UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

FORMULATION OF NEW MODELS OF PASSENGER BEHAVIOR IN PUBLIC TRANSPORT USING FARE COLLECTION DATA

TESIS PARA OPTAR AL GRADO DE
DOCTORA EN SISTEMAS DE INGENIERÍA

JACQUELINE GRACE ARRIAGADA FERNÁNDEZ

PROFESORA GUÍA:
MARCELA MUNIZAGA MUÑOZ

PROFESOR CO-GUÍA:
ANGELO GUEVARA CUÉ

MIEMBROS DE LA COMISIÓN:
ODED CATS
SONG GAO
CARLO PRATO
DANIEL SCHWARTZ PERLROTH

Este trabajo ha sido parcialmente financiado por ANID-PFCHA/Doctorado Nacional/2017-21170750, ANID/Proyecto Redes 170103, ANID/FONDECYT 1191104 e Instituto de Sistemas Complejos de Ingeniería (ANID - PIA - FB0816).

SANTIAGO DE CHILE
2022

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE DOCTOR EN SISTEMAS DE INGENIERÍA
POR: JACQUELINE GRACE ARRIAGADA FERNÁNDEZ
AÑO: 2022
PROFESORA GUÍA: MARCELA MUNIZAGA MUÑOZ
PROFESOR CO-GUÍA: ANGELO GUEVARA CUÉ

FORMULACIÓN DE NUEVOS MODELOS DE COMPORTAMIENTO DE PASAJEROS DE TRANSPORTE PÚBLICO USANDO DATOS DE TRANSACCIONES DE PAGO

Los sistemas de transporte público son fundamentales para reducir la congestión y la contaminación atmosférica en las ciudades de todo el mundo. Por ello, entender las preferencias de los pasajeros de transporte público es esencial para mejorar, promover y evaluar las políticas públicas en este ámbito. Los modelos de elección de rutas son utilizados para entender las preferencias de los pasajeros, y por tanto, para mejorar el diseño de los sistemas de transporte. En este contexto, uno de los mayores desafíos es representar el comportamiento de elección de ruta de los pasajeros de forma realista y recoger, integrar y procesar datos que puedan apoyar el desarrollo de políticas públicas informadas.

En la última década, varios estudios de elección de rutas de transporte público han utilizado tarjetas inteligentes y datos de GPS para obtener las elecciones de ruta de los pasajeros y una gran cantidad de información sobre los viajes, como por ejemplo: el tiempo de viaje y el número de transbordos. Una de las principales ventajas de los datos pasivos de transporte es la cantidad de datos recogidos y la exactitud de la información de movilidad. Aunque muchos estudios han intentado representar de forma realista el comportamiento de elección de rutas de los pasajeros, todavía quedan muchos desafíos por resolver. El objetivo general de esta tesis doctoral es formular un nuevo marco de modelación de transporte público que permita una comprensión más realista del comportamiento y las percepciones de los pasajeros de un sistema de transporte público multimodal a gran escala y con disponibilidad de datos pasivos. Esto se traduce en las siguientes contribuciones principales de esta tesis: i) desarrollar y aplicar métodos que capturen la heterogeneidad de la estrategia de elección de ruta entre los pasajeros; ii) proponer y aplicar una metodología para evaluar diferentes enfoques para abordar el problema del conjunto de consideración en los modelos de elección de ruta de transporte público; iii) proponer y aplicar un método para incorporar el proceso de aprendizaje de los pasajeros en un modelo de elección de ruta mediante el uso de datos de tarjetas inteligentes; iv) proponer y aplicar un método para evaluar el efecto de incentivos económicos y mensajes de cooperación para motivar a los pasajeros de transporte público a compartir información sobre las condiciones del sistema de transporte público utilizando una aplicación *crowdsourcing*.

Esta tesis muestra que datos pasivos de transporte público pueden utilizarse para comprender las preferencias de los pasajeros mediante la estimación de modelos de elección de rutas y que los datos recogidos a partir de tecnologías *crowdsourcing* pueden utilizarse para complementar los datos pasivos de transporte. Por último, los resultados de esta tesis ayudan a las autoridades de transporte a confiar en el uso de modelos de elección de ruta generados con datos pasivos de transporte y tecnologías de *crowdsourcing* para recoger datos de movilidad.

ABSTRACT OF THE THESIS TO APPLY
TO THE DEGREE OF DOCTOR EN SISTEMAS DE INGENIERÍA
BY: JACQUELINE GRACE ARRIAGADA FERNÁNDEZ
YEAR: 2022
ADVISOR: MARCELA MUNIZAGA MUÑOZ
COADVISOR: ANGELO GUEVARA CUÉ

FORMULATION OF NEW MODELS OF PASSENGER BEHAVIOR IN PUBLIC TRANSPORT USING FARE COLLECTION DATA

Public transport systems are key for reducing congestion and air pollution in cities worldwide. Therefore, understanding public transport passengers' preferences is essential to improve, promote, and assess public policies in this area. Route choice models are widely used to understand the public transport passengers' preferences and, therefore, to improve the public transport design. In this context, one of the greatest challenges is to represent public transport passengers' route choice behaviour in a realistic manner and to collect, integrate and process data that can support the development of informed public transport policies.

In the last decade, several public transport route choice studies have used smart cards and GPS data to obtain passengers' route choices and a large amount of information about trips, such as the in-vehicle travel time, out-of-vehicle travel time, and the number of transfers. One of the principal benefits of passive transport data, such as smart card data, is the amount of data collected and the accuracy of information about the choices made by public transport users. Even though many studies have tried to represent passengers' route choice behavior realistically, still many challenges remain on this subject. Thus, the general aim of this doctoral thesis is to formulate a new public transport modeling framework that allow for a more realistic understanding of public passengers' behavior and perceptions in the context of a large-scale multimodal public transport system and the availability of passive data. Based on a review of state-of-the-art research about studies that have used smart card data to estimate public transport route choice models different research gaps were identified. This results in the following main thesis' contributions: i) develop and apply methods that capture route choice strategy heterogeneity across passengers using smart card data; ii) propose and apply a methodology to assess different feasible approaches to address the consideration set problem in public transport route choice models; iii) propose and apply a method to incorporate the day-to-day learning process of passengers into a route choice model by using smart-card data; iv) propose and apply a method to evaluate the effect of economic incentives and cooperation messages to encourage public transport passengers to share information about public transport system conditions using a crowdsourcing app.

This thesis shows that passive public transportation data can be used to understand passenger preferences by estimating route choice models. However, the lack of some travel information still limits route choice passengers' behavior analysis using only passive transport data. In this line of analysis, this thesis proposes that data collected from transport-oriented crowdsourcing technologies can be used to complement some of the missing information in the passive transport data. Finally, the results of this thesis help transportation policymakers be confident in using route choice models generated with passive public transportation data and crowdsourcing technologies to collect mobility data.

*To my friends, my supervisors, my sisters, my nephew, my mother, my father and, especially,
to my husband and my little daughter.*

Table of Content

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Data in public transport passenger behavior modeling

Understanding the behavior and perception of public transport (PT) passengers is of great importance, as it allows transportation authorities to design or improve the system according to the needs and preferences of users. In this area of study, data from stated and revealed preference surveys has been widely used by researchers to model the behavior and perception of users of PT systems.

In stated preference (SP) surveys, people must respond to questions regarding hypothetical situations which may or may not be associated with a real experience. This data collection methodology has the advantage of being relatively inexpensive; however, it introduces hypothetical bias, since the user does not face an actual experience. Some authors who have used SP data to understand traveler behavior and/or traveler perceptions are Khattak et al. (1996); Dell'Olio et al. (2011); Grison et al. (2017); Vrtic & Axhausen (2002).

On the other hand, revealed preference (RP) surveys have the advantage of reflecting true information about the choices of users in real situations; however, traditional RP surveys involve recording data in the field at prohibitively high costs making it impractical to implement in large samples. Some authors who have used traditional RP data to understand traveler behavior and/or traveler perceptions are Eluru et al. (2012); Anderson et al. (2017); Z. Guo & Wilson (2011); Hoogendoorn-Lanser & Van Nes (2004); Raveau et al. (2011).

In the last decade, several authors have dealt with traditional RP data collection problems using smart card (SC) data (Schmöcker et al., 2013; Jánošíková et al., 2014; Kim et al., 2020; Nassir et al., 2018; Rui, 2016; Yap et al., 2020). SC data is available in cities that have introduced automated fare collection (AFC) systems for the PT system. The main purpose of SC is to collect PT revenue; however, as an additional benefit, they register a large quantity of very detailed data about the choices made by PT users at significantly lower costs and with few practical limitations, unprecedented granularity, and scalability (Pelletier et al., 2011). In summary, some analyses about passenger behavior on public transportation that previously required data from traditional RP surveys can now be performed using SC data, which overcome the shortcomings of traditional RP surveys in many aspects.

One of the principal benefits of SC data is the amount of data collected. SC data provides far more data at the spatial and temporal levels than traditional RP surveys (Bagchi & White, 2005). SC data allows the researcher to very accurately observe the time and location of boarding, the number of passenger using the PT system, and their movements, in a continuously dynamic way. To illustrate this, we compared SC data from five weekdays in April 2015 from the PT system of Santiago with RP data collected via a traditional Origin-Destination survey (EOD 2012) carried out in Santiago between July and November 2012. Focusing on peak morning hours (6:30-8:30), the EOD 2012 recorded 5,299 trip observations, while the SC database recorded 2.7 million valid trip observations[1]. In relative terms, the valid trip observations using SC data represent an increase of 509.5%. It is important to note that the evaluated EOD is the most recent and available one in Santiago, and this is from 10 years ago. On the other hand, any week of a year can be evaluated with SC data, which gives more availability and continuous trip information.

Figure 1.1 shows the number of observed trips initiated every 15 minutes using the SC data, while Figure 1.2 shows the same information based on observed trips obtained from the EOD 2012 survey. These figures reveal a better resolution of the trips collected in the SC database, which opens a door to new and diverse studies that allow us to more precisely understand the behavior of passengers in public transportation.

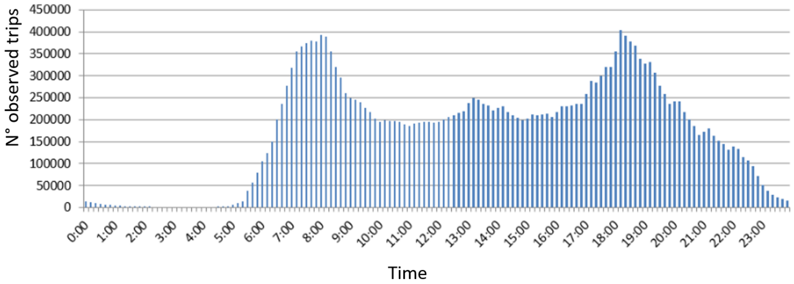


Figure 1.1: Observed trips with smart card data

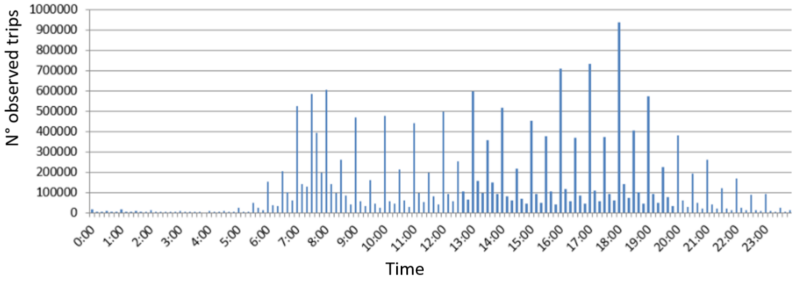


Figure 1.2: Observed trips with traditional EOD data

[1] SC database recorded 3.6 million trips, of which 2.7 million can be considered valid trip observations, since they can be processed to estimate the alighting time and alighting stop with the methodology developed by Munizaga & Palma (2012).

However, Bagchi & White (2005) indicate that SC data is a complement to traditional survey methods and not a replacement, since it provides fragmented information regarding passenger travel behavior. This type of data lacks sociodemographic information, trip purpose, identifications of transfer versus activities, quality of service perceived by users, among others. One of the greatest challenges at present is to develop methodologies to complete missing and relevant information in order to make SC data a more useful source of information for planners, operators and researchers. In this line, some authors have worked in the following lines:

- Combination of SC data with data obtained from travel surveys (Kusakabe & Asakura, 2014; Long & Thill, 2015).
- Estimation of alighting time and stop for each transaction (T. Li et al., 2018; Zhao et al., 2007; Cui, 2006; Trépanier et al., 2007; Munizaga & Palma, 2012; Polson & Sokolov, 2017; Nam et al., 2017; YU & YANG, 2006), since in some PT systems, the passenger is only required to tap-in, while tap-out is not required or available.
- Identification of transfers and activities in order to group trip stages into a complete trip (Munizaga & Palma, 2012; Seaborn et al., 2009; Devillaine et al., 2012; Gordon et al., 2013; Nassir et al., 2015).
- Identification of trip purpose, which is mainly based on activity duration limit and activity location (Kusakabe & Asakura, 2014; Devillaine et al., 2012; Lee & Hickman, 2014).
- Identification of area of residence, based on repetitive observations of the first transaction of the day (Amaya et al., 2017)

The efforts made by several authors to fill the information missing in SC transaction data present ways to include valuable information to understand PT passenger behavior, evaluate traditional travel behavior models, and formulate new models of passenger behavior. This research contributes to this area of study by building more realistic and accurate models of PT passenger behavior using SC data from Santiago (Chile), which is a city with a very large transit network. In doing it so, this research also addresses technical challenges on how to construct information for analysis of trips, such as the identification of origin/destinations zones of each trip, the identification of different attractive and available alternative routes to travel between a certain origin and destination, and the inference of the attributes of the alternative routes.

1.2 Previous studies of public transport passenger behavior using SC data

The data sources generated by AFC in PT systems, and the efforts of researchers to fill information missing from this type of data, have caused an improvement in data quantity and quality, creating the possibility to analyze passengers behavior at different levels of aggregation. PT literature has reported a wide range of studies that use SC data to aggregate the behavior of passengers at both spatial and temporal levels. For example, several authors have used SC data to build origin-destination (OD) matrices (Zhao et al., 2007; Trépanier et al., 2007; Seaborn et al., 2009; Wang et al., 2011; Nassir et al., 2011; Munizaga & Palma, 2012;

Gordon et al., 2013; Alsger et al., 2016; Yap et al., 2018), which consist of aggregating all journeys by OD zone. A group of studies, mainly using data mining techniques, has focused on day-to-day variability of passengers travel patterns to classify passengers considering the spatial and temporal regularity of their trips (Morency et al., 2007; Espinoza et al., 2017; Ma et al., 2017). Other studies have generated methodologies to evaluate the operation of PT systems, at a spatial and temporal level, such as level of crowding, average travel time, number of trip stages, number of transfers, leakage of users, among others (Núñez et al., 2015; Gschwender et al., 2016; Chapleau et al., 2011).

One of the main contributions of the introduction of SC data is the possibility of modeling the passengers' behavior on a disaggregated level by identifying the behavior of each individual. With respect to the disaggregated level (routes or trajectories used by individuals), there are some authors who have used railroad or Metro system data from SC transactions to infer the chosen route by passengers using shortest path algorithms (Kusakabe et al., 2010; Van Der Hurk et al., 2015) or probabilistic models (Zhao et al., 2017). Others studies have focused on the variability of chosen PT routes. Tao et al. (2014) used a day of SC data from the PT bus system in Brisbane, Australia and the coefficient of variation of trajectories in a given origin-destination pair to find that the chosen routes by passengers are more dispersed during the day than in the morning. Unlike Tao et al. (2014), Kurauchi et al. (2014) analyzed the variability in the routes chosen by individuals, and not by a given origin-destination pair. To do so, they used SC data from the PT system in London, and an n-step Markov model, where each state represents the service selected on the evaluated day and in days prior to the evaluation. They found that only 17% of passengers in London use the same transit line every morning (considering 4 days of evaluation). However, if they consider common lines, the variability in the chosen routes decrease, indicating that much of the variation in the route choice is due to common lines that could be part of the users' strategy. On the other hand, J. Kim et al. (2017) used six months of SC data from the Brisbane, Australia PT system to analyze the regularity with which users use different routes. This study used the metric called the "Stickiness Index" to quantify the range of preference, from users who always select the same route (high "Stickiness Index") to users whose chosen route selections are more varied (low "Stickiness Index"). J. Kim et al. (2017) analyzed the variability in route choice given an origin-destination pair and the route variability per individual.

It should be noted that the studies mentioned in the previous paragraph evaluate the variability in passenger route choices, but not the variability of the consideration set, which is the set of routes considered attractive to passengers. From the review here, very few studies have focused their analysis on the variability of the set of attractive routes or the consideration set. One of these studies is Nassir et al. (2017), which used SC data from the Queensland, Australia PT system and a statistical algorithm to infer the set of attractive routes based on a given OD pair.

Understanding the preferences of passengers on public transportation is essential for transportation authorities to design or improve the PT system according to the needs and preferences of users. One of the most-used models to understand traveler preferences is the route choice model. Specifically, route choice models allow transportation authorities to evaluate transportation planning performance, assess new transport policies, and predict travel behavior in new transport contexts. However, most of the literature regarding route choice models has been focused on private transport systems rather than PT systems due to the paucity of historical observed route choice data. SC data allows for filling this gap because it is longitudinal data. Using SC data, it is possible to observe the historical travel choices of PT passengers and answer questions related to patterns or habits of travelers, which require multiple observations of the same individual. The main contribution of the introduction of SC data is the possibility of modeling the passengers' route choice behavior more realistically. Some authors have developed route choice models using SC data (Schmöcker et al., 2013; Nassir et al., 2018; Jánošíková et al., 2014; Kim et al., 2020; Yap et al., 2020; Rui, 2016) to evaluate the passengers' perception about some important trip attributes (in-vehicle travel time, transfer walking time, waiting time, number of transfer, and crowding level).

Schmöcker et al. (2013) developed a discrete choice model for selecting a hyperpath, which is a set of paths that could be optimal to reach a destination from a given origin. The approach considers two levels. The first level represents the choice of consideration set using a Multinomial Logit (MNL) model that incorporates travel time, waiting time, number of transfers, and size of the consideration set. The second level represents the choice of transit line, based on the assumption that the user takes the first transit line that arrives at the stop (a deterministic choice). To test the model, Schmöcker et al. (2013) used SC data from a bus operator in a Japanese city and estimated the proposed model using three origin-destination pairs with direct routes. They found that the route choice behavior varies among different groups of passengers. Elderly people dislike waiting time more than other age groups, and people use smaller, more restrained consideration sets compared with the set proposed by the model. Finally, they indicate that the likelihood function is non-concave, so the parameters they obtained vary depending on the starting point.

Unlike the model proposed by Schmöcker et al. (2013), which possesses a combinatorial nature of conformation of possible consideration sets, which is a disadvantage in dense transportation systems, Nassir et al. (2018) propose a methodology to calibrate a route choice model with SC data using arc enumeration; i.e., using recursive formulas that represent the boarding, alighting, and transfer utilities at each arc along the route. The utility functions are represented by the attributes travel time, waiting time, and walking distance in transfers. The choices of taking a transit line, alighting at a stop, and transferring are modeled using a MNL model. The benefit of this methodology is that, by avoiding the explicit enumeration of routes, it does not generate a large computational cost. This model was applied to two origin-destination pairs in PT SC data from Brisbane, Australia. They found that the attraction of direct routes is very similar to the observed percentage of users using them. However, the percentage of users using routes with transfers is higher than the result estimated by the model. This could be explained because the model could estimate some quick activities as transfers.

Jánošíková et al. (2014) estimated a MNL model with SC data from a city in Slovakia to represent PT route choice. The studied transit network was relatively small, with eight trolleybus and 10 transit lines. For the model, they defined the consideration set as the routes observed in the SC data, and used the variables in-vehicle travel time, transfer walking time, number of transfers, and transit line headway as attributes of the model's utility function. Following the same line of research, Kim et al. (2020) also applied a MNL model, adding new variables, such as travel time reliability and path circuit index, using SC data of the Seoul

Metropolitan Area. This is a very large multimodal transit network with 405 transit lines and 12 Metro-rail lines. Both studies found that all variables are significant with the expected sign, showing that the route choice and the values of attribute coefficients can be inferred from SC data in both small- and large-scale transit networks.

Previous studies have used the basic MNL model, which assumes independence of alternatives, implying that the correlation due to overlapping route segments is ignored. Some authors have dealt with this problem using the analytical approach of Path Size Logit (PSL) models, which account for the correlation by adding a deterministic term that reduces the utility function of overlapped routes (Yap et al., 2020; Rui, 2016). Yap et al. (2020) used SC data from The Hague in the Netherlands, which has 12 tram lines and eight bus lines, to evaluate the influence of crowding on passengers' travel experience. On the other hand, Rui (2016) developed a path size route choice model using SC data from Singapore and, in contrast to the previous authors, implemented six practical consideration set generation approaches: the labeling approach, the link elimination approach, the K shortest paths approach, the simulation approach, the branch and bound approach, and the nested labeling and Link elimination approach.

1.3 Analytical framework and research gaps for the analysis of passengers' route choice behavior using passive transport data

To understand PT passengers' travel behavior in a realistic way, it is necessary to obtain large amounts of information and estimate a route choice model to reflect how passengers perceive different trip attributes. From the literature review of state-of-the-art research about PT passenger behavior using SC data, this thesis identifies and proposes a framework to measure the PT passengers' perceptions using passive data in five steps. This analytical framework is shown in Figure 1.3, where prior to the first step, passive transport data must be processed. This processing data process aims to infer the destination associated with each transaction, in systems where passengers must only tap-in (e.g. Trépanier et al., 2007; Munizaga & Palma, 2012), determine which transactions form one trip (e.g. Seaborn et al., 2009; Gordon et al., 2013), and remove incomplete transactions due to system errors or when some missing information is impossible to infer. As a first step, trip origins and destinations must be identified from the SC data. It is important to note that the SC data does not provide information regarding the actual origin and destination of trips; instead, it provides the first stop or station of the trip, which can be understood as the origin, and the last stop or station of the trip, which can be understood as the destination. While some authors have worked with stop-to-stop pairs to represent trip origins and destinations (Rui, 2016; Jánošíková et al., 2014; Schmöcker et al., 2013), others have aggregated stops and stations within walkable and transferable distances into one representative node (Kim et al., 2020; Yap et al., 2020). In this way, passengers departing from or arriving at different stops that belong to the same representative node should be considered as trips with the same origin-destination zone.

Once the OD pairs have been identified, the second step is to define the alternative routes. Public transport studies using passive data to understand passengers' perceptions define a route as the unique sequence of boarding locations (stop or zone), alighting locations (stop or zone), and intermediate lines or a combination of lines. The group of studies that do consider specific lines between stops or zones build the consideration set of passengers with itineraries, where all lines belonging to a set of relevant alternatives are regarded as different options (Jánošíková et al., 2014). On the other hand, most studies that work with combinations of lines between stops or zones consider that alternative routes which share the same geographical path as a single alternative (e.g. Yap et al., 2020; Kim et al., 2020).

After defining alternative routes, the third step is to obtain the attribute values for each route alternative. Combining AVL, GTFS and SC data, it is possible to estimate important trip attributes, such as in-vehicle travel time, transfer walking time, waiting time, number of transfers, and crowding level (Rui, 2016; Yap et al., 2020; Kim et al., 2020; Jánošíková et al., 2014). Once alighting stops and trip sequences are observed or inferred, the in-vehicle travel time for each trip can be obtained from AVL data. The number of transfers can be directly obtained from the sequence of trips, transfer walking time can be inferred assuming a standard walking speed, and waiting time can be inferred assuming a specific distribution (S. Guo et al., 2011; Ingvardson et al., 2018).

The fourth step to measure PT passengers' perceptions using passive data is to identify the consideration set for each OD pair identified in step 1. The consideration set generation process is complex, since there usually are countless feasible alternative routes in a transport network, especially in a dense multimodal network (C. G. Prato, 2009). Additionally, evidence suggests that the number of route alternatives known and considered by the passenger is substantially smaller than the total number of available alternatives (Hoogendoorn-Lanser & Van Nes, 2004). Two ways of identifying the consideration set in practice can be distinguished in the applied literature: it can be built using an algorithm or heuristic that emulates how individuals may build their own consideration set (C. G. Prato, 2009), or it can be imputed using historical data. Most of the studies using SC data to measure the PT passengers' perceptions have imputed the consideration set from historical SC data (Yap et al., 2020; Kim et al., 2020; Jánošíková et al., 2014), while one study has built it using an algorithm or heuristic that emulates how individuals may build their own consideration set (Rui, 2016).

The last step, after identifying the consideration set, is to determine the PT passengers' perception of trip attributes. For this, most studies using SC data to understand PT passengers' route choice behavior use the MNL discrete choice model (Schmöcker et al., 2013; Nassir et al., 2018; Jánošíková et al., 2014; Kim et al., 2020), while a small number of studies have developed models that capture the correlation between routes (Yap et al., 2020; Rui, 2016). To the best of our knowledge, no study has captured heterogeneity between passengers. It is worth noting that most of these studies were applied to relatively small-scale transit networks (Schmöcker et al., 2013; Nassir et al., 2018; Yap et al., 2020; Jánošíková et al., 2014).

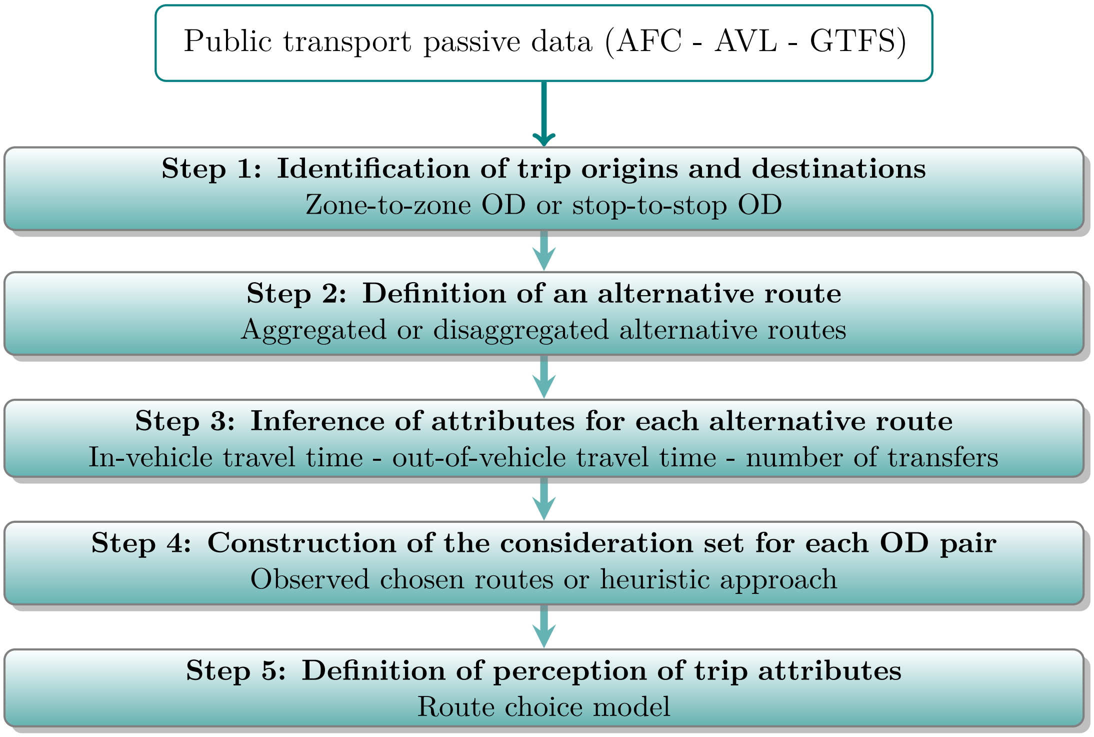

Public transport passive data (AFC - AVL - GTFS)
Step 1: Identification of trip origins and destinations
Zone-to-zone OD or stop-to-stop OD
Step 2: Definition of an alternative route
Aggregated or disaggregated alternative routes
Step 3: Inference of attributes for each alternative route
In-vehicle travel time - out-of-vehicle travel time - number of transfers
Step 4: Construction of the consideration set for each OD pair
Observed chosen routes or heuristic approach
Step 5: Definition of perception of trip attributes
Route choice model

Figure 1.3: Framework to measure the PT passengers' perceptions using passive data

Based on the literature review of state-of-the-art research about PT passenger behavior using SC data and the analytical framework in Figure 1.3, the following research gaps can be identified:

1. As the literature suggests, passengers can adopt different strategies to choose a route to reach a destination (Raveau & Muñoz, 2014; Spiess & Florian, 1989). This aspect is particularly important for the second step in Figure 1.3, where the definition of an alternative route depends on the assumption of the passengers' route choice strategy. However, to the best of our knowledge, there are no studies on PT passengers' route choices that use real, observed data to evaluate which type(s) of strategies are considered by PT passengers nor studies that consider this heterogeneity within a single model. Some models assume that all passengers choose disaggregated alternative routes (itineraries), while others assume that all passengers choose aggregated alternative routes (geographical aggregation). This gap generates a sub-optimal standard of analysis where PT authorities do not consider the heterogeneity between passengers; therefore, the transport policies may not consider the needs and preferences of all PT passengers.
2. Route choice modeling requires identifying the consideration set (step 4 in Figure 1.3), which is unknown to the researcher when working with RP data. Route choice literature presents different practical approaches to build the consideration set. Most studies using SC data have used the historical approach, which is based on intuition, while other studies have used different heuristic approaches to emulate the passenger behavior (C. G. Prato, 2009). To the best of our knowledge, no comprehensive assessment of consideration set generation approaches has been carried out in a PT context on public transport. Furthermore, little is known about the impact of the composition of the consideration set in the case of public transport route choice modeling. This gap is essential for transportation authorities as a comprehensive assessment of consideration set generation methods allows them to identify which approach would generate less bias in the estimation results of a route choice model and could help to understand passenger preferences better. In summary, identifying the consideration set generation method

with the best performance allows PT authorities to improve the system design more efficiently.

3. The route decision process usually involves an evaluation of different route attributes. This process is particularly important for step 3 in Figure 1.3, where researchers should consider that some route attributes can be fixed over time, such as the number of transfers; however, there are other attributes that represent uncertainties in the trip process, such as waiting time and in-vehicle travel time, among others. Information about these uncertain attributes can come from past travel experiences between the same OD pair (reinforcement learning) or from the description of travel information (cognitive learning). Most studies about transport route choice consider travel time and travel time variability (if considered) as static attributes; therefore, they assume that all travelers at all time points possess the same knowledge regarding the travel time distribution, ignoring the relation between choices and past experiences.

   SC data has the potential to provide the historical PT passengers' behavior, allowing observation of the chosen alternative route and the variability of route attributes. Particularly, it is possible to observe the travel time experienced by each individual when they travel in the system. Although SC data provides the opportunity to use real data, we have not been able to find any study that considers how to include the relationship between choices, past experiences, and descriptive information to capture passengers' learning processes. Understanding the learning process of PT passengers is especially important for transport policies which are oriented toward delivering travel information. Suppose transport policymakers do not understand how passengers learn from descriptive and experienced travel information; in that case, they may not be assertive in implementing communication channels to explain travel information to passengers to improve their experience in the system.

4. Although passive PT data sources, such as SC and AVL data, provide highly accurate travel and mobility information, they still present the substantial challenge of incorporating information that cannot yet be recorded passively. Among this missing information is data to characterize PT passengers, such as socioeconomic data, and trip attributes relevant for modeling the behavior of passengers, such as trip origin, trip destination, and transport infrastructure quality (both vehicles and stops). This missing information could be used in step 3 of Figure 1.3 to infer alternative route attributes or traveler characteristics. In addition, due to this lack of information, PT authorities may ignore relevant characteristics of the passengers and infrastructure problems which could guide the authorities' effort to improve the passengers' experience in the transportation system.

1.4 Research Objective, Questions and Scope

1.4.1 Research objective

Public transportation is recognized as an efficient transport mode since it reduces both congestion and air pollution. For this reason, encouraging people to use the PT system instead of traveling by private vehicle is the principal challenge for PT authorities and researchers. To this end, route choice models play a fundamental role in understanding passengers' behavior and perceptions. This information is usually used to improve the PT system design, thereby increasing the attractiveness of public transportation. For decades, PT authorities have used SP or RP traditional surveys to develop PT route choice models, which can be expensive and impractical for ensuring representativeness in a large-scale PT system. SC data provides a solution to these problems by representing the passenger choices in a more realistic manner at a low cost to the researcher. Although some studies have used SC data to estimate PT route choice models, a review of state-of-the-art research shows there are several research gaps that must be addressed. The main purposes of this thesis are: (i) to formulate a new public transport framework modeling that allow for a more realistic understanding of PT passengers' behavior and perceptions in the context of a large-scale multimodal public transport system; (ii) to propose methods to capture the current missing important information in the passive transport data.

1.4.2 Research questions

This thesis presents the following four research questions to address the main objective of this study.

1. **Do public transport passengers use different route choice strategies?**

The first research question is focused on the second level of Figure 1.3 and aims to fill research gap 1 by recognizing that there are different ways to model the passengers' route choice strategy. In particular, almost all route choice models use consideration sets composed of itineraries, while PT assignment models for strategic analysis mostly use a version of the common lines approach. We postulate that these modeling ways are correct but for different passengers, and therefore heterogeneity exists in the route choice strategy between users. Therefore, we propose the classification of possible route choice behavioral strategies in two groups: disaggregated strategies, where alternatives correspond to itineraries and common lines are not considered, and aggregated strategies, where common lines are considered as part of the same alternative. The purpose is to verify if there is heterogeneity in the route choice strategy, both between users and across different contexts.

2. **How do different consideration set generation practical approaches impact estimation and prediction in a PT route choice model?**

The second research question is focused on the consideration set generation process (step 4 in Figure 1.3), which is a substantial challenge for route choice modelers. The aim is to address research gap 2 by assessing different practical approaches, in order to verify if the method most often used by studies utilizing SC data, which we call the Historical/Cohort approach (in which the consideration set is constructed with observed choices), outperforms other typical practical approaches.

3. **How can public transport passengers' past experiences be integrated into a route choice model to incorporate the uncertain nature of in-vehicle travel time in the public transport system?**

The third research question aims to improve the representation of passengers' perceptions of uncertainty attributes (steps 3 and 5 in Figure 1.3) and relates to research gap 3. The research question focuses on the relationship between the past experiences of passengers and their current choices. It considers that the PT route choice context is a choice decision under uncertaint conditions, where passengers commonly use information from prior experiences, and therefore that the perception of some attributes (e.g. in-vehicle travel time) can vary across time. This learning process is especially important in new PT contexts, such as the implementation of a new bus service or Metro line.

4. **How can passengers be encouraged to provide mobility and transport information through crowdsourced mobile public transport applications?**

The fourth question is focused on the third level of Figure 1.3 and seeks to address research gap 4 in order to complement and improve the process of measuring trip attributes and passengers' characteristics using passive transport data. Based on the high level of penetration of smartphones worldwide, in this study, we argue that crowd-sourced mobile public transport applications can be an important channel to obtain mobility information to complement passive transport data. However, a high participation level is essential, and therefore, it is important to motivate travelers to share information. In this context, this research question focuses on showing that both low-cost economic incentives and cooperative messages can encourage crowd-sourcing mobile application users to report PT information.

1.4.3 Research scope

Based on the formulated research purpose and research questions, this thesis focuses on understanding the route choice behavior of PT passengers in Santiago, Chile. Santiago is the capital of Chile and has a population of over 7 million inhabitants. The public transport system serves roughly 50% of motorized trips, and it is operated by headway scheduling; therefore, lines do not have fixed time schedules. The fare system is almost fully integrated, with a flat fare between urban buses, Metro, and some rail services, allowing up to three trip legs within a two-hours time window. In a typical week, 3 million passengers use the system to make 25.5 million trips. The network includes 7 metro lines, more than 300 bi-directional transit lines, and one rail service.

In particular, to estimate route choice models, we use observations from passengers traveling during morning peak periods (6:30-8:30 AM) on weekdays who stay in their destination locations for at least two hours. These restrictions are aimed at capturing trips to regular activities such as working or studying. This means that other time periods, such as the afternoon peak period, and other types of trip purposes, such as shopping, are not the focus of this research.

Since the analysis for this thesis was carried out using the Santiago, Chile public transport network, this research considers Metro lines, bus lines, and one rail line. This means that while multimodal transport modes are considered, other modes of transport, such as bicycles or scooters, are not the focus of this research.

The contributions of this thesis increase our knowledge of the route choice behavior of PT passengers. These findings can be extended to those PT systems that operate similarly to the PT system in Santiago, which is operated by headway scheduling (lines do not have fixed-time schedules). It is important to note that passengers' route choice behavior in other contexts, e.g., when passengers travel in a schedule-based transit network, exceeds the scope of the present thesis.

1.4.4 Basic definitions

The following list provides a formal definition for nine terms that we frequently use.

- Node: it can be a bus stop or a Metro/train station.
- Trajectory: a spatial sequence of a public transport vehicle position.
- Transit line/ line: a group of public transport vehicles that operate between an initial node and a final node within the transit network. All vehicles on the same line travel within the network by always following the same sequence of network links and nodes. Each transit line provides a public transportation service defined by a sequence of stops, trajectory, and vehicle frequency.
- Route/ alternative route: a path that a passenger can follow within the transit network to travel between an origin and a destination zone. It can be identified by a sequence of nodes, with the first being located in the origin zone of the trip, the final being located in the destination zone of the trip, and where all intermediate nodes represent transfer points.
- Route section: a portion of a route between two consecutive initial, final and transfer nodes. Each route section is associated with a transit line or a group of transit lines.
- Itinerary/ disaggregated alternative: a route defined by a specific sequence of initial-transfer-final nodes with one specific line between each pair of consecutive nodes.
- Aggregated alternative: a route defined by a specific sequence of initial-transfer-final nodes with a group of common lines between each pair of consecutive nodes, aggregated into a single alternative.
- Common lines: the set of transit lines that altogether minimize the total expected travel time between two nodes in a network; i.e., if the passenger takes the first bus from this set of lines, the summation of the expected waiting time and the in-vehicle travel time will be minimal.
- Transfer: the movement of a passenger from one public transport vehicle to another.

1.5 Research Contribution

1.5.1 Scientific contribution

1. **Unveiling route choice strategy heterogeneity from smart card data in a large-scale public transport network**

This research offers theoretical, methodological, and empirical contributions. From a theoretical perspective, this study proposes a classification of route choice behavioral strategies as either aggregated or disaggregated, depending on whether passengers do or do not (respectively) consider common lines as part of the same alternative. In addition, this study integrates theories from PT assignment models (which usually use the aggregated strategy approach) and PT route choice models (which usually use the disaggregated strategy approach) into a single framework for route choice modeling. From a methodological perspective, we developed an indicator function to validate whether heterogeneity exists in the route choice strategy between users and across contexts. Then we proposed a methodology based on the integrated discrete choice and latent class approach to study the heterogeneity of route choice strategies using smart card data. This approach involves estimating path-size logit models built with alternatives taken from disaggregated and aggregated strategies and a latent class model built from a combination of both approaches. Finally, this study makes an empirical contribution by applying the proposed methodology in Santiago, Chile, where we found evidence of heterogeneity, suggesting a need to rethink the modeling approaches often used in the field.

2. **Assessing feasible approaches for building the consideration set in public transport route choice modeling using smart card data**

The contributions of this study are theoretical, methodological, and empirical. From a knowledge perspective, this research contributes to understanding the effect of consideration set composition in a PT route choice model. We have shown that the Historical/Cohort approach, which can be obtained directly from SC data (avoiding the need to use assumptions), allows for a better representation of PT passengers' behavior than practical methods traditionally used to generate the consideration set. Additionally, we revisit the explanation of Guevara (2022) about theoretical conditions under which the Historical/Cohort approach would recover population parameters. From a methodological perspective, this research proposes a methodology to assess different feasible approaches to address the consideration set problem in PT route choice models. This methodology is based on the estimation and prediction performance of different route choice models integrating different consideration set generation approaches. Finally, we present an empirical contribution, applying the proposed methodology with real observed data obtained from the PT payment system of Santiago, Chile.

3. **Evaluating the role of experience in the route choice context using smart card data in a large-scale public transport network**

This study offers methodological and empirical contributions. From a methodological perspective, this research proposes a method to incorporate the day-to-day learning processes of passengers into a PT route choice model. This methodology is based on integrating discrete choice and instance-based learning theories. The instance-based learning theory captures the recency of experiences in the human memory and is based on the power law of forgetting. The instances-based learning model we have applied has been previously used in route choice models in private transport and in simulated or laboratory data. Therefore, this study also offers testing of a proposed methodology in a new context, namely the public transportation system. Additionally, we present an empirical contribution, applying the proposed methodology to a case study consisting of 12 weeks of revealed preferences constructed from SC data from Santiago, Chile, coming from a subset of individuals that faced the opening of a new Metro line in the PT system. Finally, this research contributes to understanding the learning process of PT passengers in a new context of the PT system.

4. **The effect of economic incentives and cooperation messages on user participation in crowdsourced public transport technologies**

This study offers methodological and empirical contributions. This study shows that transport-oriented crowdsourcing applications provide an opportunity to improve SC data quality by providing missing information. We show that an economic incentive and a cooperation message (less effectively) encourage PT passengers to share mobility information. This research also offers a methodological contribution as it uses a large randomized field experiment providing internal and ecological validity. In this context, we examined the effect of economic incentives (a lottery for free trips) and cooperation messages (asking users to help the community) to encourage passengers to share information about bus stop conditions using a crowdsourcing app. We estimated a zero-inflated negative binomial model to examine users' contribution levels and a logit model to examine the effect of each experimental condition on the participation rate. Finally, we present an empirical contribution by applying the proposed methodology to a transport-oriented crowdsourcing mobile application for users in Santiago, Chile.

1.5.2 Societal contribution

The studies of this thesis show that SC data can be used to understand passenger preferences, and it is possible to focus the use of traditional mobility surveys on capturing only those variables/aspects which are not captured by SC data. It is important to note that this study shows that SC data can be applied to the two stages of a PT route choice model and that there is no need to execute algorithms or heuristics, which can be computationally expensive in a large-scale transit network such as Santiago's transit network. Consequently, this work helps PT authorities (regulators and operators) confidently use route choice models generated with passive public transport data in transport policy contexts to understand the most relevant trip attributes to passengers. In summary, this thesis helps PT authorities reduce monetary and time costs to obtain information on passengers' route choice behavior and improve the PT system design to make the travel experience of PT passengers better.

The findings of this research can be used to suggest some guidelines for public transport models that can help steer decisions regarding where to implement new transit routes and how to improve existing ones. Some important guidelines to improve the travel experience of PT passengers are: implementing lines with an overlap in high-demand sectors will effectively allow passengers to reduce their waiting time, identifying OD pairs that require transfers can focus planning efforts to reduce the number of onerous transfers, and high-demand transfer points should be carefully designed to avoid walking.

Also, this study finds that descriptive information about routes is essential at the beginning of a new PT context, such as a new Metro line. Therefore, trip information should be enforced in these situations. Delivery of travel information allows PT authorities to improve PT passengers' travel experience when there is a change in the design of the PT system.

Finally, this study suggests that PT authorities should complement passive data sources with data collected from transport-oriented crowdsourcing applications. This type of data allows PT authorities to improve the PT design, particularly the PT infrastructure (such as bus stops), and therefore improve the travel experience of PT passengers. Additionally, mobility data collected from community-oriented passenger information technologies reduce the monetary and time costs of processes related to obtaining some mobility information, particularly physical inspections to monitor PT infrastructure.

1.6 Outline

This research is divided into four parts. Chapter 2 evaluates the route choice strategies of public transport users and estimates a Multiple Indicator Solution (MIS) latent class model that captures heterogeneity between passengers. Chapter 3 evaluates the performance of different approaches to generate the consideration set for public transport route choice models. Chapter 4 captures passengers' learning processes in the context of a new Metro line, combining an Instance-Based Learning model with a PT route choice model. Chapter 5 examines the use of incentives to encourage crowdsourcing public transport application users to share information about the status of the transport system. Finally, chapter 6 presents the main conclusions from this thesis, together with implications for PT authorities and recommendations for future research.

Chapter 2

Unveiling route choice strategy heterogeneity from smart card data in a large-scale public transport network

This chapter is the first component of this thesis, which aims to answering the the first research question (see Section 1.4.2): Are public transport passengers using different route choice strategies?. This study contribution begin by proposing the classification of possible route choice behavioral strategies in two groups: disaggregated strategies and aggregated strategies. In the former, the alternatives correspond to itineraries, which are fixed sequences of stops and PT lines. In the latter, common line alternatives are considered, which are combinations of itineraries defined under given criteria. Almost all route choice models use consideration sets composed only of itineraries, while PT assignment models for strategic analysis mostly use a version of the common lines approach. We postulate that this dichotomy is inappropriate and that, instead, heterogeneity exists in the route choice strategy, both between users and across contexts. With the aim of verifying this hypothesis, we first propose an indicator function constructed as the difference between expected and observed trips for a given behavioral assumption. We apply then the indicator to a case study based on SC data from the city of Santiago, Chile, from which we find evidence of heterogeneity. We identify individuals that follow either an aggregated or a disaggregated strategy, as well as others who seem to be using a combination of both strategies. We further analyze the heterogeneity hypothesis using an integrated discrete choice and latent class approach, which we apply to the same case study. This approach involves estimating PSL models built with alternatives from disaggregated and aggregated strategies, as well as a latent class model built from a combination of both. It also addresses methodological challenges related to the definition of the consideration set and the correction of endogeneity. Results confirm the heterogeneity hypothesis, suggesting that 51.2% of passengers use a disaggregated strategy for route choice, while the rest use an aggregated one. We also find that travelers in the class considering aggregated alternatives appear to prefer bus over metro, while travelers in the class considering disaggregated alternatives appear to prefer metro over bus. The fact that waiting time is relatively more burdensome for travelers who consider the aggregated strategy class is in line with these travelers' preference for common lines. Walking time and bus crowding are more burdensome for passengers who consider the disaggregated strategy class, in line with their observed modal preferences.

This chapter was published in the following article:

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. Transportation Research Part C: Emerging Technologies, 134, 103467.

Author's contribution

Jacqueline Arriagada: corresponding author, conceptualization, methodology, software, formal analysis, writing-original draft, funding acquisition; **Marcela Munizaga:** conceptualization, methodology, resources, data curation, writing-review & editing, supervision, funding acquisition; **Angelo Guevara:** conceptualization, methodology, resources, writing-review editing, supervision, funding acquisition; **Carlo Prato:** conceptualization, methodology, resources, writing-review editing, supervision.

2.1 Introduction

Modeling the route choices of passengers in PT systems is a well-known complex problem, which is relevant to the improvement of PT network planning, design, and assessment. In this area, two modeling approaches can be distinguished, depending on whether the analysis focuses on route choice behavior modeling, or PT assignment models for strategic analysis.

Most studies on PT passengers' route choice behavior have used the MNL discrete choice model (Grison et al., 2017; Z. Guo, 2011; Jánošíková et al., 2014; Nassir et al., 2018; Raveau & Muñoz, 2014; Raveau et al., 2014, 2011; Vrtic & Axhausen, 2002; Kim et al., 2020), with consideration sets composed of itineraries, where all lines belonging to a set of relevant alternatives are regarded as different options. The use of the MNL model for this problem implies that the correlation due to overlapping route segments is ignored, as MNL assumes independence of alternatives. To address this limitation, the analytical approach of PSL models have often been adopted, which accounts for the correlation by adding a deterministic term that reduces the utility of overlapped alternatives (Anderson et al., 2017; Bovy & Hoogendoorn-Lanser, 2005; de Grange et al., 2012; Hoogendoorn-Lanser et al., 2005; O. A. Nielsen et al., 2021; Tan et al., 2015; Yap et al., 2020). In this study, we call this approach the disaggregated strategy.

Transit assignment models make use of the concept of common lines and hyperpaths (Chriqui & Robillard, 1975; Cominetti & Correa, 2001; De Cea & Fernández, 1993; Nguyen & Pallottino, 1988; Spiess & Florian, 1989). The common lines problem was originally described by Chriqui & Robillard (1975) as the strategy of identifying a subset of PT lines that minimizes the total expected travel time. According to this principle, passengers will take the first line of the common lines set that arrives at the bus stop. Chriqui & Robillard (1975) defined the common lines set in a portion of the route between two stops. Later, the definition of common lines is extended to a network with the name of shortest hyperpath by Nguyen & Pallottino (1988), where the first line that arrives to the stop defines the possible following characteristic of the route, e.g., the number of transfers. The common line and the hyperpath approaches have been successfully implemented for public transit assignment models for strategic analysis in software packages such as EMME/2 (INRO, 1996) and ESTRAUS (De Cea et al., 2003). We call this approach aggregated strategy.

The dichotomy between the disaggregated and aggregated approaches to, seemingly, the same problem seems inappropriate. For example, one possible implication of the concept of common lines for the modeling of discrete choices is that common lines may be perceived as part of the same alternative for passengers that face a route choice decision. However, to the best of our knowledge, only two studies have explicitly considered common lines between consecutive stops as part of the same alternative in a discrete choice framework. The first is Raveau & Muñoz (2014), who developed a MNL model using common lines as part of the same route alternative, and the second is Schmöcker et al. (2013), who developed a bi-level route choice model using hyperpaths as the consideration set. On the other hand, Kim et al. (2020) and Tan et al. (2015) use a similar idea that considers that overlapping sections between two consecutive stops of lines are regarded as one alternative. Their approach captures partially the common lines principle as it does not include those common lines that do not overlap their routes.

There is no theoretical or empirical justification for not using common lines in route choice behavior modeling. Raveau & Muñoz (2014) asked a sample of PT users to self-report the route choice strategies that they use, finding that some passengers stated choosing between itineraries, but others declared using common line or hyperpath strategies. In this study, we confirm and deepen the empirical findings by Raveau & Muñoz (2014) by using a novel methodology that makes use of massive revealed preference data gathered from SC, circumventing the problem of response bias that is inherent in self-reported data. The analysis is performed by estimating three types of models: i) a PSL model with the consideration set composed of itineraries; ii) a PSL model with the consideration set composed of common lines as part of the same alternative; and iii) a latent class model with consideration sets for two types of passengers: those who use itineraries and those who use common lines as part of the same alternative. We then use the results of these models to answer three research questions: Are PT passengers using different route choice strategies? Is it possible to capture heterogeneity in route choice strategy by using latent class models estimated from SC data? Can different behavioral parameters be captured?

Regarding the availability of data to observe route choice in PT, in recent years some studies have generated data by observing the transactions registered in automatic fare collection systems through SC. If the card ID is available, panel data with multiple choice situations can be built from SC transactions. Unlike traditional data obtained from surveys, this type of passive data has negligible costs and allows the collection of large volumes of personal travel data over long periods of time (Bagchi & White, 2005). These advantages make SC data attractive for route choice studies. Previous studies that use SC data for route choice modeling include the estimation of stop-to-stop route choice models in different types of networks: i) a railway or metro network (Kusakabe et al., 2010; Van Der Hurk et al., 2015; Zhao et al., 2017); ii) a bus network, using a reduced number of origin-destination pairs (Nassir et al., 2018; Schmöcker et al., 2013) and iii) a large scale multimodal transit network (Jánošíková et al., 2014; Tan et al., 2015; Yap et al., 2020; Kim et al., 2020). The current study provides a new case study, using SC data to observe the route choice of passengers in a large-scale multimodal network with more than 20,000 origin-destination pairs.

From the SC data, it is possible to obtain a large amount of information about a trip, such as the travel time in-vehicle, travel time out-of-vehicle, and the number of transfers. Additionally, combining AVL data, SC data, and GTFS data, it is possible to estimate the walking time in transfers and the waiting times at the beginning of the trips and the transfer points. Most studies that use SC data in route choice modeling assume that users maximize a utility function consisting of, mainly, the in-vehicle travel time, out-of-vehicle travel time, number of transfers, and an error term (Schmöcker et al., 2013; Jánošíková et al., 2014; Tan et al., 2015; Yap et al., 2020; Kim et al., 2020). The error component plays an important role in the route choice of passengers since it allows to explain the variation of the chosen route through the passengers, even when they follow the same strategy. This is because the error term captures the non-added attributes such as the level of crowding, the type of buses, the preference for traveling with other users, availability of seats, induvial characteristic, among others.

The remainder of this section is organized as follows. The next section presents a literature review on route choice strategy. An analysis of the observed passengers' travel behavior for the current case study follows. Then, the study discusses the proposed methods used. The model estimation results are then introduced. The last section draws conclusions and discusses policy and research implications.

2.2 Problem description

Every day, commuters and other travelers face the problem of selecting a route to arrive at a destination from a certain origin. To represent this choice, deterministic and probabilistic modeling approaches can be used. The deterministic approach is based on the minimum cost function, while the probabilistic approach incorporates unknown effects by adding an error term to the cost or utility function.

In the group of deterministic route choice models, the most common is the shortest-path heuristic (Gallo & Pallottino, 1988), where individuals are assumed to choose the itinerary with the lowest cost from the origin to destination. In the more complex hyperpath approach (Nguyen & Pallottino, 1988), passengers are assumed to consider a set of alternative routes that minimizes the total expected travel time and use the first line (from that set) that arrives at the bus stop.

Probabilistic modeling approaches are based on the Random Utility Maximization (RUM) framework, which has been widely used for representing route and mode choice behavior in the transport area (McFadden, 2000). The utility of each alternative is defined as depending on the valuation of a set of attributes through a set of coefficients that adjust to the perception of users, and a random component that represents the heterogeneity of preferences across individuals and additional factors unknown to the analyst.

C. G. Prato (2009) states that the major challenges in route choice modeling are the generation of a choice set of alternative routes and the estimation of discrete choice models.

The first challenge is to identify the actual choice set considered by passengers, which is unknown to the researcher. The number of alternative routes is usually large and increases in a combinatorial dimension in large-scale networks. Therefore, identifying realistic choice sets is not a trivial task. The second challenge is to represent the actual behavior of travelers while appropriately considering the correlation between alternatives that share links and stops, and to estimate the coefficients of the different attributes using real data. As mentioned in the introduction, most literature on PT passengers' route choice behavior has used the MNL model (Grison et al., 2017; Z. Guo, 2011; Jánošíková et al., 2014; Nassir et al., 2018; Raveau & Muñoz, 2014; Raveau et al., 2014, 2011; Vrtic & Axhausen, 2002; Kim et al., 2020). Only a handful of studies have actually considered similarity across routes via a Path Size logit (Anderson et al., 2017; Bovy & Hoogendoorn-Lanser, 2005; de Grange et al., 2012; Hoogendoorn-Lanser et al., 2005; O. A. Nielsen et al., 2021; Tan et al., 2015; Yap et al., 2020), possibly because of the complexity as well as the unimodality that reduces significantly the similarity across alternatives (especially for studies focusing only on rail or metro modes of PT).

2.2.1 Route Choice Strategy

The literature suggests that passengers can adopt different strategies to choose a route to reach a destination (Raveau & Muñoz, 2014; Spiess & Florian, 1989). As we will see, the strategies can be used in both probabilistic and deterministic approaches.

We propose to classify the strategies in two types: (i) a disaggregated strategy, where passengers choose a specific sequence of initial-transfer-final stops while considering specific lines between them (itineraries); and (ii) an aggregated strategy, where passengers choose a specific sequence of initial-transfer-final stops while aggregating the common lines (i.e., a set of lines that minimizes the total expected travel time) of each route section into a single alternative.

A graphical representation of a simple network that will help us to illustrate the two strategies is provided in Figure 2.1, where six transit lines operate to serve one origin-destination (OD) pair with one transfer stop (T). The transit lines are shown in different colors with their respective frequencies [buses/hour] and travel time [min].

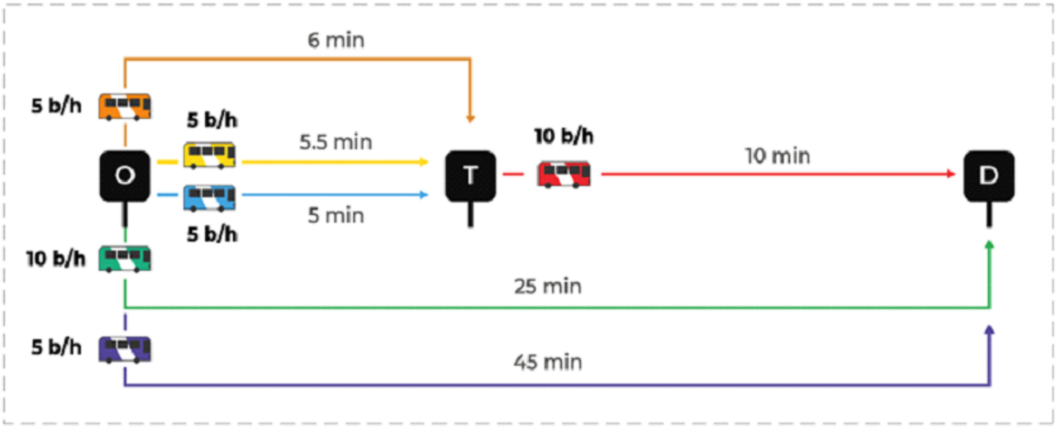


Figure 2.1: Example of the route choice problem

A disaggregated strategy in Figure 2.1 would be: "Take the blue line from stop O to stop T, transfer to the red line and alight at stop D". As can be seen, there are five itineraries or disaggregated alternatives for reaching destination D from origin O, which are described in Table 2.1. For each bus line, if the headway follows an exponential distribution, the waiting time is calculated as 60 minutes divided by the corresponding frequency. In-bus travel times are shown in Figure 2.1. The itinerary O-green line-D minimizes the total travel time; therefore, in a deterministic approach, passengers will choose this alternative. In contrast, in a stochastic approach each itinerary has a non-zero probability of being chosen because of non-observed considerations represented by the error term in the utility function.

Table 2.1: Route alternatives using a disaggregated strategy

Disaggregated alternative	**In-bus travel time [min]**	**Expected waiting time [min]**	**Total travel time [min]**
O-Orange line-T-Red line-D	16	18	34
O-Yellow line-T-Red line-D	15.5	18	33.5
O-Blue line-T-Red line-D	15	18	33
O-Green line-D	**25**	**6**	**31**
O-Purple line-D	45	12	57

An aggregated strategy for the problem in Figure 2.1 would be: "Take the next bus from either the orange, yellow, or blue lines from stop O to stop T, transfer to the red line and alight at stop D". As can be seen in Table 2.2, there are three aggregated alternatives in this network. The first route is O-orange/yellow/blue line-T-red line-D, where in the first stage there are three common lines, and passengers take the first line that arrives at stop O. The waiting time of the first stage is 60 minutes divided by the summation of the frequencies of the common lines (15 b/h); in the second stage, it is 60 minutes divided by the frequency of the red line (10 b/h), resulting in a total waiting time of 10 minutes. The in-bus travel time in the first stage is calculated as a weighted average of the line frequencies (5.5 min); in the second stage, it is 10 minutes, as shown in Figure 2.1. For the second and third alternatives, the waiting time and in-bus travel time are calculated in the same way as the disaggregated alternatives. It can be noticed that the alternatives that connect stops O and D, are not aggregated into a single alternative, because the green line and the purple are not common lines for this OD pair. The lines that do not belong to the common lines set are part of the consideration set, but they are aggregated into a different alternative to the common lines alternative.

As the first alternative has the lowest total travel time, in a deterministic approach all passengers will choose that option, and a proportion of 5/15 of passengers will be assigned to the orange line, 5/15 of passengers will be assigned to the yellow line, and 5/15 of passengers will be assigned to the blue line. In contrast, in a stochastic approach, each route has a non-zero probability of being chosen because of non-observed considerations represented by the error term in the utility function. Therefore, if the probability of choosing route 1 is p_1, the probability of taking each line (orange, yellow, or blue) is $p_1 * 5/15$.

Table 2.2: Route alternatives using an aggregated strategy

Aggregated alternative	**In-bus travel time [min]**	**Expected waiting time [min]**	**Total travel time [min]**
O-Orange/Yellow/Blue line-T-Red line-D	**15.5**	**10**	**25.5**
O-Green line-D	25	6	31
O-Purple line-D	45	12	57

In general, discrete route choice models in the literature have used a probabilistic disaggregated strategy approach. On the other hand, transit assignment models have used a deterministic aggregated strategy approach. To the best of our knowledge, only Raveau & Muñoz (2014) and Schmöcker et al. (2013) have considered aggregated alternatives within a probabilistic route choice model.

The hyperpath strategy can be seen as an extension of the aggregated strategy. Raveau & Muñoz (2014) found that only 4% of passengers followed the hyperpath strategy, evidencing a low usage frequency of this strategy. Consequently, the hyperpath strategy is not included in this work.

2.3 Data description and analysis

The analysis for this study was carried out using the Santiago, Chile multimodal PT network, known as Transantiago. Santiago is the capital of Chile and has a population of over 7 million inhabitants. The PT system serves roughly 50% of motorized trips and it is operated by headway scheduling; therefore, lines do not have fixed-time schedules. The fare system is almost fully integrated, with a flat fare between urban buses, Metro, and some rail services, allowing up to three trip legs within a two-hour time window. In a typical week, 3 million passengers use the system to make 25.5 million trips. The network includes 7 metro lines, more than 300 bi-directional transit lines, and one rail service. In this work, we use observations from frequent passengers that travel during morning peak periods (6:30-8:30 AM) on weekdays. Specifically, we select users that travel 15 days or more on the PT system during May 2018, and that stay in the destination locations for at least two hours. In this way, we try to capture trips to regular activities such as work or study. The morning peak period is the most congested, more than 700 thousand trips per day can be observed, and hence an interesting travel period from a behavioral and planning perspective.

Very detailed demand information was obtained from the automatic fare collection (AFC) system in Transantiago (Gschwender et al., 2016). The SC bip! is the only accepted payment method. Passengers must validate when boarding a bus or entering a Metro station. No alighting validation is required for bus or Metro trips. Around 27% of passengers evade fare payment on buses. Bus stops with particularly high demand have an off-vehicle payment system called zona paga (payment zone), where passengers validate when they enter the bus stop area and then board any bus without further validation. The data is already processed to estimate the boarding and alighting position for all validations and the trips (stages) associated with an origin-destination journey using the Munizaga & Palma (2012) methodology.

Additionally, in the Santiago PT system, all buses are equipped with GPS devices that record the time and position every 30 seconds and we have used those data (AVL data) to obtain the observed frequency on transit lines.

2.3.1 Common lines analysis

To determine if passengers are using common lines, we identified, for each OD pair, all route sections of the routes registered in the trips database. For those route sections, we identified the list of transit lines passing through and then applied the greedy heuristic (Chriqui & Robillard, 1975) to identify the set of common lines for each route section.

The optimization problem to find common lines is shown in Equation 2.1, where the first term of this equation corresponds to the waiting time and the second term corresponds to the average of the in-vehicle travel times. Spiess & Florian (1989) explain that the waiting time parameter can take a value of 1, when assuming an exponential distribution of interarrival times of vehicles, or a value of 0.5 when assuming constant headways of vehicles. In this study, we use the waiting time parameter equal to 1 because previous studies have found irregular headways in the vehicles of the PT system of Santiago (Arriagada et al., 2019; Godachevich & Tirachini, 2021), and because the exponential distribution has been found to fit real waiting time data well (S. Guo et al., 2011). Specifically, in Equation 2.1, it is assumed that buses arrive following a Poisson distribution with arrival rates being the sum of the observed frequency of lines, f_l being the observed frequency of line l, t_l being the in-vehicle travel time of line l, L being the number of lines serving the evaluated route section, and x_l taking value 1 when line l belongs to the common lines set and taking value 0 in other cases. Therefore, the optimal strategy for a user would be to take the first bus from this set of lines that arrives at the stop. In summary, the process to identify the common line set considers an evaluation of waiting time and in-vehicle travel time to classify a line as attractive (belongs to the common line set) or non-attractive. Therefore, common lines may or may not overlap in their trajectory, there are cases without overlap, with partial overlap, or with 100% overlap between lines that belong to the common line set.

$$\min \frac{1}{\sum_{l \in L} f_l} + \sum_{l \in L} \frac{f_l t_l x_l}{\sum_{j \in L} f_j x_j} \tag{2.1}$$

Following this approach, we found 45,089 route sections with at least one trip. Among these, we analyzed 30,385 pairs, avoiding those with extra-vehicle payment systems (zona paga) and Metro stations, where the specific trajectory used is unknown because of the way in which it is inferred from the SC data (Munizaga & Palma, 2012). For 95.8% of the route sections, all transit lines passing through were common lines and concentrated 468,275 of the observed trips. In the remaining 4.2% of route sections, which included 44,612 trips, some lines did not belong to the common lines set; in those cases, 79.9% of the trips were made using a line from the common lines set and 20.1% took buses from lines that did not belong to the common lines set.

There are many reasons why a user would not follow the common lines strategy and end up considering instead a seemingly "slower" line for a trip. This could result from the omission of variables that are relevant for the user, such as (i) level of crowding, (ii) seat availability,

(iii) the type of bus, as some lines may have newer buses or buses with air-conditioned or more comfortable seats, (iv) the bus operator, as some of them are more reliable. It could also result from behavioral conditions, such as (v) that the waiting time disutility can be higher than travel time disutility, (vi) or group behavior when passengers consider the utility of another passenger as part of the decision process. Any of these, or a combination of them, can make users prefer a slower line. We use the traditional version of the common lines approach that considers the summation of waiting and travel time, not assigning different weights to them. Possible extensions are left for further research, including the way in which the common lines set is built, that could consider a different valuation of travel time and waiting time, or incorporate other attributes.

As an illustrative example, we analyzed two route sections in the system in greater detail: one with common lines only (Figure 2.5), and another with two non-common lines (Figure 2.3). We compared the observed trip distributions with the expected trip distributions. For the expected trip distribution we assume that (i) passengers would board the first line belonging to the common lines set that arrived at the bus stop, and (ii) they would not board the lines that did not belong to the common lines set. The expected number of trips in a pair of stops or in a route section r along transit line l is expressed in Equation 2.2, where OT_r is the observed total number of trips in route section r and P_{lr} is the probability of boarding line l to travel along the route section r. P_{lr} is shown in Equation 2.3, where L_r is the number of common lines serving route section r.

$$ET_{lr} = P_{lr}OT_r \tag{2.2}$$

$$P_{lr} = \frac{f_l}{\sum_{i \in L_r} f_i} \tag{2.3}$$

In the first route section, shown in Figure 2.5, the origin bus stop is connected to a Metro station by three common lines (348, H08, and H03) with in-vehicle travel times between 11.7 and 12.8 minutes and waiting times between 13 and 14.2 minutes. Table 2.3 shows the observed distribution and the expected distribution of trips among the three common lines. The expected distribution according to the common lines theory is not significantly different from the observed distribution (Pearson's chi-squared test), so in this example, we can conclude that passengers took lines according to the common lines principle. An implication of this finding is the possibility that passengers consider common lines as part of the same alternative since they take the first line that arrives at the stop.

In the case of the route section shown in Figure 2.3 that has two lines, the origin bus stop is the first stop for two transit lines and is located in a suburban area. The only options to reach a Metro station are the local bus line (B18) or the express bus line (B18e). The optimal strategy is to wait for the express line (B18e), therefore line B18e belongs to the common lines set while B18 does not. According to the common lines model, all users should take the express line. However, Table 2.3 shows the observed distribution of trips, where 95.2% of trips are made with the express line and 4.8% are made with the local line. The 4.8% of trips made with the slower travel option were made by one passenger, who traveled in the analyzed OD pair 15 times, taking the local route 9 times and the express route 6 times. A possible explanation for this seemingly odd behavior might be any of the reasons explained above. This simple example suggests that lines that do not belong to the common lines set according to the model can still be attractive to some passengers.

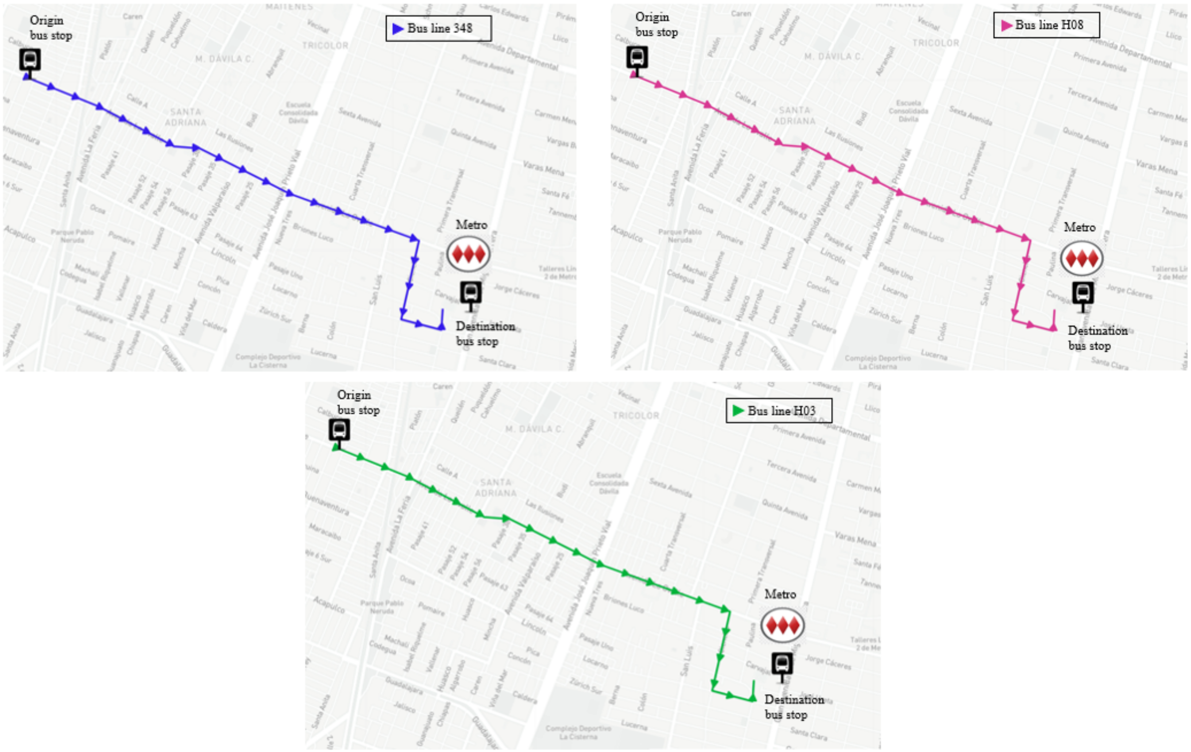


Figure 2.2: Trajectory of the three common lines (348, H08 and H03)

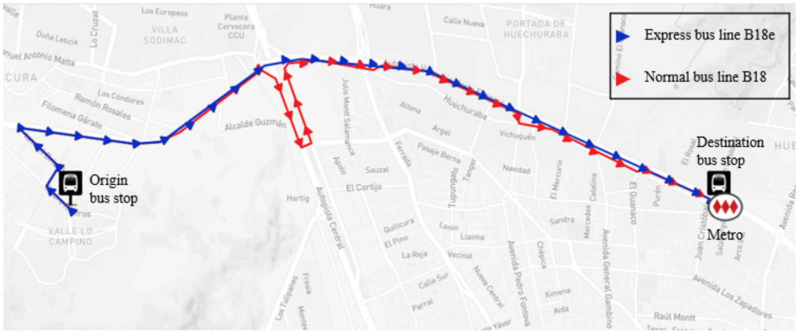


Figure 2.3: Trajectory of the two non-common lines (B18e, express bus line and B18, normal bus line)

Table 2.3: Statistics of lines in specific cases of common lines and non-common lines

Case	Line	Travel time	Expected waiting time	Observed trips	Expected trips
Common lines	348	13.3 min	12.8 min	40 (26.0%)	49 (31.8%)
	H08	13 min	12.7 min	47 (30.5%)	49.4 (32.1%)
	H03	14.2 min	11.3 min	67 (43.5%)	55.6 (36.1%)
Non-common lines	B18	35.8 min	12.2 min	9 (4.8%)	0 (0%)
	B18e	17.5 min	7.8 min	179 (95.2%)	188 (100%)

In order to capture the behavior of passengers regarding the use of common lines within the database as a whole, we propose a disaggregated line usage indicator q_n. This indicator is based on the difference between the expected number of trips (proportional to the observed frequency of the common lines and equal to zero for transit lines not belonging to the common lines set) and the observed trips in a route section.

The disaggregated line usage indicator is shown in Equation 2.4, which has the purpose to identify if the number of trips made by user n follows the common lines principle. The first level of the equation sums overall route sections visited by the user and the second level sums over each common line belonging to the route section to compare the observed trips and expected trips in each common line. In Equation 2.4, R_n is the number of the route section visited by passenger n, L_r is the number of common lines serving route section r, and π_{lrn} is defined in Equation 2.5. The elements in Equation 2.5 are: the expected number of trips for user n on line l along the route section r (ET_{lrn}) and the observed number of trips for user n on line l along route section r (OT_{lrn}). The disaggregated line usage indicator takes a value of 1 when all trips are concentrated in one of the common lines, or when they are made in some of the non-common lines. It takes a value of 0 when the number of expected trips matches the number of observed trips. The number of observed trips is obtained from SC data, and the number of expected trips is calculated with Equation 2.3.

$$q_n = \sum_{r=1}^{R_n} \sum_{l=1}^{L_r} \frac{\pi_{lrn}}{\sum_{r=1}^{R_n} L_r} \tag{2.4}$$

$$\pi_{lrn} = \frac{|ET_{lrn} - OT_{lrn}|}{max(ET_{lrn}, OT_{lrn})} \tag{2.5}$$

Using real data, we calculated q_n for the users that traveled in route sections that do not begin in a zona paga or in a Metro station. We analyzed 28,896 cards (each corresponding to an individual user) and, as can be seen in Figure 2.4, 20.9% of them showed a value of q_n equal to or less than 0.1, which means that their observed behavior is consistent with boarding the first bus of the common lines that arrive at the stop. Around 5% of them showed a value of 1.0, which means that they take one specific bus line, even when there are other common lines available. In summary, the analysis q_n shows that some users use common lines, others use a subset of common lines, and others do not use common lines.

With this analysis, we have provided empirical evidence that passengers show heterogeneity in their route choice behavior. This finding has significant implications for route choice models, since these should incorporate heterogeneity across passengers in the use of alternative choice strategies.

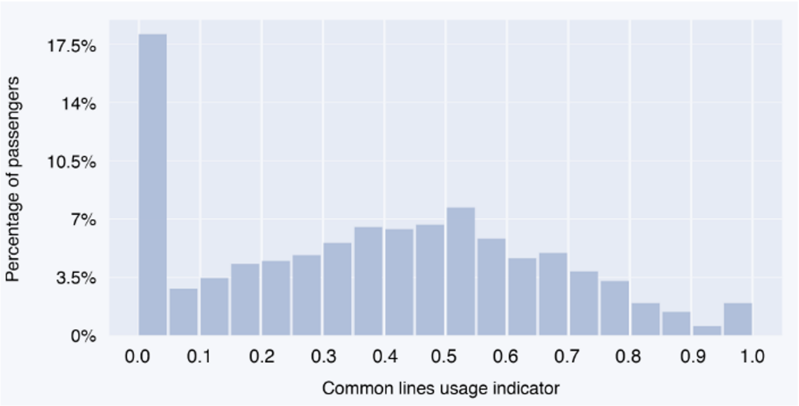


Figure 2.4: Distribution of disaggregated line usage indicator

2.4 Assessment of discrete choice modeling approaches

In this section, we study the research hypothesis by developing random utility maximization (RUM) models that incorporate disaggregated and aggregated route choice strategies.

2.4.1 Consideration set construction

The first challenge in the formulation and estimation of a PT route discrete choice model from SC data corresponds to the construction of the consideration set. For the consideration sets, we first defined the origin and destination locations, which we considered as the areas within a 100-meter radius of the origin and destination stops of the trip. It is important to note that a sensitivity analysis was performed to define the radius of the origin and destination zones and we found that 100-meter radius zones generated the most consistent results in the models. In particular, larger radius zones generated unexpected results in the vehicle travel time coefficient. For each OD pair, we constructed the consideration set by combining all available and observed routes used by every passenger traveling within the defined origin and destination areas in the study period (6:30 to 8:30). It means that for each OD pair the same consideration set is used for all passengers.

The consideration set was built depending on the behavioral assumption of the respective model. As was explained before, we considered two possible approaches: (i) an aggregated strategy that implies aggregating the common lines set in a single alternative and taking the first bus that arrives at a stop and (ii) a disaggregated strategy that implies choosing only one line in each route section.

Figure 2.1 illustrates an hypothetical OD pair, Table 2.1 shows the consideration set that would be built under a disaggregated strategy, and Table 2.2 shows the consideration set that would be built under an aggregated strategy. It is shown that if we construct the consideration set with aggregated alternatives, there are three possible routes. For route section O-T the orange line, yellow line and grey line belong to the common lines set, so they are part of the same alternative, and therefore there is only one alternative in the sequence O-T-D. For segment O-D there are two non-common lines, and therefore two alternatives. If we construct the consideration set with disaggregated alternatives, there are five possible routes in the consideration set.

2.4.2 Route choice modeling

As mentioned in the introduction, most route choice models assume that one strategy applies to all passengers within the network. However, this is not consistent with what we observe in the data, since the analysis using the disaggregated line usage indicator, in section 2.3.1, indicates passengers present different route choice strategies, some of them consider common lines and take the first incoming bus and others prefer to wait for a specific transit line. Therefore, our suggestion is to consider heterogeneity in the strategies used by passengers within route choice models. A latent class model (Walker, 2001) is appropriate for this route choice model because it incorporates different underlying behavioral rules. In our case, these behavioral rules dictate different ways of identifying the choice set.

The latent class model has two components: a class membership model and a class choice model. In our context, there are two classes: passengers who consider common lines, and therefore use an aggregated strategy (class 1), and passengers who do not consider common lines, and therefore use a disaggregated strategy (class 2). The membership probability is modeled using a binary logit model, where the deterministic utility function v_n is interpreted as the propensity of passenger n to belong to the disaggregated strategy class. The propensity of the aggregated strategy class is fixed to zero for normalization purposes. The representation of the propensity function depends on a constant and on the dispersion γ_n of transit lines used by the passenger, as shown in Equation 2.6. The dispersion can be measured with indicators based on the chi-square test, the Kolmogorov-Smirnov test, the Gini coefficient, and the statistical inference algorithm proposed by Nassir et al. (2017). In this study, we use the Gini coefficient (Glasser, 1962), which varies between 0 and 1. The value 0 means that the passenger made the same number of trips on each transit line that belongs to the common lines set, and the value 1 means that the passenger made all his/her trips on one line. Passengers with a Gini coefficient equal to zero have a higher probability of using common lines as a route choice strategy. The Gini coefficient is shown in Equation 2.7, where R_n is the number of route sections visited by passenger n, L_r is the number of common lines serving route section r, OT_{lrn} is the observed number of trips for user n on line l along route section r, and μ_{rn} is defined in Equation 2.8. The probability $\Lambda_n(\overline{CL})$ of belonging to class 2 is shown in Equation 2.9, and the probability that passenger n selects alternative i is presented in Equation 2.12, which is expressed using the total probability theorem.

$$v_n = \beta_{\overline{CL}} + \beta_\gamma \gamma_n \tag{2.6}$$

$$\gamma_n = \frac{\sum_{r=1}^{R_n} \frac{1}{2\mu_{rn} L_r (L_r - 1)} \sum_{l=1}^{L_r} \sum_{j=1}^{L_r} |OT_{lrn} - OT_{jrn}|}{R_n} \tag{2.7}$$

$$\mu_{rn} = \frac{1}{L_r}\sum_{l=1}^{L_r} OT_{lrn} \tag{2.8}$$

$$\Lambda_n(\overline{CL}) = \frac{exp(v_n)}{exp(v_n) + 1)} \tag{2.9}$$

Since we do not know if a passenger uses common lines due to the latent nature of the behavior (i.e., it is not observed), we try to approximate this behavior through the indicator γ_n, which cannot be put directly into the systematic utility function because it would result in endogeneity due to the correlation between γ_n and the error term (Guevara, 2015). This problem can be solved with the latent class approach (Guevara, 2015). However, it requires formulating a structural equation and resolving the problem through maximum simulated likelihood, which is a memory-intensive process and requires more time than any other part of the algorithm that is already computationally very expensive. To avoid this problem, we address the endogeneity problem using the Multiple Indicator Solution (MIS) method that was adapted to discrete choice models by Guevara & Polanco (2016).

The MIS method is proposed as an accessible way to address endogeneity when it is difficult to gather traditional instruments, but two (or more) indicators of the omitted variable causing the endogeneity are available. In that case, it can be shown (see Guevara & Polanco (2016)) that the problem will be solved if the researcher adds one of the indicators as an auxiliary variable of the model and uses the other indicators as instruments for applying the Control Function method. The intuition behind the MIS method is that by including the first indicator as an auxiliary variable one eliminates the endogeneity produced by the omitted attribute, but in turn, causes a different type of endogeneity resulting from the use of an improper proxy. In this modified model, the only remaining source of endogeneity resides in a measurement error for the included indicator, which can be solved using the other indicators as instruments in a Control Function application.

The MIS method is applied using the disaggregated line usage indicator q_n (Equation 2.4) as the instrument, and the Gini coefficient as the indicator, in two stages. The first comprises an OLS regression of the Gini coefficient on its instrument, the disaggregated line usage indicator q_n (Equation 2.10). The second stage consists of adding δ_n, the residual of Equation 2.10, to the propensity equation to correct for endogeneity, as shown in Equation 2.11. Finally, the choice model is estimated using the choice probability shown in Equation 2.12 but using the systematic propensity v_n shown in Equation 2.11.

$$\gamma_n = \beta_{intercept} + \beta_q q_n - - > \delta_n \tag{2.10}$$

$$v_n = \beta_{\overline{CL}} + \beta_\gamma \gamma_n + \beta_\delta \delta_n \tag{2.11}$$

$$P_n(i) = P(i|CL)(1 - \Lambda_n(\overline{CL})) + P(i|\overline{CL})\Lambda_n(\overline{CL}) \tag{2.12}$$

Given the consideration set C_n for passenger n, we adopt a RUM framework, where we associate a utility for each route alternative and choose the route with the highest utility value. The utility has two components: the random component, which is assumed to be Gumbel distributed; and the deterministic component, which is specified in Equation 3.1.

$$V_i = \sum_m \beta_{tt_m} tt_{mi} + \sum_c \beta_{t_c} t_{ci} + \sum_s \beta_{tr_s} tr_{si} + \beta_{PSC} PSC_i + \beta_{CR} CR_i BUS_i \tag{2.13}$$

In this equation, i represents the alternative route, m is the PT mode ($m \in bus, metro$), tt_{mi} is on board travel time, it is included in the sum over all modes of the alternative i, t_{ci} is other time components ($c \in \{$initial waiting time (IWT), transfer waiting time (TWT), and transfer walking time(TWaT)$\}$, it is included in the sum over all components times of the alternative i, tr_{si} represents the penalties for different transfer types ($s \in \{$bus to bus (BB), Metro to bus (MB), bus to Metro (BM)$\}$, it is included in the sum over all transfer types of the alternative i, BUS_i is a binary variable that takes a value of 1 when the alternative contains the bus mode, and CR_i is the level of crowding at line-stops that are used within the alternative. PSC_i is the path size correction term to capture the correlation due to overlapping between alternative routes.

The level of crowding of a line in a route section is calculated as the number of passengers inside of buses divided by the capacity of the buses. The level of crowding in an alternative route is the average of the crowding of each route section that belongs to that alternative.

Path size correction introduces a negative factor (PSC_i) that decreases the deterministic utility of alternative routes that have correlation with other routes. We have used the expression according to Bovy et al. (2008) in Equation 2.14, where L_r is the length of the route section r, L_i is the length of route i, ζ_i is the set of route sections belonging to route i, and δ_{rk} is the route section-route incidence number, which takes a value of 1 if route k uses route section r and a value of 0 otherwise.

$$PSC_i = \sum_{r \in \zeta_i} \frac{L_r}{L_i} \ln \frac{1}{\sum_{k \in C_p} \delta_{rk}} \tag{2.14}$$

The waiting time included in t_{ci} in Equation 3.1 is obtained assuming an exponential distribution as one divided by the observed frequency of the line (or sum of frequencies, in the case of common lines). The exponential distribution (or the gamma, which is the sum of exponentials) has been widely used for modeling waiting time in transportation systems (e.g. Nguyen & Pallottino, 1988; Raveau & Muñoz, 2014; Schmöcker et al., 2013). In aggregated alternatives, the time components and level of crowding per route section are a weighted average (based on frequency) of the values corresponding to the transit lines that conform each route section.

Because of the Gumbel distribution of the error term, the probability of passenger n choosing itinerary i given consideration set C is expressed as a PSL model (Bovy et al., 2008). Equation 2.15 presents the probability of choosing itinerary i in the case of a passenger who does not use common lines and Equation 2.16 presents the probability of choosing itinerary i

in the case of a passenger who uses common lines, where $V_{(j(i))}$ is the systematic utility of the aggregate alternative route j which contains itinerary i, ζ_i is the set of route sections that belong to alternative route i, L_r is the set of common lines that belong to route section r, and $f_{(l(i),r)}$ is the observed frequency of transit line l which passes through route section r and is part of itinerary i.

In the last case, it is necessary to multiply the probability of choosing the aggregated alternative with the probability that the line, in each route section, arrives first at the stop. As an example, if we want to know the probability of choosing itinerary alternative O-orange line-T-red line-D in Figure 2.1, and the passenger uses common lines, we should multiply the probability of choosing aggregated alternative O-orange/yellow/blue line-T-red line-D by the probability that the orange line arrives first at stop O, which is proportional to its observed frequency.

$$P(i|\overline{CL}) = \frac{\exp V_i}{\sum_{k \in C} \exp V_k} \tag{2.15}$$

$$P(i|CL) = \frac{\exp V_{(j(i))}}{\sum_{k \in C} \exp V_k} \prod_{r \in \zeta_{j(i)}} \frac{f_{(l(i),r)}}{\sum_{k \in L_r} f_{k,r}} \tag{2.16}$$

The closed-form logit formula of the PSL allows for a simple estimation of the fixed coefficients by maximizing the likelihood function.

2.4.3 Consideration sets results

The alternative routes were generated using historical choices for each OD pair. As explained in previous sections, we worked with two types of alternatives: disaggregated alternatives were used to construct consideration sets without common lines, and aggregated alternatives were used to construct choice sets with common lines. We worked with OD pairs that featured between 2 and 10 disaggregate alternatives; OD pairs with fewer or more disaggregate alternatives were discarded. Using this filter, we obtained 20,871 OD pairs with an average of 3.45 available alternatives within the consideration set with disaggregated options, and 1.38 available alternatives with aggregated options.

The overlap between alternative routes was evaluated at the stop level, which means that the alternatives that used the same route sections and transport mode were considered as correlated elements. It is important to note that for both, disaggregated and aggregated alternatives, the Path Size correction term was obtained with the same procedure, described in equation 2.14. In order to compare the correlation of each link with other alternative routes, we derived transit line trajectories from GTFS data (for Metro, we did not know the trajectory used by passengers, so we assumed that passengers took the route with the lowest travel time). The Path Size correction term (PSC_i) takes a value of 0 when there is no overlap between the evaluated alternative and others, and the value decreases as the level of route correlation increases; that is, the smaller the PSC_i value, the higher the correlation of route i. Choice sets without common lines have a mean PSC_i of -0.85 and choice sets with common lines have a mean PSC_i of -0.16. The lower value (more negative) in PSC_i for disaggregated alternatives is expected because the conformation of alternatives with common lines captures a part of correlation which is considered by PSC_i in disaggregated alternatives.

2.4.4 Model estimates

PSL models and latent class models were estimated with a sample of 150,430 observations, which corresponds to trips made during 15 workdays during the morning peak period (between 6:30 and 8:30 AM). The PSL was estimated in two cases, with aggregated and disaggregated alternatives routes. The latent class models were estimated with the MIS approach correcting for endogeneity and the two types of choice sets.

The specification of the deterministic utility function considers in-vehicle travel time for bus and Metro, waiting time at the beginning of the trip and during transfers, walking time during transfers, transfer penalties for bus to bus, bus to Metro, and Metro to bus, and the level of bus crowding. Metro to Metro transfers cannot be incorporated into the model because the information about the route that the passenger used inside the Metro network is not available. Table 2.4 shows the minimum, mean, and maximum values for each attribute considered in the models. Aggregated alternatives present lower values for waiting time because common lines are considered as part of the same alternative.

Table 2.4: Statistics of attributes used in PSL and latent class models

Attribute	**Disaggregated alternatives**			**Aggregated alternatives**		
	Min	**Mean**	**Max**	**Min**	**Mean**	**Max**
Travel time in bus	0	16.55	89.53	0	16.55	90.26
Travel time in metro	0	17.68	72.73	0	17.68	72.73
Initial waiting time	1.7	8.33	17.47	0.81	4.18	17.46
Transfer waiting time	0	2.90	17.5	0	2.38	17.40
Transfer walking time	0	0.78	9.79	0	0.82	9.79
Transfer bus-bus	0	0.08	2	0	0.08	2
Transfer Metro-bus	0	0.05	1	0	0.05	1
Transfer bus-Metro	0	0.62	1	0	0.62	1
Total transfer	0	0.75	2	0	0.75	2
Bus crowding level	0	0.21	1.2	0	0.21	0.95

Table 2.5 shows the estimated parameters and rates of substitution for PSL models using the consideration set with disaggregated and aggregated alternatives. The parameters are statistically significant and with the expected sign, except for the term PSC, which is not statistically significant in the aggregated alternatives model and is statistically significant but with the unexpected sign in the disaggregated alternatives model. Given that PSC belongs to the interval $(-\infty, 0]$ and implies a reduction in the systematic utility of correlated routes, we expect a positive sign in the coefficient. The rates of substitution suggest that transfers represent between 11.5 and 16 minutes of Travel Time in Bus (TTB), a 10% increase in crowding level represents 1.3 minutes of TTB, travel time in Metro is more burdensome than travel time in bus, and initial waiting time, transfer waiting time and transfer walking time are more burdensome than travel time in bus.

Table 2.5: PSL estimates (t tests) and rates of substitution in case of disaggregate alternatives and aggregate alternatives

Parameters	**Model estimates**		**Rates of substitution**	
	Disaggregated alternatives	**Aggregated alternatives**	**Disaggregated alternatives**	**Aggregated alternatives**
Travel time in bus	-0.063 (-43.8)	-0.062 (-37.2)	1	1
Travel time in metro	-0.067 (-34.5)	-0.070 (-32.6)	1.06 (3.3)	1.13 (6.4)
Initial waiting time	-0.098 (-82.0)	-0.139 (-57.8)	1.56 (13.9)	2.24 (16.4)
Transfer waiting time	-0.082 (-24.2)	-0.080 (-11.3)	1.30 (5.0)	1.29 (2.4)
Transfer walking time	-0.309 (-44.6)	-0.091 (-9.3)	4.91 (27.0)	1.47 (2.9)
Transfer bus-bus	-1.017 (-19.3)	-0.727 (-12.1)	16.14 (16.2)	11.73 (10.3)
Transfer bus-Metro	-0.732 (-18.6)	-0.673 (-15.0)	11.62 (16.6)	10.84 (13.4)
Transfer Metro-bus	-0.948 (-6.7)	-1.000 (-6.8)	15.05 (6.2)	16.13 (6.3)
Bus crowding level	-0.807 (-17.6)	-0.711 (-7.7)	12.81 (15.0)	11.47 (7.1)
PSC	-0.132 (-4.6)	0.071 (1.3)		
N° observations	150,430	150,430		
Log-likelihood	-157050.3	-35978		
Adjusted rho-square	0.044	0.083		

All columns show t-values between parentheses. t-tests are against zero for the model estimates and against one for the rates of substitution. For the rate of substitution, we followed the procedure proposed by Daly et al. (2012).

We estimated a simple latent class model using Equation 2.6 for the propensity of the disaggregated strategy, which adds the Gini coefficient to the utility specification. The results of this model showed unexpected results, such as a positive sign for the walking transfer time coefficient and transfer bus to bus coefficient in the class aggregate alternatives. These biased parameters can be explained because of the endogeneity, given that the Gini coefficient is correlated with the error term of the propensity function. For this reason, we applied the MIS method including the Gini coefficient in the membership model and disaggregated line usage indicator as the instrument, as described in Equation 2.11. This model, shown in Table 2.6, is better than the endogenous model, since the coefficient of walking transfer time variable and transfer bus to bus variable have a negative sign. The fit of the MIS model is slightly better than that of the simple latent class model. This allows us to conclude that the combination of the Gini coefficient with the disaggregated line usage indicator provides more information than the latent class model without any indicator or with the Gini coefficient alone.

We tested the application of the MIS method including the disaggregated line usage indicator in the membership model and the Gini coefficient as the instrument; however, it obtained a non-statistically significant coefficient for the disaggregated line usage indicator, which can be explained because the Gini indicator presents poorer performance than the disaggregated line usage indicator (Guevara et al., 2020).

The rates of substitution in the MIS latent Class model (Table 2.6) confirm some findings within the PSL models, but there are also important differences between them. In contrast to the PSL models, we found the following results in the MIS latent Class model: both the initial waiting time and the transfer waiting time are perceived as less unpleasant than TTB for the disaggregated alternatives class, increasing the difference in perception of waiting time between both classes of passengers, given that individuals of the aggregated class find initial and transfer waiting times more unpleasant than TTB; bus-to-bus transfers are more unpleasant for passengers belonging to the aggregated alternative class; and bus-to-Metro transfers are more unpleasant for passengers belonging to the disaggregated alternative class.

The most unpleasant transfer for the disaggregated strategy class is Metro-to-bus, which represents 15.9 minutes of TTB, while for the aggregated strategy class the most unpleasant is bus-to-bus, representing 18.9 minutes of TTB. With respect to the bus crowding level variable, a 10% level of bus occupation represents 1.2 minutes of TTB for the class without common lines, and 0.9 minutes for the common lines class. On the other hand, waiting time at the beginning of the trip and during transfers is more burdensome for passengers who use aggregated alternatives, which is to be expected, as people consider common lines with the purpose of decreasing waiting time. Bus crowding levels and walking during transfers are more burdensome for passengers who use disaggregated alternatives; therefore, we can suggest that these passengers may not be using all common lines because they try to avoid walking and crowding situations. The membership class model parameters have the expected signs and are statistically significant, showing the heterogeneous route choice strategies of passengers.

The adjusted rho-square of the models varies between 0.03 and 0.08. It should be reminded that rho-square in discrete choice models does not have a direct interpretation as explanatory power. It has been shown, e.g., by Bohara et al. (2007), it is possible to accept a model with a rho-square value of around 0.05. To confirm the validity of our models we performed a likelihood ratio test that rejected the null hypothesis that the coefficients are equal to 0, which shows that there is a gain in estimating the decisions of passengers using the proposed models.

Table 2.6: Multiple Indicators Solution latent class estimates (t tests) and rates of substitution in case of disaggregate alternatives and aggregate alternatives

	Model estimates		**Rates of substitution**	
Parameters	**Disaggregated alternatives**	**Aggregated alternatives**	**Disaggregated alternatives**	**Aggregated alternatives**
Travel time in bus	-0.085 (-42.7)	-0.05 (-19.1)	1	1
Travel time in Metro	-0.081 (-30.0)	-0.056 (-16.9)	0.95 (-2.1)	1.1 (3.4)
Initial waiting time	-0.041 (-24.6)	-0.099 (-29.9)	0.5 (-22.6)	2 (7.7)
Transfer waiting time	-0.043 (-9.0)	-0.058 (-6.3)	0.5 (-8.6)	1.2 (0.8)
Transfer walking time	-0.396 (-39.6)	-0.08 (-5.8)	4.7 (25.1)	1.6 (2.1)
Transfer bus-bus	-0.974 (-12.0)	-0.944 (-11.5)	11.5 (10.3)	18.9 (9.1)
Transfer bus-Metro	-0.938 (-14.0)	-0.722 (-12.1)	11 (12.6)	14.4 (10.4)
Transfer Metro-bus	-1.349 (-6.8)	-0.85 (-3.4)	15.9 (6.3)	17 (3.2)
Bus crowding level	-1.016 (-15.8)	-0.472 (-3.9)	12 (14.0)	9.4 (3.5)
PSC	-0.339 (-7.5)	0.189 (2.6)		
Membership class				
Constant disaggregated strategy	-19.25 (-7.4)			
Gini indicator	50.22 (7.5)			
Residual	-108.3 (-7.4)			
N° observations	150,418			
Log-likelihood	-155825.7			
Adjusted rho-square	0.031			

All columns show t-values between parentheses. t-tests are against zero for the model estimates and against one for the rates of substitution. For the rate of substitution, we followed the procedure proposed by Daly et al. (2012).

2.5 Discussion

This section is devoted to the analysis of PS and MIS latent class models parameters with the purpose of understanding route choice behavior in a multimodal large-scale network.

2.5.1 Travel time

The negative coefficients of travel time variables show a disutility of travel time for passengers. In the models, we estimate travel time by mode to understand the sensitivity of travel time for individuals per mode. The results in both path size logit models indicate that passengers find travel time in Metro to be more burdensome. This is not surprising, as in Santiago, the Metro offers high frequency and reliable service (stable headways) but the crowding levels inside the cars are significant, reaching up to seven passengers per square meter during peak hours, making even short trips uncomfortable. The opposite was found by Eluru et al. (2012) in Montreal, and by Anderson et al. (2017) in Copenhagen, where the unobserved factors are different. In those cases, the Metro network, besides having a high frequency and stable operations, provides protection from extreme weather conditions, especially during winter. Therefore, users prefer to travel underground.

Furthermore, the latent class models indicate that there are differences between the two classes of passengers in terms of how they perceive travel time in bus and Metro. Passengers who choose disaggregated alternatives find travel time in bus slightly more unpleasant than travel time in a Metro, while passengers who choose aggregated alternatives find travel time in Metro slightly more unpleasant than travel time in a bus. A possible explanation for this difference in the effect of travel time is related to the physical discomfort in Metro. There is a need for future research to examine this difference in detail.

2.5.2 Waiting time

The models indicate that alternative routes with lower waiting times are preferred. In the models we estimated sensitivity to waiting time at the beginning of the trip as well as at any transfer stages. The results in all models (except the disaggregated alternatives in the MIS latent class model) indicate that passengers find waiting time at the beginning of the trip to be more onerous than waiting time during transfers. Few studies have reported this difference in the perceived waiting time, which can be explained because individuals feel particularly anxious to start the trip, while during transfers, this anxiety decreases since the trip has already began.

Differences in waiting time perception are shown in the MIS latent class model, where passengers that choose aggregated alternatives find the waiting time more onerous than passengers who choose itineraries. This is expected, as passengers use common lines to decrease their waiting time. On the other hand, surprisingly, the initial waiting time is found to be less onerous than the in-vehicle time for disaggregated strategy class. It might be the case that users who care less about waiting time, are more willing to wait for a specific bus line of their preference (disaggregated strategy), which reflects in a smaller coefficient for waiting time than in-vehicle travel time. Likewise, it might be the case that users who use disaggregated alternatives are keener to use mobile transport applications that provide bus arrival time that allow them to further reduce their waiting time because they can arrive at the bus stop just in time to take the bus. As the model calculates the waiting time as a function of the observed frequency, yielding the same value for all users, the waiting time for the users who use mobile transport applications might be overestimated. If they are concentrated in disaggregate alternatives class, this may bias our results for that class, reflecting in a smaller coefficient for waiting time than in-vehicle travel time.

2.5.3 Walking time

Routes with lower walking time during transfers are preferred. All models show that walking time is more onerous than travel time in bus, and the MIS latent class model indicates that individuals from different classes show a substantial difference in the effect of the walking time variable. For passengers who choose itineraries, walking time is more unpleasant than it is to passengers in the other class, which can explain why those passengers prefer a specific itinerary instead of all lines that belong to the common lines set.

2.5.4 Transfer

The alternatives used in this work involve transfers. When there is a potential waiting time and walking time for each transfer, we incorporated the influence of transfers on alternative routes in multiple ways: transfers per mode, transfer waiting time, and transfer walking time. As expected, alternatives with fewer transfers, lower waiting times, and lower walking time per transfer are preferred. As can be seen in the MIS latent class models, both types of passengers, from the disaggregated and aggregated strategy class, prefer the bus-metro transfer to other types of transfers (bus-bus, metro-bus). This can be explained because the metro is more regular than the buses, and then all passengers prefer to transfer to a more regular service. On the other hand, the most disliked type of transfer for passengers who choose disaggregated alternatives is Metro-to-bus; for passengers who use common lines, the most disliked type of transfer is bus-to-bus. The largest difference between classes occurs in the bus-to-bus transfer that jumps from 11.5 min in the disaggregated strategy class to 18.9 min for the aggregated strategy class. We speculate that this may occur because bus-to-bus transfers are prone to a larger uncertainty, which generates an extra difficulty to plan a trip when this type of transfer is considered as an alternative, especially in the case of passengers that plan a trip using aggregated strategy. This may result in a larger bus-to-bus transfer penalty for that class. This must be further analyzed with additional information.

2.5.5 Path Size Term

In the PSL model with disaggregated alternatives, as with the disaggregated alternatives class in the MIS latent class model, the results showed that the path size correlation has a negative sign, which is not expected. Nevertheless, Anderson et al. (2017) obtained the same result for the correlation term, explaining this phenomenon as the fact that passengers prefer more opportunities to reach their destination from their point of origin. In this study, the correlation term was positive, as expected, in the aggregated alternatives.

It should be noted that the negative value for the path size is not uncommon in the literature about PT traveler's route choice. In fact, although the path size is supposed to correct for overlapping routes by reducing the utilities of overlapping routes, negative estimates for path-size terms have been found (Anderson et al., 2017; de Grange et al., 2012), most likely because of the additional utility of travelers having more opportunities to reach their destination from their origin, or because travelers might value the availability of a large number of en-route alternative options over the uniqueness of the route (Anderson et al., 2017). A very recent study (O. A. Nielsen et al., 2021) has even found non-significance of the path size estimate once transfer related variables were inserted explicitly in the utility function, suggesting that the inclusion of those variables captures the similarity explicitly.

2.5.6 Membership model of the latent class model

Using the Bayesian estimator of class membership in the MIS latent class model, we determined that 51.2% of passengers belong to the disaggregated strategy class[1] and 48.8% of passengers belong to the aggregated strategy class. Figure 5 shows a histogram for the predicted probabilities of belonging to both latent classes. It can be seen that around 41% of individuals have a probability higher than 0.98 of belonging to disaggregated strategy class and other 38% have a probability higher than 0.98 of belonging to aggregated strategy class, which means that there is a similar proportion of passengers that wait for specifics transit

[1] This value is higher than the value reported for passengers in London. Kurauchi et al. (2014) found that only 17% of passengers in London use the same transit line every morning (considering 4 days of evaluation). However, if they consider common lines, the variability in the chosen routes decrease, indicating that much of the variation in the route choice is due to common lines that could be part of the users' strategy.

lines and that consider common lines as part of the same alternatives.

In order to understand if the distance of the trip generates an effect in the route choice strategy of passengers, we calculated the probability of class membership separately for those passengers who make short trips (10 km or less of Euclidean distance) and who make long trips (more than 10 km of Euclidean distance). The results of this analysis do not show evidence that trip distance influences the choice of the route choice strategy. In this line, further work needs to be done to establish whether some characteristics of the OD pairs affect the route choice strategy of passengers. One example is the level of overlapping between alternative routes, which can allow distinguishing users who consider complex aggregated strategies (common lines with a small level of overlapping) versus users who consider less complex aggregated strategies (common lines with a high level of overlapping).

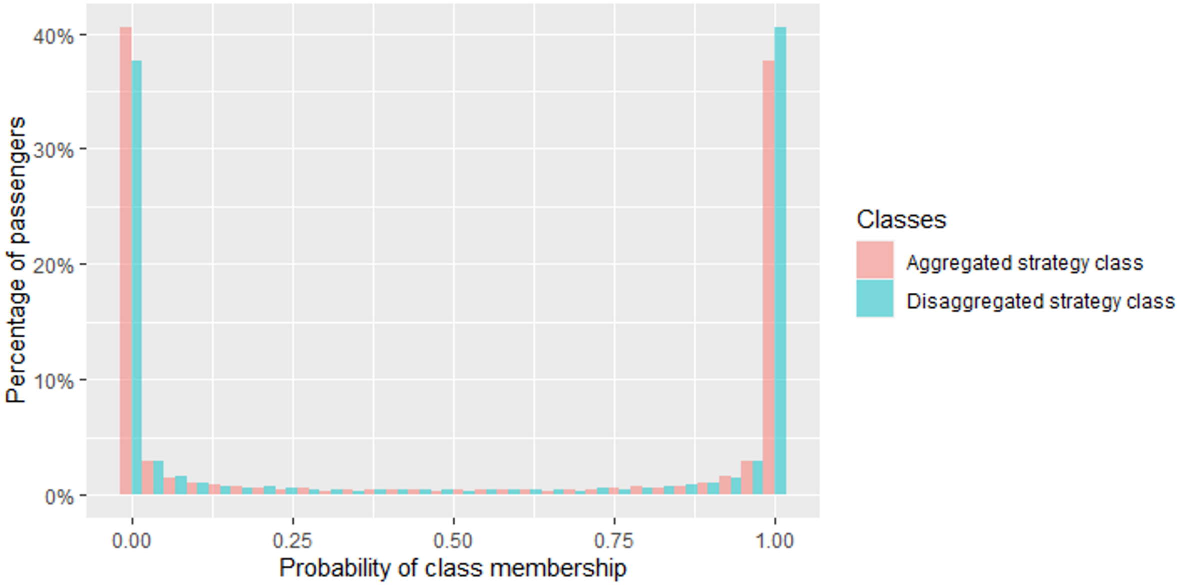


Figure 2.5: Histogram for the estimated class membership probabilities for the MIS latent class model with Gini coefficient as the indicator and disaggregated line usage index as the instrument. N=18,466

2.6 Conclusions

This study uses SC data from Santiago, Chile to evaluate the route choice strategies of PT system users. We identified two types of strategies in the existing literature: (i) a disaggregated strategy and (ii) an aggregated strategy.

To analyze the use of common lines (aggregated strategy), we used an indicator based on the difference between the number of expected and observed trips for each passenger within lines that connect the route sections used by the passenger to make trips. The results of this analysis showed evidence for heterogeneity within route choice strategies. Some passengers use common lines in each stage of their trip, while others use one specific line, even when there is more than one common line between their origin and destination. Another group is not well-defined, because it uses a subset of common lines, as defined by Chriqui & Robillard

(1975).

Because our data showed that a proportion of passengers uses common lines in each stage of their trip, we estimated three RUM models: a PSL model with disaggregated alternatives or itineraries, which assumes that all passengers use a disaggregated strategy; a PSL model with aggregated alternatives, which assumes that all passengers use an aggregated strategy; and a MIS latent class model, which takes in consideration the endogeneity problem. The MIS latent class model contains two classes, disaggregated alternatives and aggregated alternatives, assuming that some passengers use common lines in trip stages and other passengers use a disaggregated strategy. One of the more significant empirical findings to emerge from this study is that passengers present differences in their travel behavior when they choose a route to arrive to the destination: some consider common lines, while others wait for a specific line.

It is interesting to note that both groups of passengers have differences in their perceptions of some route attributes. The MIS latent class model contributes to our understanding of those differences, which would not be possible to obtain with a simple multinomial logit model or with a PSL model. With the MIS latent class model, we see that passengers using the aggregated alternatives strategy prefer to travel in bus rather than Metro, while passengers in using the disaggregated strategy have no preference for one transport mode over the other. Waiting time is more burdensome for passengers who use common lines, lending support to why they consider more than one line as part of the same alternative. Walking time and bus crowding is more burdensome for passengers who use disaggregated alternatives, suggesting that they prefer specific lines in order to avoid walking and/or crowding. With the estimation of the membership class parameters in the latent class model we found that the percentage of passengers that use itineraries as route alternatives is similar to the percentage of passengers who use common lines. We speculate that a high proportion of passengers use disaggregated strategy because a proportion of them do not know which transit lines belong to the common line set or they use real-time information to reduce their waiting time choosing the transit line that they prefer. This suggests that in order to reduce waiting times and improve passenger perception of PT, it may be important that transport authorities make real-time information channels available, allowing users to know which common line alternatives allow them to reach their destination.

In this study, we use observations from frequent passengers that travel during morning peak periods for at least 15 times during a month, and that stay in the destination locations for at least two hours. In this way, we try to capture trips to regular activities such as work or study. A natural progression of this work is to analyze the route choice strategies using data from less frequent users, who probably do not know the transport network as well and, we speculate, may use more disaggregated strategies than aggregated strategies. In other words, future research is required to determine if the frequency of PT use affects the route choice strategy of passengers. Other interesting extensions of this work would be to analyze other periods of the day, such as the afternoon peak period and off-peak hours, where one would expect that passengers have different preferences. We speculate that, for those periods, the disutility of the waiting time and travel time might be smaller since users do not have a tight schedule to reach their destination, as it happens in morning peak hours with trips to work or study.

Since the PT system in Santiago operates by headway scheduling, without fixed-time schedules, the analysis of route choice strategy undertaken here has extended our knowledge of route choice behavior of passengers that travel in frequency-based transit networks. In this context, it is important to address the common lines problem, because users face a choice between multiple transit lines from one stop to the next transfer station, and some of them prefer to board the first bus that arrives to minimize the total expected travel time. Several questions remain to be answered about the route choice strategy used by passengers that travel in other contexts. For example, passengers that travel in a schedule-based transit network, where users can optimize their trips by planning their arrival time to the bus stop. This can also be the case when users have real-time information about buses' arrival times. These contexts allow passengers to save waiting time (if the transit lines are punctual or the online information is accurate) while choosing their preferred lines. We might expect that in these cases, people may tend to use more the disaggregated strategy.

Another possible line of research is related to the improvement of the process of constructing common lines by considering different valuations for travel and waiting times, and/or by the incorporation of other attributes. An intuitive first step toward this effort could be to use for that the attributes' valuation that resulted from the choice modelling process. That approach would imply building an iterative process of uncertain convergence and huge computational costs but, more importantly, theoretical and empirical evidence suggests against it. Swait & Ben-Akiva (1987) show that it would be inappropriate to assume that attribute valuation obtained from the choice modeling stage is the same as those regarding the choice set formation, which is the role played by the common lines in our framework. Instead, if one would want to modify the valuation of the attributes of the common lines approach, the correct way to address it would have to be built within the framework of a classic four stages transportation model in which the PT assignment is affected by the car assignment, and the overall impact of the proposed changes is accounted for. The type and amount of data, as well as the computational burden needed to address such problem, exceeds the scope of the present research.

Further research should focus on the revision of the assumption of the exponential distribution that is used for the waiting time. Although it has been found to fit real data well (S. Guo et al., 2011), recent efforts of improvement in this line have been explored in recent literature, including, among other things, mixing distributions (see e.g. Ingvardson et al., 2018), and loglogistic, gamma, and erlang distributions (Q. Li et al., 2015). It should be borne in mind that using those distributions may significantly complicate the problem by requiring, e.g., a complete enumeration of transfers and stops, although some approaches to that problem have been proposed in recent literature (Q. Li et al., 2015).

Finally, the findings of our research can be used to suggest some guidelines for PT models that can help steer decisions regarding where to implement new routes and how to improve existing ones. Firstly, given that a percentage of passengers use common lines, implementing lines with overlap in high-demand sectors will effectively allow passengers to reduce their waiting time. Secondly, as the model showed that bus-to-bus and Metro-to-bus transfers are more unpleasant than bus-to-Metro transfers, identifying OD pairs that required transfers would focus planning efforts to reduce the number of onerous transfers. Thirdly, given that our models show that transfers requiring walking generate larger disutility levels for passengers, high-demand transfer points should be carefully designed to avoid this problem.

Chapter 3

Assessing feasible approaches for building the consideration set in public transport route choice modeling using smart card data

This chapter is the second component of this thesis, which aims to answering the second research question (see Section 1.3.2): How do different consideration set generation practical approaches impact the estimation and predictions in a PT route choice model?. This study contribution is to propose and apply a methodology to assess different feasible approaches to address the consideration set problem in PT route choice models, both from a theoretical perspective and using four weeks of revealed preferences constructed from SC data from Santiago, Chile. The approaches under study are K-shortest paths, Labeling, Link elimination, Link penalty, Simulation, Combined (mix of all previous approaches), and the Historical/Cohort method. The first six methods emulate heuristics that individuals may use for building the consideration set, while the seventh is originally based on intuition but can also be fully justified from a theoretical viewpoint by reinterpreting the theorem of estimation and sampling of alternatives. For the empirical assessment, the first three weeks of SC data are used for estimation and to assess the fit and behavioral coherence attained with the feasible approaches under study. The fourth week of data is used for out-of-sample prediction analysis. The analysis shows that the Historical/Cohort method outperforms all other feasible approaches analyzed in all measures considered. This strong empirical evidence supporting the Historical/Cohort approach is in line with the theoretical results supporting it and suggests the convenience of using this approach, whenever feasible, beyond the PT route choice context. However, theoretical and practical challenges remain to be addressed, especially regarding its applicability in forecasting.

This chapter is under review in the following article:

Arriagada, J., Guevara, C. A., & Munizaga, M. A. (2022). Assessing feasible approaches for building the consideration set in public transport route choice modeling using smart card data. Transportation Research Part C: Emerging Technologies (under review).

Author's contribution

Jacqueline Arriagada: corresponding author, conceptualization, methodology, software, formal analysis, writing-original draft, funding acquisition, investigation; **Angelo Guevara**: conceptualization, formal analysis, methodology, resources, writing-review and editing, supervision, funding acquisition, investigation; **Marcela Munizaga**: conceptualization, methodology, resources, data curation, writing-review and editing, supervision, funding acquisition, investigation.

3.1 Introduction

To properly estimate and forecast with discrete choice models, it is crucial to identify the set of alternatives that were truly considered by an individual when making a choice. The problem is exacerbated when the number of alternatives is excessively large, as is the case in route choice models, but it is always an issue in any choice situation. The composition of the consideration set assumed by the researcher may importantly impact the model estimates and the predicted choice probabilities (Bliemer & Bovy, 2008; C. Prato & Bekhor, 2007), since a wrong assumption implies a potentially severe misspecification. Although various feasible approaches have been proposed to build consideration sets for route choice modelling, to date, no comprehensive assessment of these approaches has been possible because of data and methodological limitations. Furthermore, little is known about the impacts of the composition of the consideration set in the case of PT route choice modeling. This study contributes to close this gap by proposing and applying a methodology to assess seven feasible approaches to address the consideration set problem in PT route choice models, using four weeks of revealed preferences constructed from SC data from Santiago, Chile.

The methods proposed to address the consideration set problem can be broadly classified in two categories. The first uses a single-stage approach to model the consideration and the choice, while the second uses a two-stage approach. In the first category, all available alternatives are implicitly considered in the utility function, but some alternatives are penalized when attributes violate certain parameters, up to a point in which they cannot be chosen (Castro et al., 2013; Martínez et al., 2009). These methods require identifying all possible alternatives, which is feasible in scenarios with a small number of alternatives, but becomes infeasible for larger choice-sets, such as those involved in route choice modeling. Because of this, and because evidence has shown that the consideration set and the choice from considered alternatives involve distinct mental processes, it has been suggested that it is preferable to explicitly model the consideration and the choice stages separately (C. G. Prato, 2009).

There is a vast literature that discusses the identification of the consideration set prior to the route choice model. Three types of literature within this realm can be distinguished. The first group corresponds to theoretical contributions related to discrete choice models that explicitly include the construction of the consideration set in their analysis. The seminal work in this regard corresponds to Manski (1977), who proposes an approach where the consideration set is treated as a latent variable. Even though Manski (1977)'s approach theoretically solves the issue, in practical terms, it has two problems. The first is that is unclear how to appropriately define a practical function for the probability of considering a given latent set. The second is that the method involves enormous computing costs to enumerate all the combinations of possible consideration sets, an almost impossible task in the case of route choices in dense transit networks. To solve these practical problems, Swait & Ben-Akiva (1987) and Ben-Akiva & Boccara (1995) propose using individual characteristics (e.g., income or driver's license ownership) or restrictions (e.g., distance) to develop expressions for the consideration-set probability, and to consider simplifying assumptions to reduce the number of potential latent sets. This approach, although intuitively appealing, is still an ad-hoc solution to the problem. Furthermore, although this approach may be feasible for choice-sets with a reduced number of alternatives, it quickly becomes infeasible for problems involving many options, such as route choice models.

A second group of studies separating the consideration set problem from the choice stage are empirical contributions, mainly in marketing, that investigate the size of the consideration set and the factors that may influence alternative consideration (Brown & Wildt, 1992; Hauser, 2014). Similarly, in transportation, Hoogendoorn-Lanser & Van Nes (2004) contrast the consideration set stated by the individual with an "objective set" that was built by first enumerating all feasible alternatives within a space-time window and then reducing the set using a branch and bound algorithm based on logical constraints. The authors conclude that, while the number of alternatives available to the travelers may be very large, the set of alternatives perceived is substantially smaller and even fewer alternatives are finally considered. In this regard, Villalobos Zaid (2018) show results that suggest that using stated preferences to collect information on the consideration set may be prone to severe hypothetical bias, reflected in the fact that the size of the consideration set gathered from such tools depends systematically on the experimental setting.

A third group of studies, mainly focused on route choice modeling, aims to develop practical methods for generating the consideration set. A practical, or feasible, method is usually an algorithm or heuristic that tries to reproduce or emulate the behavior that individuals may follow when building their consideration set. Our research effort falls within this category, with an emphasis on the PT case, for which we have extensive revealed preferences observations constructed from SC data.

The existing literature on PT users' route preferences has applied both stated preference (SP) data and revealed preference (RP) data. The studies that have used SP data (e.g. Eluru et al., 2012; Grison et al., 2017; Vrtic & Axhausen, 2002) obtain the information from a survey that asks people to choose a route from a set of alternatives in a hypothetical route choice situation, which may or may not be based on a real trip situation. This methodology to obtain the data has the advantages of being relatively inexpensive and allowing the researcher to know the true consideration set faced by the respondent, but it introduces hypothetical bias, since the user does not experience an actual trip.

On the other hand, the use of RP data has the advantage of reflecting true information about the users' choices in real situations, but, in this case, the researcher does not know the true consideration set, and data collection costs are higher. PT route choice studies that work with RP data may use traditional travel survey methods to capture choices and attributes, but these are not only expensive but also impractical for large samples (Z. Guo, 2011; Z. Guo & Wilson, 2011; Raveau et al., 2011, 2014; Raveau & Muñoz, 2014; Ton et al., 2020; Vrtic & Axhausen, 2002). Recently, some authors have dealt with this problem using SC data. The main purpose of SCs is to collect PT fares but, as a side benefit, they also collect a large quantity of very detailed data regarding the choices made by PT users at significantly lower costs when compared with traditional survey methods, with few practical limitations, and with unprecedented granularity and scalability (Pelletier et al., 2011).

In the context of PT route choice studies that use RP data, the consideration set generation process is a complex task, since, usually, there are countless feasible alternative routes in a transport network, especially in a dense multimodal one. Two ways to identify the consideration set in practice can be distinguished in the applied literature: build it using an algorithm or heuristic that emulates how individuals may build the consideration set, or impute it from historical data. Most of the heuristics used to build feasible consideration sets in practice are based on iterating some type of deterministic shortest path (C. G. Prato, 2009). Some typical heuristic approaches are the K-shortest path, Labeling (Ben-Akiva et al., 1984), Link elimination, Link penalty, Stochastic path methods, and combination of previous methods. On the other hand, in the last decade to tackle the consideration set problem has been to impute it from historical data, i.e., previous choices made in a similar situation (Yap et al., 2020; Kim et al., 2020; Jánošíková et al., 2014; Raveau et al., 2011, 2014). We denominate this method as the Historical/Cohort approach, which can be formally defined as building a practical consideration set from some collection of all observed choices made by the traveler in some timeframe prior to the instance under analysis, or the collection of observed choices of other users in the same cohort in cross-section data. Our research implements and assesses six typical heuristic approaches, together with the Historical/Cohort approach, which has become feasible with the advent of SC data.

Beyond the consideration set problem, which is the focus of this research, another important challenge in PT route choice modeling is the high level of correlation between route alternatives that share various links (e.g. C. G. Prato, 2009), which alters the choice probabilities of overlapping routes. Therefore, route choice models must properly represent the correlation structure among alternative routes. Most studies of PT passenger route choice behavior have used the Multinomial Logit (MNL) discrete choice model (Schmöcker et al., 2013; Nassir et al., 2018; Jánošíková et al., 2014; Kim et al., 2020; Grison et al., 2017; Z. Guo, 2011; Raveau et al., 2011, 2014; Raveau & Muñoz, 2014; Vrtic & Axhausen, 2002), which assumes independence of alternatives, and therefore ignores the correlation problem due to overlapping route segments. To address this limitation, the analytical Path Size Logit (PSL) approach has been proposed, which accounts for correlation by adding a deterministic term that reduces the utility of overlapped alternatives (Yap et al., 2020; Rui, 2016; Anderson et al., 2017; Bovy & Hoogendoorn-Lanser, 2005; de Grange et al., 2012; Hoogendoorn-Lanser et al., 2005; O. A. Nielsen et al., 2021). Our research implements both types of models, the basic MNL model and the PSL model.

In this study we assess the relative performance of seven feasible approaches to address the consideration set problem for PT route choice modeling using revealed preferences data. The revealed preference data is built using three weeks of SC transaction data, representing the actual route choice behavior of passengers who used Santiago, Chile's dense PT network. For the analysis, we use two specifications, the multinomial logit model and the path size logit model. We first evaluate the impact of different consideration set generation approaches by assessing the plausibility of the model estimates and the in-sample fit attained by each approach using different statistics and criteria. We conclude by studying the out-of-sample prediction performance attained by each method using the fourth week of data.

The remainder of this chapter is organized as follows. Section 2 presents a formal demonstration of why the Historical/Cohort approach can be used to obtain consistent estimators of the model parameters. Section 3 describes the case study, the Santiago, Chile transit network. Section 4 discusses the proposed methods used for the consideration set and route choice models. Section 5 introduces the estimation and prediction results. Section 5 concludes and discusses policy and research implications.

3.2 On The Theoretical support for feasible approaches to building the consideration set

Despite the increasing sophistication of feasible approaches to building the consideration set, methods that are based on repeatedly varying some type of deterministic shortest path are just heuristics that attempt to emulate a supposed consideration behavior. There is no theoretical support for them beyond the assumption that the individual's choices are rational, according to the researcher. Whenever the true consideration behavior varies from the rational behavior assumed by the researcher, the practical results may be poor. Therefore, empirical assessment of any of proposed method under a common framework is critical; however, it has been scarcely applied. Villalobos & Guevara (2021) presented Monte Carlo evidence along these lines. In the next section, we present, for the first time, a contribution along these lines using real PT route choice information extracted from SC data.

The only feasible consideration set approach for which a formal theoretical support has been given recently corresponds to variations of what we call the Historical/Cohort approach, which builds the consideration set from past choices made by the same individual, or from the choices of other individuals in the same context. Crawford et al. (2021) provide a theorem that justifies achieving consistency when considering what they call "sufficient sets" (our Historical/Cohort approach) by reinterpreting McFadden (1978) result for the estimation and sampling of alternatives. However, the result obtained by Crawford et al. (2021) is inaccurate regarding one important detail. It is sustained on the premise that this consistency would be achieved if any subset of the true consideration set were used for estimation, but McFadden (1978) shows that, in general, a sampling correction depending on the sampling protocol is needed. Following Guevara (2022), we derive the required sampling correction when the consideration set is constructed from past choices and formalize the conditions under which such a correction would satisfy the uniform condition property and thus can be ignored when constructing practical estimators.

Consider a Random Utility Model (RUM) setting in which the utility U_{in} that an individual n receives from alternative i can be written as the sum of a systematic part V_{in} and a random error term ε_{in} as shown in Equation 3.1, where V_{in} depends on attributes x_{in} and population parameters β^*.

$$U_{in} = V_{in} + \varepsilon_{in} = V_{in}(x_{in}, \beta^*) + \varepsilon_{in} \tag{3.1}$$

Then, if ε_{in} is distributed iid Extreme Value $(0, \mu)$, the probability that n will choose alternative i will correspond to the Logit model shown in Equation 3.2, where C_n is the true consideration set of J_n elements from which individual n selects one alternative. The scale μ in Equation 3.2 is not identifiable and is thus usually normalized to equal 1 to grant identification.

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} \tag{3.2}$$

Consider that the true consideration set C_n is latent to the researcher, but that she can observe the R choices that occurred in past instances. These observations could correspond to choices made by the same individual n or, assuming group homogeneity, choices made by diverse individuals who faced the same choice situation. In practice, this type of data can be gathered, for example, from a series of supermarket purchases or, in the case of a commuting mode or route, from a series of passive data records across weekdays. In cross-sectional data, this information may be obtained by observing trips that shared an OD pair, period, trip purpose, income group, household composition, etc. As is described in the next section, in the case study considered in this study, based on passive SC data, we consider that a cohort corresponds to trips taken between the same OD pair, defined at a zonal level based on the boarding and alighting bus stops, and period of the day, within the previous three weeks of available data. As was explained before, we denominate this practical choice-set as the Historical/Cohort consideration set. Since the number of previous choices within the previous three weeks may differ by OD pair, formally R shall depend on the OD pair but, for notational simplicity, we will avoid adding an OD subscript.

The researcher is interested in modeling the choice occurring at the instance R+1. To do so, she builds a practical consideration set that includes all the alternatives that were observed in the previous R instances, plus the alternative chosen at instance R+1, if it was not already included. We assume the invariability of attributes and choice sets across the R+1 instances. This implies that neither the x_{in} nor the consideration set change across time. Regarding the internal validity of the assumption of invariability of attributes, this could reasonably hold for the PT system in a normal context, but may be more questionable when there is a disruption to the system, resulting in a loss of stability (Yap et al., 2017; Malandri et al., 2018). The assumption of the invariability of the consideration set may most likely hold in practice for habitual choices, like commuting using PT or when buying staple goods in the supermarket. These invariability assumptions may become more questionable when using past cohort choices instead of one individual's own historical choices. The degree of external validity of these assumptions may only be tested in the field using real data. This is one of the purposes of our case study.

Under these assumptions, the key towards demonstrating the consistency of the Historical/Cohort approach lies in noting that the R previous choices could be understood as draws with replacement from the true consideration set, with a sampling protocol defined by the choice probability.

Formally, the problem originally tackled by McFadden (1978) was that the true consideration set C_n was too large to be processed in practice by the researcher, which was solved by building a reduced practical set $D_n \subseteq C_n$ for estimation, using some known sampling protocol conditional upon the chosen alternative. Formally, $\pi(D_n|j)$ corresponds to the conditional probability that the researcher would sample the set D_n, given that alternative j was chosen by individual n. Under this setting McFadden (1978) showed that by maximizing a pseudo-loglikelihood using the choice probabilities shown in Equation 3.3, one can obtain consistent estimators of the model parameters.

$$P_n(i|D_n) = \frac{e^{V_{in} + ln\pi(D_n|i)}}{\sum_{j \in C_n} e^{V_{jn} + ln\pi(D_n|j)}} \tag{3.3}$$

The merit of the model depicted in Equation 3.3 is that it only depends on the alternatives of the reduced set D_n, reducing a problem of possibly millions of alternatives to just a few dozen. The resulting model has a closed form that corresponds to a simple Logit model with a correction term by alternative $ln\pi(D_n|j)$ that only depends on the sampling protocol. This result holds thanks to the IIA property of the Logit model but was extended to more flexible models like GEV (MEV), RRM, and Logit Mixture by Guevara et al. (2016); Guevara & Ben-Akiva (2013a,b), respectively.

Despite the simplicity of Equation 3.3, it cannot be directly used to solve the consideration set problem using Historical/Cohort choices because the correction term $ln\pi(D_n|j)$ depends on the choice probability, which is obviously unknown to the researcher. However, in some cases, when the "uniform condition"[1], as McFadden (1978) called it, holds, the sampling correction $ln\pi(D_n|j)$ cancels out across alternatives and can therefore be ignored. This occurs, for example, when the protocol used to build D_n corresponds to drawing the chosen alternative and then adding a given number of nonchosen alternatives randomly drawn from C_n. If we can prove under which circumstances the uniform condition also holds for the Historical/Cohort approach, the problem will be solved.

As stated before, the Historical/Cohort approach to the consideration set problem can be seen as a problem of estimation with importance sampling of alternatives with replacement, a feature that was studied by Ben-Akiva & Lerman (1985) and revisited by Ben-Akiva (1989). Using the results from Ben-Akiva (1989), which assumes the multinomial distribution for the probability $\pi(D_n|j)$, it can be shown that, for the Historical/Cohort approach, the sampling correction $\pi(D_n|j)$ corresponds to the expression shown in Equation 3.4, where z_j is the number of times alternative j was chosen in the $R+1$ instances.

$$\pi(D_n|i) = \frac{z_i}{P_n(i)} \frac{R!}{\prod_{j \in D_n} P_n(j)^{z_j}} = \frac{z_i}{P_n(i)} K_D \approx \frac{z_i}{{}^{z_i}/_{R}} K_D = RK_D \tag{3.4}$$

The demonstration continues by first noting that the term K_D that does not change across alternatives and that, as the number of instances R grows, the choice probability $P_n(i)$ will become closer to ${}^{n_i}/_{R}$, resulting in the sampling correction approximating RK_D, a term that does not depend upon the alternative. Since the sampling correction in Equation 3.4 enters

[1] Uniform condition: the probability of sampling the set D_n given i is the same to the probability of sampling the set D_n given j, for all i and $j \in D_n$.

the choice probability shown in Equation 3.3, this means that the correction cancels out and can simply be ignored in the likelihood. Based on this demonstration, it can be affirmed that, under the invariability assumption, the Historical/Cohort approach can obtain the same or better performance estimation and prediction results than other heuristic methods commonly used to identify the consideration set. The case study presented in the next section is aimed at studying this hypothesis.

3.3 Case study: data sources and research methodology

In this section we describe the implementation of the case study, which consists of i) describing the available data set, ii) defining the urban modeling network, iii) applying seven consideration set generation techniques to the Santiago, Chile urban network; iv) estimating the route choice models using the consideration sets generated in (ii), and (v) evaluating the model performance for the different consideration sets.

3.3.1 Route choice data from smart cards in Santiago, Chile

The analysis for this study was carried out using the Santiago, Chile multimodal PT network, known as Transantiago. Santiago is the capital of Chile and has a population of over seven million inhabitants. The PT system serves roughly 50% of motorized trips and is operated by headway scheduling; therefore, lines do not have timetables . The fare system is fully integrated, with an almost flat fare between urban buses, Metro, and one rail service , allowing up to three trip legs within a two-hour time window. In a typical week, three million cards (passengers) use the system to take 25.5 million trips. The network includes seven Metro lines, more than 300 bi-directional transit lines, and one rail service. In this study, we use observations from passengers that travel during morning peak periods (6:30-8:30 AM) on weekdays. Specifically, we select users that remain at their destination locations for at least two hours. In this way, we aim to capture regular activity-based trips, such as work or study. The morning peak period is the most congested - more than 700 thousand trips per day can be observed - and is therefore the most interesting travel period from a planning perspective.

Very detailed demand information is available from the automatic fare collection (AFC) system in Transantiago (Gschwender et al., 2016). The SC bip! is the only accepted payment method, and passengers must validate when boarding a bus or entering a Metro station. The data is already processed to estimate the boarding and alighting position for all validations and the trip legs (stages) associated with each observed origin-destination journey, using the Munizaga & Palma (2012) methodology.

We selected trips taken in the PT system during May 2018. Specifically, we use three weeks (15 weekdays) to estimate the models and the fourth week (5 weekdays) to evaluate the prediction accuracy of the models.

3.3.2 Network representation

To represent the urban transit network, we use the network representation proposed by Cepeda et al. (2006); Spiess & Florian (1989), which is a frequency-based network formulation, through a direct graph $G = (N, A)$, where N represents the nodes of the network and A represents the arcs of the network. This representation contains two subsets of nodes. The stop nodes, which are used to represent bus stops, Metro stations, and train stations; the line nodes, which are used to represent transit lines (bus, Metro, or train). All nodes that represent the same transit line are connected by an on-board arc. When a transit line, represented by a line-node A, serves a bus stop, represented by a stop-node B, both nodes, A and B, are connected by a boarding arc and an alighting arc. When two stop nodes are separated by a walkable distance (100 m), they are connected by a walking arc. Additionally, each arc $a \in A$ is characterized by t_a, which is a nonnegative travel time and f_a, which is a nonnegative frequency. This frequency is used to estimate the waiting time, and it is assumed that the arrivals of the different transit lines are independent and exponentially distributed. This assumption has been adopted by several authors (Cepeda et al., 2006; Spiess & Florian, 1989), since the exponential distribution (or the gamma, which is the sum of exponentials) has been found to fit real data well (S. Guo et al., 2011). Alighting, walking, and on-board arcs do not have waiting time, and they are assigned infinite frequencies.

Therefore, waiting time for the boarding links is obtained assuming a Poisson process, yielding an average value of one over the observed frequency of the transit line. The observed frequency is obtained from Automatic Vehicle Location (AVL) data. In-vehicle travel time for the on-board links is obtained from a combination of AVL and General Transit Feed Specification (GTFS) data. For a link that involves a transit line with frequent headways, the process considers all dispatches along the transit line and takes an average of the travel time for the link. The travel time of the link per expedition is obtained from AVL data, and if it is not available, it is obtained from GTFS data. Since AVL data is not available for the Metro or the train, the travel time for their links is obtained from the operational parameters provided by the service operator. The walking time for walking links is obtained assuming that passengers walk at a speed of 4 km/hour, which is a standard value within transportation studies, and considers the Manhattan distance between nodes, since some studies have found that it is a better substitute for the network distance than the Euclidean distance (Mora-Garcia et al., 2018; Tien et al., 2011). The total network used in this study contains 46,583 nodes and 117,816 links.

3.3.3 Seven approaches for generating the consideration set

For the consideration sets, we first define the origin and destination location of each journey, which we consider as the area within a 100-meter radius of the corresponding origin and destination stops. An alternative path is characterized by the boarding stops, the transit lines used in each stage of the journey, and the alighting stop of the last stage of the journey.

We randomly selected 258 OD pairs to evaluate. For each OD pair, we construct the consideration set by using seven different techniques: the Labeling approach, Link elimination approach, Link penalty approach, K-shortest path approach, Simulation approach, Combined approach, and Historical/Cohort approach.

The Labeling approach is based on the behavioral assumption that users consider different objectives to select the considered route alternatives. Each label corresponds to a different objective function for which a given path is optimal (Ben-Akiva et al., 1984). In this study, this approach is applied using six different cost functions. Labels 1 through 3 use only one variable (path attribute) in the cost function while labels 4 through 6 are built by weighting multiple variables. Label 1 generates the route with the minimum in-vehicle travel time between the origin and destination zones. Label 2 generates the route with the minimum waiting time between the origin and destination zones. Label 3 generates the route with the minimum number of transfers between the origin and destination zones. Label 4's cost function adds in-vehicle travel time, waiting time, and walking transfer time, assuming the same weight for each of them. Label 5 penalizes waiting time by a factor of 1.6 and walking transfer time by a factor of 3, using values obtained from Arriagada et al. (2022). Label 6 uses the same penalization from label 5 for waiting time and walking transfer time and adds a penalization of 13 minutes to bus-to-bus, Metro-to-bus, and bus-to-Metro transfers, using values obtained from Arriagada et al. (2022).

The Link elimination approach repetitively searches for the minimum cost path after removing a link from the optimal path. This approach follows the stages: (a) identifying the generalized minimum cost path, (b) eliminating the link from the generalized minimum cost path that is closest to the origin and has not been removed previously, and (c) identifying the generalized minimum cost path. We followed the procedure used by Rui (2016). When all links along the first generalized minimum cost path have been eliminated, the iteration will move to the next generated path. The procedure ends when there are no remaining paths that reach the destination or where the maximum number of alternative paths (N) is reached. In this study, we use N=20 since the maximum size of observed paths from SC data is 18. The cost function includes in-vehicle travel time, waiting time, walking time, and transfer penalty.

The Link penalty approach also repetitively searches for the minimum cost path, but unlike the previous method, it imposes a penalty on the cost of all links that form the optimal path, instead of removing a link. This approach follows the stages: (a) identifying the generalized minimum cost path, (b) penalizing the generalized minimum cost path links by a factor of 1.05, and (c) identifying the generalized minimum cost path. We followed the procedure used by C. Prato & Bekhor (2007). The procedure ends when a maximum number of alternative paths (N) is reached. In this study, we use N=20 since the maximum size of observed paths from SC data is 18. The cost function includes in-vehicle travel time, waiting time, walking time, and transfer penalty.

The K-shortest path approach consists of the identification of the best K paths according to the link cost function. The behavioral assumption behind this approach is that passengers choose from a limited-size consideration set, avoiding costly alternatives. In this study we use the algorithm proposed by Yen (1971) and K is set to 20, since the maximum size of observed paths from SC data is 18. The cost consists of in-vehicle travel time, waiting time, walking time, and transfer penalty.

The Simulation approach searches for the minimum cost path for each random draw of link cost functions from a truncated normal distribution with the mean equal to the original cost of the link and the standard deviation equal to 20% of the original value. 50 draws of randomized cost functions were performed for each OD pair. The cost function includes in-vehicle travel time, waiting time, walking time, and transfer penalty.

The Combined approach builds the consideration set mixing the alternatives found in the Labeling approach, the Link elimination approach, the Link penalty approach, the K-shortest path approach, and the Simulation approach.

The Historical/Cohort approach consists of identifying all alternatives recorded and observed for each OD pair in the past by any traveler during the study period (in our case, SC data observations from 6:30 to 8:30 on weekdays). Each alternative is characterized by the stops and transit lines used by the passengers that traveled between the OD pair. The premise of this approach is that all travelers for the same OD pair might share the same consideration set and thus, the historical choices made by the individuals from the same cohort (OD pair) would necessarily belong to the true consideration set. This approach has been used by Yap et al. (2020); Kim et al. (2020); Jánošíková et al. (2014).

All approaches, except the Historical/Cohort approach, require the specification of a transit network and can generate alternative paths that contain walking stages at the beginning and/or end of a trip. As this type of path cannot be observed in the SC data, we specified for all six heuristic approaches that the first link could not be walking. The same consideration set was used for all passengers between an OD pair.

3.3.4 Specification of route choice models

In this study, we use two types of RUM models, the MNL discrete choice model, and the PSL model. The MNL model is the basic model, which has been used for most studies of PT passenger route choice behavior (Kim et al., 2020; Ton et al., 2020; Nassir et al., 2018; Grison et al., 2017; Jánošíková et al., 2014; Raveau et al., 2011, 2014; Raveau & Muñoz, 2014; Schmöcker et al., 2013; Z. Guo, 2011; Vrtic & Axhausen, 2002). Since route choice models present a correlation between alternatives due to overlapping route segments, it is necessary to correct the MNL model, which assumes the independence of alternatives. To address this problem, the analytical approach of PSL models have often been adopted, which account for the correlation by adding a deterministic term that reduces the utility of overlapped alternatives (Arriagada et al., 2022; Yap et al., 2020; Rui, 2016; Anderson et al., 2017; Bovy & Hoogendoorn-Lanser, 2005; de Grange et al., 2012; Hoogendoorn-Lanser et al., 2005; O. A. Nielsen et al., 2021).

The deterministic component of the MNL model is specified in Equation 3.5, where i represents the alternative route, TT_i is in-vehicle travel time, IWT_i is the initial waiting time, TWT_i is the transfer waiting time, $TWalT_i$ is the transfer walking time, and TR_i is the number of transfers between vehicles. The PSL model is presented in Equation 3.6, which contains all attributes introduced in the MNL model and adds the path size correction (PSC_i) term to capture the correlation due to overlapping between alternative routes. Path size correction introduces a negative factor that decreases the deterministic utility of alternative routes that have correlation with other routes. We have used the expression according to Bovy et al. (2008) in Equation 3.7, where L_r is the length of the route section r, L_i is the length of route i, ζ_i is the set of route sections belonging to route i, and δ_{rk} is the section-route incidence number, which takes a value of 1 if route k uses route section r and a value of 0 otherwise.

$$V_i = \beta_{TT}TT_i + \beta_{TWT}TWT_i + \beta_{TWalT}TWalT_i + \beta_{TR}TR_i \tag{3.5}$$

$$V_i = \beta_{TT}TT_i + \beta_{TWT}TWT_i + \beta_{TWalT}TWalT_i + \beta_{TR}TR_i + \beta_{PSC}PSC_i \tag{3.6}$$

$$PSC_i = \sum_{r\in\zeta_i} \frac{L_r}{Li} \ln \frac{1}{\sum_{k\in C}\delta_{rk}} \tag{3.7}$$

Because of the Gumbel distribution of the error term, the probability of passenger n choosing an alternative i given consideration set C is expressed as in Equation 3.2. The closed-form logit formula of the Logit model allows for simple estimation of the fixed coefficients by maximizing the likelihood function.

3.3.5 Evaluation of methods

The purpose of the evaluation process is to analyze the performance of the different consideration set generation approaches both qualitatively and quantitatively. Since the route choice models constructed with different consideration set generation approaches are not comparable between each other using statistics with likelihood, we use the cross-validation method, which consists of evaluating the prediction performance of the models for a part of the data set that was not used for model estimation. In our case, the cross-section units are periods of time, since the set of observations are split into two subsamples, the first subsample being a period to estimate the models and the second subsample being a period to evaluate the predictive capacity of the models. In the validation sample, we use the First Preference Recovery (FPR) index, which is the proportion of observations that use the route alternative with the highest chosen probability and the Average Likelihood (AL), which is shown in Equation 3.8. In this equation, $\lambda_n(i) = 1$ if path i is chosen by observation n (0 otherwise), N is the number of observations, C_n is the consideration set of observation n, and $P_n(i)$ is the calculated probability of observation n choosing path i.

$$AL = \frac{\sum_{n=1}^{N}\sum_{i\in C_n} P_n(i) * \lambda_n(i)}{N} \tag{3.8}$$

3.4 Results

3.4.1 Consideration set generation approaches

The objective of a consideration set generation approach is to emulate the actual behavior of passengers and obtain the maximum percentage of observations for which a path generation approach reproduces the actual behavior. In this study, we use three coverage indicators. The first indicator is Trip Coverage (TP), which is the percentage of trips in which the chosen alternative is included in the generated consideration set. Equation 3.9 shows the formulation for TP, where P is the set of OD pairs, OB is the set of alternatives recorded by the SC database for the OD pair p, CS refers to a specific consideration set generation approach, CS_p is the set of alternatives generated by the consideration set generation approach CS for the OD pair p, T_{ap} is the number of trips taken in alternative a in the OD pair p, and $\delta_{a,CS}$ takes value 1 if alternative a belongs to the set of alternatives generated by the consideration set approach CS for the OD pair p. It represents the percentage of trips that can be modeled with each approach and indicates the effectiveness of the consideration set approach. The second indicator is Efficient Coverage (EC), which is the percentage of the generated alternatives actually observed in the SC data (Rui, 2016). Equation 3.10 shows the formulation for EC, where γ_{a,OB_p} takes value 1 if the alternative a belongs to the set of alternatives recorded in the SC database for the OD pair p. This indicates the efficiency level of the consideration set approach to produce the paths used by passengers and to avoid producing unused paths. The third indicator is Passenger Path Coverage (PPC), which is the percentage of observed paths included in the generated consideration set (Rui, 2016). Equation 3.11 shows the formulation for PPC, which indicates the comprehensiveness level of the generated consideration set regarding the observed alternatives. Trip coverage is the most important indicator since it affects the second stage of a route choice model. Table 3.1 shows the three coverage indicators for each consideration set generation approach in the estimation sample and in the prediction sample. As the Historical/Cohort approach is built using the observed paths in the estimation sample; therefore, in this period, its coverage indicators are equal to 1. However, in the prediction sample, the observed paths can be different from the estimation paths, so coverage indicators may vary.

$$TP_{CS} = \frac{\sum_{p\in P}\sum_{a\in OB_p} T_{ap}\delta_{a,CS_p}}{\sum_{p\in P}\sum_{a\in OB_p} T_{ap}} \tag{3.9}$$

$$EC_{CS} = \frac{\sum_{p\in P}\sum_{a\in CS_p} \gamma_{a,OB_p}}{\sum_{p\in P}\sum_{a\in CS_p} 1} \tag{3.10}$$

$$PPC_{CS} = \frac{\sum_{p\in P}\sum_{a\in OB_p} \delta_{a,CS_p}}{\sum_{p\in P}\sum_{a\in OB_p} 1} \tag{3.11}$$

Table 3.1: Performance of each consideration set generation approach

Consideration set	Estimation sample			Prediction sample		
generation approach	TP	EC	PPC	TP	EC	PPC
Historical/Cohort	1	1	1	0.99	0.81	0.95
K-shortest paths	0.78	0.44	0.68	0.78	0.43	0.74
Labeling	0.72	0.53	0.63	0.72	0.53	0.67
Link elimination	0.79	0.33	0.69	0.79	0.32	0.74
Link penalty	0.85	0.14	0.78	0.85	0.13	0.82
Simulation	0.67	0.31	0.55	0.67	0.32	0.60
Combined	0.85	0.08	0.79	0.86	0.07	0.83

TC: Trip Coverage, EC: Efficient Coverage, PPC: Passenger Path Coverage.
These indicators were constructed using 15,594 observed trips in the SC database.

As Table 3.1 shows, the Historical/Cohort approach obtains the highest coverage indicators among all heuristics in the prediction sample. The Labeling approach obtained the highest efficient coverage (EC) results out of the six heuristic approaches, which means that it generates the most paths that are actually used by passengers. This finding was also reported by (Rui, 2016). Regarding the alternatives that were observed in the trip database but were not chosen by the Labeling approach, in general, these paths have costs similar to those generated by the labels that minimize total travel time; however, they are not captured by any label. The generated alternatives used by passengers are captured mainly by the label that minimizes total travel time, while the label that minimizes in-vehicle travel time generates paths with many transfers, and the labels that minimize waiting time and the number of transfers generate alternatives with long in-vehicle travel times.

Contrary to the Labeling approach, the Combined approach obtained the lowest EC results, which means that it generates more paths not used by passengers than any other approach included in the analysis. However, excluding the Historical/Cohort approach, the Combined approach most effectively produced the observed alternatives, since it obtained the highest TC and PPC results.

The Simulation approach obtained the poorest TP and PPC results, even though the number of draws (50) was higher than the K value for the K-shortest path approach and the N value for the link elimination and link penalty approaches. The low level of the passenger path coverage and trip coverage means that the Simulation approach generated paths with the lowest coverage of observed paths. These problems occur because many unattractive paths are generated, specifically with more transfers than the observed paths.

To evaluate the composition of the choice set, Figure 3.1 shows a boxplot that illustrates the number of alternative paths generated by each consideration set approach. The Labeling approach generated 4.2 alternative paths on average, which is similar to the number of path sets generated by the Historical/Cohort approach (3.8 alternative paths on average). The K-shortest path, Link elimination, Link penalty, and Simulation approaches generated a larger consideration set, on average, than the Historical/Cohort approach. One explanation for this is that these approaches include explicit constraints on the maximum number of alternative paths, which is set to 20 for the deterministic methods (K-shortest path, Link elimination, Link penalty) and 50 for the stochastic method (Simulation). Only the Link penalty approach always generates 20 path alternatives for each OD pair, as this method finds a different alternative for each iteration. The Combined approach generated the largest consideration set, which is to be expected, since it jointly considers the different alternatives from the five heuristics.

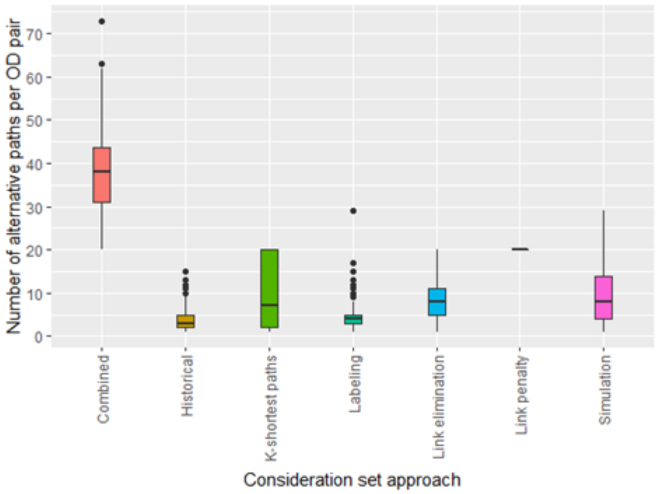


Figure 3.1: Distribution of alternative paths per consideration set approach

For each consideration set approach, we calculated the average of each attribute that characterized the alternatives for each OD pair. In Table 3.2, we show the mean and standard deviation of the average of path attributes for each generated consideration set approach. We use the Path Size (PSC) Term presented in Equation (7) to evaluate the capacity of each approach to generate diverse paths. A PSC term equal to zero means that the generated alternative paths included in the consideration set are not overlapped (the alternatives are all different). A more negative PSC term means that the generated alternative paths included in the consideration set have a higher degree of overlap (the alternatives are more similar). The results in Table 3.2 show that the Combined approach generated the highest degree of overlap between alternatives. This is because, since the Combined approach includes more alternatives, there is a higher probability that they share links. Without considering the Combined approach, the Simulation and Link penalty approaches generate the most heavily overlapped alternative paths, while the Historical/Cohort, Labeling, and Link elimination approaches generate the most diverse alternative paths.

Analyzing the number of transfers for each approach, the Historical/Cohort approach generates the fewest transfers on average, followed by the K-shortest path approach. Therefore, the K-shortest path approach generates fewer irrelevant path alternatives, on average, than the other heuristic approaches. The waiting time shows similar results, where the waiting times generated by the K-shortest path approach are quite similar to those of the Historical/Cohort approach. The Labeling approach generates paths with longer waiting times and walking transfer times, on average. This is expected, since the Labeling approach applies some labels that do not take waiting and walking transfer times into consideration. Regarding in-vehicle travel time, the K-shortest path approach and the Simulation approach obtained the results closest to the Historical/Cohort approach, while the Link penalty approach produced routes with the longest in-vehicle travel times.

Table 3.2: Statistics of alternatives path attributes for each consideration set approach

Consideration set approach	Size		Path Size		# of transfers		Waiting time		In-vehicle travel time		Walking time in transfer	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Historical	3.8	2.5	-1.2	0.7	0.8	0.9	7.7	3.3	25.5	13.4	1.0	1.4
K-shortest paths	10.0	7.7	-2.2	1.1	1.2	0.7	8.5	3.0	24.4	13.1	0.6	0.7
Labeling	4.2	2.8	-1.7	0.7	2.4	2.1	22.8	24.2	27.2	17.0	1.2	1.4
Link elimination	8.7	4.5	-1.7	0.6	1.5	0.9	14.1	12.0	29.1	16.4	1.1	0.8
Link penalty	20	0	-2.6	0.4	1.7	0.7	12.4	3.3	29.9	15.1	1.1	0.6
Simulation	9.5	6.4	-2.3	1.0	2.1	1.7	14.5	7.9	25.6	14.6	1.1	0.7
Combined	38.0	9.1	-3.4	0.5	2.1	0.9	15.7	5.7	29.4	15.3	1.2	1.1

3.4.2 Route choice models

MNL models and PSL models were estimated using a sample of 15,594 observations, which correspond to trips taken during the 15 business days comprising the first three weeks of data. Both types of logit models were estimated using the Historical/Cohort, K-shortest paths, Labeling, Link elimination, Link penalty, Simulation, and Combined approaches.

The specification of the deterministic utility function considers in-vehicle travel time, waiting time at the beginning of the trip and during transfers, walking time during transfers, and the transfer penalty, which considers bus-to-bus, bus-to-Metro, and Metro-to-bus transfers. Metro-to-Metro transfers cannot be incorporated into the model because the route that the passenger uses inside the Metro network cannot be observed from the data.

Table 3.3 shows the estimated parameters for the MNL models for each consideration set generation approach. The model that uses the Historical/Cohort reported parameters that are all are statistically significant (at the accepted 5% threshold) with the expected sign. These results are similar to the models that use a heuristic approach (Labeling, Link elimination, Link penalty, Simulation, K-shortest paths, and Combined approaches), where the parameters are statistically significant and with the expected sign, with the exception of the transfer walking time parameter, which had a positive sign for all six heuristic approaches.

Table 3.4 show the estimated parameters for the PSL models for each consideration set generation approach. The PSC term is statistically significant in all models. Given that PSC lies within the interval (-,0] and implies a reduction in the systematic utility of correlated routes, the positive sign in the coefficient obtained for all models is expected. All models maintain the results shown in Table 3.3, in terms of statistical significance and the sign of the parameters.

Comparing the results of the MNL models and the PSL models, the PSL models present better model fit in all cases. Consequently, the inclusion of the PSC term in the specification of the models allows for greater explanatory power compared to the MNL models. In summary, these results show that, all consideration set approaches can represent the perception of passengers with respect to the alternative path attributes, except for the transfer walking time attribute, which obtained an unexpected sign for heuristic approaches.

3.5 Discussion

This section analyzes the PSL model parameters to understand the differences between each evaluated consideration set approach.

The negative coefficients of the travel time and waiting time variables show a disutility of travel time for passengers. All models indicate that alternative routes with shorter in-vehicle travel time and waiting times are preferred. In the models, we estimated sensitivity to waiting time at the beginning of the trip as well as at any transfer stages. Table 3.5 shows that the results of the rates of substitution of initial waiting time with respect to in-vehicle travel time vary between 1 and 1.4. Focusing on transfer waiting time, the rates of substitution obtained with most of the consideration set approaches are around 2, which is in line with the PT route choice literature (Nassir et al., 2018; Rui, 2016; Raveau & Muñoz, 2014). However, the Link elimination approach and the K-shortest paths approach generate a rate of substitution of transfer waiting time value close to 3.

The disutility of the walking transfer time variable can only be represented by the model that uses the Historical/Cohort approach, which generates a rate of substitution of around 1.2 (see Table 3.5). This attribute cannot be evaluated for the other approaches, since their models generate a positive parameter. The trade-offs between in-vehicle travel time and transfer walking time obtained with the Historical/Cohort approach are in line with some studies that have reported a value between 1 and 2 (Jánošíková et al., 2014; Nassir et al., 2018). (Raveau & Muñoz, 2014) reported a value of around 3 minutes for this trade-off; the model that uses the Historical/Cohort approach is closest to this value.

The negative coefficient of the transfer variable shows a disutility of the number of transfers for passengers. All models indicate that passengers prefer alternative routes with the lowest number of transfers. As shown in Table 3.5, the results of the rates of substitution of the number of transfers with respect to in-vehicle travel time vary between 13 and 54 minutes. The smallest value is reported by the model that uses the Historical/Cohort approach and the highest value is reported by the model that uses the Simulation approach. Previous studies have shown that one transfer is perceived by a typical passenger as equivalent to a number that varies between 3.6 min and 16 min of in-vehicle time (Z. Guo & Wilson, 2011; Nassir et al., 2018; Raveau & Muñoz, 2014; Rui, 2016). Therefore, the model that uses the Historical/Cohort approach seems to best represent the perception of passengers regarding transfers.

For a prediction analysis, the trip dataset is split into two parts. The parameters obtained from the estimation are used to predict the paths chosen in the second part of the dataset, which corresponds to 4,685 weekday observations from the fourth week of data. We use the First Preference Recovery (FPR) and the Average Loglikelihood (AL) to evaluate the performance of each consideration set approach. Table 3.6 shows the FPR and AL values for the models constructed with each consideration set approach. The Historical/Cohort approach resulted in the best prediction performance.

In summary, the results of this study suggest that the Historical/Cohort approach obtained the highest level of prediction accuracy, and it is the only approach that returns, for all attributes evaluated in this study, similar rates of substitution reported in the PT route choice modelling literature.

Table 3.3: MNL model estimates (t tests) using each generation consideration set approach

Parameters	Consideration set approach						
	Historical/Cohort	Simulation	Labeling	Link penalty	Link elimination	K-shortest paths	Combined
Travel time in vehicle	-0.144 (-47.3)	-0.082 (-24.1)	-0.181 (-58)	-0.197 (-74.9)	-0.145 (-55.6)	-0.113 (-34.3)	-0.184 (-78.2)
Initial waiting time	-0.162 (-39.5)	-0.036 (-7)	-0.116 (-25.1)	-0.192 (-42.6)	-0.152 (-33.2)	-0.073 (-15.2)	-0.19 (-43.1)
Transfer waiting time	-0.304 (-23.5)	-0.269 (-22.5)	-0.196 (-12.8)	-0.434 (-33.4)	-0.383 (-28.5)	-0.395 (-22.3)	-0.29 (-30.6)
Transfer walking time	-0.152 (-12.8)	1.081 (38)	1.101 (37.3)	1.209 (54.2)	1.294 (42.1)	1.557 (49.4)	0.961 (45.9)
Transfer	-1.374 (-16.3)	-3.77 (-39.9)	-3.463 (-33)	-5.285 (-61.2)	-4.364 (-45.3)	-5.167 (-49.7)	-5.757 (-72.3)
# of observations	15,594	15594	15594	15594	15594	15594	15594
Log-likelihood	-16736.7	-11493.8	-12943.4	-16497.0	-14467.9	-13525.8	-17695.9
Adjusted rho-square	0.123	0.61	0.40	0.65	0.55	0.57	0.68
AIC	33483.3	22997.7	25896.8	33003.9	28945.8	27061.6	35488.3

All columns show t-values between parentheses. AIC = Akaike information criterion.

Table 3.4: PSL model estimates (t tests) using each generation consideration set approach

Parameters	Consideration set approach						
	Historical/Cohort	Simulation	Labeling	Link penalty	Link elimination	K-shortest paths	Combined
Travel time in vehicle	-0.119 (-37.7)	-0.096 (-21.5)	-0.129 (-32.9)	-0.173 (-62.3)	-0.111 (-36)	-0.135 (-37)	-0.181 (-72.9)
Initial waiting time	-0.131 (-30.7)	-0.137 (-20.6)	-0.147 (-25.8)	-0.193 (-40.3)	-0.144 (-28.7)	-0.132 (-25.7)	-0.236 (-52.2)
Transfer waiting time	-0.257 (-18.7)	-0.181 (-12.6)	-0.205 (-10)	-0.366 (-26.3)	-0.32 (-22.8)	-0.393 (-19.4)	-0.262 (-26.7)
Transfer walking time	-0.144 (-12.1)	0.478 (8.7)	0.624 (13.4)	0.693 (22.5)	0.924 (24.6)	1.323 (38.1)	0.596 (22)
Transfer	-1.527 (-18)	-5.177 (-32.2)	-4.118 (-26.6)	-5.155 (-49)	-4.261 (-42.4)	-5.128 (-43.6)	-5.846 (-63.7)
PSC	1.085 (29.5)	1.753 (46.2)	1.841 (57)	1.274 (50.4)	1.234 (55.9)	0.735 (32.5)	0.834 (43.6)
# of observations	15,594	15594	15594	15594	15594	15594	15594
Log-likelihood	-16180.1	-8086.6	-9707.4	-15066.4	-12739.4	-13133.1	-16736.2
Adjusted rho-square	0.152	0.73	0.55	0.68	0.60	0.58	0.70
AIC	32372.1	16185.3	19426.9	30144.9	25490.8	26278.2	33484.5

All columns show t-values between parentheses. AIC = Akaike information criterion.

Table 3.5: Rates of substitution of parameters obtained with the PSL model using each generation consideration set approach

Parameters	Consideration set approach						
	Historical/Cohort	Simulation	Labeling	Link penalty	Link elimination	K-shortest paths	Combined
Travel time in vehicle	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Initial waiting time	1.10	1.43	1.14	1.11	1.30	0.98	1.31
Transfer waiting time	2.16	1.88	1.59	2.11	2.90	2.91	1.45
Transfer walking time	1.21	-4.98	-4.85	-4.00	-8.36	-9.80	-3.30
Transfer	12.86	53.93	32.01	29.77	38.53	37.99	32.38

Table 3.6: Prediction results

Indicator	Consideration set approach						
	Historical/Cohort	Simulation	Labeling	Link penalty	Link elimination	K-shortest paths	Combined
First preference recovery	0.5266	0.4603	0.5057	0.4827	0.4941	0.4556	0.4954
Loglikelihood of validation sample	0.4451	0.2572	0.3500	0.3692	0.3569	0.3433	0.3476

3.6 Conclusions

Modeling route choice behavior requires the identification of non-chosen paths considered attractive by travelers to reach a destination. This set of alternatives is call the consideration set and is usually unknown to researchers working with revealed preference data. Most PT route choice studies that work with this type of data have identified the consideration set through ad-hoc heuristics, such as shortest path algorithms (Rui, 2016) or using historical data (Jánošíková et al., 2014; Kim et al., 2020). We use a variation of the historical data approach, which we term the Historical/Cohort approach, to impute the past observed choices made by the traveler, or by other users in the same cohort in cross section data, as the consideration set. In this study, we first present the theoretical conditions under which the Historical/Cohort approach would recover the population parameters. We then use a case study to assess the performance of different consideration set generation approaches, which are commonly used in the transport route choice literature, in terms of estimation and prediction. The results show that the Historical/Cohort approach surpasses the other methods with regards to all statistics considered.

The proof of the Historical/Cohort approach is based on an adaptation of the theorem of sampling of alternatives (McFadden, 1978), in which prior choices are understood as draws from the true consideration set, and the sampling correction cancels out when there are many observations and invariability conditions hold. In this context, we hypothesize that route choice models that use the Historical/Cohort approach to identify the consideration set obtain the same or better results than any other consideration set generation approach.

To evaluate this research hypothesis, we use data from the PT system in Santiago, Chile to estimate route choice models with different consideration set generation approaches: the Historical/Cohort approach and six approaches based on shortest path heuristics: the Labeling, Link elimination, Link penalty, K-shortest paths, Simulation, and Combination (of all prior) approaches. To do so, we split the database in two parts, the first corresponding to three weeks of weekday data and the second corresponding to weekday data from the fourth week. Using the first part of the database, we estimated two RUM models using each consideration set generation approach: a MNL model, which is the basic and most used model to represent route choice in PT systems, and a PSL model, which captures correlation due to overlapping between alternative paths. The results show that all PSL models obtained a better fit than the MNL models. Focusing on the PSL models, the estimated parameters suggest that all consideration set generation approaches can well-represent the perception of passengers for all attributes, except for the transfer walking time, which is well-represented only by the Historical/Cohort approach. Regarding the rates of substitution with respect to in-vehicle travel time, the Historical/Cohort approach is the only model that returns values previously reported in the PT literature for all attributes. The only attribute for which all consideration set generation approaches returned a rate of substitution reported in previous studies was the initial waiting time. For other attributes, one or more heuristic approaches reported non-expected values. The evidence from this analysis supports the idea that the Historical/Cohort approach to identify the consideration set accurately estimates the population parameters. In addition, the comparison of prediction accuracy across different consideration set generation approaches suggests that the Historical/Cohort approach estimates models with better predictive abilities with respect to the choices of passengers in the prediction sample.

This research shows theoretically and practically that SC data can be used to generate more realistic models of passenger behaviour. In this sense, SC data can be used to estimate route choice models using the Historical/Cohort approach to identify the consideration set, and can also provide empirical information to estimate the values of path attributes. Therefore, it is not necessary to use heuristics, which are generally computationally expensive, to emulate the passengers' consideration set.

Finally, this study benefits decision-makers at large-scale transport systems by providing a methodology to understand passenger perception and behaviour without using expensive survey-oriented resources.

Chapter 4

Evaluating the role of experience in a route choice context using smart card data in a large-scale public transport network

This chapter is the third component of this thesis, which aims to answer the third research question (see Section 1.4.2): How can the past experiences of PT passengers be integrated into a route choice model to incorporate the uncertain nature of in-vehicle travel time within the PT system? This study contributes by applying a passengers learning model to asses the relationship between past experiences and current route choices in the context of the beginning of operations of a new Metro line in the PT system. To do so, we used three months of revealed preferences constructed from smart-card data from Santiago, Chile, the first of which corresponds to the period right before the opening of the new Metro line, and the later two after that event. The learning model applied is an instance-based learning (IBL) model (Tang et al., 2017), which allows the researcher to represent the perceived in-vehicle travel time considering the recency of experiences and that the passenger's memory decreases with the power law of forgetting. For empirical assessment, smart-card data from one month prior to and two months after the launch of a new Metro line are used to calculate the perception of in-vehicle travel time. Two months of smart-card data after the launch of the new Metro line are used to estimate and assess the fit and behavioral coherence attained with the model using the perceived in-vehicle travel time of each passenger and with the model using the mean in-vehicle travel time (baseline model). The analysis shows that the route choice model considering the passengers' learning process outperforms the baseline model when estimated using data from five weeks after the new Metro line was opened. This empirical evidence supports the idea that passengers use knowledge from their past experiences to make a route choice when they have gained some experience in the new PT system context.

This chapter is currently a working paper

4.1 Introduction

Every time they make a trip, PT passengers have to decide which route to take to reach their destination from their origin. To achieve this, passengers consider a specific set of available and attractive routes, called the consideration set, and then choose a travel route. Traditional discrete choice models based on the RUM principle consider that, somehow, the individuals have perfect information about the attributes of the alternatives, an assumption that is questionable, especially when the attributes vary. In reality, the route-choice decision making process is more complex: it is a dynamic process that usually involves an evaluation of different route attributes, which can be fixed over time, such as the number of transfers, or uncertain, such as wait time and in-vehicle travel time, among others. These uncertainties come from unpredictable incidents or special events along the travel route. The information about the route's attributes can come from passengers' past travel experiences in the same origin-destination (OD) pair (reinforcement learning) or from descriptive travel information (cognitive learning).

When passengers consider descriptive travel information (pre-trip or en route) to choose a route, they are following a decision type called deciding from description. In this type of decision, the outcomes are specified to the decision-maker before they decide upon a certain alternative (Rakow & Newell, 2010). A purely description-based route choice context is very rare, except when a person arrives for the first time in a new city, does not have any previous experience in the PT system, and looks for travel information to obtain first-time knowledge about available routes. Route choice is a typical case of decision from experience, where a traveler makes a choice, carries out that choice, gains experience, forms an updated perception of the chosen route, and makes a choice again based on prior experience. In summary, a decision from experience requires passengers to explore an environment and learn the outcomes associated with different options (Rakow & Newell, 2010), and plays an important role in addressing the issue of understanding passengers' route-choice behavior.

Temporally variant and situation-dependent route attributes, such as in-vehicle travel time and wait time, explain the importance of including the passenger learning process in route choice models. Several PT route-choice models have been estimated considering cross-sectional data, which does not allow for the inclusion of passengers' route attribute learning processes (Kim et al., 2020; Z. Guo, 2011; Raveau et al., 2011, 2014; Raveau & Muñoz, 2014; Anderson et al., 2017; O. A. Nielsen et al., 2021). Other studies have estimated PT route-choice models using longitudinal data (Schmöcker et al., 2013; Yap et al., 2020; Nassir et al., 2018; Jánošíková et al., 2014). Although this type of data allows the researcher to incorporate the relationship between passengers' past experiences and their current choice, they ignore the individuality of the relationship, not accounting for the uncertainty in the route attributes, and instead considering them as static and assuming that all passengers have the same knowledge of the travel time distribution. The current study fills this gap by incorporating passengers' perceived travel time, which can vary across time and passengers, into a PT route choice learning model.

Many studies of the learning process in private transport literature can be found. Some are theoretical studies that propose models using an average or a weighted-average approach of previous time periods to infer the perceived travel time (Horowitz, 1984; Cascetta &

Cantarella, 1991) and others have estimated route-choice models that capture the travelers' learning process using laboratory experimental or simulated data (Bogers et al., 2007; Lu et al., 2014; Mahmassani & Liu, 1999). Ben-Elia & Avineri (2015) offers a detailed literature review of studies on the behavioral response to different types of travel information, such as experiential information, descriptive information, and prescriptive information. Up to present, far too little attention has been paid to analyze the effect of travelers' learning processes on route choice models using real world data. The reason for this is that real panel data on the learning process is virtually impossible to collect with traditional methods. The current research provides a case study with revealed preference panel data collected from smart-card usage, where we estimate a route choice model considering the learning process of PT passengers. This type of data is not only real, granular and massive, but can also be applied to study the impact on route choices of all sorts of events that occurred in the past and, therefore, constitutes a new unrivaled source for behavioral analysis.

Most studies of travelers' learning and choice behaviors under uncertain conditions assume ”recency”, that the effect of previous experience decreases at an exponential rate, where more recent experiences have a greater impact on the travelers' memory. Therefore, the updated perceived travel time on day t is a convex combination of the perceived travel time on day $t-1$ and the latest experienced travel time (Lu et al., 2014; Cascetta & Cantarella, 1991; Bogers et al., 2007). However, behavioral decision making literature has suggested that the human memory decreases following a power function instead of at an exponential rate (Lejarraga et al., 2012). Following this theory, Tang et al. (2017) proposed and applied an Instance Based Learning (IBL) model in a route choice model, which captures the recency effects and the power law of forgetting present in travelers' day-to-day learning processes. They applied this model to an experimental data set collected in an hypothetical private route choice laboratory experiment, showing that the IBL model achieves a better fit than a baseline learning model. Our research extends the study of Tang et al. (2017), implementing the IBL model in a PT route choice model using RP data inferred from SC transactions.

In this section, we use the case of Line 6 (L6) of the Santiago (Chile) Metro subway system, which was inaugurated on November 3, 2017. Line 6 has an extension of 15 km (9.3 miles) of track and its 10 Metro stations serve seven municipalities within Greater Santiago (Pineda & Lira, 2019). A new Metro line offers a particularly good opportunity to asses the learning process of PT passengers. A modification to the PT system - such as the addition of a new bus or Metro line - adds a new alternative route that can be attractive for travelers, requiring a new reinforcement learning process. Using revealed preference data, we assess the effect of the new Metro line on passenger behavior and apply an IBL model in a PSL model that accounts for the correlation among PT routes. The revealed preference data is built using three months of SC transaction data: one month prior to the opening of the new Metro line and two months after the opening, representing the actual route choice behavior of passengers who used the dense PT network in Santiago.

The remainder of this section is organized as follows. Subsection 4.2 describes the case study - the Santiago, Chile transit network - and presents an analysis of the observed passengers' behavior. We then discuss the proposed method used and introduce the model estimation results. The last subsection draws conclusions and discusses policy and research implications.

4.2 Data description and analysis

4.2.1 Data description

This study was carried out using passive transport data from the multimodal PT network in Santiago, Chile. This system serves roughly 50% of motorized trips and it is operated by headway scheduling; therefore, lines do not have fixed time schedules. The fare system is fully integrated, with an almost flat fare between urban buses, Metro, and one rail service, allowing up to three trip legs within a two-hour time window. In a typical week, 3 million passengers use the system to make 25.5 million trips. The network includes 7 Metro lines, more than 300 bi-directional transit lines, and one rail service.

Very detailed demand information was obtained from the automatic fare collection (AFC) system in Transantiago (Gschwender et al., 2016). A SC, called bip!, is the only accepted payment method. Passengers must validate when boarding a bus or entering a Metro station, but no alighting validation is required for bus or Metro trips. Around 27% of passengers evade fare payment on buses. Bus stops with particularly high demand have an off-vehicle payment system called *zona paga* (payment zone), where passengers validate when they enter the bus stop area and then board any bus without further validation.

Additionally, in the Santiago PT system, all buses are equipped with GPS devices that record a timestamp and position data every 30 seconds. We have used this Automatic Vehicle Location (AVL) data to obtain the observed frequency of transit lines. Combining SC data and AVL data, the boarding and alighting positions are already estimated for all validations and trips (stages) associated with an origin-destination journey using the methodology proposed by Munizaga & Palma (2012). Using this processed data, it is possible to obtain a large amount of information about a trip, such as the in-vehicle and out-of-vehicle travel time and the number of transfers.

As explained in Section 4.1, in this study we focus on assessing the learning process of passengers after the implementation of the Metro Line 6. As can be seen in Figure 4.1, Metro Line 6 is connected by a transfer station with three other Metro lines, as well as the rail service and in some sections it serves areas of the city that did not have a direct access to the Metro system before. This Metro line was a fully new technology for the city at that time (autonomous trains, air conditioning, platform-edge doors). The analysis of this chapter considered a group of SC data, which correspond to the SC observations from frequent passengers who traveled during morning peak periods (6:30-8:30 AM) on weekdays between origin and destination zones where the new metro line was observed as part of an alternative route. Specifically, we select users that traveled 20 days or more on the PT system during October 2017 (the month prior to the opening of L6), November 2017 (the month that L6 opened), and December 2017 (one month after the opening of L6), and who remained in their destination location for at least two hours. Through this selection, we filter for trips that are probably made to regular, daily activities such as the workplace or school. The morning peak period is the most congested period in Santiago (more than 700 thousand trips per day can be observed during this period), and is therefore an interesting travel period from a behavioral and planning perspective.

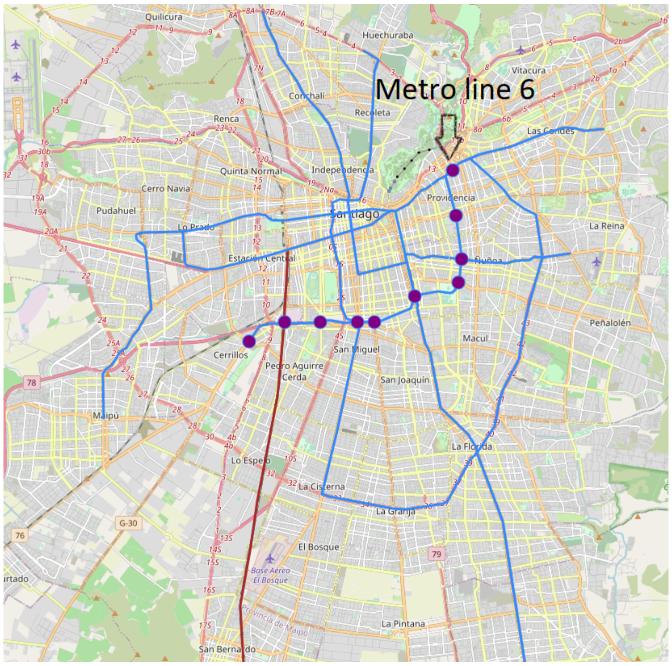


Figure 4.1: Metro network (blue lines) and the rail service (red line) in Santiago, Chile.

4.2.2 Effect of the new Metro line on mobility within Santiago and passengers' behavior

In order to understand the effect of the opening of Metro Line 6 on passengers' journeys, we selected those OD pairs where L6 is part of at least one possible alternative route to travel between them. The identification of these OD pairs was carried out by selecting trips that used L6 in any of trip stage and identifying their origin and destination zones, which we considered to be the area within a 100-meter radius of the trip's origin and destination stops. Once the origin and destination zones where L6 could be a travel alternative were identified, the trips observed in these OD zones during the three months of the study were obtained. Table 4.1, presents the number of observed trips that used L6, and trips that did not use L6, during October, November, and December 2017. No trips used L6 during October, since it was not open at that time. In total, we obtained 99,884 observed trips that used L6, and 84,864 observed trips that did not use L6 in the selected OD pairs during the analysis period. It is important to note that observed trips increased considerably by 157.7% from October to December.

Figure 4.2 shows, for six specific days, the origins of observed trips that used L6. This shows that the passengers who used L6 initiated their trips mainly in the southern sections of the city and that a cluster of more distant origins in the north of the city became present within two days of the opening of L6 (Nov 9th) but were not present by the last day of December. This is probably because the passengers in those areas tried the new alternative, but found it not to be a better alternative than their prior route choices.

Table 4.1: Number of trips made during October, November, and December 2017 in OD zones where L6 can be an alternative route.

Month	Trips that used L6	Trips that not used L6
October	0	31,694
November	44,505	26,870
December	55,379	26,300

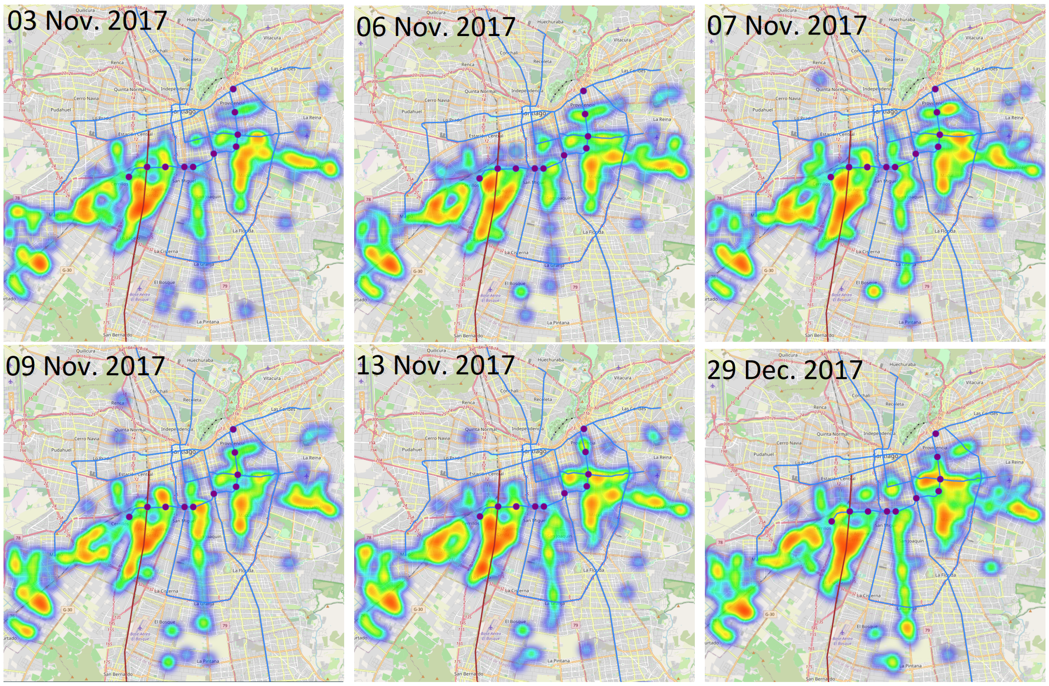


Figure 4.2: Origins of trips that use an alternative including L6

The opening of L6 had a huge impact on travel time savings. In the OD pairs where L6 was observed as an alternative, average travel time (including transfer waiting time, in-vehicle travel time and transfer walking time) decreased from 48.2 min during October (prior to the opening of L6) to 35.7 min in December (one month after the opening of L6). This represents a relative reduction of 25.9%. Figure 4.3 shows the travel time statistics for each week of data during the three months used in this study. L6 started to operate during the fifth week of data, and thereafter, travel time of observed travelers began to decrease throughout the weeks until it reached roughly 35.5 min in the last three weeks of December.

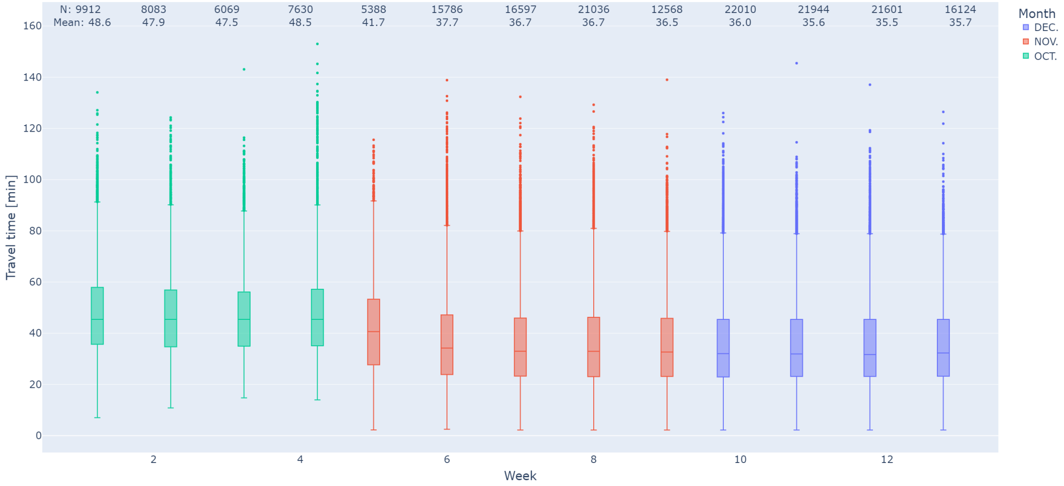


Figure 4.3: Travel time between OD pairs where L6 is observed as part of an alternative route

A similar analysis by transport mode demonstrates that, in the OD pairs where L6 was observed as an alternative, Metro travel time decreased from 27.1 min on average during October to 24.7 min on average in December. In relative terms, this is a reduction of 8.9%. However, the greatest impact was found for bus travel time, which decreased from 18.0 min on average during October to 8.9 min on average in December, a relative reduction of 50.6%. Figure 4.4 shows the bus travel time statistics for each week of data during the evaluated three months. The bus travel time started to decrease in week five until it reached around 9 min in the last three weeks of December. Three reasons can explain this effect: i) after introducing the new metro line, trips using only the metro increased by 404.7% (see Table 4.2), thus increasing the number of trips in which the bus travel time equals zero; ii) after the introduction of the new metro line, trips considering a combination of bus and metro increased by 74.2% (see Table 4.2). Additionally, this type of trip decreased the average bus travel time by 21.2%. The fact that the new metro line came to serve some areas of the city that were not served by any other metro line explains the bus travel time reduction. Since in these cases, passengers must travel less time by bus to access the metro network; iii) focusing on those trips made only by bus, we can observe that the average bus travel time decreased from 41.4 minutes in October to 36.8 minutes in December (see Table 4.3). This can be explained by the fact that the transportation system accompanied the extension of the metro network with changes in the trajectory of bus services to connect the new metro stations more efficiently. These results demonstrate a considerable impact of the new Metro line regarding travel time savings, especially on bus travel time.

It is also important to note that the between-week variation in observed average travel time decreases during the last three weeks of December. A pairwise comparison indicates that the observed average travel times between those weeks are not statistically different (all $ps > 0.5$). Therefore, we can hypothesize that after the opening of Line 6, Metro passengers began to try new travel alternatives before stabilizing their travel behavior during the last three weeks of December.

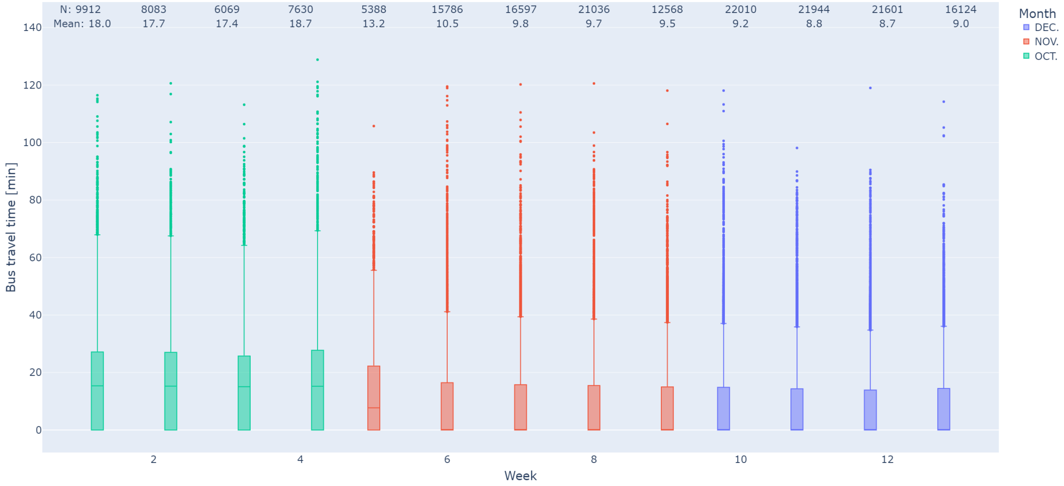


Figure 4.4: Bus travel time between OD pairs where L6 is observed as part of an alternative route

Table 4.2: Number of trips, per transport mode (only bus, only metro, and metro and bus), made during October, November, and December 2017 in OD zones where L6 can be an alternative route.

Month	Bus	Metro	Metro and bus
October	2,471	8,571	20,652
November	2,392	36,726	32,257
December	2,451	43,260	35,968

Table 4.3: Average bus travel time, per transport mode (only bus, and metro and bus), made during October, November, and December 2017 in OD zones where L6 can be an alternative route.

Month	Bus	Metro and bus
October	41.4	22.6
November	38.9	19.5
December	36.8	17.8

Furthermore, in the OD pairs where L6 was observed as an alternative, average total trip stages decreased from 1.7 in October to 1.5 in December, which represents a relative reduction of 11.7%. This is an interesting result, as it indicates the potential of a new Metro line to decrease the number of transfers made by PT passengers. Please note that transfers between Metro lines are not considered in these statistics.

Out of 12,733 PT passengers observed in the database traveled in the OD pairs where L6 was observed as an alternative (at least two observed trips use the Metro line 6), 60.5% used

L6 at least once. Out of that 60.5% of all passengers, 47.3% tried the new Metro line within five days of it opening (see Figure 4.5). Focusing on OD pairs where the route alternative(s) including L6 involve travel time savings compared with alternatives without L6, out of 4,687 PT passengers observed traveling in those OD pairs, 61.0% used L6 at least once. Out of that 61.0% of passengers, 42.7% tried the new Metro line within five days of it opening. On the other hand, focusing on OD pairs where the route alternative(s) including L6 involve increased travel time or no travel time savings occur, compared with alternatives without L6, out of 5,253 PT passengers observed traveling in those OD pairs, 15.6% used L6 at least once. Out of that 15.6% of passengers, 32.8% tried the new Metro line within five days of its opening. These results show that most passengers that tried the new Metro line traveled in those OD pairs where using L6 implied a travel time savings, and most passengers tried the new Metro line during its first week of operation.

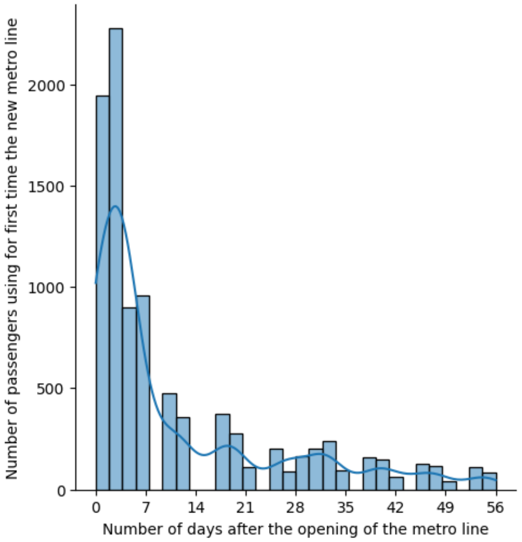


Figure 4.5: Number of passengers using the new Metro line for the first time

4.3 Model Specification, Estimation and Results

In this section, we study the passengers' learning process by developing different discrete choice models that consider a random utility maximization (RUM) approach incorporating the perceived travel time of passengers.

4.3.1 Consideration set construction

Building the consideration set is a significant challenge for formulation and estimation of a PT route discrete choice model from RP data. To build the consideration sets, we first define the origin and destination locations, which we considered as the areas within a 100-meter radius of the origin and destination stops of the observed trip. We used the Historical/Cohort approach to build the consideration set for each OD pair, which corresponds to all available and observed routes used by every passenger traveling within the defined origin and destination areas in the study period (6:30 AM to 8:30 AM). This means that, for each OD pair, the same consideration set was used for all passengers.

An alternative is defined by the boarding stop, transport mode (bus or Metro), and the last alighting stop. Combining different transit lines that serve the same route section into a single transport mode allowed us to consider only those alternative routes that could be clearly distinguished from each other. For example, if two bus transit lines serve the same route section, then they were considered as part of the same alternative.

4.3.2 Route choice modeling

As mentioned in the introduction, most route choice models assume the invariability of path attributes across time, ignoring that route choices are decisions under uncertainties, which generate an important relationship between past experiences and the current choice of passengers.

Therefore, we suggest considering PT passengers' perceptions formed based upon past experiences. SC data offers the opportunity to observe travelers' repeated choices, which allows us to infer the passenger's perception after each experience. With this purpose in mind, we adapt the IBL model proposed by Tang et al. (2017) for a route choice model in a large-scale PT network. The current study considers that the perceived in-vehicle travel time is the only attribute that evolves over time; other attributes (waiting time, number of transfers, path size correction term, and bus constant) are assumed to be constant over time.

The IBL model is based on instance-based learning theory, which was proposed to describe decision making in complex dynamic decision contexts (Lejarraga et al., 2012). In particular, this theory characterizes learning by storing in memory a sequence of actions-outcomes (instances) produced by past experiences, which are more active in memory when they are more recent and frequent (Lejarraga et al., 2012). In a PT route choice context, an instance is a past experience traveling along a route section using a specific transport mode (bus or Metro) and its associated outcome, which can be a set of transport mode attributes.

An alternative route is made up of one or more route sections. Therefore, it is necessary to specify the perceived in-vehicle travel time for each route section. This is because two different routes could share a route section. In this case, if a passenger travels along one of the alternatives, the experienced in-vehicle travel time for the route section common to both alternatives affects the perception of travel time for the non-chosen alternative.

Equation 4.1 shows, for the current day t, the relative weight of the experienced in-vehicle travel time along route section r for passenger n during a past day t'. The denominator is the summation of the activation of all the past experiences in the route section r. In particular, the term $(t-t')^{-\delta}$ measures the recency of the experienced in-vehicle travel time, and captures the rate of forgetting following a power law. Once the relative weight of the experienced in-vehicle travel time has been calculated along route section r for the passenger n during all days prior to t, it is possible to calculate the perceived in-vehicle travel time of route section r on day t for passenger n using Equation 4.2. This formula is a weighted average of experienced in-vehicle travel time of all past days when the passenger traveled along the route section. Finally, the perceived in-vehicle travel time for an alternative route i for the passenger n on day t is calculated as the sum of the perceived in-vehicle travel times along all route sections of the alternative, as shown in Equation 4.3.

$$W_{nr}(t,t') = \frac{a_{nr}(t')(t-t')^{-\delta}}{\sum_{\tau \in H_{nr}(t)} (t-\tau)^{-\delta}} \tag{4.1}$$

Where:

t: current day
t': a previous day
δ: decay parameter that captures the rate of forgetting
$H_{nr}(t)$: the set of all days before day t when the passenger n traveled along the route section r
$a_{nr}(t')$: a binary indicator equal to 1 if passenger n traveled along the route section r on day t', and 0 otherwise

$$T_{nr}(t) = \sum_{t' \in H_{nr}(t)} W_{nr}(t,t') X_{nr}(t') \tag{4.2}$$

Where:

$X_r(t')$: experienced travel time along route section r for passenger n on day t'

$$PT_{ni}(t) = \sum_{r \in \zeta_i} T_{nr}(t) \tag{4.3}$$

Where:

ζ_i: set of route sections belonging to route i

As an illustrative example, Table 4.4 shows an application of the IBL model using an observed passenger (smart card id 23145602) in the studied database. During the period of analysis, this passenger traveled during 5 work days along a specific route section using a bus transit line. The weight of experienced in-vehicle travel time for each day is calculated following Equation 4.1, and the perceived in-vehicle travel time for the route section for each day is calculated using Equation 4.2. Both formulas are calculated in four cases: i) $\delta = 0$, ii) $\delta = 0.5$, and iii) $\delta = 2$, and iv) $\delta = 3$. The initial perception of the bus transit line is gleaned with its first use, on November 9, i.e. 12.45 min. On November 16, the passenger had one past instance experienced, and her perception was 12.45 min for the three values of the rate of forgetting. On November 20, there were two instances, 12.45 min and 10.22 min, with a rate of forgetting equal to 0 resulting in an average value (11.33) between two previous experiences, while the a rate of forgetting equal to 3 resulting in a value (10.32) closer to the last instance. This situation is also observed on December 11, 18, and 20, where the δ value equal to 2 or 3 obtained the perceived in-vehicle travel time closer to the last experience. These results can be explained because, as it can be seen in Table 4.5, a smaller δ values translate to higher memory activation, where more distant experiences are more relevant, while greater δ values translate to lower memory activation, where more recent experiences are more relevant.

Table 4.4: Application of the IBL Model in the observed trips during 6 days of a passenger

		Perceived in-vehicle travel time			
Day	Experienced travel time	$\delta = 0$	$\delta = 0.5$	$\delta = 2$	$\delta = 3$
09-11-2017	12.45				
16-11-2017	10.22	12.45	12.45	12.45	12.45
20-11-2017	11.03	11.33	11.06	10.48	10.32
11-12-2017	9.63	11.23	11.18	11.05	10.99
18-12-2017	9.93	10.83	10.54	9.81	9.68
20-12-2017	11.40	10.65	10.28	9.93	9.93

Table 4.5: Weight of experienced travel time to calculate the perceived travel time on 21-12-2017 of a specific passenger

	Weight of experienced travel time on day 21-12-2017			
Day	$\delta = 0$	$\delta = 0.5$	$\delta = 2$	$\delta = 3$
09-11-2017	17%	6%	0%	0%
16-11-2017	17%	7%	0%	0%
20-11-2017	17%	7%	0%	0%
11-12-2017	17%	13%	1%	0%
18-12-2017	17%	24%	10%	4%
20-12-2017	17%	42%	89%	96%

A Path Size Logit (PSL) model accounts for correlation between alternatives due to overlapping route segments (Ben-Akiva & Bierlaire, 1999) and is adopted in this study. The deterministic component of the model is specified in Equation 4.4, where i represents the alternative route, n represents the passenger, t the evaluated day, $TT_i(t)$ is the in-vehicle travel time for day t, WT_i is the waiting time (including the waiting time at the beginning of the trip and during transfers), TR_i is the number of transfers between vehicles, and PSC_i is the path size correction term.

$$V_{ni}(t) = \beta_{TT}TT_{ni}(t) + \beta_{WT}WT_i + \beta_{TR}TR_i + \beta_{PS}PS_i \tag{4.4}$$

We estimated two types of PSL models. A base model assumes that the passenger uses knowledge about the in-vehicle travel time from descriptive information only, and represents the in-vehicle travel time with the average in-vehicle travel time of the alternative, obtained from observed trips prior to the evaluated trip, $MEANT_i(t)$, neglecting the learning process. In contrast, our proposed model represents the in-vehicle travel time with the perceived in-vehicle travel time of the alternative by the passenger considering the learning process. This second type of model, called IBL-PSL model, assumes that the passenger has a perceived in-vehicle travel time when she has tried the alternative route in the past. If she has no previous experience using the alternative route i, she will look for descriptive information, which can be represented with the average in-vehicle travel time, up to the day t, $MEANT_i(t)$. Following this assumption, in-vehicle travel time can vary between days and is calculated following Equation 4.5. $PT_{ni}(t)$ is the perceived in-vehicle travel time, calculated using Equation 4.3, $MEANT_i(t)$ is the average in-vehicle travel time of the alternative, calculated using all observed trips up to day t with this alternative in the SC database, not only those performed by individual n. Besides, $D_{ni}(t)$ takes a value of 1 if passenger n ever traveled on alternative i prior to t and a value of 0 otherwise.

$$TT_{ni}(t) = PT_{ni}(t) * D_{ni}(t) + MEANT_i(t) * (1 - D_{ni}(t)) \tag{4.5}$$

For the path size correction term, we use the expression according to Bovy et al. (2008) in Equation 4.6, where L_r is the length of route section r, L_i is the length of route i, ζ_i is the set of route sections belonging to route i, and δ_{rk} is the route section-route incidence number, which takes a value of 1 if route k uses route section r and a value of 0 otherwise.

$$PSC_i = \sum_{r \in \zeta_i} \frac{L_r}{L_i} \ln \frac{1}{\sum_{k \in C_p} \delta_{rk}} \tag{4.6}$$

Waiting time is estimated from the observed interval between buses with an assumed exponential distribution. Therefore, it is calculated as 1 divided by the observed frequency of the line (or sum of frequencies, in the case of more than one line traveling through the same route section). The exponential distribution (or the gamma, which is the sum of exponentials) has been widely used for modeling wait time in transportation systems (e.g. Nguyen & Pallottino, 1988; Raveau & Muñoz, 2014; Schmöcker et al., 2013).

With a Gumbel distribution assumed of the error term of the utility function, the probability of passenger n choosing alternative i during day t, given the consideration set C, is expressed as shown in Equation 4.7. The closed-form of this formula allows for a simple estimation of the fixed coefficients by maximizing the likelihood function.

$$P_{ni}(t) = \frac{\exp V_{ni}(t)}{\sum_{k \in C} \exp V_{nk}(t)} \tag{4.7}$$

4.3.3 Consideration set results

The alternative routes were generated using the Historical/Cohort approach for each OD pair. We worked with OD pairs that featured more than one alternative; OD pairs with fewer alternatives were discarded. Using this filter, we obtained 988 OD pairs with an average of 3.11 available alternatives within the consideration set.

The overlap between alternative routes was evaluated at the stop level, which means that the alternatives that used the same route sections and transport mode were considered as correlated elements. It is important to note that the Path Size correction term was obtained with equation 4.6. In order to compare the correlation of each link with other alternative routes, we derived transit line trajectories from GTFS data. In the case of Metro, since we do not know the trajectory used by passengers, we had to assume that passengers took the route with the lowest travel time. The Path Size correction term (PSC_i) belongs to the interval $(-\infty, 0]$, takes a value of 0 when there is no overlap between the evaluated alternative and others, and the value decreases as the level of route correlation increases; that is, the smaller the PSC_i value (more negative), the higher the correlation of route i. Choice sets have a mean PSC_i of -0.61.

In order to evaluate how long it takes for passengers to get a stable consideration set after trying the new Metro line, we define the *stability* as the point in which the consideration set C_t of all alternatives used by an individual n up to time t, does not differ to the consideration set C_{t-1} up to time $t-1$. It is needed to constraint this analysis to those individuals that traveled on enough occasions, until reaching some type of *stability.* For this reason, we first select only those passengers who traveled 20 times or more within the same OD pair, and who traveled at least five times prior to the opening of L6. Using this filter, 1,573 passengers were obtained, and 599 of them tried the new Metro line. To evaluate the composition of the consideration set across time, for each passenger and OD pair, we defined the following terms:

- Period 0: Last day that the passenger traveled in the transport system prior to the opening of Metro Line 6.
- Period 1: First day that the passenger traveled in the transport system after the opening of Metro Line 6.
- Period m: $m-th$ day that the passenger traveled in the transport system after the opening of Metro Line 6.

For each passenger n, OD pair o, and period p, the observed consideration set consists of the observed alternatives used by passenger n in period p, and in the last four trips prior to period p. In this line, for each day, we used the prior four observed trips and the observed trip during that day. Then, we calculate the indicator I_{nop} for each passenger n, OD pair o, and period p. This indicator is shown in the Equation 4.8, and takes a value of 1 if the consideration set in the evaluated period is equal to the consideration set in the previous period and a value of 0 otherwise.

$$I_{nop} = \begin{cases} 1\ if\ C_t\ =\ C_{t-1} \\ 0\ otherwise \end{cases} \tag{4.8}$$

We assume that a passenger obtains a stable consideration set in the first period with indicator I_{nop} equal to 1, and that this value never changes in periods after the evaluated period. This process can be represented executing the Algorithm 1 for each passenger n and OD pair o. If the algorithm returns the value of P, which is the day of the last observed trip, this means that the passenger could not obtain a stable consideration set within the study period. Table 4.6 shows an illustrative example with a hypothetical case where a passenger traveled 5 times prior to the opening of L6 (Days -4, -3, -2, -1, 0) using alternative A or alternative B. After the opening of L6, the passenger traveled 10 times, and only used alternative C. Using the algorithm to find the period where the passenger obtained a stable consideration set, we obtain the period 5, which means that the user defined a stable consideration set during the first 5 trips after the implementation of the new Metro line.

As can be seen in Figure 4.6, out of the 599 frequent passengers who used L6, 43.7% identified a stable consideration set during the first 5 trips after the opening of L6, and 64.9% did so during the first 10 days of travel after the opening of L6.

Algorithm 1 Stabilization period of consideration set

```
procedure
    $P \leftarrow$ Number of periods where the passenger has traveled
    $i \in \{0, 1, ..., P\}$
    $i \leftarrow 1$
loop:
    $product \leftarrow \prod_{j=i}^{P} I_{noj}$
    if $product = 1$ then
        $Return\ i - 1$
        end procedure;
    else
        $i \leftarrow i + 1$
        go to loop.
    end if
end procedure
```

Table 4.6: Example of a passenger traveling between a specific OD pair prior to and after the opening of L6

Day	Period	Observed alternative	Observed consideration set	Indicator I_{nop}
-4	-	A	-	-
-3	-	A	-	-
-2	-	B	-	-
-1	-	B	-	-
0	0	B	{A, B}	-
1	1	C	{A, B, C}	0
2	2	C	{B, C}	0
3	3	C	{B, C}	1
4	4	C	{B, C}	1
5	5	C	{C}	0
6	6	C	{C}	1
7	7	C	{C}	1
8	8	C	{C}	1
9	9	C	{C}	1
10	10	C	{C}	1

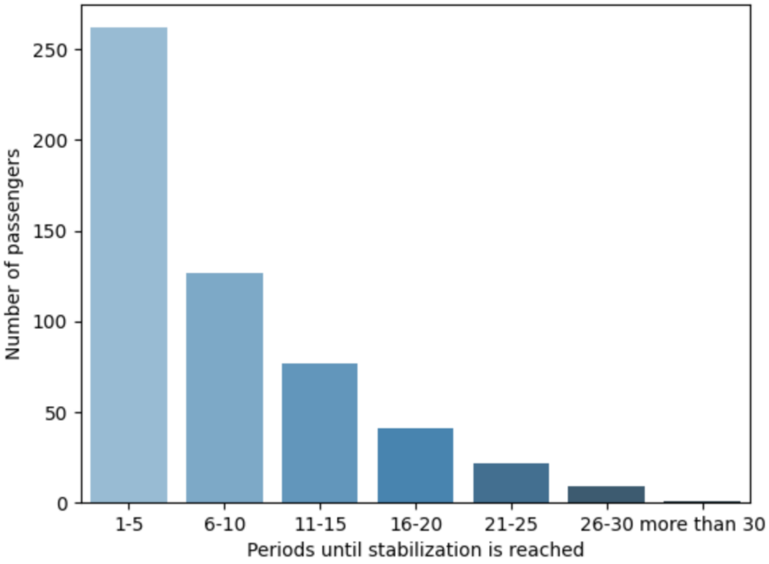


Figure 4.6: Number of days required for passengers to obtain a stable consideration set, after trying the new Metro line

4.3.4 Model estimates

PSL models were estimated using observations from 8 weeks: 4 weeks from November and 4 weeks from December. These observations correspond to trips made during 35 weekdays (exceding holidays, Saturdays, and Sundays) after the opening of Metro Line 6 during the morning peak period (between 6:30 and 8:30 AM) in November and December 2017. These trips were made by 1,826 frequent passengers who traveled 20 times or more between the same OD pairs during October, November, and December. It is important to note that this research works with trip data from passengers that traveled between origin and destination zones where the new metro line was observed as a part of an alternative route. The PSL model was estimated in two cases, using the IBL model to represent the perceived in-vehicle travel time, which captures the passengers' learning process, and using the mean in-vehicle travel time (ignoring the passengers' learning process). It is important to note that for an specific trip observation, the perceived in-vehicle travel time and the mean in-vehicle travel time, for each route alternative in the consideration set, is calculated using all observed trips (in the same route alternative) prior to the day where the evaluated trip is observed[1]. For this purpose, all trip experiences from October, November, and December 2017 are considered.

The specification of the deterministic utility function considers in-vehicle travel time, including bus and Metro; waiting time, considering the waiting time at the beginning of the trip and during transfers; and number of transfers, considering bus to bus, bus to Metro, and Metro to bus transfers. Metro to Metro transfers cannot be incorporated into the model because the information about the route that the passenger used inside the Metro network is

[1] It is important to note that when a trip observation has a consideration set with a route alternative that has not been previously used by any passenger this observation is not used for the estimation of the model, since in this case the mean in-vehicle travel time with previous observations can not be calculated.

not available. Table 4.7 shows the mean, standard deviation, minimum, and maximum values for each attribute considered in the models.

Table 4.7: Statistics of attributes used in PSL models

Attribute	Mean	Std	Min	Max
In-vehicle travel time [min]	46.0	17.4	5.4	116.3
Wait time [min]	7.9	4.1	2.1	30.5
Number of transfers	0.9	0.5	0.0	2.0
Path Size term	-0.5	0.5	-2.1	0.0

Table 4.8 shows the estimated parameters for the PSL model and IBL-PSL model using observations from the four weeks of November. The model that uses the perceived in-vehicle travel time (IBL-PSL model) with data from the week right after the opening of the new metro line obtained parameters that are all statistically significant (except for the path size correction term). However, surprisingly, the sign of the rate of forgetting parameter is not coherent with the instance-based learning theory, which is based on the idea that more recent experiences are more active in the human memory and consequently the rate of forgetting parameter should have a positive sign. It might be the case that people tend to increase exploration in a new context of the PT system, which moved choice behavior toward random choices (Erev & Barron, 2005), which can be reflected in a non-expected sign for the rate of forgetting parameter. Likewise, it might be the case that the IBL model does not suit well for a few instances. During the first week of November, passengers did not have many experiences in the new metro line, and therefore the IBL model can not well capture the effect of previous passengers' experiences on their route choice behavior. For these reasons, The first week right after the opening of the new metro line was not used for the following behavior analysis.

The other models that uses the perceived in-vehicle travel time (IBL-PSL model) reported parameters that are all statistically significant (at the habitual 5% threshold) with the expected sign, with the exception of the rate of forgetting parameter, which is not statistically significant in the models of the second and third weeks, and the path size correction term parameter, which is not statistically significant in the models of all weeks evaluated. These results are similar to the models that use the mean in-vehicle travel time, where the parameters are statistically significant and with the expected sign, with the exception of the path size correction term parameter, which is not statistically significant in the PSL models of all evaluated weeks.

Comparing the results of the IBL-PSL models and the PSL models, they present a similar model fit for all evaluated weeks of November. In particular, for week 2, the PSL model presents a slightly better fit than the IBL-PSL model, while for week 4, the IBL-PSL model presents a slightly better fit than the PSL model. Consequently, the inclusion of the perceived in-vehicle travel time in the specification of the models does not result in greater improvement explanatory power compared with models that do not consider the variation in passengers' perception of in-vehicle travel time. Considering this result and that the rate of forgetting parameter is not statistically significantly different from zero during the second and third evaluated weeks in November, we can argue that during the first weeks after the implementation of a new Metro line in the PT system, passengers mainly use descriptive information, which is represented by the mean in-vehicle travel time, to make a route choice decision. However, during the last week of November, the rate of forgetting parameter starts to be statistically significant (at the accepted 5% threshold). Considering this result and that the IBL-PSL model presents a slightly better fit than the PSL model for this week, we can argue that during the last week of November passengers start to consider the perceived in-vehicle travel time to make a route choice decision.

Table 4.9 shows the estimated parameters for the PSL model and IBL-PSL model using observations from each week in December. The model that uses the perceived in-vehicle travel time (IBL-PSL model) reported parameters that are statistically significant (at the accepted 5% threshold) with the expected sign, with the exception of the path size correction term parameter, which is not statistically significant for the model of three first week in December, and it is statistically significant with a negative sign in the model of week 7. These results are similar to the models that use the mean in-vehicle travel time, where the parameters are statistically significant and with the expected sign, with the exception of the path size correction term parameter, which is statistically significant with a negative sign in the PSL models of week 7.

Comparing the results of the IBL-PSL models and the PSL models, during all weeks of December, the IBL-PSL model presents better model fit. Consequently, in this month, the inclusion of the perceived in-vehicle travel time in the specification of the models does result in greater explanatory power compared with models that do not consider the variation in passengers' perception of in-vehicle travel time. In summary, these results allow us to hypothesize that during the month after the implementation of a new Metro line in the PT system, passengers mainly use their perception from past experience to make a route choice decision.

In summary, these results suggest that during the first weeks of a new context in a PT system, the passengers' route choice behavior is different from their route choice behavior for some weeks afterward, when they start to consider experience-based information to make route choice decisions. To prove this, we have estimated a constrained model, in which the parameters of the seven evaluated weeks are forced to be the same, and an unconstrained model, in which the parameters of the second and third weeks of November can be different from the parameters of the last five evaluated weeks. The results are shown in Table 4.10, where the IBL-PSL and PSL models reported similar results to the previous analysis. Then, using the log-likelihood of the constrained IBL-PSL model and the log-likelihood of the unconstrained IBL-PSL model, the application of a formal likelihood ratio test largely rejects the null hypothesis that the models are equal ($p - value < 1\%$). This shows that during the second and third weeks of November passengers evaluate all their experiences in the same level of importance for a route choice decision, while for the last five evaluated weeks, more recent experiences became more relevant for the PT passengers' route choice decisions.

Table 4.10 also shows that using all evaluated weeks, the IBL-PSL model is superior to the PSL model by type on in-sample, confirming that, the inclusion of the perceived in-vehicle travel time in the specification of a route choice model results in greater explanatory power compared with models that do not consider the variation in passengers' perception of in-vehicle travel time.

It is important to note that the negative coefficients of the in-vehicle travel time, initial waiting time, transfer waiting time and number of transfers parameters indicate that these attributes are perceived as disutility, which is in line with previous PT route choice studies (Nassir et al., 2018; Schmöcker et al., 2013; Raveau et al., 2011, 2014; Rui, 2016; Jánošíková et al., 2014; Z. Guo, 2011). The memory decay parameter (δ) is estimated between 1.3 and 3.6 during the last four evaluated weeks. These values are higher than values reported by Tang et al. (2017) using experimental data in a car route choice context. This means that PT passengers present a smaller activation in memory than car drivers. Therefore, more recent experiences are more relevant for PT passengers' perception than for car drivers' perception.

On the other hand, it should be noted that the negative value for the path size is not uncommon in the literature about PT users' route choices. In fact, although the path size is supposed to correct for overlapping routes by reducing the utilities of overlapping routes, negative estimates for path-size terms have been found (Anderson et al., 2017; de Grange et al., 2012), most likely because of the additional utility of travelers having more opportunities to reach their destination from their origin, or because travelers might value the availability of a large number of en-route alternative options over the uniqueness of the route (Anderson et al., 2017).

Table 4.8: PSL model and PSL-IBL model estimates (t tests) using four weeks of data from November 2017

Parameters	Week 1		Week 2		Week 3		Week 4	
	IBL-PSL model	PSL model	IBL-PSL model	PSL model	IBL-PSL model	PSL model	IBL-PSL model	PSL model
In-vehicle travel time	-0.114(-15.1)	-0.113(-14.8)	-0.1 (-17.5)	-0.103 (-17.7)	-0.107 (-22.7)	-0.11 (-22.6)	-0.115 (-21.7)	-0.118 (-21.7)
Rate of forgetting (δ)	-3.198(-2.7)		0.956 (1.8)		-0.1 (-0.2)		1.256 (3.9)	
Waiting time	-0.05(-4.2)	-0.043(-3.6)	-0.074 (-8.3)	-0.072 (-8.1)	-0.096 (-12.9)	-0.094 (-12.8)	-0.094 (-11.6)	-0.093 (-11.5)
N° transfers	-2.447(-11.3)	-2.36(-11.2)	-1.897 (-12.5)	-1.926 (-12.6)	-2.056 (-16.6)	-2.05 (-16.6)	-2.204 (-15.7)	-2.235 (-15.9)
PSC	0.114(0.6)	0.112(0.6)	0.209 (1.5)	0.242 (1.8)	0.009 (0.1)	0.03 (0.3)	-0.005 (0)	0.015 (0.1)
Bus constant	-0.891(-3.1)	-0.974(-3.4)	-1.599 (-6.4)	-1.551 (-6.2)	-1.318 (-6.9)	-1.311 (-6.9)	-1.233 (-5.8)	-1.187 (-5.6)
Log-likelihood	-856.344	-865.802	-1569.088	-1568.13	-2356.907	-2356.789	-1997.979	-2007.316
Adjusted rho-square	0.521	0.517	0.488	0.489	0.492	0.492	0.500	0.498
AIC	1724.688	1741.606	3150.176	3146.261	4725.814	4723.577	4007.957	4024.633
BIC	1758.055	1769.411	3186.413	3176.458	4764.371	4755.708	4045.6	4056.002
N° observations	1922		3101		4565		3920	
N° of days	4		4		5		4	
N° of passengers	894		1203		1451		1497	

All columns show t-values between parentheses. AIC = Akaike Information Criterion and BIC = Bayesian Information Criterion. Week 1: November 06, 07, 09, and 10. Week 2: November 13, 14, 15, and 16. Week 3: November 20, 21, 22, 23, and 24. Week 4: November 28, 29, and 30, and December 1.

Table 4.9: PSL model and PSL-IBL model estimates (t tests) using four weeks of data from December 2017

Parameters	Week 4		Week 5		Week 6		Week 7	
	IBL-PSL model	PSL model	IBL-PSL model	PSL model	IBL-PSL model	PSL model	IBL-PSL model	PSL model
In-vehicle travel time	-0.125 (-23.5)	-0.124 (-22.7)	-0.138 (-28.4)	-0.123 (-25.8)	-0.149 (-29.3)	-0.133 (-27.2)	-0.152 (-26.7)	-0.124 (-23.3)
Rate of forgetting (δ)	2.183 (5.7)		3.561 (5.8)		2.35 (7.8)		2.939 (7.1)	
Waiting time	-0.106 (-13.4)	-0.104 (-13.3)	-0.107 (-14.9)	-0.099 (-14.4)	-0.112 (-16)	-0.106 (-15.7)	-0.124 (-15.2)	-0.113 (-14.8)
N° transfers	-2.291 (-16.9)	-2.288 (-17)	-2.432 (-20)	-2.425 (-20.4)	-2.641 (-21.5)	-2.612 (-21.9)	-2.538 (-19)	-2.557 (-19.8)
PSC	-0.004 (0)	0.002 (0)	-0.1 (-1)	-0.182 (-1.8)	0.122 (1.2)	0 (0)	-0.25 (-2.1)	-0.4 (-3.6)
Bus constant	-1.243 (-5.9)	-1.228 (-5.8)	-1.172 (-5.9)	-1.209 (-6.2)	-0.92 (-4.7)	-0.966 (-5)	-1.195 (-5.2)	-1.225 (-5.5)
Log-likelihood	-2085.221	-2123.883	-2607.526	-2753.545	-2605.122	-2757.049	-2033.505	-2220.127
Adjusted rho-square	0.505	0.496	0.514	0.487	0.519	0.491	0.534	0.492
AIC	4182.442	4257.766	5227.052	5517.089	5222.243	5524.097	4079.01	4450.253
BIC	4220.42	4289.414	5266.42	5549.896	5261.61	5556.903	4117.064	4481.965
N° observations	4145		5226		5225		4198	
N° of days	4		5		5		4	
N° of passengers	1561		1667		1690		1651	

All columns show t-values between parentheses. AIC = Akaike Information Criterion and BIC = Bayesian Information Criterion. Week 5: December 4, 5, 6, and 7. Week 6: December 11, 12, 13, 14, and 15. Week 7: December 18, 19, 20, 21, and 22. Week 8: December 26, 27, and 29.

Table 4.10: Constrained and unconstrained models estimates (t tests)

	Constrained model		Unconstrained models			
Parameters	From week 2 to 8		Weeks 2 and 3		Last five weeks	
	IBL-PSL model	PSL model	IBL-PSL model	PSL model	IBL-PSL model	PSL model
In-vehicle travel time	-0.126(-64.9)	-0.12(-61.3)	-0.104(-28.8)	-0.107(-28.8)	-0.135(-58.3)	-0.125(-54.2)
Rate of forgetting (δ)	2.379(15.8)		0.247(0.8)		2.587(14.5)	
Waiting time	-0.103(-35.2)	-0.098(-34.5)	-0.087(-15.2)	-0.085(-15)	-0.108(-31.8)	-0.103(-31.2)
N° transfers	-2.306(-46.8)	-2.326(-47.5)	-1.99(-20.7)	-1.999(-20.8)	-2.428(-41.9)	-2.438(-42.7)
PSC	-0.022(-0.5)	-0.061(-1.5)	0.087(1)	0.11(1.3)	-0.044(-0.9)	-0.114(-2.4)
Bus constant	-1.212(-15.4)	-1.207(-15.3)	-1.425(-9.4)	-1.401(-9.3)	-1.138(-12.3)	-1.147(-12.4)
Log-likelihood	-15352.67	-15822.29	-3930.74	-3928.37	-11364.81	-11874.39
Adjusted rho-square	0.507	0.492	0.490	0.491	0.514	0.492
AIC	30717.34	31654.58	7873.49	7866.77	22741.63	23758.79
BIC	30767.27	31696.19	7915.15	7901.49	22789.81	23798.94
N° observations	30380		7666		22714	
N° of days	31		9		22	
N° of passengers	1826		1482		1823	

All columns show t-values between parentheses. AIC = Akaike Information Criterion. BIC = Bayesian Information Criterion.

4.4 Conclusions

Modeling route choice decisions is a typical example of an experience-based decision making process, where travelers choose a route alternative, gain experience with the route, and update their perception of uncertain attributes such as travel time. Even when this learning process based on past experiences plays an essential role in the understanding of passengers' behavior, most PT route choice studies ignore this process and assume that uncertain attributes do not vary across time. We try to fill this literature gap, using the Instance Based Learning (IBL) model proposed by Tang et al. (2017) to represent the perceived in-vehicle travel time of passengers in a large-scale multimodal PT system.

In this study, we use the implementation of Metro Line 6 in the Santiago Metro to asses passengers' learning process within a new PT system context. The aim of this study is to understand passengers' learning process when they face a new alternative route between an OD pair where they usually travel. To evaluate this, we used data from the PT system in Santiago, Chile, from October, November, and December 2017; one month prior to the launch of Metro Line 6, and two months after it was opened. These observations were used to calculate route alternative attributes, such as the mean in-vehicle travel time, perceived in-vehicle travel time, and wait time. Additionally, these observations were used to build the consideration set following the Historical/Cohort approach for each OD pair. Only observations from November (after November 3) and December were used in the estimation process, since we aimed at capturing travelers' behavior after the opening of the new Metro line. To do so, we split the database in eight parts: four groups of observations in November and four groups of observations in December (one group for each week in each month). Using each week of observations, we analyzed two models: (i) the IBL-PSL model, which considers the perceived in-vehicle travel time and mean in-vehicle travel time, assuming that passengers make route choice decisions based on experience information, if they have previously used the alternative, and descriptive information, if they have not previously used the alternative, and (ii) the PSL model, which considers in-vehicle travel time as a static attribute and assumes that passengers make route choice decisions using purely descriptive information.

The results suggest that both types of models can be used to represent the perception of passengers for in-vehicle travel time, wait time and number of transfers. The IBL-PSL model obtained a better fit than the PSL model when examining data from the last week of November and during all weeks of December. When using data from the second week after the implementation of Metro Line 6, the PSL model obtained a better fit than the IBL-PSL model. These results suggest that during the first weeks of a new context in a PT systems, the passengers mainly use descriptive information to select a route, while some weeks afterwards, they start to consider experience-based information to make route choice decisions. In summary, one of the more significant empirical findings to emerge from this study is that passengers use information from past experiences to select a route alternative after some initial experiences, and they maintain recent experiences more active in their memory compared with older experiences.

Finally, we use all evaluated weeks to estimate the IBL-PSL model and the PSL model. The results confirm that the inclusion of the perceived in-vehicle travel time in the specification of a route choice model results in greater explanatory power compared with models that do not consider the variation in passengers' perception of in-vehicle travel time.

This research provides a new case study, using real world data from the PT system of Santiago, Chile, to apply the IBL model, which is a travelers' learning model. Still, many challenges remain in this line of research, which are discussed below. First, the IBL model can be applied to the formation of the consideration set. This means that the Historical/Cohort approach to generate the consideration can be improved by considering the passengers' learning process to include or remove some alternatives. Second, the learning model applied in this research assumes that passengers use only one type of information to make a route choice, either experience information or descriptive information; however, it is interesting to evaluate if passengers consider both types of information in the same route choice situation. In this context, some questions should be answered by future studies, such as what type of real-time information PT passengers use to select a route. Depending on the answer to this question, a follow-up question is how descriptive information can be combined with experience-based knowledge to form the passengers' perception of travel time.

Chapter 5

The effect of economic incentives and cooperation messages on user participation in crowdsourced public transport technologies

The main contribution of this chapter is to show that community-oriented passenger information technologies can be a tool to capture mobility information that is missing in passive PT data. This study contributes to the fourth research question (see section 1.4.2): How can passengers be encouraged to provide mobility and transport information through crowdsourcing applications oriented to PT?. The contribution of this study is to examine the effect of economic incentives (a lottery for free trips) and cooperation messages (asking users to help the community) to encourage users to share reports about bus stop conditions using a crowdsourcing app. We found that offering an economic incentive increased the participation rate almost three times compared to a control group, which did not receive any message. This positive effect lasted for several weeks but decreased over time, especially for users who had not made reports prior to the experiment. This incentive also increased the number of reports shared by users as well as the coverage of bus stops. Using a cooperation message, with or without the economic incentive, also increased the participation rate compared to the control group, but adding a cooperation message decreased the effect of a standalone economic incentive.

This chapter was published in the following article:

Arriagada, J., Mena, C., Munizaga, M., & Schwartz, D. (2022). The effect of economic incentives and cooperation messages on user participation in crowdsourced public transport technologies. Transportation, 1-28.

Author's contribution

All authors contributed to the formulation of the study goal, methodology, formal analysis, and the research and investigation process. **Jacqueline Arriagada and Claudio Mena** performed the data collection process and creation of the initial draft. **Claudio Mena and Daniel Schwartz** performed the design and implementation of the computer code for the analysis. The revision and edition of the final draft were performed by **Jacqueline Arriagada, Marcela Munizaga, and Daniel Schwartz**.

5.1 Introduction

A key aspect for planning and operating transport systems is the availability of mobility data, which is essential for network design, operation optimization, coverage assessment, and service quality, among other essential tasks. For many years, both transport planners and transport researchers have relied mainly on traditional survey data to collect information about travel patterns and user perceptions, as well as physical inspections to monitor infrastructure. However, these methods are generally expensive, and they do not achieve adequate spatiotemporal coverage, which requires a significant undertaking. To deal with these disadvantages, in recent years, there has been an increased interest in new transport data collection methods based on sources such as GPS devices and smartphone devices (Bonnel & Munizaga, 2018). In particular, crowdsourcing applications have become a significant data source based on information shared by users to make transport information available for commuters and transport system planners (Nandan et al., 2014; Hong et al., 2020; Mondschein, 2015). These applications typically gather automatic location data to provide bus arrival times (Lau et al., 2011; Zhou et al., 2012; Steinfeld et al., 2011) and add user-reported information about the PT system (Steinfeld et al., 2011; Faber & Matthes, 2016). For example, crowdsourcing mobile applications, such as Moovit, Tiramisu, and Transapp, provide bus arrival times, and request their users to report bus overcrowding levels and whether buses and bus stops are in poor condition and in need of repair.

Crowdsourcing applications, which are voluntary participatory information systems, require a critical mass of users willing to provide information to be useful. However, these applications generally suffer from low participation rates that sometimes hover close to zero (Ling et al., 2005). For example, in 2013, Waze, the worldwide car crowdsourcing app, had 50 million users globally, but only 0.01% sent reports about detours or other traffic information (Weitzenkorn, 2013). This is also a problem for PT crowdsourcing systems, in which planners require widespread active participation (Zimmerman et al., 2011).

The phenomenon of low participation rates in voluntary information systems was summarized by J. Nielsen (2006), who defined the 90-9-1 rule. This rule states that 90% of users behave as lurkers –they benefit from the contributions of others, but never contribute themselves–, 9% contribute occasionally, and 1% actively contribute. This distribution implies that a very small fraction of users not only generates most of the contributions, but also leads to a skewed representation of the users. This is problematic for many voluntary information systems, such as crowdsourcing apps, online communities and online review of products and services. For example, if a crowdsourcing information system for PT receives user feedback regarding buses and bus stop conditions, and only a small self-selection of users contribute, it is probable that large areas of the city will lack information, making the platform less useful both for users and PT planners.[1]

[1] Other examples include Wikipedia, in which only 0.2% of active US visitors are active contributors

Due to this problem, a handful of crowdsourcing transport applications have tried to increase participation and contribution rates using elements of gamification, such as avatars and badges (Faber & Matthes, 2016). Most of this research has been conducted using survey or lab studies, with very small research samples and with qualitative measures (Hamari et al., 2014), limiting its application to broader information systems. Other studies have used "quid pro quo" techniques to limit app usage to those who contribute. For example, Tomasic et al. (2014) motivated users to share information about bus arrival times and onboard conditions, such as seat availability, by making such information available only to contributors. This study found that, despite increasing contribution, a "quid pro quo" approach increased the likelihood of users abandoning the crowdsourcing app altogether. On the other hand, simply asking users to contribute did not increase participation rates. In general, research on transport-oriented crowdsourcing applications has offered little discussion regarding how to encourage new users to participate in these new data collection technologies.

The current research aims to motivate contribution in a transport crowdsourcing technology using economic rewards and cooperation messages. First, economic incentives have been used in many public policy domains in order to motivate socially beneficial behaviors, such as donating blood (Lacetera et al., 2014) or recycling (Schwartz et al., 2021; Córdova et al., 2021). In the field of transportation, economic incentives have been used to promote more sustainable transport modes (Bamberg & Schmidt, 2001; Jakobsson et al., 2002; Rosenfield et al., 2020; Thøgersen & Møller, 2008), to motivate car drivers to avoid rush hours (Ben-Elia & Ettema, 2011b,a), and to collect mobility data with traditional surveys (Zumkeller et al., 2011; Hoogendoorn-Lanser et al., 2015). However, economic incentives have also been shown to have backfiring effects. For example, Hilton et al. (2014) showed that offering an economic incentive may reduce preferences for taking the most environmentally friendly mode of transportation for an intercity trip.[2]

Second, previous research has also shown that individuals are willing to cooperate with others even if they could free ride. This has been explained by altruistic preferences, and in particular by warm-glow altruism; i.e., people have been shown to have altruistic preferences in order to feel good about themselves (Andreoni, 1990, 1993; Andreoni & Miller, 2002). For example, contributors to Wikipedia have reported that one of the most relevant reasons to cooperate is due to altruistic factors (Nov, 2007). In transport, in the context of environmental problems, previous research has evaluated different ways to promote more sustainable transport modes by providing information on carbon dioxide emissions (see e.g., Rose & Ampt (2001); Avineri & Waygood (2013); Waygood & Avineri (2016, 2011)). They seek to increase awareness about the impact on the environment and others, so people can decide to cooperate through more sustainable travel decisions.

Our research contributes to the described literature by assessing the use of economic

(J. Nielsen, 2006). In this case, even though few contributors may provide high-quality information, there is concern about inequality (e.g., gender) (J. Nielsen, 2006; Torres, 2016). Similarly, only a small fraction of buyers provide an online review despite the fact that online shoppers highly value online reviews of products from many different consumers (NationMaster, 2019).

[2] One reason that economic incentives may backfire is the so-called "crowding-out of intrinsic motivation", in which economic incentives reduce the chance of a desired behavioral change by undermining people's intrinsic motivation – i.e., their desire to perform a task for its own sake without any economic reward (Frey & Oberholzer-Gee, 1997; Gneezy & Rustichini, 2000; Schwartz et al., 2015, 2020).

incentives and cooperation messages to increase participation of users to report bus stop conditions through a crowdsourcing application, and by doing so, contribute to improving the PT system. Even though bus stops are part of the trip experience and play a key role in customer satisfaction and efficient PT operations and maintenance (Eboli & Mazzulla, 2007), scant research has covered this portion of PT amenities and conditions. More broadly, with this intervention, we overcome the scarce attention that the use of economic incentives has received in research involved in collecting data for crowdsourcing information systems, and how such incentives have been combined with a cooperation message in this domain.[3] We also examine how economic incentives and a cooperation message affect different types of users to better represent a larger base of PT travellers and the transport network they use. The study also offers a methodological contribution as it uses a large randomized field experiment providing internal and ecological validity.

The remainder of this chapter is organized as follows. Section 2 describes the experiment developed using a public transport-oriented crowdsourcing smartphone app. Section 3 describes the results. Finally, Section 4 discusses the results and relates them to the existing literature.

5.2 Background information and method

5.2.1 Background information

We collaborated with a widely-used crowdsourcing smartphone application, Transapp (Arriagada & Munizaga, 2017), based in Santiago (Chile). Santiago is a large and congested city, with an integrated PT system that serves over 4.5 million trips per day. In a typical week, 3 million passengers use the system to make 25.5 million trips. Transapp allows users to easily access real-time information about bus arrival times, driver behavior, overcrowding, bus conditions, bus stop conditions, and bus bunching, among other factors. This information is publicly available to all users that have downloaded the application. In addition, the app allows users to indicate whether certain information is true or false, creating a self-regulated environment. As of September 2019, Transapp was downloaded 144,917 times since launching and had 47,320 active users who used the application at least once during September 2019 and accessed the app 719,545 times, mainly to check wait times.

The reporting feature, in which users can share information about buses and bus stops, requires a critical mass and widespread contribution from users. However, it suffers from the low participation problem described above. In fact, when studying the contributions of active users who used the app at least once in the year before this study, only 16.73% sent at least one report; the remaining 83.27% did not share any reports (i.e., they would be considered *lurkers*). Even more, 48.32% of all reports were contributed by only 1% of users, and the remaining 51.68% of reports were contributed by 15.73% of users, following Nielsen's Rule reasonably closely. Since reports are verified by the user community, if more users validate the veracity of reports, the data is more reliable, and users will consider it so. In addition, higher rates of user participation can capture a broader set of information both spatially and temporally, making the data on user experience, system operations, and infrastructure

[3] The literature on the effect of economic incentives on socially desirable behavior has been mixed, showing that their effect may depend on how incentives are structured and delivered (Gneezy et al., 2011; Kamenica, 2012; Schwartz et al., 2019).

status more efficient and complete. For example, users' reports can detect problems in the maintenance of bus stops. As a reference, Santiago's public transport system has more than 11,000 bus stops, making it practically impossible for dedicated inspectors to routinely carry out a thorough visual inspection of all assets.

5.2.2 Participants

Transapp provided a database that contained all users and reports sent in the app. We selected all active users during September 2019, considered as those who used the app to at least look at some information about bus time arrivals, resulting in a database of 46,516 users [4]. Then, we classified these users into two categories according to the number of reports they had shared in the previous two months: *"Previous Contributors"* and *"Previous Lurkers"*. Table 5.1 shows that those users who sent at least one report represented 17.5% of all active users, and those who never sent a report represented the remaining 82.5%. This classification was made in order to evaluate if messages and incentives had different effects depending on a user's past behavior.

Table 5.1: Classification of users before the campaign

Users	Reports sent	N	Percentage
Previous Contributors	1 or more	8,136	17.5%
Previous Lurkers	None	38,380	82.5%
Total	-	46,516	100%

5.2.3 Experimental design and procedure

We sent smartphone push notifications inviting users to participate in a three-day campaign to provide information on bus stops (e.g., if they need repair). We sent out one notification reminder once a day during the campaign at specific times, using historical data on periods of high user activity.[5] To examine the effect of incentives and messages on participation rates, we randomly assigned users (N = 46,516) into four experimental conditions using a block randomization procedure based on users' previous reporting behavior prior to the experiment, such that each condition had the exact same proportion of Previous Contributors. The experimental conditions were: (1) Economic incentive condition, (2) Cooperation message, (3) Both economic incentive and cooperation message condition, and (4) Control.

For the economic incentive condition, the message indicated that users who shared a report about bus stop conditions would be participating in a drawing for three-$13.95 reloads on their PT smart-card.[6] In other words, those users who received a message with the economic incentive and shared a report about a bus stop participated for one of the three rewards

[4] We used almost the entire database of active users at the time of the experiment, and excluded only a few hundred users who participated in a pilot test.

[5] Even though push notifications could be received even if the app was not being used, sending them when users were more likely to use the app increased the chance that they had some information to share.

[6] The messages showed the amounts in Chilean pesos (CLP), but we show them here in U.S. dollars (USD) using the prevailing conversion rate at the time of the experiment.

distributed as a lottery. The fare structure in Santiago's public transport system requires users to use a SC to pay for every trip made in the system (there is no multiple-ride pass or monthly ticket available). The economic incentive represents 4.81 times the value of the average SC reload, and allows users to make up to 13 one-way trips in non-peak hours. While the economic incentive is high for PT users, it is a low expenditure for PT authorities. For the cooperation message condition, users were reminded that sending reports about bus stops would help other passengers and contribute to improving the PT system. The third condition combined the economic incentive and the cooperation message. Users assigned to the control group did not receive any notification, representing the baseline scenario. If users opened the push notification, they accessed the message section of the app, which repeated the text from the push notification and included instructions on how to share a report about a bus stop. Users could also see the notification and report later on (see Appendix 5.5.1 for all materials used in the experiment). [7]

5.2.4 Empirical strategy

In this section we describe the empirical strategy to evaluate the participation rate and the level of contribution.

To examine the effect of each experimental condition on the participation rate, we estimate a logit model for the probability of participating using:

$$Y_i = B_0 + \sum_j B_j * D_{ij} + e_i \tag{5.1}$$

where Y_i indicates whether user i sent at least one report during the three-day campaign (=1, 0 if the user did not engage), D_{ij} is a dummy variable indicating whether user i was assigned to condition $j \in \{Economic, Cooperation, Both\}$ (=1, 0 if not). Therefore, all estimates use the control condition as the baseline. ε_i is the error term. Because users were randomly assigned, β_j will provide an unbiased estimate of the average treatment effects (Rubin, 1974). Additionally, we estimated a linear regression model (see Appendix 5.5.3) to facilitate the interpretation of results.

To examine users' contribution levels, i.e. the number of reports shared by users, we ran a zero-inflated negative binomial model. This model is well-suited for data distributions with an excess of zeros. Its central idea is that participation and report counts are generated by separate processes. In this case, the excess of zeros is attributable to users who did not receive or see the notification (e.g., push notifications were not allowed, or were deactivated, on some phones), and to users who may have automatically disregarded the push notification without reading it, or saw it but decided not to report. Across conditions, 96% of participants did not report during the campaign. This model is shown in equations 5.2 and 5.3:

[7] We oversampled the experimental conditions with an economic incentive based on the results from a pilot (Appendix 5.5.2 provides information about the sample size and the statistical power analysis).

$$Pr(Y_i = j) = \begin{cases} \pi_i + (1-\pi_i)g(y_i = 0|\mu_i) & if\ j = 0 \\ (1-\pi_i)g(y_i|\mu_i) & if\ j > 0 \end{cases} \tag{5.2}$$

$$log(E(y_i) = \mu_i) = B_0 + \sum_j B_j * D_{ij} + e_i \tag{5.3}$$

where π_i is the logistic function, which associates individuals who do not participate with probability π_i, and users who contribute with probability $1 - \pi_i$. Therefore, π_i can be interpreted as the probability of observing users not reporting. $g(y_i)$ is the negative binomial distribution, since the assumption is that the report count is generated according to this distribution with μ_i as the expected value of the negative binomial component. Its regression equation is presented in Equation 5.3, where y_i indicates the number of reports made by user i, D_{ij} is a dummy variable indicating whether user i was assigned to condition $j \in \{Economic, Cooperation, Both\}$ (=1, and 0 if not) and ε_i is the error term.

5.3 Results

In this section, we show the results of the randomized experiment. In particular, we focus on the participation rate (both overall and disaggregated by type of user), level of contribution, and the effect over time.

5.3.1 Participation rate

Figure 5.1 shows the participation rate across groups during the campaign, and the first column of Table 5.2 shows the results using Equation 5.1. Only 1.34% of users in the control condition reported (this is the baseline, as these users did not receive any type of message). The likelihood of reporting substantially increased to 5.26% when users were offered an economic incentive ($OR = 4.10$; $p < 0.001$), which represents a relative increase of 294%. Similarly, users who were sent a cooperation message or a combined economic incentive and cooperation message also increased their likelihood of reporting to 2.30% ($OR = 1.74$; $p < 0.001$; a relative increase of 72%) and 4.02% ($OR = 3.10$; $p < 0.001$; a relative increase of 201%), respectively. A pairwise comparison indicates that participation rates between treatments are all statistically different (all $ps < 0.001$). This means that there is a detrimental effect when a cooperation message is included with the economic incentive.[8] In comparison, a demanding "quid pro quo" approach (Tomasic et al., 2014) increased the participation rate in a transport crowdsourcing app in the US by 3.6 percentage points (a 23% relative increase from their baseline). In addition, in Transapp, the natural proportion of contributors in the two weeks previous to the experiment were 1.55% and 1.53%, respectively, a relative reduction of 1.3% in the number of contributors between the two weeks. Compared with the increase in the participation rate during the campaign (up to 3.92 percentage points, more than 200% in

[8] We also found that the percentage of users who uninstalled the app during the campaign was small and very similar across conditions: economic incentive (0.58%), cooperation message (0.54%), both (0.47%), and control (0.57%).

relative terms), this shows that an economic incentive can strongly boost participation rates, despite being a low-cost tool for the system.

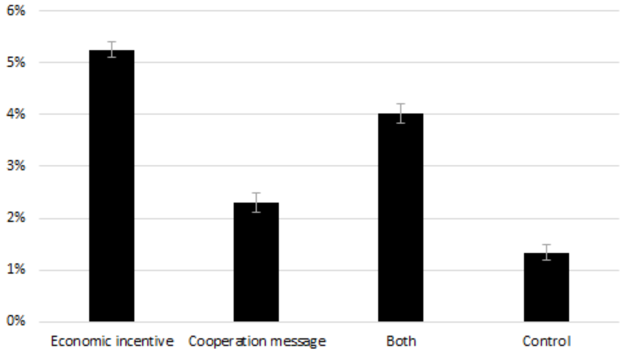


Figure 5.1: Percentage of users that sent a bus stop report during the campaign. Error bars represent ± 1 standard error.

A similar analysis by user type, shown in the last two columns of Table 5.2, demonstrates that users who had reported prior to the experiment (Previous Contributors) increased their participation rate. They increased their participation from 4.4%, in the control group, to 12.3% when they were offered an economic incentive alone ($OR = 3.05$; $p < 0.001$), to 10.6% when the message also included the cooperation message ($OR = 2.58$; $p < 0.001$). In relative terms, this is an increase of 180% and 141%, respectively. Previous Contributors' participation rate was 7.9% when only a cooperation message was used (a relative increase of 80%; $OR = 1.87 = 12.1$; $p < 0.001$).

However, the largest relative effect was found for the Previous Lurkers group (those who had never reported in the app prior to the experiment). For these users, the baseline control is 0.7%, implying that only a tiny fraction of users would have reported without the campaign. Users' participation in the economic incentive condition was 3.8% ($OR = 5.65$; $p < 0.001$), a 447% relative increase. For the both condition, Previous Lurkers' participation rate was 2.6% ($OR = 3.89$; $p < 0.001$), a relative increase of 280%. Finally, the cooperation message condition had a participation rate for this group of 1.1% ($OR = 1.63$; $p = 0.019$)), a relative increase of 62%. These results demonstrate a small effect of cooperation messaging for Previous Lurkers, which is consistent with this group's lack of previous (intrinsic) motivation to send reports. The results with Previous Lurkers are also notable as they indicate a potential expansion of the contributor base of the crowdsourcing app. Appendix 5.5.3 shows these results using a linear probability model.

Table 5.2: Estimation of the effect of each experimental condition on participation rate using a logit model

	All	Previous Contributors	Previous Lurkers
Economic incentive	1.411***	1.115***	1.731***
	(0.112)	(0.153)	(0.169)
	[4.099]	[3.049]	[5.647]
	<0.001	*<0.001*	*<0.001*
Cooperation message	0.555***	0.624***	0.488*
	(0.136)	(0.182)	(0.210)
	[1.741]	[1.867]	[1.629]
	<0.001	*<0.001*	*0.019*
Both	1.130***	0.948***	1.359***
	(0.118)	(0.162)	(0.177)
	[3.095]	[2.580]	[3.892]
	<0.001	*<0.001*	*<0.001*
Constant	-4.302***	-3.080***	-4.972***
	(0.108)	(0.145)	(0.165)
	[0.014]	[0.046]	[0.007]
	<0.001	*<0.001*	*<0.001*
P-values for pairwise comparisons			
Economic incentive vs Cooperation message	*<0.001*	*<0.001*	*<0.001*
Economic incentive vs Both	*<0.001*	*0.058*	*<0.001*
Cooperation message vs Both	*<0.001*	*0.014*	*<0.001*
Log-likelihood	-7658.923	-2635.653	-4626.176
Observations	46516	8136	38380

$^{+}$p<0.10, *p<0.05, ** p<0.01, *** p<0.001.

All columns show standard errors between parentheses, p-values in italics, and odd-ratios between brackets.

5.3.2 Level of user contribution

The previous sub-section focused on participation rates (i.e., an extensive margin analysis). In the following analysis, we examine the number of reports shared by users (i.e., intensive margin). Figure 5.2 shows the average number of shared reports per user and per day, conditional on participation. Users who were offered an economic incentive, with or without a cooperation message, made 19.5% more reports when compared to users in the control group, from 0.80 to 0.96 daily reports per user (both significant only at the 10% significance level; $d = 0.2$ for the both condition)[9]. This means that in the economic incentive group, users who reported sent almost three reports, on average, during the campaign. There is a much smaller difference for users who were sent a cooperation message; they made 0.86 average reports, on average, which represents an increase of 6.7% from the control group, and is not sizably different from the daily reports shared by the control group ($p = 0.54$; $d = 0.1$).

[9] We excluded outlier observations with an extremely high number of reports – over the 99.5th percentile – to avoid a strong influence from very few observations. For completeness, in the Appendix 5.5.4, we show an analysis that includes these observations.

These results suggest that the economic incentive not only increased the active user base, but incentivized that user base to interact (i.e., share reports) slightly more frequently. While all treatments showed a positive effect regarding both extensive and intensive margins, the economic incentive was the most successful treatment with regards to both margins.

The contribution-level analysis by user type shows that Previous Contributors increased the number of shared reports conditional on participation. They increased their contribution from 0.82 average daily reports, in the control group, to 1.12 average daily reports when they were offered an economic incentive with or without the cooperation message, a relative increase of 36% ($p = 0.013$ for the economic incentive and $p = 0.03$ for both, $d = 0.4$). Previous Contributors' contribution level was 0.99 daily average reports when only a cooperation message was used (a relative increase of 20.6%; $p = 0.18$; $d = 0.3$). For the Previous Lurkers group, there were no sizably significant differences in shared reports compared to the control group – these were users who started to report for the first time, so it is hard to expect that their intensive margin was any greater than those in the control group. Overall, these results indicate that economic incentives, mainly, but also cooperation messaging, increased the likelihood of participation for all users, with a larger increase for users who never reported before, and also increased the number of reports shared by those who had experience reporting with the app.

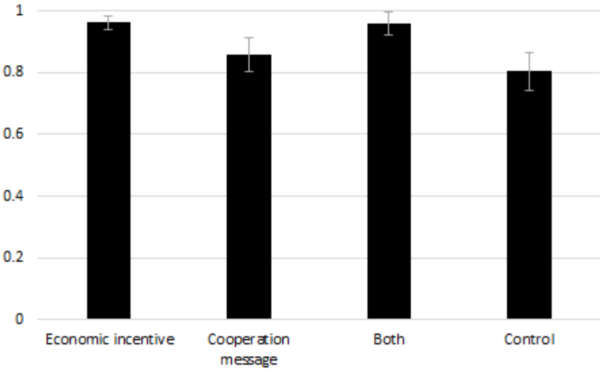


Figure 5.2: Average number of daily reports shared by users who participated during the campaign. Error bars represent ± 1 standard error.

The previous analysis must be taken with caution because of the change in participation likelihood across conditions. In this regard, it is remarkable that even though all treatments increased participation, they also increased the level of contribution – one may expect that these new users would report less frequently compared with users in the control group. Nevertheless, as explained in Section 5.2.4, we use a zero-inflated negative binomial model to account for the decision to participate and how many reports to share. Table 5.3 shows the results. The bottom section of the table shows the log-odds of not reporting under each treatment, using the control group as the baseline. Consistent with the previous analysis, users are more likely to report under all treatments. Overall, being in the economic incentive treatment (vs. in the control group) decreases the odds of not participating by a factor of 0.27 ($e^{-1.297}$), $p < 0.001$. In other words, the economic incentive increases the participation rate. The top section of the table shows the effect on the number of reports for those who share at least one report. Here, the economic incentive and the both treatments increase the number of reports compared to the control group. For example, for someone in the economic incentive condition, the number of reports increases by a factor of 1.34 ($e^{0.293}$), $p = 0.04$. Table 5.3 shows the same analysis for Previous Contributors and Previous Lurkers, with results consistent with the previous analysis. For robustness, in Appendix 5.5.4 we show the analysis with alternative specifications (e.g., using a Poisson model or using bus reports as the outcome variable, which were not part of the campaign and were not expected to affect participation). The robustness of these analyses is consistent with the previous results.[10]

Table 5.3: Estimation of the effect of each experimental condition on the level of user contribution using a zero-inflated negative binomial model

Negative binomial			
	All	Previous Contributors	Previous Lurkers
Economic incentive	0.293*	0.461*	0.158
	(0.139)	(0.182)	(0.204)
	0.035	*0.011*	*0.439*
Cooperation message	0.110	0.286	-0.282
	(0.172)	(0.210)	(0.295)
	0.523	*0.173*	*0.338*
Both	0.292*	0.458*	0.110
	(0.147)	(0.192)	(0.216)
	0.047	*0.017*	*0.611*
Constant	0.281+	0.507**	0.023
	(0.145)	(0.179)	(0.238)
	0.053	*0.005*	*0.924*
Zero-inflated logit			
	All	Previous Contributors	Previous Lurkers
Economic incentive	-1.297***	-0.953***	-1.668***
	(0.130)	(0.176)	(0.196)
	<0.001	*<0.001*	*<0.001*
Cooperation message	-0.491**	-0.491*	-0.634*
	(0.158)	(0.208)	(0.259)
	0.002	*0.018*	*0.014*
Both	-1.015***	-0.785***	-1.316***
	(0.137)	(0.186)	(0.206)
	<0.001	*<0.001*	*<0.001*
Constant	3.715***	2.707***	4.130***
	(0.135)	(0.172)	(0.226)
	<0.001	*<0.001*	*<0.001*
$Ln\alpha$	0.170	-0.426**	0.629**
	(0.123)	(0.147)	(0.225)
	0.169	*0.004*	*0.005*
Log-likelihood	-10690.824	-4059.634	-6243.081
Observations	46438	8082	38356

$^{+}$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

All columns show standard errors between parentheses and *p*-values in italics.

[10] Compared to the negative binomial model, the Poisson distribution does not assume overdispersion of the count data. In our case, there is overdispersion as the unconditional mean number of reports is much lower than its variance for each experimental condition.

5.3.3 Effect over time

The experiment was conducted during the second week of October 2019. One day after the campaign ended, a series of massive demonstrations and severe riots known as the "Social Outburst" (*Estallido Social*) occurred throughout Chile.[11] These events paralyzed the PT system due to the burning of buses and Metro stations. The system started working again at partial capacity one week later.[12] Because of this force majeure event, we were doubtful about the lasting impacts of the campaign, given this vast disruption to any habit-forming behavior. Nevertheless, we examined how participation rates varied over time for several weeks after the campaign ended. Figure 5.3 presents the percentage of users that reported (participation rates) in each treatment group over time.

The period analyzed was the three-day campaign-treatment period and the five weeks after the campaign ended. During the first post-treatment week, many people stayed at home due to disruptions in the public transportation system associated with the demonstrations and the declaration of a curfew and state of emergency. In the second post-treatment week, even though participation rates and levels of user contribution started to decrease after the campaign, and a major disruption, the participation rate of users in the economic incentive condition was 39% greater than those in the control condition (a 0.7 percentage points increase from 1.7%; $OR = 1.40$; $p = 0.001$). Afterward, participation rates further decreased, reaching a level of 1.9% in the fifth week after treatment, which was very similar to the participation rate of the control group at the same point (1.7%). A similar pattern was observed for the other experimental condition groups. To examine these differences, we conducted a statistical analysis using the same model from Section 5.2.4 for each period described below.

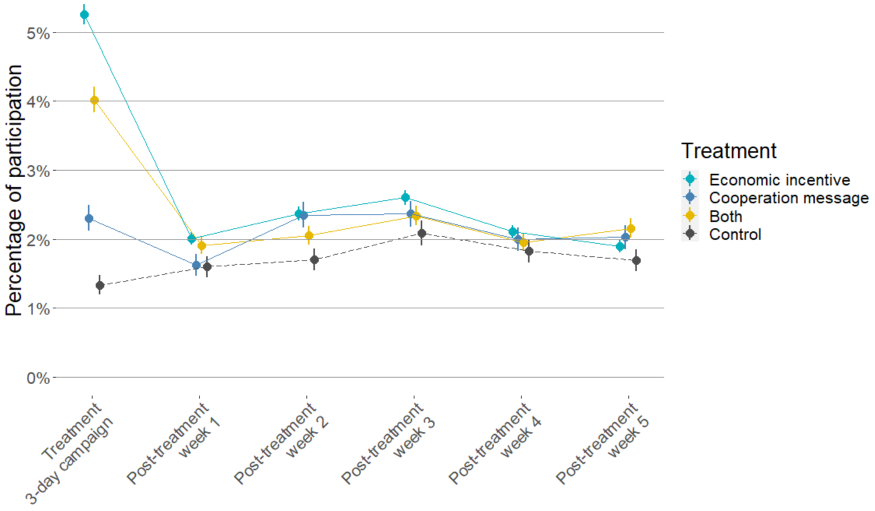


Figure 5.3: Participation rates over time. Error bars represent ± 1.96 standard error.

Table 5.5 in Appendix 5.5.5 shows the results obtained from the zero-inflated negative binomial model for each period: a two-week pre-treatment period (for which we expected

[11] https://www.ciperchile.cl/2019/10/27/el-reventon-social-en-chile-una-mirada-historica/

[12] https://www.interior.gob.cl/noticias/2019/10/28/informacion-oficial-del-gobierno-de-chile-con-las-medidas-para-enfrentar-la-situacion-de-emergencia/

there to be no effect), the three-day campaign-treatment period, and each of the five weeks following the campaign. The bottom panel shows the estimation for the zero-inflation portion of the model, where negative coefficients indicate a decrease in the probability of obtaining zero reports. This portion of the model indicates that, except for the first *particular* post-treatment week, the economic-incentive group's participation rate was still higher than that of the control group for three weeks after the campaign ended, with the effect decreasing until it was not sizably different from zero during the fourth post-treatment week. The other treatment conditions showed a similar pattern, but their trends were slightly more erratic. Therefore, despite the disruption during the third week of October, the campaign was able to change users' participation behavior for several weeks. The top panel shows the estimation for the negative binomial portion of the model, where positive coefficients indicate an increase in the number of reports generated. This non-zero portion of the analysis shows that it was not possible to observe significant differences in the number of reports shared compared with users in the control group (conditional on reporting). This result is to be expected, since people who began to report after receiving the push notification would not be expected to report frequently, as they did not show an intrinsic motivation to participate prior to the campaign. In Appendix 5.5.5, we also conduct this analysis for Previous Contributor and Previous Lurkers. It shows that the positive impact of participation in the post-treatment periods was more attributable to Previous Lurkers and an increased number of reports for two weeks after the treatment was driven by Previous Contributors.

5.4 Discussion and conclusions

Crowdsourcing PT applications allow commuters to report PT system conditions along with their level of satisfaction regarding the service provided. These data collection systems face three principal challenges: (i) motivate as many users as possible to contribute information, (ii) motivate users to deliver as much information as possible, and (iii) obtain information that covers most of the transport network both spatially and temporally. In this context, we evaluated the effectiveness of economic incentives and cooperation messages to motivate users to report key information about bus stop conditions.

Our results show that economic incentives and cooperation messaging increased participation rates and the number of reports shared by users. The relative increases compared to the control group were 294% for users who received an economic incentive, 72% for those who received a cooperation message, and 201% for users who received a combination of both. The economic incentive condition increased the participation rate most effectively, especially for users who had not reported prior to the campaign, and also increased the number of reports conditional on participation.[13] Furthermore, we found that offering an economic incentive helped to encourage lurkers - those users who had not made prior contributions - to participate, thereby increasing the contributor base, which is one of the most important goals of crowdsourcing applications in transportation.

The cooperation message had a positive impact on the participation rate compared to not

[13] Regarding report quality, only a small percentage of reports may be considered dubious (i.e., more users rejected the report instead of confirming it) out of all the reports made during the campaign: economic incentive (5.5%), cooperation message (8.4%), both (6.9%), and control (11.8%).

sending any message (i.e., control group), but its impact was significantly less than offering an economic incentive. As the crowdsourcing app is inherently a platform based on providing and receiving contributions, with no economic recompense, the cooperation message most likely did not change the status quo. In other words, for most people, the cooperation message probably acted simply as a reminder with a short-lasting effect. Previous research has shown that when people are reminded of something they are already aware of, behavioral effects are short-lived (Schwartz et al., 2013). Interestingly, combining the cooperation message with an economic incentive reduced the participation rate compared to offering only an economic incentive. This result suggests that some of these users paid more attention to the first section of the notification ("Help and participate for bip! reloads of $13.95"), reducing their chance to share a report, or they resisted mixing an emphasized cooperation activity with an economic (more self-centered) motivation (Heyman & Ariely, 2004). In line with our results, recent research has found that monetary incentives work better to encourage behavior when offered without combining them with an emphasis on cooperation (Lacetera et al., 2012; Niessen-Ruenzi et al., 2014; Lacetera et al., 2014; Schwartz et al., 2021, 2020).

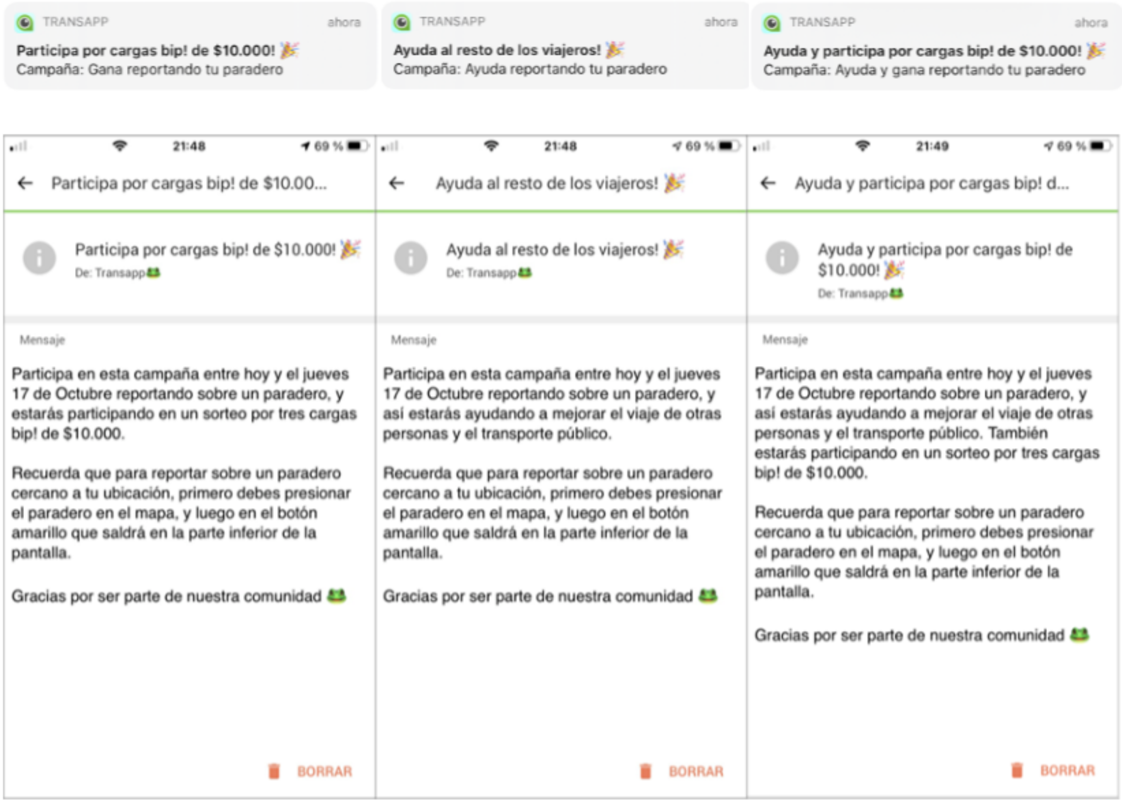
We found that providing an economic incentive can have positive impacts on participation and contribution for several weeks after a campaign ends. However, the positive effects rapidly decrease compared to the initial impact. Future research may consider whether a long-lasting effect is possible in the majority of cases, given that in the case of this experiment, a major disruption to the PT network occurred one week after the campaign, which may have dampened the campaign's lasting impact, or whether using multiple messages for several weeks can strengthen collaborative habits on crowdsourcing platforms. In addition, even though a cooperation message had less of an impact than a standalone economic incentive, future research may examine whether framing the cooperation message as a personal characteristic (e.g., "those who report help improve the system") or normalizing the behavior (e.g., "many people collaborate by sending reports") can encourage cooperation behavior. The interplay of economic incentives and altruistic behavior has puzzled researchers in recent decades. Our research should help deepen the understanding of the role that economic incentives can play by providing evidence for transport planners, crowdsourcing information managers, and government authorities to more effectively increase the use of crowdsourcing systems that benefit the wider community.

Finally, the results of this study suggest that it is possible to enrich current PT databases, used for the understanding of PT passenger behavior and the evaluation of PT service, using crowdsourcing applications to provide detailed information about the system. Currently, passive data, such as Automatic Fare Collection (AFC) data, Automatic Vehicle Location (AVL) data, and Automatic Passenger Counter (APC) data, are widely used by PT authorities and researchers to understand the demand and the operation of PT systems (Bagchi & White, 2005; Munizaga & Palma, 2012; Gschwender et al., 2016; Devillaine et al., 2012). Unlike traditional data obtained from surveys, passive data allows the collection of large volumes of travel data over long periods of time. However, it lacks relevant user information, such as information related to infrastructure maintenance, which is essential for improving the PT system. This study shows that it is possible to encourage transport-oriented crowdsourcing applications users to share the currently missing information from passive databases using cost-effective monetary incentives.

5.5 Appendices

5.5.1 Notifications received by users on their phones

Messages sent to users. From left to right: Economic incentive, Cooperation message and Both conditions. Users had to press the notification (at the top of the figure) in order to read the full message.[14]



TRANSAPP ahora
Participa por cargas bip! de $10.000! 🎉
Campaña: Gana reportando tu paradero

21:48 69 %

← Participa por cargas bip! de $10.00...

Participa por cargas bip! de $10.000! 🎉
De: Transapp 🐸

Mensaje

Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y estarás participando en un sorteo por tres cargas bip! de $10.000.

Recuerda que para reportar sobre un paradero cercano a tu ubicación, primero debes presionar el paradero en el mapa, y luego en el botón amarillo que saldrá en la parte inferior de la pantalla.

Gracias por ser parte de nuestra comunidad 🐸

BORRAR

TRANSAPP ahora
Ayuda al resto de los viajeros! 🎉
Campaña: Ayuda reportando tu paradero

21:48 69 %

← Ayuda al resto de los viajeros! 🎉

Ayuda al resto de los viajeros! 🎉
De: Transapp 🐸

Mensaje

Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público.

Recuerda que para reportar sobre un paradero cercano a tu ubicación, primero debes presionar el paradero en el mapa, y luego en el botón amarillo que saldrá en la parte inferior de la pantalla.

Gracias por ser parte de nuestra comunidad 🐸

BORRAR

TRANSAPP ahora
Ayuda y participa por cargas bip! de $10.000! 🎉
Campaña: Ayuda y gana reportando tu paradero

21:49 69 %

← Ayuda y participa por cargas bip! d...

Ayuda y participa por cargas bip! de $10.000! 🎉
De: Transapp 🐸

Mensaje

Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. También estarás participando en un sorteo por tres cargas bip! de $10.000.

Recuerda que para reportar sobre un paradero cercano a tu ubicación, primero debes presionar el paradero en el mapa, y luego en el botón amarillo que saldrá en la parte inferior de la pantalla.

Gracias por ser parte de nuestra comunidad 🐸

BORRAR

Details of the notifications and messages sent to users, translated to English:

[14] The economic incentive condition had two possible specific messages, which could be seen only if people *opened the economic incentive notification* – one with the economic incentive specific message and another with the both specific message. Because few people likely read the messages in the app, we found no sizable differences between messages for people who received the economic incentive notification, so we decided to present the results based on the notifications only.

	Economic incentive	Cooperation message	Both
Notification	Participate for CLP$10,000 in bip! credits!	Help other commuters!	Help and participate for CLP$10,000 in bip! credits!
	Campaign: Win by reporting your bus stop	Campaign: Help by reporting your bus stop	Campaign: Help and win by reporting your bus stop
Condition-specific message	Participate in this campaign between [dates] by sending reports about a bus stop, and you will be participating in a draw for three bip! credits of CLP$10,000	By participating in this campaign between [dates] by sending reports about a bus stop, you will be helping to improve other peoples' trips, as well as the PT system	Participate in this campaign between [dates] by sending reports about a bus stop, and you will be helping to improve other peoples' trips and the PT system. You will also be participating in a draw for three bip! charges of CLP$10,000
Common message	Remember that to report a bus stop near your location, you must first click on the bus stop on the map, and then on the yellow button that will appear at the bottom of the screen. Thanks for being part of our community.		

5.5.2 Sample size

To identify the sample size required for each treatment, we perform a statistical power analysis based on the results of a pilot. This pilot showed that the cooperation message had the smallest effect compared to the control (0.8 p.p. from 1.1%). Therefore, we required approximately 6,000 individuals in each of these experimental conditions to have a 95% statistical power (we added a few hundred people because some phones may have changed or not be working). The rest of the sample was evenly distributed to detect differences between the economic incentive and both conditions, and to be able to split the economic incentive condition into two additional conditions for people who *opened the economic incentive notification*. For the latter, the message section in the app either repeated the text from the notification (i.e., offering an economic incentive) or also included the text from the cooperation message. We expected the difference to be small, if detectable, because few people may use the message section in the app. Therefore, the final sample was 11,164 for each of these three groups (the Both condition and the two inside the economic incentive one). Consistently, Table 5.4 shows no significant difference between the texts in the message section for the economic incentive condition ($p > 0.2$ for all models).

Table 5.4: Estimation of the effect of each experimental condition on participation rate using a logit model

	All	Previous Contributors	Previous Lurkers
Economic incentive	1.446***	1.134***	1.778***
	(0.116)	(0.160)	(0.173)
	<0.001	*<0.001*	*<0.001*
Economic incentive (adding cooperation text in the app)	1.374***	1.096***	1.682***
	(0.116)	(0.160)	(0.174)
	<0.001	*<0.001*	*<0.001*
Cooperation message	0.555***	0.624***	0.488*
	(0.136)	(0.182)	(0.210)
	<0.001	*<0.001*	*0.02*
Both	1.130***	0.948***	1.359***
	(0.118)	(0.162)	(0.177)
	<0.001	*<0.001*	*<0.001*
Constant	-4.302***	-3.080***	-4.972***
	(0.108)	(0.145)	(0.165)
	<0.001	*<0.001*	*<0.001*
P-values for pairwise comparisons			
Economic incentive (adding cooperation text in the app)	*0.231*	*0.697*	*0.216*
Log-likelihood	-7658.204	-2635.577	-4625.409
Observations	46516	8136	38380

$^{+}$p<0.10, *p<0.05, ** p<0.01, *** p<0.001.

All columns show standard errors between parentheses and p-values in italics.

5.5.3 Estimation of the effect of each experimental condition on participation rate using a linear probability model

	All	Previous Contributors	Previous Lurckers	All with interactions
Economic incentive	0.039***	0.079***	0.031***	0.024***
	(0.003)	(0.010)	(0.003)	(0.003)
	<0.001	*<0.001*	*<0.001*	*<0.001*
Cooperation message	0.010**	0.035**	0.004	-0.002
	(0.003)	(0.013)	(0.003)	(0.004)
	0.005	*0.005*	*0.169*	*0.538*
Both	0.027***	0.062***	0.019***	0.013***
	(0.003)	(0.011)	(0.003)	(0.003)
	<0.001	*<0.001*	*<0.001*	*<0.001*
Contributors-economic incentive				0.085***
				(0.003)
				<0.001
Contributors-cooperation message				0.068***
				(0.006)
				<0.001
Contributors-both				0.080***
				(0.005)
				<0.001
Constant	0.013***	0.044***	0.007**	0.013***
	(0.002)	(0.009)	(0.002)	(0.002)
	<0.001	*<0.001*	*0.002*	*<0.001*
p-values for pairwise comparisons				
Economic incentive vs Cooperation message	*<0.001*	*<0.001*	*<0.001*	*<0.001*
Economic incentive vs Both	*<0.001*	*0.043*	*<0.001*	*<0.001*
Cooperation message vs Both	*<0.001*	*0.016*	*<0.001*	*<0.001*
R^2	0.006	0.008	0.006	0.027
Observations	46516	8136	38380	46516

$^{+}$p<0.10, *p<0.05, **p<0.01, ***p<0.001

All columns show standard errors between parentheses and *p*-values in italics.

5.5.4 Estimation of the effect of each experimental condition on the level of user contribution using different models and different analyses

Negative binomial			
	Without exclusion	With Poisson	With bus reports
Economic incentive	1.231**	0.235*	0.322
	(0.408)	(0.112)	(0.430)
	0.003	*0.036*	*0.453*
Cooperation message	0.517^{+}	0.088	0.685
	(0.289)	(0.139)	(0.693)
	0.074	*0.523*	*0.323*
Both	1.526*	0.234*	0.226
	(0.590)	(0.119)	(0.639)
	0.010	*0.048*	*0.724*
Constant	-2.43***	0.752***	0.108
	(0.231)	(0.108)	(1.039)
	<0.001	*<0.001*	*0.916*
Zero-inflated			
Economic incentive	-18.1***	-1.33***	-0.28
	(0.348)	(0.118)	(0.106)
	<0.001	*<0.001*	*0.239*
Cooperation message	-0.63^{+}	-0.50***	0.106
	(0.371)	(0.143)	(0.164)
	0.089	*<0.001*	*0.739*
Both	-1.21^{+}	-1.05***	0.164
	(0.656)	(0.124)	(3.695)
	0.065	*<0.001*	*0.588*
Constant	0.345	4.196***	3.695
	(0.317)	(0.113)	(1.030)
	0.277	*<0.001*	*<0.001*
Ln alpha	3.965***		2.454*
	(0.125)		(0.037)
	<0.001		*0.037*
alpha	52.751		11.641
	(6.637)		(13.72)
Log-likelihood	-11865.09	-11057.26	-2431.907
Observations	46,516	46,438	46,516

$^{+}$p<0.10, *p<0.05, **p<0.01, ***p<0.001

All columns show standard errors between parentheses and *p*-values in italics.

5.5.5 Estimation of the effect of each experimental condition on the level of user contribution using a Zero-inflated negative binomial model

Table 5.6: Estimation of effects of each experimental condition on the level of Previous Contributors' contributions over time

Negative binomial							
	Pre-treatment (2 weeks)	Treatment (3 days)	Post-treatment Week 1	Post-treatment Week 2	Post-treatment Week 3	Post-treatment Week 4	Post-treatment Week 5
Economic incentive	-0.201	0.461*	0.455*	0.308$^+$	0.260	-0.167	0.175
	(0.162)	(0.182)	(0.204)	(0.165)	(0.195)	(0.211)	(0.183)
	0.214	*0.011*	*0.026*	*0.063*	*0.182*	*0.430*	*0.339*
Cooperation message	0.169	0.286	0.859**	0.457*	0.412$^+$	0.019	0.054
	(0.203)	(0.210)	(0.264)	(0.218)	(0.239)	(0.286)	(0.236)
	0.405	*0.173*	*0.001*	*0.036*	*0.084*	*0.947*	*0.820*
Both	-0.153	0.458*	0.327	0.077	-0.124	-0.312	0.168
	(0.174)	(0.192)	(0.226)	(0.193)	(0.235)	(0.236)	(0.192)
	0.380	*0.017*	*0.148*	*0.689*	*0.597*	*0.187*	*0.380*
Constant	0.657**	0.507**	0.234	0.505**	0.541**	0.679**	0.642**
	(0.236)	(0.179)	(0.251)	(0.194)	(0.206)	(0.256)	(0.197)
	0.005	*0.005*	*0.352*	*0.009*	*0.009*	*0.008*	*0.001*
Zero-inflated							
	Pre-treatment (2 weeks)	Treatment (3 days)	Post-treatment Week 1	Post-treatment Week 2	Post-treatment Week 3	Post-treatment Week 4	Post-treatment Week 5
Economic incentive	-0.286$^+$	-0.953***	-0.075	-0.203	-0.022	-0.247	-0.153
	(0.160)	(0.176)	(0.189)	(0.169)	(0.162)	(0.175)	(0.174)
	0.074	*<0.001*	*0.691*	*0.229*	*0.893*	*0.157*	*0.381*
Cooperation message	-0.023	-0.491*	0.230	-0.033	0.045	-0.133	-0.302
	(0.187)	(0.208)	(0.231)	(0.207)	(0.196)	(0.218)	(0.212)
	0.903	*0.018*	*0.318*	*0.874*	*0.817*	*0.541*	*0.154*
Both	-0.240	-0.785***	-0.056	-0.189	-0.032	-0.239	-0.297
	(0.175)	(0.186)	(0.208)	(0.189)	(0.188)	(0.196)	(0.187)
	0.174	*<0.001*	*0.787*	*0.317*	*0.867*	*0.222*	*0.111*
Constant	0.559$^+$	2.707***	2.187***	2.243***	1.910***	2.122***	2.411***
	(0.333)	(0.172)	(0.250)	(0.200)	(0.199)	(0.243)	(0.193)
	0.093	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*
$Ln\alpha$	1.597***	-0.426**	0.792**	0.618*	0.736**	0.961**	0.504*
	(0.287)	(0.147)	(0.287)	(0.247)	(0.255)	(0.295)	(0.239)
	<0.001	*0.004*	*0.006*	*0.012*	*0.004*	*0.001*	*0.035*
Log-likelihood	-6339.622	-4059.634	-2641.881	-3002.615	-3314.733	-2885.729	-2835.970
Observations	8122	8082	8129	8126	8124	8129	8125

$^+$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

All columns show standard errors between parentheses and *p*-values in italics.

Table 5.7: Estimation of effects of each experimental condition on the level of previous lurkers' contributions over time

Negative binomial							
	Pre-treatment	Treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment
	(2 weeks)	(3 days)	Week 1	Week 2	Week 3	Week 4	Week 5
Economic incentive	0.121	0.158	0.006	-0.160	-0.374$^+$	0.249	-0.216
	(0.169)	(0.204)	(0.334)	(0.278)	(0.212)	(0.281)	(0.220)
	0.491	*0.439*	*0.986*	*0.566*	*0.078*	*0.374*	*0.325*
Cooperation message	0.093	-0.282	0.707	-0.346	0.002	0.540$^+$	-0.208
	(0.204)	(0.295)	(0.484)	(0.365)	(0.288)	(0.320)	(0.304)
	0.649	*0.338*	*0.144*	*0.344*	*0.996*	*0.092*	*0.494*
Both	0.084	0.110	-0.117	-0.479	-0.666*	0.237	-0.256
	(0.223)	(0.216)	(0.353)	(0.307)	(0.279)	(0.312)	(0.250)
	0.707	*0.611*	*0.740*	*0.118*	*0.017*	*0.477*	*0.305*
Constant	-1.326	0.023	-3.526***	-0.378	0.183	-0.825	-0.266
	(1.109)	(0.238)	(0.334)	(0.600)	(0.408)	(0.568)	(0.618)
	0.232	*0.924*	*<0.001*	*0.529*	*0.654*	*0.146*	*0.668*
Zero-inflated							
	Pre-treatment	Treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment
	(2 weeks)	(3 days)	Week 1	Week 2	Week 3	Week 4	Week 5
Economic incentive	0.099	-1.668***	-2.592	-0.445*	-0.506**	0.026	-0.113
	(0.170)	(0.196)	(7.648)	(0.206)	(0.171)	(0.222)	(0.188)
	0.562	*<0.001*	*0.735*	*0.030*	*0.003*	*0.908*	*0.547*
Cooperation message	0.105	-0.634*	1.056	-0.621*	-0.129	0.261	-0.079
	(0.203)	(0.259)	(1.118)	(0.264)	(0.215)	(0.260)	(0.245)
	0.605	*0.014*	*0.345*	*0.018*	*0.550*	*0.314*	*0.746*
Both	0.209	-1.316***	-8.148***	-0.397$^+$	-0.553**	0.124	-0.284
	(0.218)	(0.206)	(1.883)	(0.235)	(0.207)	(0.245)	(0.207)
	0.338	*<0.001*	*<0.001*	*0.091*	*0.008*	*0.613*	*0.171*
Constant	1.159	4.130***	-1.292	3.216***	3.517***	2.960***	3.249***
	(1.444)	(0.226)	(1.144)	(0.608)	(0.392)	(0.579)	(0.611)
	0.422	*<0.001*	*0.259*	*<0.001*	*<0.001*	*<0.001*	*<0.001*
$Ln\alpha$	2.507*	0.629**	4.856***	1.625*	1.176*	1.449*	1.535*
	(1.236)	(0.225)	(0.134)	(0.766)	(0.550)	(0.694)	(0.774)
	0.042	*0.005*	*<0.001*	*0.034*	*0.033*	*0.037*	*0.047*
Log-likelihood	-6178.482	-6243.081	-3029.881	-3439.630	-3767.970	-3067.687	-2987.203
Observations	38380	38356	38380	38378	38380	38380	38380

$^+$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

All columns show standard errors between parentheses and *p*-values in italics.

Table 5.5: Estimation of effects of each experimental condition on the level of all users' contributions over time

Negative binomial							
	Pre-treatment	Treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment
	(2 weeks)	(3 days)	Week 1	Week 2	Week 3	Week 4	Week 5
Economic incentive	-0.095	0.293*	0.299	0.134	0.004	-0.041	0.056
	(0.140)	(0.139)	(0.188)	(0.155)	(0.149)	(0.187)	(0.144)
	0.499	*0.035*	*0.112*	*0.388*	*0.981*	*0.826*	*0.697*
Cooperation message	0.188	0.110	0.831**	0.141	0.263	0.193	-0.010
	(0.179)	(0.172)	(0.253)	(0.205)	(0.189)	(0.238)	(0.192)
	0.294	*0.523*	*0.001*	*0.491*	*0.164*	*0.418*	*0.959*
Both	-0.045	0.292*	0.166	-0.110	-0.361^{+}	-0.141	0.037
	(0.155)	(0.147)	(0.204)	(0.178)	(0.185)	(0.206)	(0.156)
	0.770	*0.047*	*0.418*	*0.535*	*0.050*	*0.495*	*0.812*
Constant	-1.300	0.281^{+}	-0.627	-0.021	0.250	-0.197	0.209
	(0.941)	(0.145)	(0.441)	(0.266)	(0.220)	(0.349)	(0.221)
	0.167	*0.053*	*0.155*	*0.937*	*0.256*	*0.572*	*0.345*
Zero-inflated							
	Pre-treatment	Treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment	Post-treatment
	(2 weeks)	(3 days)	Week 1	Week 2	Week 3	Week 4	Week 5
Economic incentive	-0.270	-1.297***	-0.107	-0.279*	-0.221^{+}	-0.164	-0.087
	(0.565)	(0.130)	(0.149)	(0.128)	(0.116)	(0.137)	(0.125)
	0.633	*<0.001*	*0.472*	*0.030*	*0.057*	*0.232*	*0.485*
Cooperation message	0.144	-0.491**	0.323^{+}	-0.260^{+}	-0.012	0.004	-0.155
	(0.367)	(0.158)	(0.184)	(0.157)	(0.143)	(0.170)	(0.156)
	0.694	*0.002*	*0.080*	*0.098*	*0.934*	*0.980*	*0.321*
Both	-0.049	-1.015***	-0.114	-0.232	-0.265^{+}	-0.126	-0.238^{+}
	(0.276)	(0.137)	(0.164)	(0.145)	(0.137)	(0.153)	(0.135)
	0.859	*<0.001*	*0.488*	*0.110*	*0.054*	*0.411*	*0.078*
Constant	-0.585	3.715***	2.479***	2.931***	2.896***	2.641***	3.143***
	(2.627)	(0.135)	(0.459)	(0.260)	(0.212)	(0.354)	(0.210)
	0.824	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*
Ln alpha	3.389***	0.170	1.886***	1.326***	1.166***	1.641***	1.063***
	(1.001)	(0.123)	(0.506)	(0.331)	(0.287)	(0.414)	(0.281)
	0.001	*0.169*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*
Log-likelihood	-10690.824	-10690.824	-10690.824	-10690.824	-10690.824	-10690.824	-10690.824
Observations	46438	46438	46438	46438	46438	46438	46438

$^{+}$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

All columns show standard errors between parentheses and *p*-values in italics.

Chapter 6

Conclusions and future research

6.1 Main findings

The main purpose of this thesis is to contribute to the understanding of PT passengers' behavior using real-world data. In particular, this research uses SC data from the PT system in Santiago, Chile, and transport-oriented crowdsourcing application data. Revealed preference data collected from different transport technologies allows us to obtain day-to-day passengers' observed choices. This research uses this opportunity to answer four research questions, which are summarized in this section.

1. **Are public transport passengers using different route choice strategies?**

Most PT passengers' route choice modeling literature have used the MNL or PSL discrete choice model, with consideration sets composed of itineraries, where all lines belonging to a set of relevant alternatives are regarded as different options. We call this approach the disaggregated strategy. On the other hand, transit assignment modeling literature use the common lines principle, which was described by Chriqui & Robillard (1975) as the strategy of identifying a subset of PT lines that minimized the total expected travel time. According to this principle, passengers will take the first line of the common-line set that arrives to the bus stop. We call this approach aggregated strategy. This means that two modeling approaches can be distinguished, depending on whether the analysis focuses on route choice behavior modeling, or transit assignment models for strategic analysis.

In chapter 2, we show that this dichotomy between the disaggregated and aggregated approaches to the same problem is inappropriate and that, instead, heterogeneity exists in the route choice strategy, both between users and across contexts. We verified this using an indicator function constructed as the difference between expected and observed trips based on the assumption that passengers follow the common line principle. We applied this indicator to a case study based on SC data from the city of Santiago, Chile, from which identify individuals that follow either an aggregated or a disaggregated strategy, as well as others who seem to be using a combination of both strategies.

In the same chapter, we analyzed the heterogeneity regarding the PT passengers' route choice strategy behavior using an integrated discrete choice and latent class approach, using SC data from the city of Santiago, Chile. The analysis was performed by estimating three types of models: i) a PSL model with the consideration set composed of itineraries (disaggregated strategy); ii) a PSL model with the consideration set composed of common lines as part of the same alternative (aggregated strategy); and iii) a MIS latent class model, which takes in consideration the endogeneity problem and considers two types of consideration sets for two types of passengers: those who use itineraries and those who use common lines as part of the same alternative. The estimation of the MIS latent class model confirmed the heterogeneity strategy route choice behavior between passengers, suggesting that 51.2% of passengers use a disaggregated strategy for route choice, while the rest use an aggregated one.

The MIS latent class model showed that passengers have differences in their perceptions of some route attributes. For example, passengers using the aggregated alternatives strategy prefer to travel by bus rather than Metro, while passengers using the disaggregated strategy prefer to travel by Metro rather than bus. Waiting time is more burdensome for passengers who use common lines, which is probably the reason why they consider more than one line as part of the same alternative. Walking time and bus crowding is more burdensome for passengers who use disaggregated alternatives, which suggests that they prefer specific lines in order to avoid walking and/or crowding.

The findings of this research can be used to suggest some guidelines for PT authorities that can help steer decisions regarding how to improve the design, planning, and operation of the PT system following the preferences of passengers. Firstly, given that a percentage of passengers use common lines, implementing lines with an overlap in high-demand sectors will effectively allow passengers to reduce their waiting time. Also, in order to reduce waiting times and improve passenger perception of PT, it may be important that transport authorities make real-time information channels available, allowing users to know which common line alternatives allow them to reach their destination. Secondly, identifying OD pairs that required transfers would focus on planning efforts to reduce the number of onerous transfers (bus-to-bus and Metro-to-bus). Thirdly, given that our models show that transfers requiring walking generate larger disutility levels for passengers, high-demand transfer points should be carefully designed or redesigned to avoid this problem.

2. **How do different consideration set generation practical approaches impact estimation and prediction in a PT route choice model?**

One of the biggest challenges in PT route choice modeling, estimated with RP data, is to identify the set of alternatives that were considered by the passengers as attractive alternatives when they made the route choice decision. This set of alternatives is called the consideration set and the difficulty of identifying it lies in the fact that there are countless alternative routes in a transport network, especially in a dense multimodal one. The consideration set composition can importantly impact the model estimates and model predictions. In the route choice literature, there are various feasible approaches proposed to build the consideration set. While some research has been carried out on the impacts of the different consideration set generation approaches in the model estimates for private transport, no single study exists which evaluates this issue for PT systems.

Two practical ways to identify the consideration set can be distinguished in the applied route choice literature: build it using an algorithm or heuristic that emulates how individuals may build the consideration set, or impute it from historical data. Most of these heuristics are based on the assumption that travelers minimize a cost function and therefore they require the specification of a transit network to iteratively apply some deterministic minimum cost path search. During the last decade, various authors have imputed the consideration set using the previous travelers' route choices. This is called the Historical/Cohort approach in this study, which allows avoiding the execution of any heuristic that can be computationally expensive for a large-scale transit network.

In chapter 3, we formally defined the Historical/Cohort approach as the consideration set built from past choices made by the same individual, or from the choices of other individuals in the same context (OD pair, period, and trip purpose). Then, following Guevara et al. (2020), we derived the required sampling correction when the consideration set is constructed from past choices and formalize the conditions under which a route choice model that uses the Historical/Cohort approach to generate the consideration set can obtain population parameters.

In the same chapter, we assessed the relative performance of the Historical/Cohort approach and six other feasible heuristics approaches: the Labeling approach, Link elimination approach, Link penalty approach, K-shortest path approach, Simulation approach, and Combined approach. Using SC transaction data from weekdays and peak morning periods from Santiago, Chile, we estimated an MNL and a PSL model for each consideration set generation approach. With three weeks of data, we evaluated the impact of different consideration set generation approaches by assessing the consistency of the model estimates. Finally, with one week of data, we studied the out-of-sample prediction performance obtained by each technique.

The results of this analysis (see chapter 3), show that all consideration set generation approaches allow to well-represent the perception of passengers regarding the waiting time, in-vehicle travel time, and the number of transfers. However, walking transfer time was well-represented only by the Historical/Cohort approach with a negative sign, while the models that used an heuristic approach obtained a non-expected positive sign for this attribute. Additionally, the prediction performance analysis shows that the Historical/Cohort method outperforms all other heuristic approaches analyzed. This empirical evidence supports the theoretical results about the ability of the Historical/Cohort approach to recover population parameters and suggests the convenience of using this approach, whenever feasible, beyond the PT route choice context.

This study benefits decision-makers in large-scale transport systems by providing a methodology to understand passengers' perceptions and behavior using passive-transport data and avoiding heuristics approaches that can be computationally expensive in a large-scale transit network.

3. **How can public transport passengers' past experiences be integrated into a route choice model to incorporate the uncertain nature of in-vehicle travel time in the public transport system?**

The route-choice decision process is a dynamic process where passengers have to evaluate the different route attributes and select just one route alternative from their consideration set. There are many uncertainties involved in this decision process since different events and incidents can occur in the trajectory of a travel route. These uncertainties are more important in those PT systems that do not consider dedicated lanes for PT vehicles. It makes that some route attributes are not fixed in time and they vary according to the situation in the PT system, such as the in-vehicle travel time. In this context, PT passengers evaluate the route attributes using knowledge from their past experience and from descriptive travel information. Consequently, the PT passengers' learning process plays a fundamental role to understand their route choice behavior.

Most PT route choice studies consider uncertainty route attributes as of static attributes and that all passengers have the same knowledge of the travel time distribution. Consequently, they ignore the relationship between passengers' learning process from their past experiences and their current choices. In chapter 4, we adressed to fill this research gap by incorporating, into a PT route choice model, the passengers' perceived in-vehicle travel time, which can vary across time. Since a modification to the PT system offers a particularly good opportunity to assess the learning process of PT passengers, we used the case of a new metro line (line 6) in Santiago, Chile, which was launched on November 3, 2017.

Using 3 months (one month previous and two months after to the launch of the metro line 6) of SC transaction data from weekdays and peak morning periods, we first analyzed the data to understand the effect of the new metro line on the total travel time, in the in-vehicle travel time, and the number of transfers. After that, we estimated two models: the IBL-PSL model, which combines the perceived in-vehicle travel time with the mean in-vehicle travel time, assuming that passengers make route choice decisions based on information from experience, if they have previously used the alternative, and descriptive information if they have not previously used the alternative, and the baseline model, the PSL model, which considers in-vehicle travel time as a static attribute and assumes that passengers make route choice decisions using purely descriptive information.

To represent the perceived in-vehicle travel time in the IBL-PSL model, we used an Instance Basel Learning (IBL) model proposed by Tang et al. (2017), which captures the recency effects and the power law of forgetting present in travelers' day-to-day learning processes. The results showed that both models can be used to represent the perception of passengers for in-vehicle travel time, waiting time, and number of transfers. In particular, the parameter of the rate of forgetting of the IBL-PSL is positive and statistically significant, which is in line with the IBL theory. Also, results suggest that during the first weeks of a new context in a PT system, the passengers use mainly descriptive information to select a route, while weeks afterward, passengers start to consider experience-based information to make route choice decisions.

The results of the chapter 4 are very useful for PT authorities since they can support the idea that incrementing the amount of descriptive information is highly recommended at the beginning of the new context in the PT system to improve the passengers' trip experience.

4. **How can passengers be encouraged to provide mobility and transport information through crowdsourced mobile public transport applications?**

One of the biggest challenges in PT route choice modeling using passive transport data is the lack of some important trip information, such as the access/egress time, trip purposes, and the state of the vehicles and buses. In this context, the usage of crowdsourcing applications applied to public transport provides the opportunity not only to give more passengers access to information about the transport system but also to have better knowledge of passengers' behavior and perceptions, which can allow complementing the missing data in the currently passive transport data. However, these crowdsourcing applications, which are voluntary participatory information systems, require a critical mass of users willing to share transport information to be useful. Even when some studies have made some effort to increase the users' participation in crowdsourcing transport applications, mainly using gamification or "quid pro quo" techniques, research on which type of incentives allows to encourage new users to participate in these new data collection technologies is scarce.

We collaborated with Transapp, a widely-used transport-oriented crowdsourcing application in Santiago, Chile. This app allows users to access real-time travel information and to share travel information such as driver behavior, overcrowding, bus conditions, and stops conditions. The reporting feature suffers from the low participation problem. Only 83.27% of active users, who used the app at least once in the year before this study, did not share any reports (called in this study *lurkers*). In chapter 5, we examine the use of economic incentives (a lottery for free trips) and prosocial messages (asking users to help the community) to encourage Transapp's users to share reports about bus stop conditions in order to increase users' level of participation and contributions. For that, we conducted a large-scale field experiment in which users received push notifications, offering an economic incentive, a prosocial message, both (economic incentive and prosocial message), and there was a control group that did not receive any message.

To examine the effect of each experimental condition on the participation rate (number of users who shared at least one report) we estimated a logit model, and to examine the level of contribution (number of reports shared by users) we estimated a zero-inflated negative binomial model, which is highly recommended for data distribution with an excess of zeros. The results show that economic incentives and cooperations messaging increased the participation rate and the level of contribution. For users in the economic incentive group, the relative increases of the participation rate compared to the control group were 294%, this value decreased to 201% when a cooperation message is combined with an economic incentive, and to 72% for those users who received only a cooperation message. This means that the economic incentive alone increase the participation rate most effectively. Regarding to the level of user contribution conditional on participation, users who were offered an economic incentive made 19.5% more reports when compared to users in the control group, while user who were offered a cooperation message did not show a significant difference with respect to the number of shared reports by the control group. These results suggest that the economic incentive not only increased the number of users that share reports, but incentivized that user share reports more frequently.

Based on this study, we show that it is possible to complement current passive-transport databases using transport-oriented crowdsourcing applications. Consequently, we recommend PT authorities to motivate PT passengers, via PT crowdsourcing applications for smartphones, to gather infrastructure data and service quality of the transport system, taking into account that while both economic and prosocial incentives are effective, the first will allow gathering more mobility data with a small financial cost as a lottery way.

6.2 Recommendations for future research

This thesis uses passive transport data to estimate PT route choice models that allow answering the first three research questions (section 1.4.2). One of the main purposes of a PT route choice model is to understand the trip preferences of passengers, which are related to the attributes of route alternatives and depend on the characteristics of passengers. Consequently, it is relevant to consider attributes of route alternatives and characteristics of passengers as inputs in the route choice modeling process. Once we understand travelers' preferences, it is possible to understand and predict the PT passengers' route choice decisions. It is important to note that after passengers make a route choice decision, they experience a trip and consequently they update their perception about the route attributes, which can be essential for the future route choice decision of the passenger. Additionally, passengers can use descriptive travel information, such as waiting times and/or total travel time, to learn about the route alternatives attributes.

Most PT route choice studies that use SC data incorporate route alternative attributes obtained from passive-transport data (mixing AVL, AFC, and GTFS data), however, very little attention has been paid to the role of PT passengers' learning process in the route choice decisions. This research considers different alternative attributes, such as in-vehicle travel time, out-of-vehicle travel time, and the number of transfers. In chapter 2 and 3, we assume that all route alternative attributes as of static values, ignoring that passengers can vary their perception of the route attributes based on their experience. In chapter 4, we incorporate this issue, proposing a methodology to integrate the PT passengers' experiences into a PT route choice modeling. However, further work needs to be done to understand how knowledge from real-time information affects the route choice behavior of PT passengers and to establish a methodology to combine PT passengers' knowledge from experiences and from real-time information into a PT route choice model.

This research is limited by the lack of passengers' characteristics information, such as their socio-demographic information and attitudes, users' perceptions, and some attributes of the trips, such as the access and egress time of the trip, which can not be directly estimated using passive-transport data only. In this line, further research needs to examine more closely the links between transport-passive data and traditional survey data with the purpose to establish a methodology that allows complementing both data sources. In chapter 5, we suggest that a transport-oriented crowdsourcing application can be a channel to collect the missing information in transport-passive data. We have shown that a small economic incentive allows motivating users to share transport information, however, several questions remain to be answered, such as which type of economic incentive obtains a higher level of users' participation, a small incentive for each contributor or a big incentive as a lottery way between contributors? and how long is the effect of each type of incentive?.

In chapters 2, 3, and 4, we use observations from frequent passengers that travel during morning peak periods, and that stay in the destination locations for at least two hours. In this way, we try to capture trips to regular activities such as work or study. A natural progression of this thesis is to analyze the PT passengers' route choices using data from other periods of the day, such as the afternoon peak period and off-peak hours, where one would expect that passengers have different preferences. We speculate that, for those periods, the disutility of the waiting time and travel time might be smaller since users do not have a tight schedule to reach their destination, as it happens in the morning peak hours with trips to work or study.

Since the PT system in Santiago operates by headway scheduling, without fixed-time schedules, the analysis of PT passengers' route choices undertaken here has extended our knowledge of the route choice behavior of passengers that travel in frequency-based transit networks. Several questions remain to be answered about the route choice strategy used by passengers that travel in other contexts. For example, passengers that travel in a schedule-based transit network, where users can optimize their trips by planning their arrival time at the bus stop. These contexts allow passengers to save waiting time (if the transit lines are punctual or the online information is accurate) while choosing their preferred lines.

Bibliography

Alsger, A., Assemi, B., Mesbah, M., & Ferreira, L. (2016). Validating and improving public transport origin–destination estimation algorithm using smart card fare data. *Transportation Research Part C: Emerging Technologies*, *68*, 490–506.

Amaya, M., Cruzat, R., & Munizaga, M. A. (2017). Estimating the residence zone of frequent public transport users to make travel pattern and time use analysis. *Journal of Transport Geography*.

Anderson, M. K., Nielsen, O. A., & Prato, C. G. (2017). Multimodal route choice models of public transport passengers in the greater copenhagen area. *EURO Journal on Transportation and Logistics*, *6*(3), 221–245.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal*, *100*(401), 464–477.

Andreoni, J. (1993). An experimental test of the public-goods crowding-out hypothesis. *The American economic review*, 1317–1327.

Andreoni, J., & Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737–753.

Arriagada, J., Gschwender, A., Munizaga, M. A., & Trépanier, M. (2019). Modeling bus bunching using massive location and fare collection data. *Journal of Intelligent Transportation Systems*, *23*(4), 332–344.

Arriagada, J., & Munizaga, M. (2017). What information can be obtained from a public transport app? In *11th international conference on transport survey methods: In the era of big data: Facing the challenges, esterel, canada.*

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. *Transportation Research Part C: Emerging Technologies*, *134*, 103467.

Avineri, E., & Waygood, E. O. D. (2013). Applying valence framing to enhance the effect of information on transport-related carbon dioxide emissions. *Transportation research part A: policy and practice*, *48*, 31–38.

Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data.

Transport Policy, *12*(5), 464–474.

Bamberg, S., & Schmidt, P. (2001). Theory-driven subgroup-specific evaluation of an intervention to reduce private car use 1. *Journal of Applied Social Psychology*, *31*(6), 1300–1329.

Ben-Akiva, M. (1989). *Lecture notes of large set of alternatives, with a correction of result in bal chapter 9.* Unpublished Manuscript, Massachusetts Institute of Technology.

Ben-Akiva, M., Bergman, M., Daly, A. J., & Ramaswamy, R. (1984). Modelling inter urban route choice behaviour. In *Papers presented during the ninth international symposium on transportation and traffic theory held in delft the netherlands, 11-13 july 1984.*

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice methods and their applications to short-term travel decisions. In R. Hall (Ed.), *Handbook of transportation science* (p. 5-34). Kluwer.

Ben-Akiva, M., & Boccara, B. (1995). Discrete choice models with latent choice sets. *International journal of Research in Marketing*, *12*(1), 9–24.

Ben-Akiva, M., & Lerman, S. (1985). *Discrete choice analysis: theory and application to travel demand* (Vol. 9). MIT press.

Ben-Elia, E., & Avineri, E. (2015). Response to travel information: A behavioural review. *Transport reviews*, *35*(3), 352–377.

Ben-Elia, E., & Ettema, D. (2011a). Changing commuters' behavior using rewards: A study of rush-hour avoidance. *Transportation research part F: traffic psychology and behaviour*, *14*(5), 354–368.

Ben-Elia, E., & Ettema, D. (2011b). Rewarding rush-hour avoidance: A study of commuters' travel behavior. *Transportation Research Part A: Policy and Practice*, *45*(7), 567–582.

Bliemer, M. C., & Bovy, P. H. (2008). Impact of route choice set on route choice probabilities. *Transportation Research Record*, *2076*(1), 10–19.

Bogers, E. A., Bierlaire, M., & Hoogendoorn, S. P. (2007). Modeling learning in route choice. *Transportation Research Record*, *2014*(1), 1–8.

Bohara, A. K., Caplan, A. J., & Grijalva, T. (2007). The effect of experience and quantity-based pricing on the valuation of a curbside recycling program. *Ecological Economics*, *64*(2), 433–443.

Bonnel, P., & Munizaga, M. A. (2018). Transport survey methods-in the era of big data facing new and old challenges. *Transportation Research Procedia*, *32*, 1–15.

Bovy, P. H., Bekhor, S., & Prato, C. G. (2008). The factor of revisited path size: Alternative derivation. *Transportation Research Record*, *2076*(1), 132–140.

Bovy, P. H., & Hoogendoorn-Lanser, S. (2005). Modelling route choice behaviour in multi-modal transport networks. *Transportation*, *32*(4), 341–368.

Brown, J. J., & Wildt, A. R. (1992). Consideration set measurement. *Journal of the Academy of Marketing Science*, *20*(3), 235–243.

Cascetta, E., & Cantarella, G. E. (1991). A day-to-day and within-day dynamic stochastic assignment model. *Transportation Research Part A: General*, *25*(5), 277–291.

Castro, M., Martínez, F., & Munizaga, M. A. (2013). Estimation of a constrained multinomial logit model. *Transportation*, *40*(3), 563–581.

Cepeda, M., Cominetti, R., & Florian, M. (2006). A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. *Transportation research part B: Methodological*, *40*(6), 437–459.

Chapleau, R., Chu, K. K. A., & Allard, B. (2011). *Synthesizing afc, apc, gps and gis data to generate performance and travel demand indicators for public transit* (Tech. Rep.).

Chriqui, C., & Robillard, P. (1975). Common bus lines. *Transportation science*, *9*(2), 115–121.

Cominetti, R., & Correa, J. (2001). Common-lines and passenger assignment in congested transit networks. *Transportation science*, *35*(3), 250–267.

Crawford, G. S., Griffith, R., & Iaria, A. (2021). A survey of preference estimation with unobserved choice set heterogeneity. *Journal of Econometrics*, *222*(1), 4–43.

Cui, A. (2006). *Bus passenger origin-destination matrix estimation using automated data collection systems* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Córdova, A., Imas, A., & Schwartz, D. (2021). Are non-contingent incentives more effective in motivating new behavior? evidence from the field. *Games and Economic Behavior*, *130*, 602-615.

Daly, A., Hess, S., & de Jong, G. (2012). Calculating errors for measures derived from choice modelling estimates. *Transportation Research Part B: Methodological*, *46*(2), 333–341.

De Cea, J., & Fernández, E. (1993). Transit assignment for congested public transport systems: an equilibrium model. *Transportation science*, *27*(2), 133–147.

De Cea, J., Fernandez, J. E., Dekock, V., Soto, A., & Friesz, T. L. (2003). Estraus: a computer package for solving supply-demand equilibrium problems on multimodal urban transportation networks with multiple user classes. In *Proceedings of the 82th transportation research board (trb) annual meeting* (pp. 1–33).

de Grange, L., Raveau, S., & González, F. (2012). A fixed point route choice model for transit networks that addresses route correlation. *Procedia-Social and Behavioral Sciences*, *54*, 1197–1204.

Dell'Olio, L., Ibeas, A., & Cecin, P. (2011). The quality of service desired by public transport users. *Transport Policy*, *18*(1), 217–227.

Devillaine, F., Munizaga, M., & Trépanier, M. (2012). Detection of activities of public transport users by analyzing smart card data. *Transportation Research Record: Journal of the Transportation Research Board*(2276), 48–55.

Eboli, L., & Mazzulla, G. (2007). Service quality attributes affecting customer satisfaction for bus transit. *Journal of public transportation*, *10*(3), 2.

Eluru, N., Chakour, V., & El-Geneidy, A. M. (2012). Travel mode choice and transit route choice behavior in montreal: insights from mcgill university members commute patterns. *Public Transport*, *4*(2), 129–149.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological review*, *112*(4), 912.

Espinoza, C., Munizaga, M., Bustos, B., & Trépanier, M. (2017). Assessing the public transport travel behavior consistency from smart card data. *Working paper*.

Faber, A., & Matthes, F. (2016). Crowdsourcing and crowdinnovation. *Digital Mobility Platforms and Ecosystems*, 36–48.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *The American economic review*, *87*(4), 746–755.

Gallo, G., & Pallottino, S. (1988). Shortest path algorithms. *Annals of operations research*, *13*(1), 1–79.

Glasser, G. J. (1962). Variance formulas for the mean difference and coefficient of concentration. *Journal of the American Statistical Association*, *57*(299), 648–654.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of economic perspectives*, *25*(4), 191–210.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *The Quarterly journal of economics*, *115*(3), 791–810.

Godachevich, J., & Tirachini, A. (2021). Does the measured performance of bus operators depend on the index chosen to assess reliability in contracts? an analysis of bus headway variability. *Research in Transportation Economics*, *90*, 101000.

Gordon, J., Koutsopoulos, H., Wilson, N., & Attanucci, J. (2013). Automated inference of linked transit journeys in london using fare-transaction and vehicle location data. *Transportation Research Record: Journal of the Transportation Research Board*(2343), 17–24.

Grison, E., Burkhardt, J.-M., & Gyselinck, V. (2017). How do users choose their routes in public transport? the effect of individual profile and contextual factors. *Transportation Research Part F: Traffic Psychology and Behaviour*, *51*, 24–37.

Gschwender, A., Munizaga, M., & Simonetti, C. (2016). Using smart card and gps data for policy and planning: The case of transantiago. *Research in Transportation Economics*, *59*, 242–249.

Guevara, C. A. (2015). Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research Part A: Policy and Practice*, *82*, 240–254.

Guevara, C. A. (2022). A note on "a survey of preference estimation with unobserved choice set heterogeneity" by gregory s. crawford, rachel griffith, and alessandro iaria. *Note*.

Guevara, C. A., & Ben-Akiva, M. E. (2013a). Sampling of alternatives in logit mixture models. *Transportation Research Part B: Methodological*, *58*, 185–198.

Guevara, C. A., & Ben-Akiva, M. E. (2013b). Sampling of alternatives in multivariate extreme value (mev) models. *Transportation Research Part B: Methodological*, *48*, 31–52.

Guevara, C. A., Chorus, C. G., & Ben-Akiva, M. E. (2016). Sampling of alternatives in random regret minimization models. *Transportation Science*, *50*(1), 306–321.

Guevara, C. A., & Polanco, D. (2016). Correcting for endogeneity due to omitted attributes in discrete-choice models: the multiple indicator solution. *Transportmetrica A: Transport Science*, *12*(5), 458–478.

Guevara, C. A., Tirachini, A., Hurtubia, R., & Dekker, T. (2020). Correcting for endogeneity due to omitted crowding in public transport choice using the multiple indicator solution (mis) method. *Transportation Research Part A: Policy and Practice*, *137*, 472–484.

Guo, S., Yu, L., Chen, X., & Zhang, Y. (2011). Modelling waiting time for passengers transferring from rail to buses. *Transportation Planning and Technology*, *34*(8), 795–809.

Guo, Z. (2011). Mind the map! the impact of transit maps on path choice in public transit. *Transportation Research Part A: Policy and Practice*, *45*(7), 625–639.

Guo, Z., & Wilson, N. H. (2011). Assessing the cost of transfer inconvenience in public transport systems: A case study of the london underground. *Transportation Research Part A: Policy and Practice*, *45*(2), 91–104.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?–a literature review of empirical studies on gamification. In *2014 47th hawaii international conference on system sciences* (pp. 3025–3034).

Hauser, J. R. (2014). Consideration-set heuristics. *Journal of Business Research*, *67*(8), 1688–1699.

Heyman, J., & Ariely, D. (2004). Effort for payment: A tale of two markets. *Psychological science*, *15*(11), 787–793.

Hilton, D., Charalambides, L., Demarque, C., Waroquier, L., & Raux, C. (2014). A tax can nudge: The impact of an environmentally motivated bonus/malus fiscal system on transport preferences. *Journal of Economic Psychology*, *42*, 17–27.

Hong, J., McArthur, D. P., & Livingston, M. (2020). The evaluation of large cycling infrastructure investments in glasgow using crowdsourced cycle data. *Transportation*, *47*(6), 2859–2872.

Hoogendoorn-Lanser, S., Schaap, N. T., & OldeKalter, M.-J. (2015). The netherlands mobility panel: An innovative design approach for web-based longitudinal travel data collection. *Transportation Research Procedia*, *11*, 311–329.

Hoogendoorn-Lanser, S., & Van Nes, R. (2004). Multimodal choice set composition: Analysis of reported and generated choice sets. *Transportation research record*, *1898*(1), 79–86.

Hoogendoorn-Lanser, S., van Nes, R., & Bovy, P. (2005). Path size modeling in multimodal route choice analysis. *Transportation research record*, *1921*(1), 27–34.

Horowitz, J. L. (1984). The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological*, *18*(1), 13–28.

Ingvardson, J. B., Nielsen, O. A., Raveau, S., & Nielsen, B. F. (2018). Passenger arrival and waiting time distributions dependent on train service frequency and station characteristics: A smart card data analysis. *Transportation Research Part C: Emerging Technologies*, *90*, 292–306.

INRO. (1996). *Emme/2 user's manual.*

Jakobsson, C., Fujii, S., & Gärling, T. (2002). Effects of economic disincentives on private car use. *Transportation*, *29*(4), 349–370.

Jánošíková, L., Slavík, J., & Koháni, M. (2014). Estimation of a route choice model for urban public transport using smart card data. *Transportation planning and technology*, *37*(7), 638–648.

Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annu. Rev. Econ.*, *4*(1), 427–452.

Khattak, A., Polydoropoulou, A., & Ben-Akiva, M. (1996). Modeling revealed and stated pretrip travel response to advanced traveler information systems. *Transportation Research Record: Journal of the Transportation Research Board*(1537), 46–54.

Kim, Kim, H.-C., Seo, D.-J., & Kim, J. I. (2020). Calibration of a transit route choice model using revealed population data of smartcard in a multimodal transit network. *Transportation*, 1–24.

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

Kurauchi, F., Schmöcker, J.-D., Shimamoto, H., & Hassan, S. M. (2014). Variability of commuters' bus line choice: an analysis of oyster card data. *Public Transport*, *6*(1-2), 21–34.

Kusakabe, T., & Asakura, Y. (2014). Behavioural data mining of transit smart card data: A data fusion approach. *Transportation Research Part C: Emerging Technologies*, *46*, 179–191.

Kusakabe, T., Iryo, T., & Asakura, Y. (2010). Estimation method for railway passengers' train choice behavior with smart card transaction data. *Transportation*, *37*(5), 731–749.

Lacetera, N., Macis, M., & Slonim, R. (2012). Will there be blood? incentives and displacement effects in pro-social behavior. *American Economic Journal: Economic Policy*, *4*(1), 186–223.

Lacetera, N., Macis, M., & Slonim, R. (2014). Rewarding volunteers: A field experiment. *Management Science*, *60*(5), 1107–1129.

Lau, J. K. S., Tham, C.-K., & Luo, T. (2011). Participatory cyber physical system in public transport application. In *2011 fourth ieee international conference on utility and cloud computing* (pp. 355–360).

Lee, S. G., & Hickman, M. (2014). Trip purpose inference using automated fare collection data. *Public Transport*, *6*(1-2), 1–20.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*, *25*(2), 143–153.

Li, Q., Chen, P. W., & Nie, Y. M. (2015). Finding optimal hyperpaths in large transit networks with realistic headway distributions. *European Journal of Operational Research*, *240*(1), 98–108.

Li, T., Sun, D., Jing, P., & Yang, K. (2018). Smart card data mining of public transport destination: A literature review. *Information*, *9*(1), 18.

Ling, K., Beenen, G., Ludford, P., Wang, X., Chang, K., Li, X., … others (2005). Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication*, *10*(4), 00–00.

Long, Y., & Thill, J.-C. (2015). Combining smart card data and household travel survey to analyze jobs–housing relationships in beijing. *Computers, Environment and Urban Systems*, *53*, 19–35.

Lu, X., Gao, S., Ben-Elia, E., & Pothering, R. (2014). Travelers' day-to-day route choice behavior with real-time information in a congested risky network. *Mathematical population studies*, *21*(4), 205–219.

Ma, X., Liu, C., Wen, H., Wang, Y., & Wu, Y.-J. (2017). Understanding commuting patterns using transit smart card data. *Journal of Transport Geography*, *58*, 135–145.

Mahmassani, H. S., & Liu, Y.-H. (1999). Dynamics of commuting decision behaviour under advanced traveller information systems. *Transportation Research Part C: Emerging Technologies*, *7*(2-3), 91–107.

Malandri, C., Fonzone, A., & Cats, O. (2018). Recovery time and propagation effects of passenger transport disruptions. *Physica A: Statistical Mechanics and its Applications*, *505*, 7–17.

Manski, C. F. (1977). The structure of random utility models. *Theory and decision*, *8*(3), 229.

Martínez, F., Aguila, F., & Hurtubia, R. (2009). The constrained multinomial logit: A semi-compensatory choice model. *Transportation Research Part B: Methodological*, *43*(3), 365–377.

McFadden, D. (1978). Modeling the choice of residential location. *Karlquist, Lundqvist, Snickers, Weibull (Eds.), Spatial Interaction Theory and Residential Location, North Holland, Amsterdam*(1), 75–96.

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

Mondschein, A. (2015). Five-star transportation: using online activity reviews to examine mode choice to non-work destinations. *Transportation*, *42*(4), 707–722.

Mora-Garcia, R.-T., Marti-Ciriquian, P., Perez-Sanchez, R., & Cespedes-Lopez, M. F. (2018). A comparative analysis of manhattan, euclidean and network distances. why are network distances more useful to urban professionals? *International Multidisciplinary Scientific GeoConference: SGEM*, *18*(2.2), 3–10.

Morency, C., Trépanier, M., & Agard, B. (2007). Measuring transit use variability with smart-card data. *Transport Policy*, *14*(3), 193–203.

Munizaga, M. A., & Palma, C. (2012). Estimation of a disaggregate multimodal public transport origin–destination matrix from passive smartcard data from santiago, chile. *Transportation Research Part C: Emerging Technologies*, *24*, 9–18.

Nam, D., Kim, H., Cho, J., & Jayakrishnan, R. (2017). A model based on deep learning for predicting travel mode choice. In *Proceedings of the transportation research board 96th annual meeting transportation research board, washington, dc, usa* (pp. 8–12).

Nandan, N., Pursche, A., & Zhe, X. (2014). Challenges in crowdsourcing real-time information for public transportation. In *2014 ieee 15th international conference on mobile data management* (Vol. 2, pp. 67–72).

Nassir, N., Hickman, M., & Ma, Z. (2017). Statistical inference of transit passenger boarding strategies from farecard data. *Transportation Research Record: Journal of the Transportation Research Board*(2652), 8–18.

Nassir, N., Hickman, M., & Ma, Z.-L. (2015). Activity detection and transfer identification for public transit fare card data. *Transportation*, *42*(4), 683–705.

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

Nassir, N., Khani, A., Lee, S. G., Noh, H., & Hickman, M. (2011). Transit stop-level origin–destination estimation through use of transit schedule and automated data collection system. *Transportation research record*, *2263*(1), 140–150.

NationMaster. (2019). *Sixty percent of shoppers are influenced by online reviews.* Retrieved 2021-10-18, from `https://www.nationmaster.com/blog/sixty-percent-of-people-are-influenced-by-online-reviews-before-making-a-purchase/`

Nguyen, S., & Pallottino, S. (1988). Equilibrium traffic assignment for large scale transit networks. *European journal of operational research*, *37*(2), 176–186.

Nielsen, J. (2006). *The 90-9-1 rule for participation inequality in social media and online communities.* Retrieved 2021-10-18, from `https://www.nngroup.com/articles/participation-inequality/`

Nielsen, O. A., Eltved, M., Anderson, M. K., & Prato, C. G. (2021). Relevance of detailed transfer attributes in route choice models for public transport passengers. *Transportation Research Part A: Policy and Practice, forthcomin.*

Niessen-Ruenzi, A., Weber, M., & Becker, D. M. (2014). To pay or not to pay? evidence from whole blood donations in germany. *Working paper*.

Nov, O. (2007). What motivates wikipedians? *Communications of the ACM*, *50*(11), 60–64.

Núñez, C., Munizaga, M., Gschwender, A., & Transantiago, C. (2015). Cálculo de indicadores de calidad de servicio del sistema de transporte público de santiago a partir de datos pasivos. In *Congreso chileno de ingeniería de transporte. anales. disponível em:¡ http://www. sochitran. cl/wp-content/uploads/acta-2013-10-06. pdf.*

Pelletier, M.-P., Trépanier, M., & Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, *19*(4), 557–568.

Pineda, C., & Lira, B. M. (2019). Travel time savings perception and well-being through public transport projects: The case of metro de santiago. *Urban Science*, *3*(1), 35.

Polson, N. G., & Sokolov, V. O. (2017). Deep learning for short-term traffic flow prediction. *Transportation Research Part C: Emerging Technologies*, *79*, 1–17.

Prato, C., & Bekhor, S. (2007). Modeling route choice behavior: how relevant is the composition of choice set? *Transportation Research Record: Journal of the Transportation Research Board*(2003), 64–73.

Prato, C. G. (2009). Route choice modeling: past, present and future research directions.

Journal of choice modelling, *2*(1), 65–100.

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*(1), 1–14.

Raveau, S., Guo, Z., Muñoz, J. C., & Wilson, N. H. (2014). A behavioural comparison of route choice on metro networks: Time, transfers, crowding, topology and socio-demographics. *Transportation Research Part A: Policy and Practice*, *66*, 185–195.

Raveau, S., & Muñoz, J. C. (2014). *Analyzing route choice strategies on transit networks* (Tech. Rep.).

Raveau, S., Muñoz, J. C., & De Grange, L. (2011). A topological route choice model for metro. *Transportation Research Part A: Policy and Practice*, *45*(2), 138–147.

Rose, G., & Ampt, E. (2001). Travel blending: an australian travel awareness initiative. *Transportation Research Part D: Transport and Environment*, *6*(2), 95–110.

Rosenfield, A., Attanucci, J. P., & Zhao, J. (2020). A randomized controlled trial in travel demand management. *Transportation*, *47*(4), 1907–1932.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688–701. Retrieved from `http://psycnet.apa.org/journals/edu/66/5/688/`

Rui, T. (2016). *Modeling route choice behaviour in public transport network* (Unpublished doctoral dissertation). National University of Singapore (Singapore).

Schmöcker, J.-D., Shimamoto, H., & Kurauchi, F. (2013). Generation and calibration of transit hyperpaths. *Transportation Research Part C: Emerging Technologies*, *36*, 406–418.

Schwartz, D., Bruine de Bruin, W., Fischhoff, B., & Lave, L. (2015). Advertising energy saving programs: The potential environmental cost of emphasizing monetary savings. *Journal of Experimental Psychology: Applied*, *21*(2), 158.

Schwartz, D., Fischhoff, B., Krishnamurti, T., & Sowell, F. (2013). The hawthorne effect and energy awareness. *Proceedings of the National Academy of Sciences*, *110*(38), 15242–15246.

Schwartz, D., Keenan, E. A., Imas, A., & Gneezy, A. (2021). Opting-in to prosocial incentives. *Organizational Behavior and Human Decision Processes*, *163*, 132-141.

Schwartz, D., Loewenstein, G., & Agüero-Gaete, L. (2020). Encouraging pro-environmental behaviour through green identity labelling. *Nature Sustainability*, 1–7.

Schwartz, D., Milfont, T. L., & Hilton, D. (2019). The interplay between intrinsic motivation, financial incentives and nudges in sustainable consumption. In *A research agenda for economic psychology.* Edward Elgar Publishing.

Seaborn, C., Attanucci, J., & Wilson, N. (2009). Analyzing multimodal public transport journeys in london with smart card fare payment data. *Transportation Research Record: Journal of the Transportation Research Board*(2121), 55–62.

Spiess, H., & Florian, M. (1989). Optimal strategies: a new assignment model for transit networks. *Transportation Research Part B: Methodological*, *23*(2), 83–102.

Steinfeld, A., Zimmerman, J., Tomasic, A., Yoo, D., & Aziz, R. D. (2011). Mobile transit information from universal design and crowdsourcing. *Transportation research record*, *2217*(1), 95–102.

Swait, J., & Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B: Methodological*, *21*(2), 91–102.

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. *Transportation Research Record*, *2538*(1), 11–18.

Tang, Y., Gao, S., & Ben-Elia, E. (2017). An exploratory study of instance-based learning for route choice with random travel times. *Journal of choice modelling*, *24*, 22–35.

Tao, S., Rohde, D., & Corcoran, J. (2014). Examining the spatial–temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap. *Journal of Transport Geography*, *41*, 21–36.

Thøgersen, J., & Møller, B. (2008). Breaking car use habits: The effectiveness of a free one-month travelcard. *Transportation*, *35*(3), 329–345.

Tien, D. N., MacDonald, T., & Xu, Z. (2011). Tdplanner: Public transport planning system with real-time route updates based on service delays and location tracking. In *2011 ieee 73rd vehicular technology conference (vtc spring)* (pp. 1–5).

Tomasic, A., Zimmerman, J., Steinfeld, A., & Huang, Y. (2014). Motivating contribution in a participatory sensing system via quid-pro-quo. In *Proceedings of the 17th acm conference on computer supported cooperative work & social computing* (pp. 979–988).

Ton, D., Shelat, S., Nijënstein, S., Rijsman, L., van Oort, N., & Hoogendoorn, S. (2020). Understanding the role of cycling to urban transit stations through a simultaneous access mode and station choice model. *Transportation research record*, *2674*(8), 823–835.

Torres, N. (2016). *Why do so few women edit wikipedia?* Retrieved 2021-10-18, from `https://hbr.org/2016/06/why-do-so-few-women-edit-wikipedia`

Trépanier, M., Tranchant, N., & Chapleau, R. (2007). Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems*, *11*(1), 1–14.

Van Der Hurk, E., Kroon, L., Maróti, G., & Vervest, P. (2015). Deduction of passengers' route choices from smart card data. *IEEE Transactions on Intelligent Transportation Systems*, *16*(1), 430–440.

Villalobos, N., & Guevara, C. A. (2021). Caracterización del conjunto de consideración en elección de ruta. *Estudios de Transporte*, *22*(1), 1–26.

Villalobos Zaid, G. N. (2018). Caracterización del conjunto de consideración en elección de ruta.

Vrtic, M., & Axhausen, K. W. (2002). The impact of tilting trains in switzerland: A route choice model of regional-and long distance public transport trips. *Arbeitsberichte Verkehrs-und Raumplanung*, *128*.

Walker, J. L. (2001). *Extended discrete choice models: integrated framework, flexible error structures, and latent variables* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Wang, W., Attanucci, J. P., & Wilson, N. H. (2011). Bus passenger origin-destination estimation and related analyses using automated data collection systems. *Journal of Public Transportation*, *14*(4), 7.

Waygood, E. O. D., & Avineri, E. (2011). Does" 500g of co2 for a five mile trip" mean anything? towards more effective presentation of co2 information. In *Proceedings of the transportation research board 90th annual meeting* (pp. 23–27).

Waygood, E. O. D., & Avineri, E. (2016). Communicating transportation carbon dioxide emissions information: does gender impact behavioral response? *Transportation Research Part D: Transport and Environment*, *48*, 187–202.

Weitzenkorn, B. (2013, Jun). Google to buy social mapping startup waze. *NBC News*. Retrieved from `https://www.nbcnews.com/id/wbna52207395`

Yap, M., Cats, O., & van Arem, B. (2020). Crowding valuation in urban tram and bus transportation based on smart card data. *Transportmetrica A: Transport Science*, *16*(1), 23–42.

Yap, M., Cats, O., van Oort, N., & Hoogendoorn, S. (2017). A robust transfer inference algorithm for public transport journeys during disruptions. *Transportation research procedia*, *27*, 1042–1049.

Yap, M., Nijënstein, S., & van Oort, N. (2018). Improving predictions of public transport usage during disturbances based on smart card data. *Transport Policy*, *61*, 84–95.

Yen, J. Y. (1971). Finding the k shortest loopless paths in a network. *management Science*, *17*(11), 712–716.

YU, J., & YANG, X.-g. (2006). Estimation a transit route od matrix using on/off data: An application of modified bp artificial neural network [j]. *Systems Engineering*, *4*, 018.

Zhao, J., Rahbee, A., & Wilson, N. H. (2007). Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. *Computer-Aided Civil and Infrastructure Engineering*, *22*(5), 376–387.

Zhao, J., Zhang, F., Tu, L., Xu, C., Shen, D., Tian, C., … Li, Z. (2017). Estimation of passenger route choice pattern using smart card data for complex metro systems. *IEEE Transactions on Intelligent Transportation Systems*, *18*(4), 790–801.

Zhou, P., Zheng, Y., & Li, M. (2012). How long to wait? predicting bus arrival time with mobile phone based participatory sensing. In *Proceedings of the 10th international conference on mobile systems, applications, and services* (pp. 379–392).

Zimmerman, J., Tomasic, A., Garrod, C., Yoo, D., Hiruncharoenvate, C., Aziz, R., … Steinfeld, A. (2011). Field trial of tiramisu: crowd-sourcing bus arrival times to spur co-design. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1677–1686).

Zumkeller, D., Chlond, B., Ottmann, P., Kagerbauer, M., & Kuhnimhof, T. (2011). Deutsches mobilitätspanel (mop)–wissenschaftliche begleitung und erste auswertungen. *Kurzbericht. Karlsruhe: Institut für Verkehrswesen, Universität Karlsruhe.*