



Public transport reliability across preferences, modes, and space

Jaime Soza-Parra¹ · Sebastián Raveau¹ · Juan Carlos Muñoz¹

Accepted: 5 March 2021 / Published online: 15 March 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Reliability is understood in public transport as the certainty travellers have regarding the level of service they will experience when travelling. The travel time, waiting time, or the comfort level they will experience inside the vehicle are some of the most important reliability attributes. Reliability is usually neglected from travel behavioural models, since there is a lack of studies addressing its impact in travellers' choices. This study proposes an in-depth analysis and characterization of travel time reliability in a public transport system using passive data only, which every day is getting more popular worldwide. The study is comprised of two parts, taking Transantiago (the public transport system of Santiago, Chile) as a study case, as it is a good representation of a frequency-based public transport system. In the first part a statistical and graphical analysis of travel times is conducted, focusing on characterizing headway and travel time reliability of public transport services as well as the effect of dedicated infrastructure, based on smartcard transactions and GPS information. In the second part an aggregate mode choice model based on revealed data is developed to analyse the effect of travel time reliability on travellers' preferences. Overall, this study provides evidence of significant differences among headway variability and in-vehicle travel time dispersion different public transport modes. The standard deviation (a measure of dispersion) can be quite high for bus trips, while in the case of metro is smaller than 4 min independent of trip length. Regarding the aggregate public transport mode choice model, the coefficient of variation of headways is a key attribute to explain modal preferences. It was found that average bus users would accept traveling ~5 min longer to completely avoid headway irregularity. All these results were obtained by analysing passive-data from AFC and AVL systems only, which represents a novel approach for choice modelling. Transport modellers should consider the impact of service reliability to improve route choice and passenger assignment models, and better represent users' perceptions and behaviour.

Keywords Travel time reliability · Public transport · Traveller's experience · Aggregate demand · Transport planning

✉ Jaime Soza-Parra
jaime.soza@uc.cl

¹ Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago, Chile

Introduction

Travel time reliability plays an important role in public transport travellers' satisfaction and their perception regarding the level of service (Soza-Parra et al. 2019), as well as in operational costs (de Jong and Bliemer 2015). Nevertheless, when planning public transport systems, travellers' behaviour has been usually modelled through traditional variables such as monetary cost, expected travel time and planned waiting time. Other elements such as crowdedness, excess waiting time, and mode/service reliability (understood as the certainty travellers have regarding their travel time, their arrival time or the comfort level they will experience inside the vehicle) (van Oort 2011) are usually neglected from these behavioural models (Petersen and Vovsha 2006; Raveau et al. 2014). This could lead to erroneous predictions of the demand for public transport alternatives.

In public transport systems, unreliability's main impact is generally the potential delay on the arrival to the destination. Travellers can handle this situation by adjusting their departure time, changing routes or changing modes (Benezech and Coulombel 2013). In general, users' preferred option is to add a safety margin to the ideal departure time (Bates et al. 2001). Travellers' reaction to unreliability has been widely studied in developed cities among the world, mainly in Europe and North America. However, in developing regions such as Latin America there is a lack of studies regarding public transport reliability. Besides, developing regions are characterized by an accelerated urbanization process and a significant percentage of urban population (Jirón 2013). This, along with poor urban planning policies, leads to a significant proportion of long trips, from the periphery of the cities to highly concentrated activity centres (García Palomares 2008; Vignoli 2008, 2012). These circumstances hinder the operation of public transport services based only on schedules, mainly because of the high frequency needed and the stochastic nature of public transport (Muñoz and Gschwender 2008). Thus, frequency-based public transport systems are the natural option. In this context, therefore, public transport reliability needs to be understood and addressed in a different way to what has been done in developed cities.

In terms of planning information, the availability of large volumes of automated data regarding the operation of public transport systems has increased over the recent years. This valuable source of detailed information, properly processed, allows analysing and understanding the system's operation (Birr et al. 2014; Bucknell et al. 2017; Cham 2006; Fadaei and Cats 2016; Furth and Muller 2006; Gschwender et al. 2016) and modelling it in a better way (Cats and Gkioulou 2017; van Oort et al. 2015; Raveau 2017; Frappier et al. 2018). This type of information usually comes from sensors strategically placed within vehicles (such as GPS systems) and smartcard data from passengers boarding and/or alighting the vehicles and allows understanding travel times in a better way.

An application of automated data to characterize public transport level of service is the study by the BRT Center of Excellence (BRT 2012), which compares the level of service of six Latin American cities: Santiago, Chile; Porto Alegre, Brazil; Guadalajara, Mexico; Mexico City, Mexico; Bogota, Colombia; Lima, Peru. For each city, a socio-economic description of the population was made, as well as a description of the characteristics of the existing public transport system (such as the number of operators, metro lines, operation, fares, payment schemes, infrastructure, vehicles, information systems, quality perception, among others). Level of service indicators of the respective public transport systems were calculated by estimating travel, wait and walk times for 400 representative trips in each city. A relevant indicator within the study relates to travel time variability in the systems. To compute this indicator, the study defined two distinct types of variability: (1) an

interpersonal variability, which accounts for the heterogeneity of the existing levels of service within the city, and (2) an intrapersonal variability, related to how variable the same trip performed repetitively by an individual is (i.e. how reliable is the level of service). This second kind of variability is called day-to-day variability (Hollander 2006; Jenelius 2012). However, the travel demand is not adequately considered when measuring the average variability, as the indicators are not weighted by the number of travellers performing each trip. Nor is there an in-depth analysis of the differences between modes and/or operating conditions.

The purpose of this study is to use automated data to perform an in-depth analysis of the impacts of travel time components unreliability in a public transport system. We seek to characterize reliability for different modes and types of infrastructure and to measure their effects of travel time reliability on travellers' mode choice decisions. For this, the case of Santiago, Chile, is considered.

Santiago's public transport system is called Transantiago, where bus and metro services are integrated in fare (Muñoz et al. 2014). This means you can transfer to a new service of the same mode without having to pay any extra and only pay the fare difference when transferring from bus to metro. Besides, the fare is fixed regardless the distance travelled. There are mainly two types of bus services: regular services, which stop in every bus stop of the route, and express services, which stop only in some of them. There are also five metro lines, where Line 1 is the most crowded one during peak periods, as it runs through the city centre. Line 1 is also the oldest one and has an average distance of 660 m between stations, the lowest average distance of the network. Furthermore, it is the only line that goes from the west to the east of the city, passing through the most important activity centres. These characteristics make Line 1's performance significantly different and thus it might be needed to study it separately.

In addition, this public transport system operates in a frequency-based scheme. Under these circumstances, passengers do not know in advance when the next vehicle (bus or metro) is going to arrive. In general terms, both metro and buses operate following an operational plan which indicates specific frequencies for each service for different periods of the day. Even though metro operates isolated, both modes suffer vehicle bunching. In fact, passengers indicate headway regularity as a critical attribute when evaluating the quality of service (Soza-Parra et al. 2019).

In Transantiago, the smartcards only record boardings. Munizaga and Palma (2012a, b) proposed a methodology to estimate a public transport trip matrix (inferring alightings) using the sequence of validations made with the smartcard and the geographical position of the buses. This trip matrix is used in this study to characterize travel time reliability for public transport routes of similar length in the city during the morning peak period. Additionally, the progressive change of travel time variability as travel length increases is analysed.

So far, the available automated data has not been used in Santiago to understand how travel time variability has an impact (if any) on user's decisions. Furthermore, to the best of the authors knowledge, revealed preferences have not been used to analyse the impact of reliability on public transport travellers' preferences. Based on the available information, this study develops an aggregate mode choice model, in which the explanatory variables are both average level-of-service indicators and indicators of their variability. This analysis further emphasizes the importance of travel time reliability. The analysis and characterization of the travel time reliability and its effects were carried out using only passive data, without the need of any survey. This represents a novel and quite promising approach for choice modelling.

This document is structured in four sections. “[Characterizing travel time reliability](#)” section describes the statistical analysis conducted to obtain travel time distributions for regular bus, express bus, and metro services, as well as the impact of segregated corridors in travel time variability. The results are supported by both statistical and graphical analysis. “[The effect of travel time reliability on mode choice](#)” section presents the methodology for estimating an aggregate behavioural model as well as its results, allowing us to understand the effect reliability has on passenger choices. Finally, “[Conclusion](#)” section presents the main conclusions of this study and elaborates on how these results should steer following studies and models for public transport planning.

Characterizing travel time reliability

In this section, the methodology applied for the travel time characterization as well as its graphical analysis are presented. To characterize travel times across the city, a statistical analysis of actual travel times of all travellers on a given week is performed. This analysis is conducted at a trip-leg level.

Travel time distributions and headway regularity

For bus trips, the data comes from smartcard transactions and GPS information. This information was extracted from a trip-leg table constructed with the methodology proposed by Munizaga and Palma (2012a, b). This data source considers the period between 07:00 and 09:00 during the morning peak period for a typical workweek in April 2017. For each day, around 210.000 validations are considered. For each smartcard validation, public transport service is recorded as well as the moment and place in which the traveller boards and alights the bus. This information is estimated from the GPS information delivered by the vehicles every 30 s, the geo-referenced bus route and the geographical position of the stops along the route.

The resulting database has the boarding and alighting time for every bus trip-leg made by at least one individual. With this information, it is possible to construct travel time distributions for any service between any pair of stops where there are trips within the network. Origin-destinations pairs without any trips are excluded for the analysis. The travel time distributions can be discretised by the in-route distance that separates the pair of stops, in order to obtain travel time histograms for trips of a given length range.

For the case of metro trips, the database of arrival and departure times for every train at every station was provided by Metro de Santiago. Just like in the case of the buses, it is possible to obtain travel time distributions by distance range for every pair of stations that belong to the same line.

In order to analyse headway distributions, a new database was used for bus trips. This database consists in the estimated arrival time of every bus at every bus stop for the same previous week of analysis. This extensive database was filtered to consider only the same services described above that cover the set of origin–destination pairs of interest. With this information, for both metro and buses it is possible to calculate consecutive headways for every stop/station in the network. Then, for the time-period of analysis, we compute the coefficient of variation of headways as a reliability indicator due to its direct incidence in the expected waiting time.

The previously described methodology treats each vehicle alike, regardless of the number of passengers travelling in them. To obtain travel time distributions from a point of view of the user’s experience, it is necessary to weight the travel time distributions associated to each origin–destination pair by the travel demand of that pair. To do this, information from smartcard transactions is used. In addition to boarding and alighting stops, and the service boarded, the trip matrix contains expansion factors for each observation. Adding all the expansion factors of those trips that had the same vehicle, boarding and alighting stops, it is possible to obtain the demand for all trips within each service.

However, the smartcard demand database used for the metro services contains information of trips between any two stations within the entire metro network (as no transfers are recorded), while the metro travel times are within lines. For this reason, the available data must be transformed so all information corresponds to metro trip-legs within lines. One way to solve this is to divide each metro trip between any pair of stations (which could use different lines) into its trip-legs. This is not straightforward, as for some pair of stations there is more than one reasonable route. To solve this issue, a choice probability to each of these routes was considered. These probabilities correspond to the travellers’ choice proportions, obtained from an appropriate route choice model (Raveau et al. 2011). This model is estimated based on an origin–destination survey conducted within the metro network. As any specific trip-leg that involves a transfer station will be part of multiple origin–destination pairs within the network, the sum of the demand of all those multiple pairs must be computed to have the actual demand of every trip-leg in the metro system.

The entire data set used for the week of analysis in this article, and its relation with the attributes studied, is presented in Fig. 1.

Graphical analysis

As a first step, we performed a graphical analysis to compare for buses and metro how does their coefficient of variation of headways evolve with the distance from the terminal (Fig. 2), as well as the distribution of travel time for trips of similar length (Fig. 3). It is important to highlight that these two indicators are not interchangeably, but they do provide different perspectives of the complex phenomenon of reliability.




 <p>Smartcard trip validations</p>	<p>Travel demand Buses’ travel time (mean and dispersion) Number of transfers Buses’ transfer times</p>
 <p>Buses’ arrival times</p>	<p>Buses’ headways (mean and dispersion)</p>
 <p>Metro data</p>	<p>Metro’s travel time (mean and dispersion) Metro’s headways (mean and dispersion) Metro’s transfer times</p>

Fig. 1 Databases for the same week of April 2017

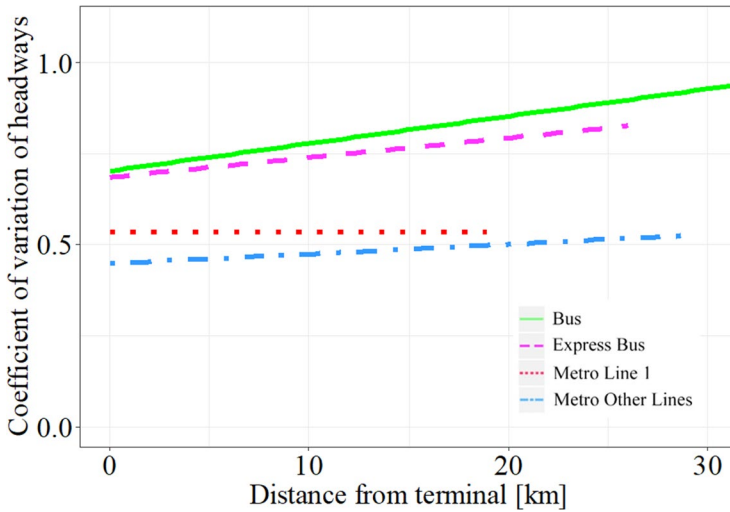


Fig. 2 Coefficient of variation of headways by distance to the terminal for buses and metro services

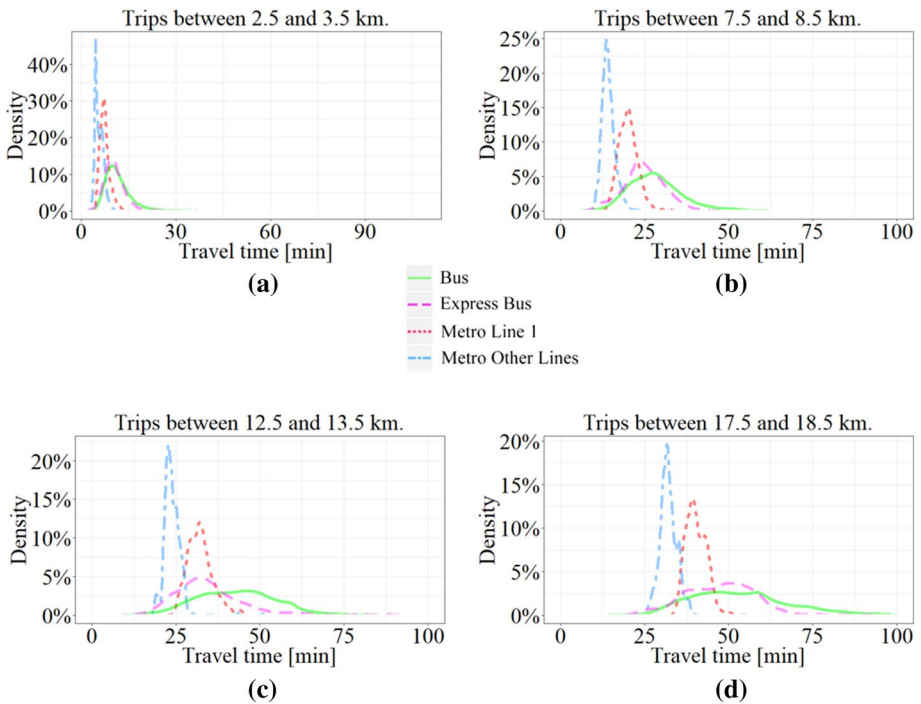


Fig. 3 Travel time distributions by mode and distance range. **a** between 2.5 and 3.5 km; **b** between 7.5 and 8.5 km; **c** between 12.5 and 13.5 km; **d** between 17.5 and 18.5 km

Regarding metro services, there is a significant difference in the speed of metro Line 1 compared to the other lines, and therefore Line 1 is shown separately from the others. There is also a broader dispersion for Line 1 in comparison to the rest of the lines, but, as mentioned in the case of buses, this will be of further analysed. Although the dispersion of the performance of express bus services shows that many of them have a level of service similar to that of regular bus services, there is a portion that resembles both the best lines of metro and Line 1. As future research, it will be important to understand what conditions make these services show such a level of service.

The relationship between travel time dispersion measures of the histograms and distance can be seen in Fig. 4. Two dispersion measures are considered for the analysis: the standard deviation of travel time and the difference between the 95th percentile and the average travel time, which in the literature has been called *Reliability Buffer Time* (RBT, Engelson and Fosgerau 2016). While the standard deviation takes into account travel times shorter and longer than average, the reliability buffer time only measures the difference between the longer travel times and the average. Thus, it is important to analyse different dispersion metrics, as they might lead to different conclusions.

Overall, we observe that both dispersion measures present similar information. This indicate that we might use them interchangeably in subsequent analysis. Besides, both of them increase with travel length for every mode except for the last segment of Line 1. This happens as the number of origin–destination pairs for every travel distance range decreases as the distance gets longer. For the longest distance range, there is only one origin–destination pair analysed, which is from one end to the other. Thus, the variability is expected to be smaller as there is no variability due to differences in demand on different pairs. Besides, the difference between regular and express bus services is higher when analysing the RBT. This indicates that the degree at how the longest travel times grow is higher for this type of buses.

The dispersion is always smaller than 2.5 min for metro (except for Line 1, whose maximum is 4 min) when the standard deviation is considered as the measure of dispersion, which could hardly be perceived by travellers. Considering the reliability buffer time, the dispersion is almost eight minutes for Line 1 and always smaller than five minutes for the rest of the lines.

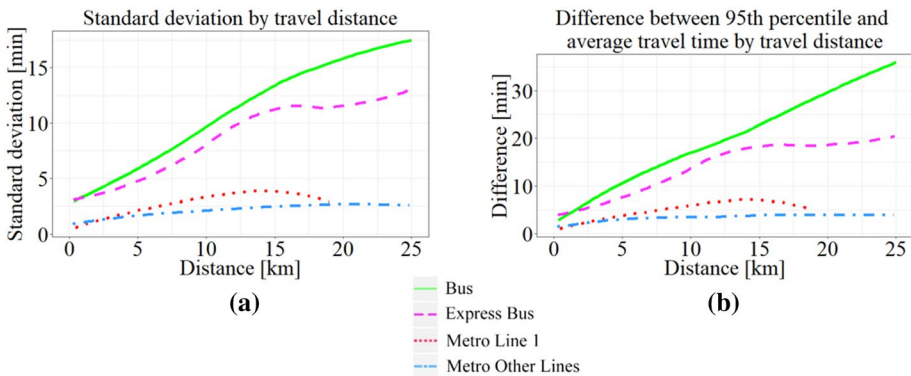


Fig. 4 Relationship between dispersion measures and travel length. **a** Standard deviation; **b** difference between 95th percentile and average travel time

The impact of segregated bus corridors on travel time reliability

Over the last decade, to palliate the effect of traffic congestion (mainly due to the increase of private car use) on the performance of public transport systems, there has been a substantial increase in the length of specialized infrastructure for bus services. According to the BRT Centre of Excellence, there are 4,900 kms of segregated bus corridors moving 32 million passengers daily in 166 cities worldwide (BRT+Centre of Excellence and EMBARQ 2018).

Buses operating in segregated corridors increase their speed in comparison to those operating in mixed traffic. It has also been observed that segregated corridors have a positive impact in avoiding bus bunching by reducing headway variability growth along the route (Danés et al. 2015). However, their impact regarding travel time reliability is less clear.

In this section we provide a graphical comparison between services running on specific segregated corridors and on parallel comparable mixed-traffic lanes based exclusively on passive data. Figure 5 shows the considered corridors. The segregated corridors analysed (dashed lines) are *Las Industrias* and *Av. Grecia* and their comparable mixed traffic corridors (solid lines) are *Vicuña Mackenna* and *Eduardo Castillo Velasco–Los Orientales–Las Parcelas (ECV-LO-LP)*, respectively. To isolate the effect of the segregated corridors, only

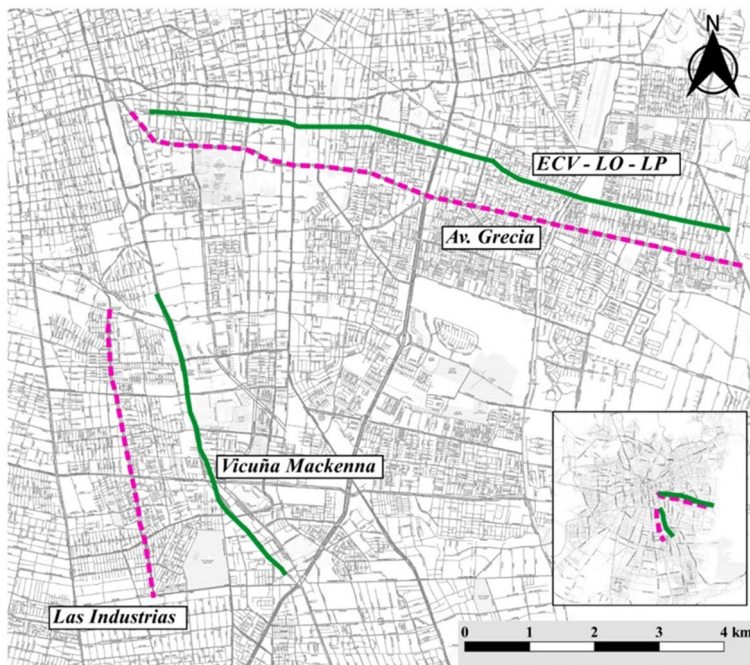


Fig. 5 Segregated corridors analysed and their comparable mixed traffic corridors

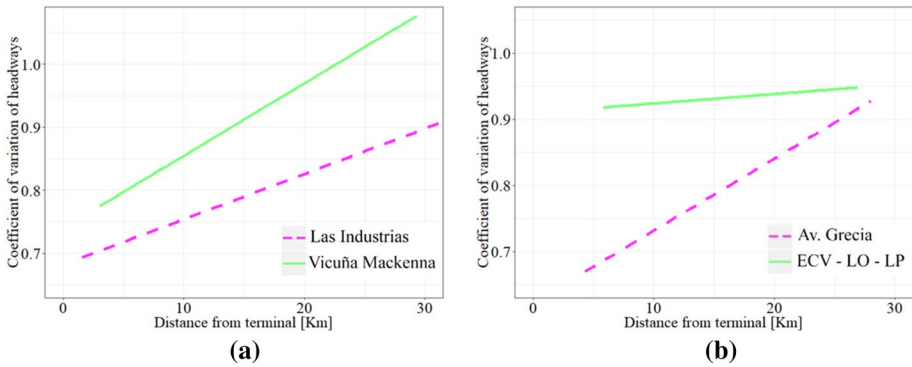


Fig. 6 Coefficient of variation of headways by distance from terminal. **a** Las Industrias & Vicuña Mackenna; **b** Av. Grecia & ECV – LO – LP

trips that started and finished at bus stops inside the analysed corridors (both segregated and mixed-traffic) were considered.

The evolution of the coefficient of variation of headways as vehicles move away from the terminal, as shown in Fig. 2, is presented in Fig. 6. For both bus corridor comparisons, the headway coefficient of variation is lower when services operate in a segregated bus corridor. The evolution of the coefficient of variation for the service operating in a segregated corridor in both cases are quite alike. However, the evolution of the headway coefficient of variation of the services operating in mixed traffic describe different stories in each situation. In (a), mixed-traffic services start their operation at a relatively low value and grows at a faster pace than the segregated corridor service indicator. In (b), mixed-traffic services start their operation with a headway coefficient of close to 1 (which is what we would observe for instance if buses arrived according to a Poisson process). This indicates that the services operating in these streets are already highly irregular, and therefore the pace at which headway variability grows is lower. Thus, it is not surprising that the increment of the headway variability would be slower than in the previous scenario, as can be confirmed in Fig. 6.

The speed distribution for both corridor comparisons is presented in Fig. 7. For both cases the average speed observed in the segregated corridor is higher than in the mixed traffic corridor. This is represented by the difference on the modes of the distributions.

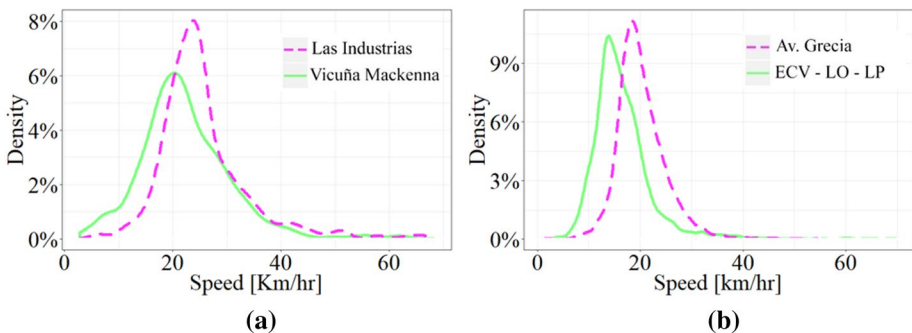


Fig. 7 Bus services' speed distribution. **a** Las Industrias & Vicuña Mackenna; **b** Av. Grecia & ECV – LO – LP

However, the difference between speed variability is not clear since the spread of both distributions seems (at least at first sight) rather similar.

As the analysis presented in Sect. 1.1, travel data was grouped in different distance ranges within each corridor. Within each distance range, average travel times and variability measures were computed. The positive impact of segregated corridors on the average speed can be seen in Fig. 8, which confirms that trips on segregated corridors are, on average, faster (as expected) (Durán-Hormazábal and Tirachini 2016). Figure 8 also shows that the reduction on travel times increases as the trips get longer.

However, as it has been argued in this study, characterizing the impact of segregated corridors should include in-vehicle travel time variability. For that purpose, three indicators of travel time variability are computed for every travel distance range: the standard deviation, the coefficient of variation and the reliability buffer time. This analysis is presented in Fig. 9. Overall, segregated corridors present a better performance in terms of in-vehicle travel time variability. The only exception would be the case of the coefficient of variation in *Av. Grecia* which is equal to the one of the comparable ways for trips longer than 1.5 kms. These figures indicate that there is evidence of a positive impact of dedicated infrastructure, as segregated corridors, not only on the average travel times but for their variability as well.

This extensive graphical analysis outperforms any statistical summary as it is possible to clearly detect the differences between these two types of public transportation services. Besides, it is possible to also visualize the rate of change for every measure studied, which gives a new dimension of analysis.

The effect of travel time reliability on mode choice

In this section, the methodology for the aggregate demand model is presented and the results obtained are discussed. This model allows understanding the effect that travel time reliability has on travellers' preferences and on the observed travel structure. To estimate the aggregate public transport mode choice model, it is necessary to build a database for that purpose. The model only considers origin–destination pairs where metro is an alternative to the buses, to analyse individuals' choices between both modes and its combination.

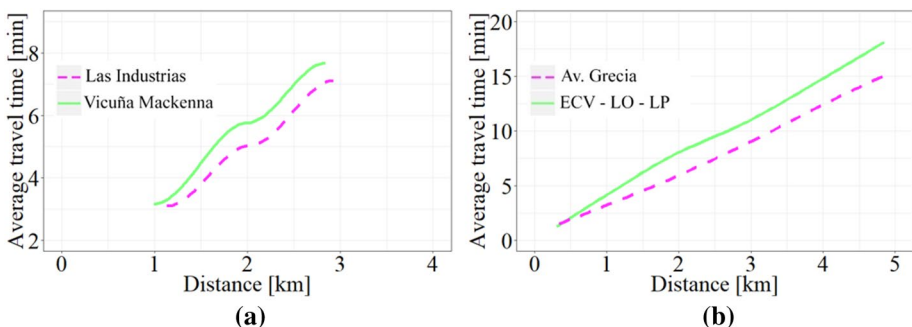


Fig. 8 Average in-vehicle travel time by travel distance. **a** Las Industrias & Vicuña Mackenna; **b** Av. Grecia & ECV - LO - LP

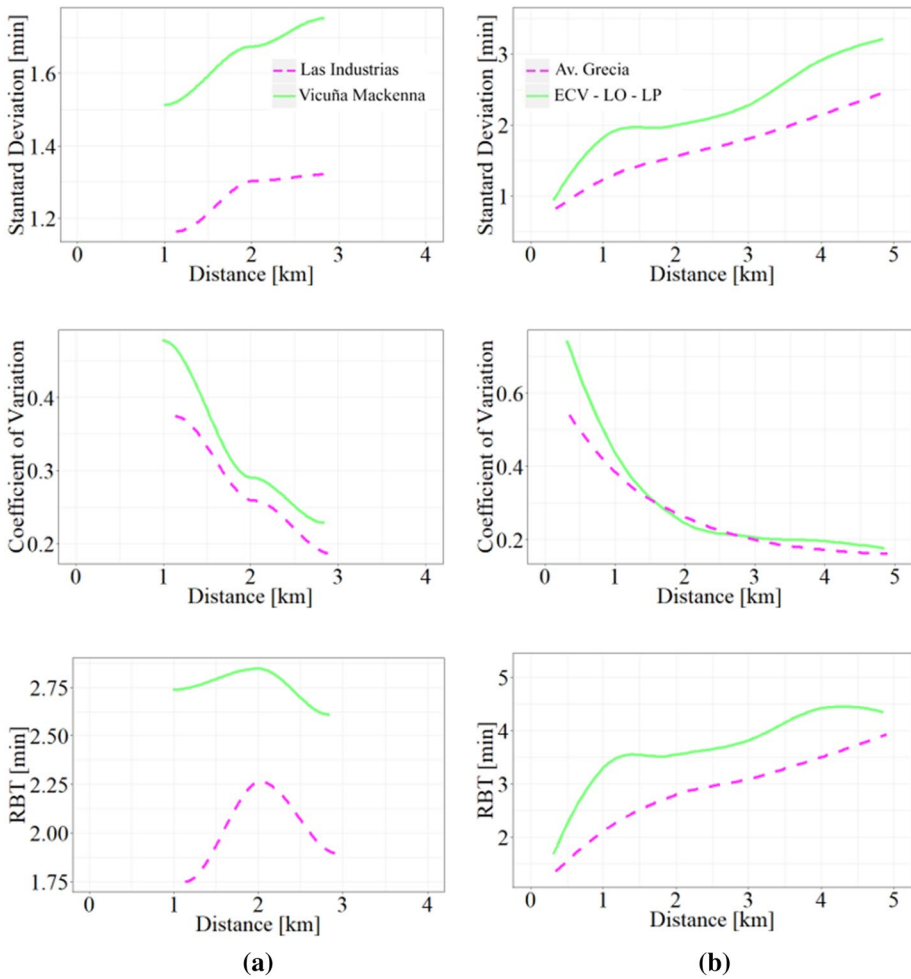


Fig. 9 Variability measures for in-vehicle travel time by travel distance. **a** Las Industrias & Vicuña Mackenna; **b** Av. Grecia & ECV – LO – LP

Origin–destination pairs

To identify the bus services that are an alternative to metro, buffers (or influence zones) of 750 m radius were defined around each metro station. All bus stops within the buffer define the origin–destination pairs for bus trips or combined bus-metro trips that could have been done only by metro. For those bus stops with more than one metro station within the range, the closest station was assigned. This way the demand for buses and combined trips can then be aggregated for each origin–destination pair. The 750 m radius was selected based on the results of Tamblay et al. (2015), which identifies that 95% of Metro users walk less than this distance to reach their metro stations. The result of this procedure is shown in Fig. 10, where the left panel shows a general view of the city with while the panel on the right shows, in a more detailed way, the circular buffers created surrounding the metro stations. With this information it is possible to create an aggregate database of public

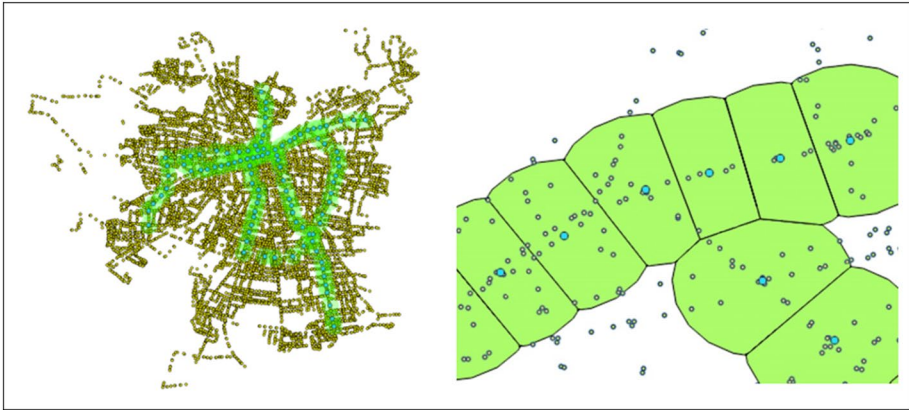


Fig. 10 Buffers of 750 m around metro stations

transport trips in bus, metro and bus-metro between selected origin–destination pairs, in order to study travellers’ mode choices.

Database creation

Once the metro stations are matched to their corresponding surrounding bus stops, it is possible to group all trips in the database based on the origin and destination metro station to understand how travel attributes impact the total demand for each mode. The analysis must be conducted at the level of full travel and not at the trip-leg level (as was the case of the analysis presented in “[Introduction](#)” section), since only this way it is possible to properly understand and explain travellers’ behaviour.

The three considered alternatives are bus, metro, and combined bus-metro. For bus services, the average travel time and its variance were obtained directly from the database used in “[Introduction](#)” section, grouping trip-legs into trips. For metro, the information was obtained based on the arrival and departure times provided by Metro de Santiago. As for some origin–destination pairs there is more than one reasonable route, the level of service was computed as a weighted average based on route probabilities calibrated in a previous study for the same network (Raveau et al. 2011). Finally, for the trips made by a combination of metro and bus services, the level of service was obtained as a combination of both procedures. To get a total variance for each trip, it was assumed that every trip-leg was independent and their travel time variances to be additive.

The behavioural analysis presented in this study focuses only on the morning peak period. As more than one observation is needed to obtain reliability indicators, origin–destination pairs with five or less trips for any mode were not considered in the analysis. Since the estimated bus arrival time database is not exhaustive, we also filtered those OD pairs without headway information for this mode. This allowed us to guarantee that every origin–destination pair considered in our analysis contains enough information to measure variability for both travel and waiting times. The number of origin–destination pairs for every criterion is displayed in Table 1.

The number of origin–destination pairs is reduced significantly when the criteria are applied. However, the remaining origin–destination pairs present, as expected, higher

Table 1 Number of origin–destination pairs selected

Criterion	Number of origin–destination pairs	Morning peak travel demand	Percentage of the total (OD pairs/demand)
At least 1 observation in metro	9082	1,330,896	100%/100%
At least 1 observation in bus or bus-metro	7328	1,289,621	80.69%/96.90%
At least 6 observations in every alternative	2315	669,232	25.49%/50.28%
Presence of headway observations for buses	2264	662,063	24.93%/49.75%

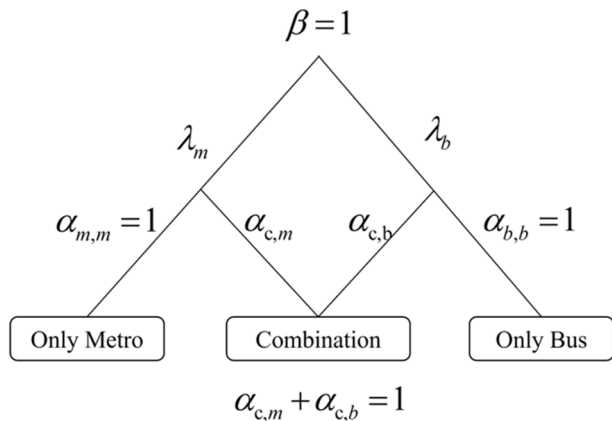
demand than the deleted pairs, which represent almost one half of the total demand. Finally, bus observations were corrected by fare evasion (Cantillo et al. 2018), which is on average 26% for the selected bus stops during the period of analysis. The final number of observations considered is 695,113.

Results

Based on a Random Utility Maximization approach (McFadden 1974; Ortúzar and Wilumsen 2011) for modelling aggregate mode choice, different specifications were tested to obtain a good fit for the data set. As for travel and waiting time variability, the standard deviation, the variance, the coefficient of variation and the reliability buffer time were tested.

As the bus-metro alternative is a combination of the other alternatives, correlation is expected. To address this issue a Cross Nested Logit model (Vovsha 1997) was calibrated. Two nests are defined, one for the *metro* alternatives and the other for the *bus* alternatives. On one side, the pure modal alternatives belong entirely to their respective nest, with each inclusion coefficient (denoted by α) equal to 1. On the other side, the combined metro-bus alternative belongs to both nests, with inclusion coefficients to be estimated. For identifiability purposes the scale parameter at the root was set equal to 1, as well as one of the nest coefficients (denoted by λ). The scale parameter estimated was the associated with the

Fig. 11 Cross Nested Logit model structure



metro alternative. The model structure, nest coefficients and inclusion coefficients can be seen in Fig. 11.

The best specification was found by the Likelihood-ratio test. Regarding travel attributes’ variability, only the variability of headways had a significant effect. From all the different variability measures, the coefficient of variation performed the best. Non-linearity was tested for both travel and waiting time. However, no significant improvement was found in this sense. The selected utility specification V_i considered is:

$$\begin{aligned}
 V_i = & ASC_i + \theta_{TT} \cdot (1 + \delta_{CV(h)} \cdot CV(h)_i) \cdot TT_i + \theta_{WT} \cdot E(W)_i + \theta_{CV(h)} \cdot CV(h)_i \\
 & + \theta_{TrT} \cdot TrT_i + \theta_{\#Tr(M \rightarrow M)} \cdot \#Tr(M \rightarrow M)_i + \theta_{\#Tr(B \rightarrow B)} \cdot \#Tr(B \rightarrow B)_i \\
 & + \theta_{\#Tr\left(\begin{smallmatrix} M \rightarrow B \\ B \rightarrow M \end{smallmatrix}\right)} \cdot \#Tr\left(\begin{smallmatrix} M \rightarrow B \\ B \rightarrow M \end{smallmatrix}\right)_i
 \end{aligned}$$

where TT_i Average travel time for alternative i , $E(W)_i$ Expected waiting time for alternative i , $CV(h)_i$ Coefficient of variation of headways for alternative i , TrT_i Average transfer time for alternative i , $\#Tr(M \rightarrow M)_i$ Average number of metro-to-metro transfers for alternative i , $\#Tr(M \rightarrow M)_i$ Average number of bus-to-bus transfers for alternative i , $\#Tr\left(\begin{smallmatrix} M \rightarrow B \\ B \rightarrow M \end{smallmatrix}\right)_i$ Average number of metro-to-bus or bus-to-metro transfers for alternative i .

Before presenting the results analysis, there are three important modelling considerations to highlight. Firstly, the combined alternative considers both specific and common parameters. For example, the average in-vehicle travel time is split by mode and share the parameters for both metro and bus. The same happens with the number of transfers, sharing bus-to-bus and metro-to-metro parameters. The coefficient of variation and the average waiting time is the same parameter than for buses.

Secondly, there is a strong correlation between ASC_{Comb} and $\#Tr\left(\begin{smallmatrix} M \rightarrow B \\ B \rightarrow M \end{smallmatrix}\right)_i$ as the combined alternative is the only one with transfers between metro and bus (and vice-versa) and the average value is close to 1. Therefore, both parameters associated cannot be estimated at the same time. As our major objective is to understand travellers’ behaviour, we decided to define the combined alternative specific constant conveniently, in order to estimate the parameter associated with the number of transfers. The specific constant is defined as follows, proportionally to the amount of in vehicle time spent in bus for the combined alternative:

$$ASC_{Comb} = ASC_{Bus} \cdot \frac{TT_{bus_{Comb}}}{TT_{bus_{Comb}} + TT_{metro_{Comb}}}$$

Finally, the database consists in 5429 observations which represent 695,113 individual revealed choices. As these choices are aggregated by origin–destination pairs, every observation has the total amount of trips for each origin–destination pair for each mode assigned. If this attribute is neglected, the model would consider that each observed modal split is equal. However, the common practice is to multiply this value so that the total sum of them is equal to the total number of observations. This is done to correctly estimate the standard deviation of the estimated parameters.

The results of the model are shown in Table 2. The parameters part of the different alternative utility functions are denoted with B, M, and C for bus, metro and combined respectively. This parameters were obtained using PythonBiogeme (Bierlaire 2016a, b). All parameters associated with travel attributes have the expected sign and are statistically

Table 2 Calibrated parameters values

Attribute	Alternative	Parameter	Estimate	<i>p</i> -value
Average travel time	B, M, C	θ_{TT}	−0.087	0.00
Marginal effect of coefficient of variation of headways in the perception of Travel time	B, M, C	$\delta_{CV(h)}$	−0.269	0.00
Average waiting time	B, C	θ_{WT}	−0.086	0.00
	M	θ_{WT}	−0.203	0.00
Coefficient of variation headways	B, C	$\theta_{CV(h)}$	−0.424	0.09
Average transfer time	B, M	θ_{TT}	−0.078	0.00
	C	θ_{TT}	−0.026	0.00
Number of transfers	B, C	$\theta_{\#Tr(B \rightarrow B)}$	−1.050	0.00
	M, C	$\theta_{\#Tr(M \rightarrow M)}$	−0.455	0.00
	C	$\theta_{\#Tr(M \rightarrow B)}$	−1.140	0.00
Alternative specific constant	B, C	ASC_{Bus}	−0.159	0.45
	M, C	ASC_{Metro}	0	Fixed
Combined inclusion coefficient – metro		$\alpha_{c,m}$	0.854	0.00
Combined inclusion coefficient – bus		$\alpha_{c,b}$	0.146	0.00
Nest coefficient – metro		λ_m	1.780	0.00
Nest coefficient – bus		λ_b	1.000	Fixed
Number of observations			5429	
Log-likelihood			−2,979.41	

significant at a 91% confidence level. Also, the combined inclusion coefficients ($\alpha_{c,m}$ and $\alpha_{c,b}$) and the scale parameter for metro (λ_m) were found to be significantly different from 1, which confirms the Cross-Nested Logit structure. Besides, the higher value of $\alpha_{c,m}$ means the combined alternative relates more with metro. This is mostly explained as in bus-metro trips, bus is mostly used as an access/egress mode, meaning that the bigger portion of the trip is done in metro.

In order to understand better the impact of each attribute, we calculated the marginal rate of substitution of them with in-vehicle travel time. Typically, this analysis is regarding the cost attribute. However, the model that considers the fare fails to pass the Likelihood-ratio test when compared with the model presented in the manuscript. This might happen as this attribute is highly correlated with the alternative specific constants, as bus fare is \$640 CLP and both metro and combined alternative’s fare is \$740 CLP. As in-vehicle travel time perception depends on headway regularity, we define two scenarios to analyse: perfect headway regularity, where $CV(h)$ equals 0 and an irregular scenario, such as arrivals follow a Poisson process, where $CV(h)$ equals 1.

In terms of average waiting time, in a regular scenario we see it is considered higher than in-vehicle travel time only for the cases of metro trips. For this case, the MRS between this attribute and in-vehicle travel time is 2.33. In an irregular scenario, the MRS rises to 1.35 for buses. If we consider the average values of $CV(h)$ for both modes, MRS equals 1.26 for buses and 2.39 for metro. This might be explained as during morning peak hours, waiting for metro might be considered worse as in some stations the amount of people waiting reach an uncomfortable level.

In relation to transfers, we observe that transfer times are lower than in-vehicle travel time under regular headways. Considering the average coefficient of variation on headways,

the MRS is 1.15 for buses, 0.92 for metro and 0.38 for the combined alternative. One possible explanation considers that transfer time for the combined alternative are related with access or egress to the metro network, which might be perceived better than travelling itself.

On the other side, the worst perceived transfer is Bus-to-Metro/Metro-to-Bus, followed by Bus-to-Bus, and lastly Metro-to-Metro. This is also in line with our current knowledge, where Metro transfers are not perceived as bad as bus transfers. Based on the MRS, people would travel on average around 17, 15, and 5 min extra to avoid a Bus-to-Metro/Metro-to-Bus, Bus-to-Bus, and Metro-to-Metro transfer respectively.

Regarding the coefficient of variation of headways, there are four different aspects to recall. Firstly, this attribute has three different effects in the utility function. It increases the expected waiting time (Osuna and Newell 1972a, b), it has a direct impact in the utility function, and it has a marginal impact in the perception of travel time. This allow the model to differentiate between two services with different frequency and headway regularity but with the same expected waiting time.

$$E(W)_i = \frac{E(h)_i}{2} \cdot (1 + CV(h)_i^2)$$

Secondly, the coefficient of variation of headways by itself only has a significant impact for bus and combined alternatives. The non-significative impact for metro can be explained as the combination of its regularity and frequency level might not be high enough to have a perceivable effect in passengers’ experience.

Thirdly, the marginal effect of the coefficient of variation of headways in the perception of in-vehicle travel time, $\delta_{CV(h)}$, was found to be significant and negative. This means than passengers are willing to travel longer in order to travel in a reliable bus service. In an irregular scenario, the perception of in-vehicle travel time is ~27% lower than in a perfectly regular scenario.

Fourthly, the MRS between $CV(h)$ and TT_{Bus} is calculated as follows:

$$\begin{aligned} &MRS(CV(h), TT_{Bus}) \\ &= \frac{\theta_{TT} \cdot \delta_{CV(h)} \cdot TT \cdot \partial CV(h) / \partial CV(h) + \theta_{WT} \cdot \partial E(W) / \partial CV(h) + \theta_{CV(h)} \cdot \partial CV(h) / \partial CV(h)}{\theta_{TT} \cdot (1 + \delta_{CV(h)} \cdot CV(h)) \cdot \partial TT_{Bus} / \partial TT_{Bus}} \\ &= \frac{\theta_{TT} \cdot \delta_{CV(h)} \cdot TT + \theta_{WT} \cdot E(h) \cdot CV(h) + \theta_{CV(h)}}{\theta_{TT} \cdot (1 + \delta_{CV(h)} \cdot CV(h))} \\ &= \frac{-0.27 \cdot TT + 0.98 \cdot E(h) \cdot CV(h) + 4.89}{1 - 0.27 \cdot CV(h)} [\text{min}] \end{aligned}$$

This means the rate of substitution grows with both the coefficient of variation of headways and the expected value (the inverse of the frequency) and decreases with travel time. If we consider the average value of $E(h)$, $CV(h)$, and TT , 6.85 min, 0.81, and 21.53 min respectively, we obtain a MRS equals to 5.84 min, which is comprised by -7.41 min in terms of travel time, 7.02 min in terms of excess waiting time and 6.23 min in terms of regularity. This means passengers, on average, are willing to increase their in-vehicle travel time in 5.08 min to have a service with perfectly regular headways. By multiplying this amount by the Chilean social value of time (MDS 2016) we obtain a monetary value of ~\$156 CLP, which is equivalent to a 24.38% of the fare at that time. The in-vehicle dependency of this marginal rate of substitution means that, under unreliable waiting time, passengers are willing to travel longer for shorter trips. If

in-vehicle travel time is 38.52 min and both $E(h)$ and $CV(h)$ remain constant, the MRS equals 0, which means passengers are not willing to travel any longer to reduce their waiting time uncertainty.

Finally, every time the coefficient of variation of headways was included in the specification function, the alternative specific constant of buses decreased its significance. As this attribute was not found to be significant for metro trips, we consider that headway variability as a key attribute in order to explain modal choices. All these results confirm the idea that public transport reliability has a significant impact on passenger's mode decisions.

Conclusions

This study provides evidence of significant differences among headway regularity and travel time dispersion (measured as the standard deviation or the difference between the 95th percentile and the average travel time) for trips of similar length on different public transport modes. The study also shows that these dispersions also increase with travel length for every mode. However, the dispersion is always smaller than 4 min for metro (when the standard deviation is considered), which could hardly be perceived by travellers, particularly on long trips.

The results were displayed allowing a clear visualization of the reliability differences between different modal alternatives. The graphs provide an intuition about why certain services could be used to a lesser extent than what is predicted by conventional models (which ignore the uncertainty in the level of service). The figures presented in this paper considered every service within certain distance range, but the methodology is of course applicable for any subset of services satisfying specific conditions (as with the case of particular corridors presented in this study). This visualization allows identifying opportunities for improvement in the system by recognising similarities in the level of service between some bus-based services and metro. For example, clustering those services whose characteristics mimic in some sense their operation with metro (such as the express services in Transantiago or those operating over a segregated corridor).

The aggregate demand analysis proved the significant impact of public transport reliability (measured as the coefficient of variation of headways) in travellers' choice between buses and metro for origin–destination pairs where both modes are available. One would have expected a significant impact in aggregate choices of travel time variability when studying travel reliability. In this study, that was not the case. We suspect this is due to the big existing differences between public transport systems operating on a frequency-based and schedule-based manner and because of the high-demand-high-frequency nature of the public transport systems in Latin America. In this context, headway regularity plays an even more crucial role. However, unreliability is not limited to travel or waiting times, also affecting average crowding and its variability. The effect of variability on these attributes should be also analysed and included in demand models, further increasing the impact of unreliability on passenger's behaviour.

It is important to emphasize that all the analysis in this study was conducted by only using passive-data, without the need of any kind of survey or external information. The data used comprises smartcard validation, buses' GPS position and trains' time schedules. Although the demand model is quite general (as no individual information, such as gender or income, is recorded in the smartcards) to the best of the authors knowledge, revealed

preferences have not been used to analyse the impact of reliability on the preferences of public transport travellers. In a world where passive-data collection technologies rapidly gain importance over former techniques, studies similar to the one presented here will help to better understand passive-data capabilities and limitations.

The aggregate demand model suggests that, in a more detailed disaggregated model (at an individual level), variability should also have a significant impact in the travellers' decisions. Such disaggregated model would require a travel survey to gather socio-demographic information, and more detailed travel information to compute reliability indicators for each individual based on their past travel experiences. Such disaggregated revealed preference model could provide further insights regarding the effect of reliability in travel demand and have higher repercussions in public policy.

As public transport time reliability has a relevant impact on travellers' decisions, it is necessary to improve it, enhancing the level of service. This paper shows that this is particularly important for bus services, which lag behind metro in this dimension. An effective way to improve bus reliability is with segregated corridors. This study shows that segregated corridors not only reduce average travel times, but also reduce travel time variability. The methodology presented in this study could be used to assess the impact that other policies and strategies (such as public transport signal priority or bus holdings at stops) have on reducing travel time variability.

Transport planners and modellers should consider these results to improve project evaluation and decision-making processes by better understanding the effects of travel time reliability on public transport travellers. Extending the behavioural models to include additional level-of-service components (such as waiting times and crowding levels) would be an interesting research subject. Further understanding the causes and effects of public transport variability has a significant impact on public policy.

Acknowledgements This research was supported by the Centro de Desarrollo Urbano Sustentable, CEDEUS (Conicyt/Fondap 15110020), the Bus Rapid Transit Centre of Excellence funded by the Volvo Research and Educational Foundations (VREF), the FONDECYT project number 11170127: Behavioural Modelling of Public Transport Systems, and the scholarship funded by CONICYT for Ph.D. studies (CONICYT-PCHA/Doctorado Nacional/2016).

Authors' contribution All authors contributed to the study conception and design. Material preparation and data collection were performed by Jaime Soza-Parra. Data analysis was performed by Jaime Soza-Parra, Sebastián Raveau, and Juan Carlos Muñoz. The first draft of the manuscript was written by Jaime Soza-Parra and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declaration

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Bates, J., Polak, J., Jones, P., Cook, A.: The valuation of reliability for personal travel. *Transp. Res. Part E Logist. Transp. Rev.* **37**(2–3), 191–229 (2001)
- Benezech, V., Coulombel, N.: The value of service reliability. *Transp. Res. Part B Methodol.* **58**, 1–15 (2013)

- Bierlaire, M.: PythonBiogeme: a short introduction, Technical report TRANSP-OR 160706 Transport and Mobility Laboratory ENAC, EPFL (2016)
- Bierlaire, M. (2016) PythonBiogeme: a short introduction. Report TRANSP-OR 160706, Series on Biogeme. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland
- Birr, K., Jamroz, K., Kustra, W.: Travel time of public transport vehicles estimation. *Transp. Res. Proc.* **3**, 359–365 (2014)
- BRT: Asesoría Experta para la Ejecución de un Estudio Comparativo Indicadores de Ciudades Latinoamericanas (2012)
- BRT+ Centre of Excellence and EMBARQ (2018) “Global BRTData.” Version 3.35. Last update: May 9, 2018. Available in: <http://www.brtdata.org>
- Bucknell, C., Schmidt, A., Cruz, D., Muñoz, J.C.: Identifying and visualizing congestion bottlenecks with automated vehicle location systems: application to Transantiago, Chile. *Transp. Res. Rec.* **2649**, 61–70 (2017)
- Cantillo, L.-A., Raveau, S., Iglesias, P., Tamblay, S., & Muñoz, J.C.: A methodology for correcting smartcard trip matrices by fare evasion. In: Conference on Advanced Systems in Public Transport and TransitData (2018)
- Cats, O., Gkioulou, Z.: Modeling the impacts of public transport reliability and travel information on passengers’ waiting-time uncertainty. *EURO J. Transp. Logist.* **6**(3), 247–270 (2017)
- Cham, L.C.: Understanding Bus Service Reliability: A Practical Framework Using AVL/APC Data. Master of Science in Transportation Thesis, MIT (2006)
- Danés, C., Muñoz, J.C., Guevara, A.: Public transport’s reliability: the case of Santiago, Chile. Presented at Conference of Advanced Systems for Public Transport. Rotterdam, 19–23 July, 2015
- de Jong, G.C., Bliemer, M.C.J.: On including travel time reliability of road traffic in appraisal. *Trans. Res. Part A Policy Pract.* **73**, 80–95 (2015)
- de Ortúzar, J., D Willumsen L.G. , L.G.: *Modelling Transport*, 4th edn. John Wiley and Sons, Chichester (2011)
- Durán-Hormazábal, E., Tirachini, A.: Estimation of travel time variability for cars, buses, metro and door-to-door public transport trips in Santiago, Chile. *Res. Transp. Econ.* **59**, 26–39 (2016)
- Engelson, L., Fosgerau, M.: The cost of travel time variability: three measures with properties. *Transp. Res. Part B Methodol.* **91**, 555–564 (2016)
- Fadaei, M., Cats, O.: Evaluating the impacts and benefits of public transport design and operational measures. *Transp. Policy* **48**, 105–116 (2016)
- Frappier, A., Morency, C., Trépanier, M.: Measuring the quality and diversity of transit alternatives. *Transp. Policy* **61**(October 2017), 51–59 (2018). <https://doi.org/10.1016/j.tranpol.2017.10.007>
- Furth, P.G., Muller, T.H.J.: Service reliability and hidden waiting time: insights from automatic vehicle location data. *Transp. Res. Rec.* **1955**, 79–87 (2006)
- García Palomares, J.C.: Incidencia en la movilidad de los principales factores de un modelo metropolitano cambiante. *Eure* **34**(101), 5–23 (2008)
- Gschwender, A., Munizaga, M., Simonetti, C.: Using smart card and GPS data for policy and planning: the case of Transantiago. *Res. Transp. Econ.* **59**, 1–8 (2016)
- Hollander, Y.: Direct versus indirect models for the effects of unreliability. *Transp. Res. Part A Policy Pract.* **40**(9), 699–711 (2006)
- Jenelius, E.: The value of travel time variability with trip chains, flexible scheduling and correlated travel times. *Transp. Res. Part B Methodol.* **46**(6), 762–780 (2012)
- Jirón, P.: Sustainable Urban Mobility in Latin America and the Caribbean. Global Report on Human Settlements. United Nations (2013). Retrieved from <http://www.unhabitat.org/grhs/2013>.
- McFadden, D.: Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (ed.) *Frontiers of Econometrics*, pp. 105–142. Academic Press, New York (1974)
- MDS: Precios Sociales Vigentes. Ministerio de Desarrollo Social. División de Evaluación Social de Inversiones. Gobierno de Chile, Santiago, 1–20 (2016)
- Munizaga, M., Palma, C.: Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. *Transp. Res. Part C Emerg. Technol.* **24**, 9–18 (2012)
- Munizaga, M., Palma, C.: Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. *Transp. Res. Part C: Emerg. Technol.* **24**(October 2012), 9–18 (2012). <https://doi.org/10.1016/j.trc.2012.01.007>
- Muñoz, J.C., Gschwender, A.: Transantiago: a tale of two cities. *Res. Transp. Econ.* **22**, 45–53 (2008)
- Muñoz, J.C., Batarce, M., Hidalgo, D.: Transantiago, five years after its launch. *Res. Transp. Econ.* **48**, 184–193 (2014)

- Osuna, E.E., Newell, G.F.: Control strategies for an idealized public transportation system. *Transp. Sci.* **6**(1), 52–72 (1972)
- Osuna, E.E., Newell, G.F.: Control strategies for an idealized public transport system. *Transp. Sci.* (1972) (August 2015)
- Petersen, E., Vovsha, P.: Directions for coordinated improvement of travel surveys and models. Proceedings of the 42nd Conference on Innovations in Travel Demand Modeling. Transportation Research Board of the National Academies, 85–88 (2006)
- Raveau, S.: Modelling travel choices on public transport systems with smart card data. In: Kurauchi, F., Schmöcker, J.D. (eds.) *Public Transport Planning with Smart Card Data*. CRC Press, Boca Raton, United States (2017)
- Raveau, S., Muñoz, J.C., de Grange, L.: A topological route choice model for metro. *Transp. Res. Part A Policy Pract.* **45**, 138–147 (2011)
- Raveau, S., Guo, Z., Muñoz, J.C., Wilson, N.H.M.: A behavioural comparison of route choice on metro networks: time, transfers, crowding, topology and socio-demographics. *Transp. Res. Part A Policy Pract.* **66**(1), 185–195 (2014)
- Rodríguez Vignoli, J.: ¿Policentrismo o ampliación de la centralidad histórica en el área metropolitana del gran santiago? evidencia novedosa proveniente de la encuesta casen 2009. *Eure* **38**(114), 71–97 (2012)
- Soza-Parra, J., Raveau, S., Muñoz, J.C., Cats, O.: The underlying effect of public transport reliability on users' satisfaction. *Transp. Res. Part A Policy Pract.* **126**(January), 83–93 (2019). <https://doi.org/10.1016/j.tra.2019.06.004>
- Tamblay, S., Galilea, P., Iglesias, P., Raveau, S., Muñoz, J.C.: A zonal inference model based on observed smart-card transactions for Santiago de Chile. *Transp. Res. Part A* (2015). <https://doi.org/10.1016/j.tra.2015.10.007>
- van Oort, N.: *Service Reliability and Urban Public Transport Design*. Ph.D. thesis series T2011/2, Delft University of Technology (2011)
- van Oort, N., Brands, T., De Romph, E., Aceves Flores, J.: Unreliability effects in public transport modeling. *Int. J. Transp.* **3**(1), 113–130 (2015)
- Vignoli, R.J.: Movilidad cotidiana, desigualdad social y segregación. *Eure* **34**(103), 49–72 (2008)
- Vovsha, P.: The cross-nested logit model: application to mode choice in the Tel Aviv metropolitan area. *Transp. Res. Rec.* **1607**, 13–20 (1997)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.