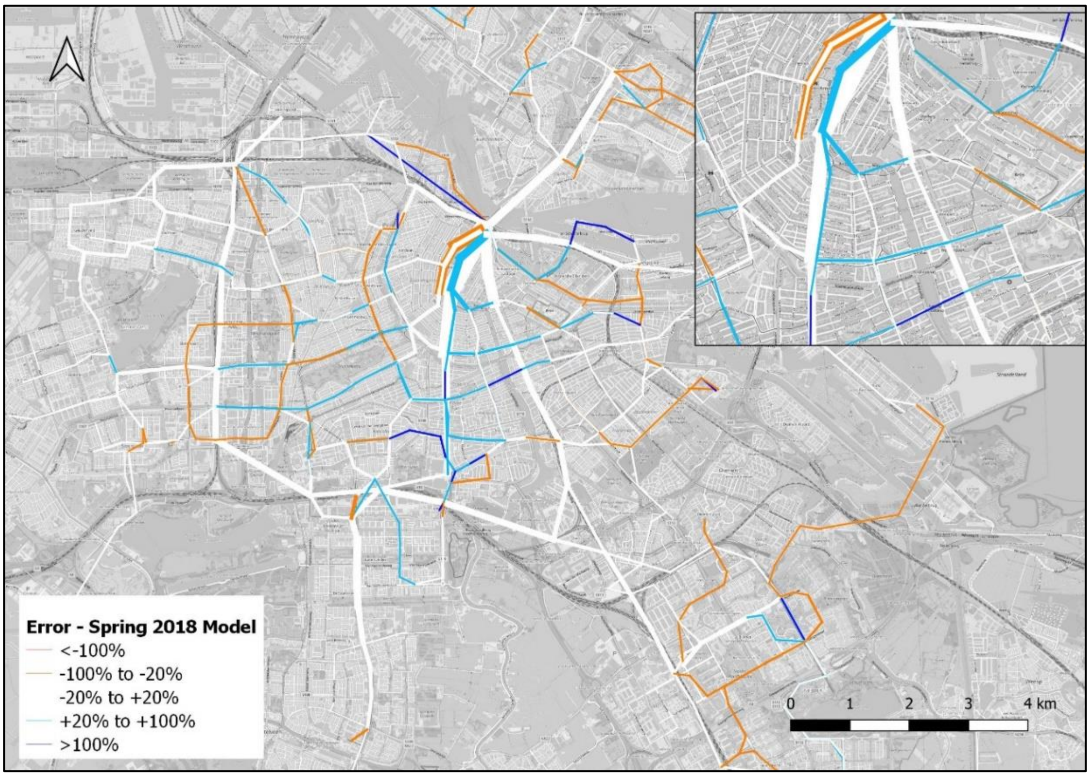


Figure 5.4. Percentage error in prediction per link for the transferred model.

Next, we compare the market shares of each transit mode combination in the data (**Table 5.4**). The predicted and observed shares are found to be close to each other with a difference of less than 1 percent for most mode combinations for both local and transferred models. As expected, the MAPE is found to be slightly higher for the transferred model than for the local model. Overall, the transferred model is found to perform close to the local model in terms of mode-share predictions as well.

Table 5.4. Observed and predicted mode shares

Mode combination	Observed share	Predicted share		Error	
		Local model	Transferred model	Local model	Transferred model
Bus only	18.4%	18.6%	18.1%	0.2%	-0.3%
Tram only	31.8%	31.6%	31.4%	-0.2%	-0.4%
Metro only	37.8%	37.8%	37.7%	0.0%	-0.1%
Bus+tram	0.4%	0.4%	0.4%	0.0%	0.0%
Bus+metro	3.3%	3.1%	3.1%	-0.2%	-0.2%
Tram+metro	8.2%	8.4%	9.3%	0.2%	1.1%
All modes	0.0%	0.0%	0.0%	0.0%	0.0%
MAPE				0.1%	0.3%

5.4.3 Impact of omitted variables

In this section, we analyse the impact of omitting/adding one or more variables on the model's predictive performance (**Table 5.5**). We test two scenarios:

1. **Generic MNL:** Generic travel time and transfer parameters as opposed to mode-specific ones as specified in Equation (5.10).

2. **Including overlap:** Path size correction logit (PSCL) model including path size correction terms as defined in Equation (5.11) to incorporate the impact of overlap between alternate routes.

Both in the estimation ('before') and the prediction ('after') contexts, the model with generic travel time and transfer parameters has the worst fit for the data, as shown by the respective log-likelihood values (even when adjusted for the number of parameters). The predictive performance is also found to suffer significantly when generic attributes are used. Conversely, when overlap is incorporated, the model fit is improved significantly in the estimation context (Likelihood ratio statistic of 838.6 exceeding the critical χ^2 value of 9.2 at 1% significance level (df=2)), but the log-likelihood for the prediction context is found to be lower than the mode-specific MNL model. In terms of predictions, there is a marginal improvement in the FPR and MAE for route-level prediction. However, the Brier score and the predictions at mode level are slightly worse.

Table 5.5. Predictive measures of transferability for alternate transferred model specifications.

Statistic	Generic MNL	Mode-specific MNL	Including overlap (PSCL)	Local model (Mode-specific MNL)
Number of parameters	7	11	13	11
Log-likelihood of estimation ('before' context)	-234,899	-233,892	-233,473	-
Log-likelihood of prediction ('after context')	-335,470	-330,299	-330,598	-328,646
First preference recovery	71.2%	71.5%	71.6%	71.7%
Brier score	0.379	0.372	0.372	0.370
MAE per route	48.2	45.4	45.2	43.2
MAPE (mode-share)	0.60%	0.30%	0.34%	0.14%

When inferring route choice using revealed preference data sources in general, and smart card data in particular, the analyst does not have any 'direct' information on which attributes were considered by the decision-maker when making the choice. Hence, the selection of attributes to be included in the model depends heavily on the judgement of the analyst, and often data availability. It is known that the omission of a relevant variable can impact the model transferability (Koppelman and Wilmot, 1986), especially when the missing variable is a confounding one. Conversely, if a simple model with fewer variables can perform just as well, then excluding variables can make the data collection as well as estimation easier. In our case, generic travel time and transfer parameters negatively impact the predictive ability of the transferred models. In contrast, including overlap does not offer a clear improvement in the predictive ability. In the end, the optimal selection of attributes depends on the purpose for which it is intended to be used. For predicting passenger flows in the regions where the tram was replaced by the new metro line, all attributes that distinguish a metro from a tram should ideally be included in the model. The mode-specific MNL shows that the travel time and transfer parameters are different for different modes. Hence, using generic travel time and transfer parameters impacts the predictive performance of the model significantly. On the other hand, correcting for route overlap typically leads to an improvement in model fit in the case of route choice models (like in our case for the estimation context). However, our results seem to suggest that it may not necessarily increase the transferability of the models, and the overlap term(s) may be context-specific and hence not as transferable.

5.5 Conclusion

This study adds to the scarce literature on the validation of travel demand models and is the first to undertake an external validation for a transit route choice model. The model was developed based on smart card data for the urban transit network of Amsterdam and was used to predict the impact of a significant network change (i.e. the introduction of a new metro line) on the route choice behaviour of travellers. Validation was conducted by comparing the parameter values and a series of statistical performance indicators for the predictions with the observed behaviour after the network change.

Our results are overall in agreement with existing literature: the conclusion regarding model transferability depends on the (statistical) test used (Koppelman and Wilmot, 1982). In our case, model parameter equality failed for most attributes, implying care should be taken in directly inferring behavioural insights from the parameter values from models such as those used in this study, specifically if they have not been thoroughly validated. However, the predictive performance of the transferred model was found to be close to the locally estimated model. When compared with the observed choices at an individual level, the model performed satisfactorily with a First Preference Recovery of 71.5%. Moreover, the predicted mode-shares were close to the observed ones, with a MAPE of 0.3%. When used for route and link level predictions, the errors were relatively larger, but the performance of the transferred model was similar to the local model (less than 10% error increase).

We also investigated the impact of omitting relevant variables on predictive performance. When the mode-specific travel time and transfer parameters were replaced by generic ones, the performance suffered significantly. Conversely, including overlap in the model specification did not offer a clear improvement in model predictions, even though it had a better fit for the base data. This suggests that overlap definition may be context specific and could perhaps be excluded when using a route choice model for predictions in favour of a parsimonious model.
When using smart card data for travel demand modelling, several assumptions are made regarding travellers' consideration choice set and perceived travel attributes. Visualizing link-level prediction errors can help indicate potential causes of errors. In our case, the assumption regarding consideration choice set may be responsible for some of the prediction errors, which are consistent between local and transferred models.

"All models are wrong, but some are useful" (Box, 1976). To establish how wrong a model needs to be to stop being useful, we need more studies undertaking validation analysis for different networks and policy scenarios. Guidelines and standards on what is considered acceptable in terms of the various transferability statistics remain yet to be defined. In the past, the cost of undertaking a model validation was high primarily due to data collection costs. The abundance of passively collected data such as the smart card provides an opportunity to validate transit route choice and assignment models at relatively low additional costs. Hence, validation must become an integral part of the development process for such models, and should be considered non-negotiable when using them for deriving any policy recommendations.

Chapter 6 - Conclusion

With an overall objective of improving public transport planning using automated sources of transit data, the research presented in this dissertation contributed to advancing transit performance assessment and route choice modeling in the context of urban multi-modal transit networks. Methods were developed with the intention of leveraging key strengths of such data sources, namely having network-wide journey data at a high spatio-temporal granularity, while adjusting to its limitations on the depth of information available such as the intention of travelers, time of arrival at the origin stop and socio-demographic characteristics. It utilized data from the multi-modal transit network of Amsterdam for developing and applying these methods. In this chapter, we synthesize the key findings from the research, implications for practice, and directions for future research.

6.1 Main Findings

Based on the research undertaken, we provide answers to the four research questions formulated in *Section 1.3* and summarize the scientific contribution of our work.

Research Question 1: How can travel time reliability for multi-modal transit journeys from a passenger perspective be quantified? (Chapter 2)

For measuring travel time reliability from a passenger perspective, the metric used needs to be sensitive to the variations in all components of a transit journey experienced by passengers including waiting, in-vehicle, and transfer times for all journey legs. Moreover, it should be comparable across different modes to enable its application in multi-modal transit networks. To achieve these requirements, we extended the existing metric Reliability Buffer Time (RBT) (Chan, 2007; Uniman et al., 2010) to include journeys with transfers. The developed metric can be calculated based on observed travel times using a combination of smart card and AVL data, which was demonstrated by applying it to the urban transit network of Amsterdam. First, we extracted the passenger journeys from raw smart card data by linking individual transactions to identify transfers. The chosen smart card system did not measure waiting time at the origin for journeys starting with bus and tram modes, unlike for those starting with metro where this was measured by having the smart card check-in at the station as opposed to the vehicle. Hence, a method was proposed to estimate the waiting time at the origin for journeys starting with these modes using the observed headway from AVL data and assuming random arrivals of passengers at the origin transit stops. This made the travel time measurements comparable across all modes. The RBT was then calculated for each transit route for each Origin-Destination pair in the data.

The proposed metric can be aggregated at multiple dimensions depending on the requirement of the analyst. We demonstrated its application at the route, transit stop, and mode level. Aggregation at a large-scale such as at a mode level can be used to analyze the impact of policies affecting one or more modes on reliability. For example, in the case of Amsterdam, the tram was found to be the least reliable for single-leg journeys. For journeys with transfers, bus to bus transfer was found to be the least reliable of all transfers. However, aggregating the reliability metric at large scale averages the variations present at the route level. Analyzing reliability at a disaggregated level such as for a specific transit stop or route can help identify at which stops or routes are the reliability improvements most needed.

Research Question 2: How can the effects of network design on distributional aspects of travel times and fares paid in the network be characterized? (Chapter 3)

Circuity of transit networks is the ratio of the network to Euclidean distance travelled by transit riders from their origin to their destination transit stop, and is a measure of directness of transit routes. The circuity of transit networks, which is a function of network design, impacts the (network) distance traveled by transit riders, and ultimately the travel times experienced by them. In the case of networks with distance-based fares, it also directly impacts the fare paid by them. Hence, to answer the research question, we analyzed how the circuity of transit routes is distributed in the network and what is its impact on travel times and fares paid. First, the circuity of all transit journeys in the network was calculated by combining the travel patterns from smart card data with the circuity of observed routes from the transit network data. Next, the neighborhood-level income data was used to assign an income profile to each transit stop. Lastly, the interactions between network design and land-use patterns on the disparity in travel times and fare paid were explored.

Our results revealed that in Amsterdam, the higher the share of high-income people living in proximity to a transit stop, the lower the circuity of journeys from the stop when controlled for the Euclidean distance covered, residential location with respect to the river, and spatial auto-correlation. The uneven distribution of circuity was found to exacerbate the disparity in distance traveled, and hence the fare paid between the income groups. Essentially, travelers from predominantly lower-income areas tended to travel longer distances and paid a higher share of fares in the network. However, the travel time per Euclidean distance favored the lower-income areas, presumably due to the circuitous routes serving these areas being compensated by higher travel speeds.

Traditionally, transit networks were designed based on efficiency and demand, while ignoring equity concerns (Soltani and Ivaki, 2011). This study highlighted the role of transit network design in determining its equity outcomes and emphasized the importance of considering equity during route and fare planning. In the case of Amsterdam, the transit network design exacerbated the disparity in fares paid. Conversely, the circuity of transit routes serving low-income areas could be improved to achieve a more (vertically) equitable distribution of fares paid. However, this may not be in line with other network planning considerations such as maximizing the ridership within a limited budget. Alternatively, changes could be made to the fare policy to compensate for the disparities in distance traveled and travel times.

Research Question 3: How can travelers' perception of overlap between alternative routes be incorporated into models of transit route choice? (Chapter 4)

To understand and incorporate travelers' perception of overlap between alternate transit routes, we compared three different types of overlap between routes and analyzed empirically how each of them is perceived by travelers. In addition to the standard definition of overlap in terms of path (including the overlap of some links, and of entire journey legs), we also proposed a definition of overlap in terms of common transfer points. Path Size Correction (PSC) Logit models (Bovy et al., 2008) were used to incorporate the three types of overlap. The path and node-based overlaps were considered separately as well as together to isolate the valuation of each of them. In addition to the path size correction terms, mode-specific travel times and transfer parameters were used to ensure mode-based overlap is also taken into account.

The overlap between transit routes was found to be valued positively when incorporated using either link-based, leg-based or transfer node-based PSC individually, with the transfer node-based PSC resulting in the best model fit for our data. When considered simultaneously, the overlap of transfer nodes was valued positively by the travelers, but the subsequent overlap of journey legs was valued negatively, implying that travelers prefer having multiple (distinct) travel options at common transfer locations. Our results indicate that there are utility benefits associated with routes having a common transfer point that can provide multiple options to the travelers, adding to the robustness of their choices. However, in the case of Amsterdam, and perhaps also in other large-scale networks, transfer stops with multiple lines are also typically larger, more aesthetically pleasing, have more amenities and a better access to real time information for travelers, which could have also contributed to their increased attractiveness to travelers.

This study contributed to advancing transit route choice modelling by improving our understanding of how overlap can be defined and is perceived by transit travelers. Majority of studies in the literature on transit route choice define overlap in terms of path only. Our results for the case of Amsterdam public transport network indicate that when estimating transit route choice models, overlap should be defined in terms of both path and transfer nodes, as skipping the latter may result in contrasting conclusions about the former.

Research Question 4: How valid are models of transit route choice estimated using smart card data? (Chapter 5)

To answer this question, we used the model(s) developed in *Chapter 4* to predict the change in travel demand in response to a major network change, i.e. the introduction of a new metro line. The new metro line was added to the network in 2018, along with the restructure of some existing transit lines. The data from before and after the network change (~3 months apart) was used for model estimation and validation, respectively. The model validation was undertaken based on two types of tests. First, the model parameters estimated based on the data before the opening of the new metro line were compared with those estimated on the data after the opening of the new line. Second, the predictive ability of the model was tested for individual predictions, as well as aggregate predictions at route, link, and mode levels. The predictions were also benchmarked against locally estimated models based on the data from after the network change.

Amongst all the attributes used in the model, only the parameters for travel time by metro, number of transfers between bus and tram, and the transfer time were found to be stable between the two contexts. There could be multiple reasons behind the differences in parameters, including change in underlying population, failure of the model to capture the causal relationship (due to presence of endogeniety and/or missing attributes) and misspecification of model. Irrespective of the reasons behind differences, our results highlight that care should be taken when deriving behavioral conclusions from such models in the absence of a thorough validation.

Despite the instability of most parameters, the predictive performance of the model at all levels was similar to the locally estimated model. Moreover, individual choices and transit mode-share predictions were found to be close to the observed ones, with a First Preference Recovery of over 70%. The errors were relatively larger for the link and route-level predictions, some of which could be attributed to the assumptions made regarding the consideration choice set given as input to the model. Visualization of prediction errors for each link in the network helped identify potential causes of errors. In our case, the aggregation of transit stops to form consideration choice set was possibly responsible for some of the prediction errors.

On comparing alternative model specifications, using generic instead of mode-specific travel attributes lead to a strong degradation in predictive performance. Conversely, a model incorporating overlap between routes, with a better model fit in the base period, did not offer a clear improvement in prediction performance, suggesting that overlap definition may be context specific, and not as transferable.

The conclusion regarding model validity would depend on the purpose for which the model is to be used. The type of validation test and corresponding test statistics should be selected accordingly, as our results show that depending on the aggregation level, the predictive performance can vary. Nonetheless, our study stressed the importance of validating transit route choice models before using them for deriving policy recommendations, especially in this data-rich age in which it can often be undertaken at a relatively low additional cost.

6.2 Implications for practice

Public transport agencies have traditionally relied on costly and unreliable data sources such as surveys for evaluating and understanding the performance of their systems (Wilson et al., 2009). Automated data collection systems have been in use for more than a decade now, and their prevalence is expected to grow further in the coming years as its potential in improving transport planning becomes increasingly recognized (ITF, 2021). In this dissertation, we further highlighted the value of such data sources by proposing several improvements to the way they can be used by transit practitioners to derive actionable insights. We demonstrated the real-world application of these improvements by applying it to the network-wide data from urban public transport network of Amsterdam. Based on our results, several implications for public transport service providers and policy makers emerge, which are described below:

Implications for transit performance assessment
The reliability measure proposed in this thesis *(Chapter 2)* can be used by transit agencies to get a better understanding of travel time reliability from a passenger perspective, specifically

for multi-modal transit networks. Based on the smart card semantic in use, the methodology can be modified and applied by transit authorities to analyze and compare reliability of different routes, lines, modes or transit stops. The calculation can be automated to enable monitoring of changes at different temporal or spatial resolutions.

Implications for transit network design
Our study has three main implications for transit network design which touch upon both efficiency and equity goals of transit networks. First, as demonstrated in *Chapter 3*, transit agencies can link automated transit data with aggregate level socio-demographic data to analyze and identify the existing inequities in travel times and fare paid in the network. The study further highlighted how network design, and specifically the circuity of routes, can be used to alter the equity outcomes. An analysis of the current disparities can enable agencies to come up with solutions to mitigate the disparities, and be able to measure the impact of any mitigations implemented.

Second, our results stress the importance of transfer hubs in making a transit network more attractive to travelers, potentially increasing its ridership. In larger networks, it is often not possible to provide direct routes for each origin-destination pair. In such cases, routes with common transfer locations but multiple distinct travel options are ideal from a travelers' perspective (*Chapter 4*). We saw that such transfer hubs not only increase the robustness of networks in case of disruptions, but also add to the attractiveness of the routes from the perspective of travelers.

Third, the mode-specific transfer penalties obtained in *Chapter 4* highlight that travelers prefer some transfers over others. Specifically, intermodal transfers within bus/tram network or within the metro network were found to be preferred over the ones between metro and bus/trams. In the case of Amsterdam, as with many other large transit networks, many bus and tram lines are intended to serve as feeders to the metro lines. However, the high transfer penalty for intermodal transfers to and from the metro indicates that more attention is needed to make these transfers more attractive to travelers.

Implications for travel demand modelling
Lastly, the results of this study could also be used for improving travel demand modelling in practice in a few ways. The passenger-oriented reliability measure proposed in *Chapter 2* can be used as an input to existing travel demand models of mode and route choice, to incorporate the impact of travel time reliability on passenger decision making and potentially improve the accuracy of such models.

Chapters 4 and 5 of this thesis were particularly dedicated to improving transit route choice models. Our insights on the definition and perception of overlap of three different types during transit route choice can be used by practitioners to improve the fit of such models in practice. Currently, overlap between alternate transit routes is considered in terms of path only. Our findings show that travelers also consider the overlap of transfer points, which should be accounted for when estimating these models. However, when using such route choice models for predictions, overlap between alternate routes could be omitted in favor of parsimony, as it did not lead to a major improvement in predictive performance in our case. The validation exercise also cautions practitioners on using models of transit route choice based on smart card data for deriving conclusions on relative valuation of travel attributes. The stability of estimates must be checked by undertaking an external validation before using them for policy analysis. However, our results show that even if the parameter estimates are unstable, such models could still have reasonable accuracy for making prediction of demand changes in response to network changes.

Implications for use of big data in transport planning
In addition to smart card data, other ‘big’ data sources in transport such as the mobile phone network and smart card app data are becoming increasingly available for use for estimating OD flows, and understanding and forecasting travel behavior. Although gaining popularity, there is still a wide variation in how transit agencies in different countries currently use these big data sources for transport planning (ITF, 2021). The privacy regulations around using such data, for example the General Data Protection Regulation (GDPR) in Europe, make the agencies more cautious in utilizing them for deriving insights for improving transport planning and management. The privacy regulations also limit the type of analysis that can be conducted with the data. Although this dissertation focused on smart card data, many of the challenges, specifically regarding privacy concerns and lack of personal information are common to other big data sources in transit. To that end, the research presented in this dissertation demonstrates the potential of such data sources, and in particular smart card data for transit performance assessment (*Chapter 2*), equity analysis (*Chapter 3*) and travel behavior analysis and predictions (*Chapter 3 & 4*), without compromising on privacy concerns.

The lack of socio-demographic information is a major limitations of using smart card data for transit planning. It is one of the many reasons why it is argued that it cannot replace traditional survey data as of now (Bonnel and Munizaga, 2018; ITF, 2021). However, collecting survey data costs time and money, and hence cannot be undertaken often whereas these new data sources can give immediate results and can be validated more often. However, to extract the most out of such data, it is recommended to have a standardized format of collecting and storing the data, which will make the methods developed for one network easily transferable to another, also enabling comparison and validation of methods used and results obtained.

6.3 Future research directions

In addition to the future research directions already discussed toward the end of each chapter (in the discussion and conclusion sections), we identify here a few key areas, based on our overall findings, in which we believe further research could be useful. These are described below:

Augmenting smart card data with survey data for travel behavior analysis
As demonstrated in this dissertation, smart card data can provide valuable insights for transit performance assessment and route choice modelling, often at a lower cost than traditional data sources. However, as discussed in *Chapters 1 and 5*, it also has several limitations that restrict its applications for transit planning. For some of these limitations, survey data could be used to augment smart card data by filling in missing information and correcting for any bias present in the data. Two such limitations are discussed here. First, from smart card data we do not know the intention behind the choices made, and assumptions need to be made in this regard. For example, because of the lack of information on origin location, we do not know which access stops were considered by the travelers when making route choice decisions. In this dissertation, we used hierarchical clustering to arrive at the access stop choice set of travelers. In this case, survey data from a limited sample can be used to provide insights into the access stop choice which can then be used to work backwards and formulate more accurate access stop choice set from observational data. Similarly, smart card data does not distinguish between the routes chosen intentionally or due to limited network knowledge/for leisure purposes/by mistake. For

such cases also, survey data could be used to inform boundary conditions for identifying irrational travel behavior such as journeys with very high circuity or beyond a maximum transfer walking distance. Another major limitation of smart card data is the lack of socio-demographic characteristics of travelers. For this, more research could focus on the ways to associate socio-demographic information with the travel data without compromising the privacy of users. A possibility could be to obtain consent from a sample of willing users and then developing methods to derive insights for the population. More research on using survey-data to fill in the gaps in information from smart card data will enable improvement in assumptions and ultimately the developed models.

Comprehensive investigation of relationship between network design and transit equity
The findings of *Chapter 3* show that network design can impact the equity outcomes in more than one way. The coverage of transit networks directly impacts the transit availability for different residents, but providing a higher coverage at the cost of higher travel times (and fares) can lead to inequity in accessibility experienced by them. Hence, our work in this regard provides only one piece of the larger puzzle of how transit network design could be optimized to reduce inequity in the network. In reality, all these different parts of the puzzle, namely different aspects of network design, land-use distribution, fare structure, budgets and equity need to be evaluated simultaneously to arrive at their optimal combination based on a transit authority's goals. More research could be undertaken to understand the trade-offs between these different factors and build a comprehensive model capturing the various interdependencies.

Examining relationship between transit network design and overlap perception
Our study was the first to define overlap in terms of transfer stops for transit route choice modelling (*Chapter 4*). For the case of Amsterdam, our results show that travelers prefer routes with common transfer stops but multiple distinct travel options. However, such transfer stops are also likely to be bigger with more amenities, better availability of information, and more aesthetically pleasing. As an extension to our work, one can investigate and isolate the impact of such factors on overlap perception. In our study, we also found that between the link-based and leg-based definition of path overlap, the latter explains transit users' route choices better. However, being relatively smaller, Amsterdam transit network has limited routes with overlapping paths. Hence, it will be interesting to perform similar research for larger networks to explore how case-specific our results are, and more broadly, how the network design and topology impact the perception of overlap by travelers. In this direction, at least two research questions could be explored. First, one could examine how the network representation on a map impacts the travelers' perception of overlap. In addition, impact of factors related to network topology (including network knowledge) as studied by Raveau et al. (2011) could also be investigated. Such an analysis can have implications for design of transit networks as well as their representation in maps. Second, one could investigate how the perception of overlap varies by the type of overlapping transit lines. Specifically, the distinction between modes where the transfer between overlapping paths is easy versus harder could be explored. For example, bus routes running on the same path can be transferred between easily, but two parallel metro lines may be grade separated making it harder to transfer.

Causal models for travel demand modeling
For accurately predicting response to a policy, the travel demand model in use must be able to capture the causal relationship between the variables that the policy impacts and the decisions of travelers. In *Chapter 5*, we used the model based on the data from before the opening of a new metro line to predict changes in demand after its opening. The results showed that the predictive performance of the model was on the higher side. This could be partly because the

changes in the network, although significant from an infrastructure perspective, were not beyond what passengers were already familiar with. For example, the new metro line was added to the exiting metro network of four other lines, with the attributes of the new line (travel times, frequencies etc.) being within the range of what travelers had experienced before. Hence, it was reasonable to assume that the perception of travelers toward the new line would be similar to their perception toward the existing ones. This may not be the case when the policy change involves something travelers have not experienced before, for example the introduction of a new transit mode. Another key point is that the estimation and validation contexts were only three months apart, during which the underlying population is unlikely to have changed significantly. However, when forecasting for a longer time horizon or for a different geographical region, omitting the impact of individual characteristics on the choices is expected to have a significant impact on the prediction performance. As a minimum requirement for building predictive models, it is necessary to include all the relevant and confounding variables for the relationship being modelled. In sum, the correct predictions in this study could be because the forecast context was quite similar to the estimation context. However, this may not be the case for other policy scenarios. In order to build models that can be used for forecasting under different policy scenarios, it is essential to build a model that can capture the causal relationships between different input and output variables. However, as of now, there is limited research on how to go about building a causal model for travel demand modelling (Brathwaite and Walker, 2018). Hence, more research in this area is needed, specifically when using observational data such as the one used for this study.

Comparing revealed preference with stated preference data for travel demand modelling
Several studies in the past have highlighted that the willingness to pay values estimated based on revealed preference (RP) data differ from the corresponding values based on stated preference (SP) data (Brownstone and Small, 2005; Li et al., 2020; M. Yap et al., 2017). When deriving behavioral insights, one would like to know the true value of these, especially when using them for cost-benefit analysis for large investments. This difference between the SP data and RP data, commonly referred to as the 'hypothetical bias' (Hensher, 2010), varies from positive to negative across studies, and sometimes between different modes within the same study (Li et al., 2020). Although RP data is typically assumed to be free of this hypothetical bias, it is limited by the variation already present in the data and cannot reveal the true value of willingness outside the range of attributes already present in the data. Moreover, in the data such as the one used in this study, the attributes considered by the traveler when making a decision are not directly known, and with the limited information available, it is possible to miss a relevant variable in the model specification. In such cases, the estimates from RP could be biased as well due to model misspecification. Hence, more research undertaking a cross-sectional comparison of SP and RP data is needed with the aim to examine and establish the bias in each of these data source.

References

Ahmed, Q.I., Lu, H., Ye, S., 2008. Urban transportation and equity: A case study of Beijing and Karachi. Transp. Res. Part A Policy Pract. https://doi.org/10.1016/j.tra.2007.06.004

AMS, 2021. De impact van de Noord/ Zuid lijn - Eindrapport [WWW Document]. URL http://nzlijn.ams-institute.org/Impactstudie-NoordZuidlijn.pdf

Anderson, M.K., Nielsen, O.A., Prato, C.G., 2017. Multimodal route choice models of public transport passengers in the Greater Copenhagen Area. EURO J. Transp. Logist. 221–245. https://doi.org/10.1007/s13676-014-0063-3

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Springer, Dordrecht.

Arriagada, J., Munizaga, M.A., Guevara, C.A., Prato, C., 2022. Unveiling route choice strategy heterogeneity from smart card data in a large-scale public transport network. Transp. Res. Part C Emerg. Technol. 134, 103467. https://doi.org/10.1016/J.TRC.2021.103467

Atherton, T., Ben-Akiva, M., 1976. Transferability and Updating Disaggregate Travel Demand Models. Transp. Res. Rec. J. Transp. Res. Board 610.

Badoe, D.A., Miller, E.J., 1995. Analysis of temporal transferability of disaggregate work trip mode choice models. Transp. Res. Rec. 1–11.

Bagherian, M., Cats, O., van Oort, N., Hickman, M., 2016. Measuring Passenger Travel Time Reliability Using Smart Card Data. TRISTAN IX Trienn. Symp. Transp. Anal. Oranjestad, Aruba.

Bandegani, M., Akbarzadeh, M., 2016. Evaluation of Horizontal Equity under a Distance-Based Transit Fare Structure. J. Public Transp. 19, 161–172. https://doi.org/10.5038/2375-0901.19.3.10

Barthélemy, M., 2011. Spatial networks. Phys. Rep. https://doi.org/10.1016/j.physrep.2010.11.002

Bekhor, S., Prato, C.G., 2009. Methodological transferability in route choice modeling. Transp. Res. Part B Methodol. 43, 422–437. https://doi.org/10.1016/J.TRB.2008.08.003

Ben-Akiva, M., Bierlaire, M., 1999. Discrete Choice Methods and their Applications to Short Term Travel Decisions, in: Hall R.W. (Eds) Handbook of Transportation Science. Springer, Boston, MA. https://doi.org/10.1007/978-1-4615-5203-1_2

Bertini, R.L., El-Geneidy, A., 2003. Generating transit performance measures with archived data. Transp. Res. Rec. https://doi.org/10.3141/1841-12
Bierlaire, M., 2020. A short introduction to PandasBiogeme. Report TRANSP-OR 200605. Lausanne, Switzerland.
Bliemer, M.C.J., Bovy, P.H.L., 2008. Impact of route choice set on route choice probabilities. Transp. Res. Rec. 2076, 10–19. https://doi.org/10.3141/2076-02
Bonnel, P., Munizaga, M.A., 2018. Transport survey methods - in the era of big data facing new and old challenges, in: Transportation Research Procedia. Elsevier, pp. 1–15. https://doi.org/10.1016/J.TRPRO.2018.10.001
Bovy, P.H.L., Bekhor, S., Prato, C.G., 2008. The factor of revisited path size: Alternative derivation. Transp. Res. Rec. 2076, 132–140. https://doi.org/10.3141/2076-15
Bovy, P.H.L., Hoogendoorn-Lanser, S., 2005. Modelling route choice behaviour in multi-modal transport networks. Transportation (Amst). 32, 341–368. https://doi.org/10.1007/s11116-004-7963-2
Box, G.E.P., 1976. Science and statistics. J. Am. Stat. Assoc. 71, 791–799. https://doi.org/10.1080/01621459.1976.10480949
Brand, J., Hoogendoorn, S., Van Oort, N., Schalkwijk, B., 2017. Modelling multimodal transit networks integration of bus networks with walking and cycling, in: 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings. https://doi.org/10.1109/MTITS.2017.8005612
Brands, T., Dixit, M., van Oort, N., 2020. Impact of a new metro line in amsterdam on ridership, travel times, reliability and societal costs and benefits. Eur. J. Transp. Infrastruct. Res. 20, 335–353. https://doi.org/10.18757/ejtir.2020.20.4.4084
Brathwaite, T., Walker, J.L., 2018. Causal inference in travel demand modeling (and the lack thereof). J. Choice Model. 26, 1–18. https://doi.org/10.1016/j.jocm.2017.12.001
Bresters, P., 2019. Toelichting Wijk- en Buurtkaart 2017, 2018 en 2019.
Brier, G.W., 1950. Verification of forecasts expressed in terms of probability, in: Monthly Weather Review. pp. 1–3.
Brown, A.E., 2018. Fair fares? How flat and variable fares affect transit equity in Los Angeles. Case Stud. Transp. Policy 6, 765–773. https://doi.org/10.1016/j.cstp.2018.09.011
Brownstone, D., Small, K.A., 2005. Valuing time and reliability: Assessing the evidence from road pricing demonstrations. Transp. Res. Part A Policy Pract. 39, 279–293. https://doi.org/10.1016/j.tra.2004.11.001
Brueckner, J.K., Thisse, J.F., Zenou, Y., 1999. Why is central Paris rich and downtown Detroit poor? An amenity-based theory. Eur. Econ. Rev. 43, 91–107. https://doi.org/10.1016/S0014-2921(98)00019-1
Bunschoten, T., Molin, E., van nes, R., 2013. Tram or bus: does the tram bonus exist?, in: European Transport Conference 2013. https://doi.org/10.1002/kin.10070
Camporeale, R., Caggiani, L., Fonzone, A., Ottomanelli, M., 2017. Quantifying the impacts of horizontal and vertical equity in transit route planning. Transp. Plan. Technol. https://doi.org/10.1080/03081060.2016.1238569
Carleton, P.R., Porter, J.D., 2018. A comparative analysis of the challenges in measuring transit equity: definitions, interpretations, and limitations. J. Transp. Geogr. https://doi.org/10.1016/j.jtrangeo.2018.08.012
Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A., 1996. A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, in: Transportation and Traffic Theory: Proc., 13th International Symposium on Transportation and Traffic Theory. Lyon, France.
Cats, O., 2014. Regularity-driven bus operation: Principles, implementation and business models. Transp. Policy 36, 223–230. https://doi.org/10.1016/J.TRANPOL.2014.09.002

Central Bureau of Statistics (CBS) Netherlands, 2020a. Percentage personen met laag inkomen - Buurten [WWW Document]. URL https://cbsinuwbuurt.nl/ (accessed 11.2.20).

Central Bureau of Statistics (CBS) Netherlands, 2020b. Mobility; per person, travel characteristics, travel motives, regions [WWW Document]. URL https://opendata.cbs.nl/#/CBS/nl/dataset/84702NED/table?ts=1608342007091 (accessed 12.1.20).

Chan, J., 2007. Rail transit OD matrix estimation and journey time reliability metrics using automated fare data. Master's thesis, MIT, Dep. Civ. Environ. Eng. June 2007 1–191.

Chorus, C.G., Kroesen, M., 2014. On the (im-)possibility of deriving transport policy implications from hybrid choice models. Transp. Policy 36, 217–222. https://doi.org/10.1016/J.TRANPOL.2014.09.001

Cole, J.P., King, C.A.M., 1968. Quantitative geography: Techniques and theories in geography. Wiley.

Currie, G., 2004. Gap Analysis of Public Transport Needs:Measuring Spatial Distribution of Public Transport Needs and Identifying Gaps in the Quality of Public Transport Provision. Transp. Res. Rec. J. Transp. Res. Board 1895, 137–146. https://doi.org/10.3141/1895-18

Currie, G., Douglas, N.J., Kearns, I., 2012. An Assessment of Alternative Bus Reliability Indicators. Australas. Transp. Res. Forum.

de Luca, S., Cantarella, G.E., 2016. Validation and Comparison of Choice Models, in: Sammer, G., Saleh, W. (Eds.), Travel Demand Management and Road User Pricing. Routledge, pp. 57–78. https://doi.org/10.4324/9781315549743-12

Delbosc, A., Currie, G., 2011. Using Lorenz curves to assess public transport equity. J. Transp. Geogr. 19, 1252–1259.

Diab, E.I., Badami, M.G., El-Geneidy, A.M., 2015. Bus Transit Service Reliability and Improvement Strategies: Integrating the Perspectives of Passengers and Transit Agencies in North America. Transp. Rev. 35, 292–328. https://doi.org/10.1080/01441647.2015.1005034

Dixit, M., Brands, T., Cats, O., van Oort, N., Hoogendoorn, S., 2019a. Impact analysis of a new metro line in Amsterdam using automated data sources, in: Transit Data 2019. Paris.

Dixit, M., Brands, T., van Oort, N., Cats, O., Hoogendoorn, S., 2019b. Passenger Travel Time Reliability for Multimodal Public Transport Journeys. Transp. Res. Rec. J. Transp. Res. Board 2673, 149–160. https://doi.org/10.1177/0361198118825459

Dixit, M., Cats, O., Brands, T., van Oort, N., Hoogendoorn, S., 2021. Perception of overlap in multi-modal urban transit route choice. Transp. A Transp. Sci. https://doi.org/10.1080/23249935.2021.2005180

Ehrlich, J.E., 2010. Applications of Automatic Vehicle Location Systems Towards Improving Service Reliability and Operations Planning in London 165.

El-Geneidy, A., Grimsrud, M., Wasfi, R., Tétreault, P., Surprenant-Legault, J., 2014. New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas. Transportation (Amst). 41. https://doi.org/10.1007/s11116-013-9508-z

El-Geneidy, A., Levinson, D., Diab, E., Boisjoly, G., Verbich, D., Loong, C., 2016. The cost of equity: Assessing transit accessibility and social disparity using total travel cost. Transp. Res. Part A Policy Pract. 302–316. https://doi.org/10.1016/j.tra.2016.07.003

Eliasson, J., Mattsson, L.G., 2006. Equity effects of congestion pricing. Quantitative methodology and a case study for Stockholm. Transp. Res. Part A Policy Pract. https://doi.org/10.1016/j.tra.2005.11.002

Erhardt, G.D., Lock, O., Arcaute, E., Batty, M., 2017. A big data mashing tool for measuring transit system performance, in: Seeing Cities Through Big Data. Springer, pp. 257–278. https://doi.org/10.1007/978-3-319-40902-3_15

European Court of Auditors, 2014. Effectiveness of EU-supported public urban transport projects. Luxembourg. https://doi.org/10.2865/30848
Farber, S., Bartholomew, K., Li, X., Páez, A., Nurul Habib, K.M., 2014. Assessing social equity in distance based transit fares using a model of travel behavior. Transp. Res. Part A Policy Pract. 67, 291–303. https://doi.org/10.1016/j.tra.2014.07.013
Feitelson, E., 2002. Introducing environmental equity dimensions into the sustainable transport discourse: Issues and pitfalls. Transp. Res. Part D Transp. Environ. https://doi.org/10.1016/S1361-9209(01)00013-X
Forsey, D., Nurul Habib, K., Miller, E.J., Shalaby, A., 2014. Temporal transferability of work trip mode choice models in an expanding suburban area. Transp. A Transp. Sci. 10, 469–482. https://doi.org/10.1080/23249935.2013.788100
Foth, N., Manaugh, K., El-Geneidy, A.M., 2013. Towards equitable transit: Examining transit accessibility and social need in Toronto, Canada, 1996-2006. J. Transp. Geogr. https://doi.org/10.1016/j.jtrangeo.2012.12.008
Fox, J., Daly, A., Hess, S., Miller, E., 2014. Temporal transferability of models of mode-destination choice for the Greater Toronto and Hamilton Area. J. Transp. Land Use 7. https://doi.org/10.5198/jtlu.v7i2.701
Fox, J., Hess, S., 2010. Review of Evidence for Temporal Transferability of Mode-Destination Models: Transp. Res. Rec. 74–83. https://doi.org/10.3141/2175-09
Fox, J.B., 2015. Temporal transferability of mode-destination models. University of Leeds.
Frejinger, E., Bierlaire, M., 2007. Capturing correlation with subnetworks in route choice models. Transp. Res. Part B Methodol. 41, 363–378. https://doi.org/10.1016/j.trb.2006.06.003
Fung, S.W.C., Tong, C.O., Wong, S.C., 2005. Validation of a conventional metro network model using real data. J. Intell. Transp. Syst. 9, 69–79. https://doi.org/10.1080/15472450590934624
Furth, P.G., Muller, T.H.J., 2006. Service Reliability and Hidden Waiting Time: Insights from Automatic Vehicle Location Data. Transp. Res. Board 1955.
Garcia-Martinez, A., Cascajo, R., Jara-Diaz, S.R., Chowdhury, S., Monzon, A., 2018. Transfer penalties in multimodal public transport networks. Transp. Res. Part A Policy Pract. 114, 52–66. https://doi.org/10.1016/J.TRA.2018.01.016
Gini, C., 1912. Variabilità e mutabilità.
Gittens, A., Shalaby, A., 2015. Evaluation of Bus Reliability Measures and Development of a New Composite Indicator. Transp. Res. Rec. J. Transp. Res. Board 2533, 91–99. https://doi.org/10.3141/2533-10
Gordon, J., Koutsopoulos, H., Wilson, N., Attanucci, J., 2013. Automated Inference of Linked Transit Journeys in London Using Fare-Transaction and Vehicle Location Data. Transp. Res. Rec. 2343, 17–24. https://doi.org/10.3141/2343-03
Gunn, H.F., Ben-Akiva, M.E., Bradley, M.A., 1985. Tests of the Scaling Approach to Transferring Disaggregate Travel Demand Models. Transp. Res. Rec. 1037, 21–30.
Guo, Z., Wilson, N.H.M., 2011. Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. Transp. Res. Part A Policy Pract. 45, 91–104. https://doi.org/10.1016/j.tra.2010.11.002
Guzman, L.A., Oviedo, D., Rivera, C., 2017. Assessing equity in transport accessibility to work and study: The Bogotá region. J. Transp. Geogr. https://doi.org/10.1016/j.jtrangeo.2016.12.016
GVB, 2016. Beter bereikbaar in een drukkere stad Concept vervoerplan bij start Noord/Zuidlijn.
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. Transp. Res. Part A Policy Pract. 89, 106–123. https://doi.org/10.1016/J.TRA.2016.05.010

Hensher, D.A., 2010. Hypothetical bias, choice experiments and willingness to pay. Transp. Res. Part B Methodol. 44, 735–752. https://doi.org/10.1016/J.TRB.2009.12.012

Hensher, D.A., Truong, T.P., Mulley, C., Ellison, R., 2012. Assessing the wider economy impacts of transport infrastructure investment with an illustrative application to the North-West Rail Link project in Sydney, Australia. J. Transp. Geogr. 24, 292–305. https://doi.org/10.1016/J.JTRANGEO.2012.03.009

Hoogendoorn-Lanser, S., Bovy, P., 2007. Modeling overlap in multimodal route choice by including trip part-specific path size factors. Transp. Res. Rec. 2003, 74–83. https://doi.org/10.3141/2003-10

Hoogendoorn-Lanser, S., van Nes, R., Bovy, P., 2005. Path Size Modeling in Multimodal Route Choice Analysis. Transp. Res. Rec. 1921, 27–34. https://doi.org/10.3141/1921-04

Hörcher, D., Graham, D.J., Anderson, R.J., 2017. Crowding cost estimation with large scale smart card and vehicle location data. Transp. Res. Part B Methodol. 95, 105–125. https://doi.org/10.1016/j.trb.2016.10.015

Huang, J., Levinson, D.M., 2015. Circuity in urban transit networks. J. Transp. Geogr. https://doi.org/10.1016/j.jtrangeo.2015.09.004

ITF, 2021. Big Data for Travel Demand Modelling: Summary and Conclusions. Roundtable Reports, No. 186.

Jánošíkova, L., Slavík, J., Koháni, M., 2014. Estimation of a route choice model for urban public transport using smart card data. Transp. Plan. Technol. 37, 638–648. https://doi.org/10.1080/03081060.2014.935570

Jenelius, E., 2018. Experienced Public Transport Service Reliability : Integrating Travel Time and Travel Conditions. 7th Int. Symp. Transp. Netw. Reliab. (INSTR), 2018.

Jenelius, E., Cats, O., 2015. The value of new public transport links for network robustness and redundancy. Transp. A Transp. Sci. 11, 819–835. https://doi.org/10.1080/23249935.2015.1087232

Kaplan, S., Popoks, D., Prato, C.G., Ceder, A. (Avi), 2014. Using connectivity for measuring equity in transit provision. J. Transp. Geogr. 37, 82–92. https://doi.org/10.1016/J.JTRANGEO.2014.04.016

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 (Amst). https://doi.org/10.1007/s11116-019-10008-8

Kim, K.M., Hong, S.P., Ko, S.J., Kim, D., 2015. Does crowding affect the path choice of metro passengers? Transp. Res. Part A Policy Pract. 77, 292–304. https://doi.org/10.1016/j.tra.2015.04.023

Kittelson & Associates, 2013. Transit Capacity and Qulity of Service Manual, 3rd Edition, TCRP Report 165.

Kittleson & Associates, I., Urbitran, I., LKC Consulting Services, I., MORPACE Consulting Services, I., Queensland University of Technology, Nakanishi, Y., 2003. TCRP PROGRAM REPORT 88: A Guidebook for Developing a Transit Performance-Measurement System, Tranrportation Research Board, Washington, D.C. https://doi.org/10.17226/23367

Koppelman, F.S., Wilmot, C.G., 1986. The effect of omission of variables on choice model transferability. Transp. Res. Part B Methodol. 20, 205–213. https://doi.org/10.1016/0191-2615(86)90017-2

Koppelman, F.S., Wilmot, C.G., 1982. Transferability Analysis of Disaggregate Choice Models. Transp. Res. Rec. J. Transp. Res. Board 18–24.

Kusakabe, T., Iryo, T., Asakura, Y., 2010. Estimation method for railway passengers' train choice behavior with smart card transaction data. Transportation (Amst). 37. https://doi.org/10.1007/s11116-010-9290-0

Lai, X., Bierlaire, M., 2015. Specification of the cross-nested logit model with sampling of alternatives for route choice models. Transp. Res. Part B Methodol. 80, 220–234. https://doi.org/10.1016/J.TRB.2015.07.005

Lee, A., van Oort, N., van Nes, R., 2014. Service Reliability in a Network Context. Transp. Res. Rec. J. Transp. Res. Board 2417, 18–26. https://doi.org/10.3141/2417-03

Lee, Y.J., Choi, J.Y., Yu, J.W., Choi, K., 2015. Geographical applications of performance measures for transit network directness. J. Public Transp. 18, 89–110. https://doi.org/10.5038/2375-0901.18.2.7

LeSage, J., 2008. Introduction to spatial econometrics, Revue d'economie industrielle. https://doi.org/10.1111/j.1467-985x.2010.00681_13.x

Levinson, D., El-Geneidy, A., 2009. The minimum circuity frontier and the journey to work. Reg. Sci. Urban Econ. 39, 732–738. https://doi.org/10.1016/j.regsciurbeco.2009.07.003

Li, Z., Hensher, D.A., Ho, C., 2020. An empirical investigation of values of travel time savings from stated preference data and revealed preference data. Transp. Lett. 12, 166–171. https://doi.org/10.1080/19427867.2018.1546806

Litman, T., 2002. Evaluating Transportation Equity. World Transp. Policy Pract. 8, 50–65.

Liu, Y., Bunker, J., Ferreira, L., 2010. Transit userś route-choice modelling in transit assignment: A review. Transp. Rev. https://doi.org/10.1080/01441641003744261

Luo, D., Bonnetain, L., Cats, O., van Lint, H., 2018. Constructing Spatiotemporal Load Profiles of Transit Vehicles with Multiple Data Sources. Transp. Res. Rec. J. Transp. Res. Board In press.

Mai, T., 2016. A method of integrating correlation structures for a generalized recursive route choice model. Transp. Res. Part B Methodol. 93, 146–161. https://doi.org/10.1016/J.TRB.2016.07.016

Martens, K., Bastiaanssen, J., Lucas, K., 2019. Measuring transport equity: Key components, framings and metrics, in: Measuring Transport Equity. https://doi.org/10.1016/B978-0-12-814818-1.00002-0

McFadden, D., 1974. Conditional Logit Analysis of Qualitative Choice Behavior, in: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.

Munizaga, M.A., Palma, C., 2012. Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. Transp. Res. Part C Emerg. Technol. 24, 9–18. https://doi.org/10.1016/j.trc.2012.01.007

Neutens, T., Schwanen, T., Witlox, F., de Maeyer, P., 2010a. Equity of urban service delivery: A comparison of different accessibility measures. Environ. Plan. A. https://doi.org/10.1068/a4230

Neutens, T., Versichele, M., Schwanen, T., 2010b. Arranging place and time: A GIS toolkit to assess person-based accessibility of urban opportunities. Appl. Geogr. https://doi.org/10.1016/j.apgeog.2010.05.006

Nielsen, O.A., Eltved, M., Anderson, M.K., Prato, C.G., 2021. Relevance of detailed transfer attributes in large-scale multimodal route choice models for metropolitan public transport passengers. Transp. Res. Part A Policy Pract. 147, 76–92. https://doi.org/10.1016/J.TRA.2021.02.010

Parady, G., Ory, D., Walker, J., 2021. The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: A review of validation practices in the transportation academic literature. J. Choice Model. 38. https://doi.org/10.1016/j.jocm.2020.100257

Pelletier, M.-P., Trépanier, M., Morency, C., 2011. Smart card data use in public transit: A literature review. Transp. Res. Part C Emerg. Technol. 19, 557–568. https://doi.org/10.1016/J.TRC.2010.12.003

Pereira, R.H.M., Schwanen, T., Banister, D., 2017. Distributive justice and equity in

transportation. Transp. Rev. https://doi.org/10.1080/01441647.2016.1257660
Poon, M.H., Tong, C.O., Wong, S.C., 2003. Validation of a schedule-based capacity restraint transit assignment model for a large-scale network. J. Adv. Transp. 38, 5–26. https://doi.org/10.1002/atr.5670380103
Prato, C.G., 2009. Route choice modeling: Past, present and future research directions. J. Choice Model. 2, 65–100. https://doi.org/10.1016/S1755-5345(13)70005-8
Prato, C.G., Bekhor, S., 2007. Modeling route choice behavior: How relevant is the composition of choice set? Transp. Res. Rec. 2003, 64–73. https://doi.org/10.3141/2003-09
Pritchard, J.P., Tomasiello, D., Giannotti, M., Geurs, K., 2019. An International Comparison of Equity in Accessibility to Jobs: London, São Paulo, and the Randstad. Transp. Find. https://doi.org/10.32866/7412
Pucher, J., 1981. Equity in transit finance: distribution of transit subsidy benefits and costs among income classes. J. Am. Plan. Assoc. https://doi.org/10.1080/01944368108976521
Ramming, M.S., 2002. Network knowledge and route choice. Ph.D. Thesis. Massachusetts Institute of Technology, Cambridge, Massachusetts.
Raveau, S., Carlos Muñoz, J., De Grange, L., 2011. A topological route choice model for metro. Transp. Res. Part A 45, 138–147. https://doi.org/10.1016/j.tra.2010.12.004
Raveau, S., Guo, Z., Muñoz, J.C., Wilson, N.H.M., 2014. A behavioural comparison of route choice on metro networks: Time, transfers, crowding, topology and socio-demographics. Transp. Res. Part A Policy Pract. 66, 185–195. https://doi.org/10.1016/j.tra.2014.05.010
Rossi, T.F., Bhat, C.R., 2014. Guide for Travel Model Transfer (No. FHWA-HEP-15-006).
Rubensson, I., Susilo, Y., Cats, O., 2020. Is flat fare fair? Equity impact of fare scheme change. Transp. Policy 91, 48–58. https://doi.org/10.1016/j.tranpol.2020.03.013
Sanko, N., Morikawa, T., 2010. Temporal transferability of updated alternative-specific constants in disaggregate mode choice models. Transportation (Amst). 37, 203–219. https://doi.org/10.1007/s11116-009-9252-6
Soltani, A., Ivaki, Y.E., 2011. Inequity in the Provision of Public Bus Service for Socially Disadvantaged Groups. J. Sustain. Dev. 4. https://doi.org/10.5539/jsd.v4n5p229
Swait, J., Louviere, J., 1993. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. J. Mark. Res. 30, 305–314.
Swierstra, A.B., van Nes, R., Molin, E.J.E., 2017. Modelling travel time reliability in public transport route choice behaviour. Eur. J. Transp. Infrastruct. Res. 17, 263–278. https://doi.org/https://doi.org/10.18757/ejtir.2017.17.2.3194
Tan, R., Adnan, M., Lee, D.H., Ben-Akiva, M.E., 2015. New path size formulation in path size logit for route choice modeling in public transport networks. Transp. Res. Rec. 2538, 11–18. https://doi.org/10.3141/2538-02
Tirachini, A., Sun, L., Erath, A., Chakirov, A., 2016. Valuation of sitting and standing in metro trains using revealed preferences. Transp. Policy 47. https://doi.org/10.1016/j.tranpol.2015.12.004
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 Behav. Soc. 13, 105–117. https://doi.org/10.1016/j.tbs.2018.07.001
Train, K., 1978. A validation test of a disaggregate mode choice model. Transp. Res. 12, 167–174. https://doi.org/10.1016/0041-1647(78)90120-X
Trépanier, M., Morency, C., Agard, B., 2009. Calculation of transit performance measures using smartcard data. J. Public Transp. 12, 76–96. https://doi.org/10.1016/j.tranpol.2007.01.001
Trépanier, M., Tranchant, N., Chapleau, R., 2007. Individual trip destination estimation in a transit smart card automated fare collection system. J. Intell. Transp. Syst. Technol. Planning, Oper. https://doi.org/10.1080/15472450601122256

UITP, 2020. Public Transport Moving Europe Forward.
Uniman, D., Attanucci, J., Mishalani, R., Wilson, N., 2010. Service Reliability Measurement Using Automated Fare Card Data. Transp. Res. Rec. J. Transp. Res. Board 2143, 92–99. https://doi.org/10.3141/2143-12
United Nations, 2018. The World's Cities in 2018 - Data Booklet.
van Lierop, D., El-Geneidy, A., 2016. Enjoying loyalty: The relationship between service quality, customer satisfaction, and behavioral intentions in public transit. Res. Transp. Econ. 59, 50–59. https://doi.org/10.1016/j.retrec.2016.04.001
van Oort, N., 2016. Incorporating enhanced service reliability of public transport in cost-benefit analyses. Public Transp. 8, 143–160. https://doi.org/10.1007/s12469-016-0121-3
Van Oort, N., 2011. Service Reliability and Urban Public Transport Design. TRAIL Thesis Series T2011/2, the Netherlands TRAIL Research School.
van Oort, N., Brands, T., de Romph, E., 2015a. Short-Term Prediction of Ridership on Public Transport with Smart Card Data. Transp. Res. Rec. 2535, 105–111. https://doi.org/10.3141/2535-12
van Oort, N., Sparing, D., Brands, T., Goverde, R.M.P., 2015b. Data driven improvements in public transport: the Dutch example. Public Transp. 7, 369–389. https://doi.org/10.1007/s12469-015-0114-7
van Oort, N., van Nes, R., 2009. Regularity analysis for optimizing urban transit network design. Public Transp. 1, 155–168. https://doi.org/10.1007/s12469-009-0012-y
van Oort, N., Yap, M., 2021. Innovations in the appraisal of public transport projects, in: Mouter, N. (Ed.), Advances in Transport Policy and Planning. Academic Press, pp. 127–164.
Walker, J., 2008. Purpose-driven public transport: creating a clear conversation about public transport goals. J. Transp. Geogr. 16, 436–442. https://doi.org/10.1016/j.jtrangeo.2008.06.005
Wei, R., Liu, X., Mu, Y., Wang, L., Golub, A., Farber, S., 2017. Evaluating public transit services for operational efficiency and access equity. J. Transp. Geogr. 65. https://doi.org/10.1016/j.jtrangeo.2017.10.010
Wilson, N., Zhao, J., Rahbee, A., 2009. The Potential Impact of Automated Data Collection Systems on Urban Public Transport Planning, in: Nuzzolo, A., Wilson, N. (Eds.), Schedule-Based Modeling of Transportation Networks - Theory and Methods. Springer, pp. 75–99.
Wood, D., 2015. A Framework for Measuring Passenger-Experienced Transit Reliability Using Automated Data. Master Thesis, Massachusetts Inst. Technol.
Yap, M., Cats, O., van Arem, B., 2020. Crowding valuation in urban tram and bus transportation based on smart card data. Transp. A Transp. Sci. 16, 23–42. https://doi.org/10.1080/23249935.2018.1537319
Yap, M., Cats, O., Yu, S., Arem, B. van, 2017. Crowding valuation in urban tram and bus transportation based on smart card data, in: Thredbo 15: Competition and Owndership in Land Passenger Transport. pp. 4305–4325.
Yap, M.D., Cats, O., Van Oort, N., Hoogendoorn, S.P., 2017. A robust transfer inference algorithm for public transport journeys during disruptions. Transp. Res. Procedia 27, 1042–1049. https://doi.org/10.1016/j.trpro.2017.12.099
Zhao, F., Ubaka, I., 2004. Transit Network Optimization - Minimizing Transfers and Optimizing Route Directness. J. Public Transp. https://doi.org/10.5038/2375-0901.7.1.4
Zhao, J., Rahbee, A., Wilson, N.H.M., 2007. Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. Comput. Civ. Infrastruct. Eng. https://doi.org/10.1111/j.1467-8667.2007.00494.x
Zhao, J., Zhang, F., Tu, L., Xu, C., Shen, D., Tian, C., Li, X.Y., Li, Z., 2017. Estimation of

Passenger Route Choice Pattern Using Smart Card Data for Complex Metro Systems. IEEE Trans. Intell. Transp. Syst. 18. https://doi.org/10.1109/TITS.2016.2587864

Zhao, Z., Koutsopoulos, H.N., Zhao, J., 2020. Uncovering spatiotemporal structures from transit smart card data for individual mobility modeling, in: Demand for Emerging Transportation Systems: Modeling Adoption, Satisfaction, and Mobility Patterns. Elsevier, pp. 123–149. https://doi.org/10.1016/B978-0-12-815018-4.00007-3

Summary

By 2050, 68% of the world population will live in urban areas (United Nations, 2018). With growing urbanization, cities today struggle to provide efficient transportation to their citizens amidst growing demand, while minimizing congestion, accidents, and pollution (European Court of Auditors, 2014). Public transport offers a potential solution for many of these problems. However, a common challenge for transit authorities is making public transport more attractive for its travelers, within budget limitations. To improve public transport services, it is important to understand what is the current state of the service, and where are the improvements most needed. Service quality has been known to impact users' perception of public transport, and ultimately their choices.

Public transport smart card has been implemented for more than a decade now, providing access to a massive amount of passively collected data on network usage, as opposed to a limited sample from the traditional data collection methods such as surveys. When combined with Automatic Vehicle Location (AVL) data, it can also enable an accurate measurement of service quality from a passenger perspective at a relatively lower cost. Together, they can be used to understand how travelers undertake public transport route choice decisions, giving insights into the relative valuation of different service quality characteristics. However, these data sources have not yet been explored to their full potential, and the methods in use are often borrowed from past research using more traditional data sources. Hence, the overarching aim of this PhD dissertation is to improve performance assessment and route choice modeling for urban multi-modal transit networks using smart card data. We use the case study of the Amsterdam public transport network to undertake this research and make several scientific and practical contributions to fulfill our overall aim, which are summarized below.

In the first part of this dissertation, we focus on transit performance assessment and address two specific gaps in the scientific literature in this domain. First, we develop a methodology for measuring travel time reliability for multimodal transit journeys from a passenger perspective using smart card data. We do so by extending the existing Reliability Buffer Time (RBT) to

incorporate journeys with transfers and include all components of the journey experienced by a passenger including the waiting time at the origin transit stop, in-vehicle times, and transfer waiting and walking times. In the case of Amsterdam, the travel time by smart card data is measured differently for different systems. Our method makes it comparable, and finally, we have a route level RBT that can be compared between routes, transit stops, and modes. The metric can be used by transit authorities to undertake continuous and even real-time monitoring of changes to reliability and a disaggregate level, and identifying routes/times that need improvement. It can also be used as an input to travel demand models such as mode, route or departure time choice.

As a second contribution toward transit performance assessment, we leverage smart card data for understanding the impact of transit network design on the equity outcomes of travel times and fares paid in a network. Circuity of a transit journey is defined as the ratio of the network to Euclidean distance traveled. Everything else being equal, a higher circuity implies longer travel times for the same Euclidean distance. Moreover, for transit networks such as Amsterdam where the fare is calculated based on the (network) distance traveled, higher circuity also means higher fare for the same Euclidean distance. This makes circuity relevant from an equity perspective. This study explores the role of transit circuity on the disparity in distance traveled by travelers' income profiles and its implications on travel times and costs for networks with distance-based fares. The analysis is based on travel patterns from smart card data for bus, tram, and metro modes, combined with neighborhood-level income data. Results reveal that in Amsterdam, the higher the share of high-income people living in proximity to a transit stop, the lower the circuity of journeys from the stop when controlled for the Euclidean distance covered and spatial auto-correlation. The uneven distribution of circuity exacerbates the disparity in distance traveled, and hence fare paid between the income groups. However, the travel time per Euclidean distance favors the low-income group, possibly due to the circuitous routes serving these areas being compensated by higher travel speeds.

In the second part of this dissertation, we attempt to improve models of transit route choice estimated using smart card data. The contributions in this area are described below.

First, we comprehensively investigate how different types of overlap between alternate transit routes are perceived by travelers. We propose a new definition of overlap in terms of common transfer nodes, which is particularly relevant for large-scale urban transit networks. We compare this new definition with the traditionally used definition of overlap in terms of the path (both in the form of links and entire journey legs). Path size correction (PSC) logit models are used to incorporate each of these types of overlap. The results indicate that the overlap between transit routes is valued positively when incorporated using either link-based, leg-based or transfer node-based PSC individually, with the transfer node-based PSC resulting in the best model fit. When considered simultaneously, the overlap of transfer nodes is valued positively by the travelers, but the subsequent overlap of journey legs is valued negatively, implying that travelers prefer having multiple (distinct) travel options at common transfer locations. This study contributes to advancing transit route choice modeling by improving our understanding of how overlap can be defined and is perceived by transit travelers.

Lastly, we undertake an external validation of the transit route choice models developed in the previous step. For this, we use smart card data from before and after the opening of a new metro line in Amsterdam, the Netherlands for model estimation and validation, respectively. The estimated parameters are checked for stability between the two time periods, and predictive ability are evaluated at different aggregation levels. Although most model parameters were found to be unstable between the two contexts, the predictive performance at aggregation levels was similar to the locally estimated model. Moreover, individual choices and transit mode-share predictions are found to be close to the observed ones. The errors were relatively larger for the link and route-level predictions, some of which could be attributed to the assumptions made regarding the consideration choice set given as input to the model. On comparing alternative model specifications, using generic instead of mode-specific travel attributes lead to a strong degradation in predictive performance. Conversely, a model incorporating overlap between routes, with a better model fit in the base period, did not offer a clear improvement in prediction performance. This study adds to the scarce literature on the validation of travel demand models and, is the first to undertake an external validation of a transit route choice model. Our results highlight the need to validate transit route choice models before using them for deriving policy recommendations, especially in this data-rich age in which it can often be undertaken at a relatively low additional cost.

Overall, this dissertation leverages smart card data to make advances in transit performance assessment and route choice modeling, specifically in the context of urban multi-modal transit networks. The prevalence of Automated data collection systems for transport planning is expected to grow further in the coming years as its potential in improving transport planning becomes increasingly recognized. In this dissertation, we further highlight the value of such data sources by proposing several improvements for attaining both scientific and practical implications.

Samenvatting

Tegen 2050 zal 68% van de wereldbevolking in stedelijke gebieden wonen (United Nations, 2018). Gezien de groeiende verstedelijking worstelen steden vandaag de dag met het bieden van efficiënt vervoer aan hun burgers, terwijl ze congestie, ongevallen en vervuiling tot een minimum proberen te beperken (European Court of Auditors, 2014). Openbaar vervoer biedt een potentiële oplossing voor veel van deze problemen. Een veelvoorkomende uitdaging voor vervoersautoriteiten is echter om het openbaar vervoer aantrekkelijker te maken voor reizigers binnen het vastgestelde budget. Om het openbaar vervoer te verbeteren, is het belangrijk om te begrijpen wat de huidige staat van de dienstverlening is, en waar de verbeteringen het meest nodig zijn. Het is bekend dat de kwaliteit van de dienstverlening een invloed heeft op de perceptie van ov-gebruikers, en uiteindelijk ook op hun keuzes.

Data van smartcards in het openbaar vervoer wordt nu al meer dan tien jaar gebruikt en biedt toegang tot een enorme hoeveelheid passief verzamelde gegevens over het netwerkgebruik, in tegenstelling tot steekproeven die uitgevoerd worden middels traditionele dataverzamelingsmethoden zoals enquêtes. In combinatie met data over de automatische plaatsbepaling van voertuigen (Automatic Vehicle Location, of AVL) kan de kwaliteit van de dienstverlening vanuit het oogpunt van de passagier nauwkeurig worden gemeten tegen een relatief lagere kostprijs. Samen kunnen ze worden gebruikt om te begrijpen hoe reizigers beslissingen nemen over de keuze van hun ov-traject, en om inzicht te krijgen in de relatieve waardering van verschillende servicekwaliteitskenmerken. Het potentieel van deze databronnen wordt echter nog niet ten volle benut, en de gebruikte methoden zijn vaak ontleend aan onderzoek uit het verleden waarin traditionele databronnen werden gebruikt. Het overkoepelende doel van dit proefschrift is dan ook het verbeteren van prestatiebeoordeling en routekeuze-modellering voor stedelijke multimodale ov-netwerken met behulp van smartcard-data. We gebruiken de case study van het Amsterdamse openbaar vervoernetwerk om dit onderzoek uit te voeren en leveren verschillende wetenschappelijke en praktische bijdragen om ons algemene doel te bereiken. Deze bijdragen worden hieronder samengevat.

In het eerste deel van dit proefschrift richten we ons op de beoordeling van de prestaties van het openbaar vervoer en pakken we twee specifieke openstaande vraagstukken in de wetenschappelijke literatuur op dit gebied aan. Ten eerste ontwikkelen we een methodologie voor het meten van reistijdbetrouwbaarheid voor multimodale ov-reizen vanuit het passagiersperspectief met behulp van smartcard-data. We doen dit door de bestaande Reliability Buffer Time (RBT) uit te breiden naar reizen met overstap en houden daarnaast rekening met alle componenten van de reis die een passagier ervaart, inclusief de wachttijd bij de halte van vertrek, de tijd in het voertuig, en de wacht- en wandeltijden bij de overstap. In het geval van Amsterdam wordt de reistijd aan de hand van smartcard-data voor verschillende systemen verschillend gemeten. Onze methode maakt het mogelijk om te vergelijken, en daarnaast hebben we een RBT op trajectniveau die kan worden vergeleken tussen routes, doorvoerhaltes, en modaliteiten. De RBT kan worden gebruikt door ov-autoriteiten om continu en zelfs real-time monitoring van betrouwbaarheidsverandering op een uitgesplitst niveau uit te voeren. Daarnaast stelt het men in staat om routes en tijden te identificeren die verbetering behoeven. Het kan tevens worden gebruikt als input voor vraagmodellen, zoals modaliteit-, route- of vertrektijdkeuze.

Als tweede bijdrage aan de beoordeling van ov-prestaties gebruiken we smartcard-data voor het begrijpen van de impact van het ontwerp van het ov-netwerk op de rechtvaardigheid van reistijden en tarieven in het netwerk. Circuity van een ov-traject wordt gedefinieerd als de verhouding van het netwerk tot de Euclidische afgelegde afstand. Als al het andere gelijk blijft, impliceert een hogere circuity langere reistijden voor dezelfde Euclidische afstand. Bovendien betekent een hogere circuity voor ov-netwerken zoals Amsterdam, waar de ritprijs wordt berekend op basis van de afgelegde (netwerk)afstand, tevens een hogere ritprijs voor dezelfde Euclidische afstand. Dit maakt circuity relevant vanuit het perspectief van rechtvaardigheid. Deze studie onderzoekt de invloed van circuity op het verschil in afgelegde afstand door de inkomensprofielen van reizigers en de implicaties daarvan op reistijden en kosten voor netwerken met afstand-gebaseerde tarieven. De analyse is gebaseerd op reispatronen van smartcard-data voor bus, tram en metro, gecombineerd met inkomensgegevens op buurtniveau. De resultaten in Amsterdam laten zien dat hoe hoger het aandeel van mensen met een hoog inkomen die in de nabijheid van een halte wonen, hoe lager de circuity van reizen vanaf de halte wanneer gecontroleerd wordt voor de Euclidische afgelegde afstand en ruimtelijke autocorrelatie. De ongelijke verdeling van de circuity vergroot het verschil in afgelegde afstand, en daarmee de betaalde ritprijs tussen de inkomensgroepen. De reistijd per Euclidische afstand is echter in het voordeel van de lage-inkomensgroep, mogelijk doordat de omslachtige routes die deze gebieden bedienen worden gecompenseerd door hogere reissnelheden.

In het tweede deel van deze dissertatie proberen we modellen van ov-routekeuze, geschat met behulp van smartcard-data, te verbeteren. De bijbehorende bijdragen worden hieronder beschreven.

Ten eerste onderzoeken we uitgebreid hoe verschillende soorten overlap tussen alternatieve ov-routes worden waargenomen door reizigers. We stellen een nieuwe definitie van overlap voor in termen van gemeenschappelijke transfernodes, die relevant is voor grootschalige stedelijke ov-netwerken. We vergelijken deze nieuwe definitie met de traditioneel gebruikte definitie van overlap in termen van het pad (zowel in de vorm van links als volledige trajecten). Path size correction (PSC) logit-modellen worden gebruikt om elk van deze vormen van overlap op te nemen. Uit de resultaten blijkt dat de overlap tussen ov-routes positief wordt gewaardeerd wanneer gebruik wordt gemaakt van op links, trajecten of transfernodes gebaseerde PSC, waarbij de op transfernodes-gebaseerde PSC resulteert in de beste modelmatch. Wanneer gelijktijdig beschouwd, wordt de overlap van transfernodes positief gewaardeerd door de reizigers, maar de daaropvolgende overlap van trajecten wordt negatief gewaardeerd, wat impliceert dat reizigers de voorkeur geven aan meerdere (verschillende) reismogelijkheden op gemeenschappelijke overstaplocaties. Deze studie draagt bij aan de vooruitgang van routekeuzemodellering in het openbaar vervoer door het verbeteren van de definitie van overlap en hoe deze wordt ervaren door ov-reizigers.

Tenslotte voeren we een externe validatie uit van de routekeuzemodellen ontwikkeld in de vorige stap. Hiervoor gebruiken we smartcard-data van voor en na de opening van een nieuwe metrolijn in Amsterdam voor respectievelijk de schatting en validatie van het model. De geschatte parameters worden gecontroleerd op stabiliteit tussen de twee tijdsperiodes, en het voorspellend vermogen wordt geëvalueerd op verschillende aggregatieniveaus. Hoewel de meeste modelparameters instabiel bleken tussen de twee contexten, was de voorspellende prestatie op aggregatieniveaus vergelijkbaar met het lokaal geschatte model. Bovendien bleken de voorspellingen van de individuele keuzes en het aandeel van de vervoerwijzen dicht bij de waargenomen keuzes te liggen. De fouten waren relatief groter voor de voorspellingen op link- en routeniveau, waarvan een deel zou kunnen worden toegeschreven aan de aannames die zijn gemaakt met betrekking tot de keuzeset die als input voor het model is gebruikt. Bij het vergelijken van alternatieve modelspecificaties, leidt het gebruik van generieke in plaats van modaliteitsspecifieke reisattributen tot een sterke verslechtering van de voorspellingsprestaties. Omgekeerd leverde een model dat overlap tussen routes integreert, met een betere model fit in de basisperiode, geen duidelijke verbetering van de voorspellingsprestaties op. Deze studie vormt een bijdrage aan de schaarse literatuur over de validatie van reisvraagmodellen en is de eerste die een externe validatie uitvoert van een ov-routekeuzemodel. Onze resultaten benadrukken de noodzaak om routekeuzemodellen voor het openbaar vervoer te valideren alvorens ze te gebruiken voor het afleiden van beleidsaanbevelingen, vooral in dit data-gedreven tijdperk waarin dit vaak kan worden gedaan tegen relatief weinig moeite en kosten.

In deze dissertatie wordt smartcard-data gebruikt om de evaluatie van ov-prestaties en routekeuzemodellen te verbeteren, specifiek binnen de context van stedelijke multimodale ov-netwerken. De aanwezigheid van geautomatiseerde dataverzamelingssystemen voor transportplanning zal naar verwachting verder toenemen in de komende jaren, omdat het potentieel ervan binnen dit gebied steeds meer erkend wordt. In dit proefschrift benadrukken we verder de waarde van dergelijke gegevensbronnen door verschillende verbeteringen aan te dragen voor het bereiken van zowel wetenschappelijke als praktische doelen.

About the author

Malvika Dixit was born in Indore, India in 1990. She completed her Bachelors in Civil Engineering from Birla Institute of Technology and Science (BITS) Pilani, India. In 2014, she received her MSc degree in Transport with distinction from Imperial College London. As part of her Master's thesis, she developed logsum measures of accessibility for London, and explored their suitability for undertaking transport equity analysis. Following her MSc, she worked as a transport planner for a private consulting firm in Dubai for three years, where she worked on various traffic, transport, and parking studies. She was also involved in tram station sizing, metro ridership data analysis, and pedestrian accessibility planning projects.

She started her PhD in 2018 with the Smart Public Transport Lab at the Transport and Planning (T&P) department, Faculty of Civil Engineering and Geosciences, Delft University of Technology. During her doctoral studies, she was involved in teaching activities and co-supervised the thesis of three graduate students. Recently, she won the Young Researcher of the Year 2022 award by the International Transport Forum (ITF) for her research on inclusive mobility conducted as part her PhD.

Since September 2022, she works at Amazon, Luxembourg contributing to the long-term planning of their end-to-end transportation network for Europe. In her leisure time, Malvika enjoys cooking, baking, learning about the connection between neurobiology and human behaviour, and exploring spirituality and its integration with the modern way of living.

List of Publications

Journal articles

1. **Dixit, M.**, Cats, O., van Oort, N., Brands, T., Hoogendoorn, S. (2022). Validation of a Multi-modal Transit Route Choice Model Using Smart Card Data. (*Under review*).

2. Brands, T., **Dixit, M.**, Zúñiga, E., van Oort, N. (2022) Perceived and actual travel times in a multi-modal urban public transport network: comparing survey and AVL data. *Public Transport*, 14, 85–103.

3. **Dixit, M.**, Cats, O., Brands, T., van Oort, N., Hoogendoorn, S. (2021). Perception of overlap in multi-modal urban transit route choice. *Transportmetrica A: Transport Science*.

4. **Dixit, M.**, Chowdhury, S., Cats, O., Brands, T., van Oort, N., Hoogendoorn, S. (2021). Examining the circuity of urban transit networks from an equity perspective. *Journal of Transport Geography*, 91.

5. **Dixit, M.**, Sivakumar, A. (2020) Capturing the impact of individual characteristics on transport accessibility and equity analysis. *Transportation Research Part D – Transport & Environment*, 87.

6. Brands, T., **Dixit, M.**, van Oort, N. (2020). Impact of a new metro line in Amsterdam on ridership, travel times, reliability and societal costs and benefits. *European Journal of Transport and Infrastructure Research,* 20(4), 335-353.

7. **Dixit, M.**, Brands, T., van Oort, N., Cats, O., Hoogendoorn, S. (2019). Passenger Travel Time Reliability for Multimodal Public Transport Journeys. *Transportation Research Record*, 2673(2), 149–160.

Peer reviewed conference publications

1. Kolkowski, L., **Dixit, M.**, Cats, O., Verma, T., Jenelius, E. (2022). Measuring Activity-Based Social Segregation using Public Transport Smart Card Data. *10th symposium of the European Association for Research in Transportation (hEART)*, Leuven, Belgium.

2. Brands, T., **Dixit, M.**, van Oort, N. (2021). Transformation to a trunk and feeder network: effects on passenger flows, travel times and reliability. *Transportation Research Days 2021*, Virtual.

3. **Dixit M.**, Cats O., Brands T., van Oort N., Hoogendoorn S. (2021). Route Overlap in Multi-modal Urban Transit Route Choice. *100th Transportation Research Board (TRB) Annual Meeting*, Virtual.

4. **Dixit M.**, Chowdhury S., Cats O., Brands T., van Oort N., Hoogendoorn S. (2021). Examining Circuity of Urban Transit Networks from an Equity Perspective. *100th Transportation Research Board (TRB) Annual Meeting*, Virtual.

5. Brands T., **Dixit M.**, van Oort N. (2019) Impact assessment of new North/South metro line in Amsterdam. *European Transport Conference*, Dublin, Ireland.

6. Brands T., **Dixit M.**, van Oort N. (2019). Impact of a new metro line on ridership, travel times, reliability and societal costs and benefits. *Transportation Research days 2019*, Gent, Belgium.

7. Brands, T., R. Veldhuijzen van Zanten, **Dixit, M.** (2019). De impact van de Noord/Zuidlijn in Amsterdam: vergelijking van reizigers en reistijden (in Dutch). *Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk 2019*, Leuven, Belgium.

8. **Dixit M.**, Brands T., Cats O., van Oort N., Hoogendoorn S. (2019). Impact analysis of a new metro line in Amsterdam using automated data sources. *Transit Data 2019*, Paris.

9. **Dixit M.**, Brands, T., van Oort N., Cats O., Hoogendoorn S. (2019). Passenger Travel Time Reliability for Multi-Modal Public Transport Journeys. *98th Transportation Research Board Annual Meeting*, Washington DC.

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.

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

Azadeh, K., *Robotized Warehouses: design and performance analysis*, T2021/30, TRAIL Thesis Series, the Netherlands

Chen, N., *Coordination Strategies of Connected and Automated Vehicles near On-ramp Bottlenecks on Motorways,* T2021/29, December 2021, TRAIL Thesis Series, the Netherlands

Onstein, A.T.C., *Factors influencing Physical Distribution Structure Design*, T2021/28, December 2021, TRAIL Thesis Series, the Netherlands

Olde Kalter, M.-J. T., *Dynamics in Mode Choice Behaviour,* T2021/27, November 2021, TRAIL Thesis Series, the Netherlands

Los, J., *Solving Large-Scale Dynamic Collaborative Vehicle Routing Problems: an Auction-Based Multi-Agent Approach,* T2021/26, November 2021, TRAIL Thesis Series, the Netherlands

Khakdaman, M., *On the Demand for Flexible and Responsive Freight Transportation Services,* T2021/25, September 2021, TRAIL Thesis Series, the Netherlands

Wierbos, M.J., *Macroscopic Characteristics of Bicycle Traffic Flow: a bird's-eye view of cycling,* T2021/24, September 2021, TRAIL Thesis Series, the Netherlands

Qu, W., *Synchronization Control of Perturbed Passenger and Freight Operations,* T2021/23, July 2021, TRAIL Thesis Series, the Netherlands

Nguyen, T.T., *Highway Traffic Congestion Patterns: Feature Extraction and Pattern Retrieval,* T2021/22, July 2021, TRAIL Thesis Series, the Netherlands

Pudāne, B., *Time Use and Travel Behaviour with Automated Vehicles,* T2021/21, July 2021, TRAIL Thesis Series, the Netherlands

Gent, P. van, *Your Car Knows Best,* T2021/20, July 2021, TRAIL Thesis Series, the Netherlands

Wang, Y., *Modeling Human Spatial Behavior through Big Mobility Data,* T2021/19, June 2021, TRAIL Thesis Series, the Netherlands

Coevering, P. van de, *The Interplay between Land Use, Travel Behaviour and Attitudes: a quest for causality,* T2021/18, June 2021, TRAIL Thesis Series, the Netherlands

Landman, R., *Operational Control Solutions for Traffic Management on a Network Level,* T2021/17, June 2021, TRAIL Thesis Series, the Netherlands

Zomer, L.-B., *Unravelling Urban Wayfinding: Studies on the development of spatial knowledge, activity patterns, and route dynamics of cyclists,* T2021/16, May 2021, TRAIL Thesis Series, the Netherlands

Núñez Velasco, J.P., *Should I Stop or Should I Cross? Interactions between vulnerable road users and automated vehicles,* T2021/15, May 2021, TRAIL Thesis Series, the Netherlands