Taking the path less travelled
Valuation of denied boarding in crowded public transport systems

Yap, Menno; Cats, Oded

DOI
10.1016/j.tra.2021.02.007

Publication date
2021

Document Version
Final published version

Published in
Transportation Research Part A: Policy and Practice

Transportation Research Part A 147 (2021) 1–13

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Taking the path less travelled: Valuation of denied boarding in crowded public transport systems

Menno Yap *, Oded Cats

Delft University of Technology, department of Transport and Planning, Delft, the Netherlands

A R T I C L E I N F O

Keywords:
Crowding
Denied boarding
Metro systems
Revealed preference
Smart card data

A B S T R A C T

Many public transport networks worldwide experience high crowding levels. Overcrowding can result in passengers not able to board the first arriving vehicle. We infer how waiting time induced by being denied boarding in crowded public transport systems is valued by passengers, based on observed passenger route choice behaviour. For this purpose, we estimate a revealed preference route choice model based on passenger and vehicle movement data. As denied boarding typically occurs only at specific locations and within strict time bands, whilst its occurrence is notoriously uncertain, we propose additional constraints to generate an appropriate choice set for which observed route choices can be used to estimate denied boarding perceptions. We found that the additional waiting time caused by denied boarding is valued 68% more negatively compared to the initial waiting time. On average, one minute of initial and denied boarding wait time are perceived as 1.62 and 2.72 min on-board an uncrowded vehicle, respectively. Not incorporating this more negative denied boarding wait time valuation can result in an underestimation of the passenger and societal impact of overcrowding in public transport systems. Moreover, it can underestimate the benefits of potential crowding relief measures and as such underestimate the benefit-cost ratio of these measures.

1. Introduction

Many public transport (PT) systems worldwide frequently experience high crowding levels, at least when considering the PT ridership levels prior to the COVID-19 pandemic. Metro systems in large, dense metropolitan areas - for example in London, Santiago, Tokyo or Singapore - are particularly prone to overcrowding, especially during peak hours. Overcrowding can have a negative impact on the perceived attractiveness of PT, relative to other modes of transportation. In addition, it can reduce PT accessibility of a metropolitan area, if overcrowding results in prolonged travel times. Understanding the impact of (over)crowding in PT networks on passengers' nominal and experienced journey times is key for public transport planning. A more accurate quantification of the nominal and perceived journey time impacts of crowding improves the calculation of expected benefits when evaluating crowding relief schemes (such as service frequency increases, station redesign or adding new PT links) in business cases or cost-benefit analyses (see for example Jenelius and Cats 2015). Furthermore, a better insight into passengers' perceptions and preferences related to crowding has the potential to improve the crowding parameters of PT network assignment models, thereby resulting in better model validation and improved route choice and passenger flow forecasts (see for example Hamdouch et al. 2011, Nuzzolo et al. 2012, Pel et al. 2014, Hänseler et al. 2020).

* Corresponding author.
E-mail address: M.D.Yap@TUDelft.nl (M. Yap).

https://doi.org/10.1016/j.tra.2021.02.007
Received 6 October 2020; Received in revised form 28 December 2020; Accepted 16 February 2021
Available online 11 March 2021
0965-8564/

Public transport crowding has three distinct passenger impacts: 1) on-board discomfort, 2) service irregularity and unreliability, and 3) denied boarding (Cats et al. 2016). The first and foremost studied effect relates to a more negative valuation of in-vehicle time on board crowded services. This effect does not influence the nominal journey times, but does affect the perceived journey times as perceived on-board discomfort increases with increasing crowding levels. Over the last couple of decades, many studies worldwide inferred how passengers' in-vehicle time is perceived as a function of the on-board load factor or standing density, often using a Stated Preference (SP) approach. SP based studies estimating the perceived in-vehicle time as a function of the average number of standing passengers per square metre were for example performed for UK train services (MVA Consultancy 2008), bus and metros in Santiago, Chile (Batarce et al. 2016, Tirachini et al. 2017) and for RER services in Ile-de-France (Kroes et al. 2014). In the latter study, SP estimation results were validated based on passenger observations, cameras and surveys. For an extensive review and *meta*-analysis of studies concerning on-board crowding valuation performed until the first decade of the 21st century, we refer to Wardman and Whelan (2011) and Li and Hensher (2011). In addition to SP based studies, there is an increasing number of studies estimating perceived in-vehicle times using Revealed Preference (RP) data. Raveau et al. (2014) assessed the crowding valuation in metros for London and Santiago, using realised route choice behaviour based on surveyed trips for selected origin–destination (OD) pairs. This is however a time consuming task to perform frequently for the entire PT network of consideration. In Batarce et al. (2015), a mixed SP/RP based model is estimated to obtain crowding valuations in Santiago, Chile, in which a SP survey is enriched with observed route choice behaviour obtained from smart card data. The availability of large-scale passenger movement data obtained from Automated Fare Collection (AFC) systems has more recently allowed for the estimation of crowding valuations entirely based on revealed preference route choices. RP studies provide an alternative to SP studies, which are intrinsically susceptible to estimation biases caused by a potential discrepancy between stated and realised behaviour. Hörcher et al. (2017) and Tirachini et al. (2016) adopted a RP based approach to infer the in-vehicle time crowding multipliers related to standing probabilities and standing density for metros in Hong Kong and Singapore, respectively. Furthermore, Yap et al. (2020) estimated on-board crowding multipliers for urban trams and buses in the Netherlands based on observed route choices. The second impact of crowding is the slowing down of the boarding and alighting process, thus resulting in longer dwell times at stops. Consequentially, this increases service irregularity and unreliability, as this reduces the headway with the trailing service. This can potentially result in the well-known bunching phenomenon, reducing the effective capacity of a line and further amplifying average crowding levels. For example, Kim et al. (2015) estimated several models to determine how passengers perceive this dwell time delay caused by crowding. The third impact of severe crowding entails the possibility of passengers being left behind (denied boarding), when PT vehicles have reached their crush capacity. For passengers, this implies an extension of their platform wait time, until they are able to board a consecutive PT service. To understand this impact, two aspects are relevant: 1) deriving the amount of additional waiting time caused by denied boarding, and 2) inferring how passengers perceive this waiting time. In relation to the former aspect, Zhu et al. (2017b) propose maximum likelihood estimation and Bayesian inference methods to infer the probability distribution of the number of left behind passengers based on AFC and Automated Vehicle Location (AVL) data. Miller et al. (2018) develop a bi-level regression model for this same purpose. Regarding the second aspect, one could arguably expect that left behind passengers perceive one minute waiting time after being denied boarding more negatively than one minute of initial waiting time, since the former contributes to delays and induces uncertainty and possibly frustration (Cats et al. 2016). In an attempt to capture this effect, Cats et al. (2016) use a delay time multiplier of 3.5 derived from Börjesson et al. (2012) as a proxy in their simulation based assignment model, thus assuming one minute waiting after being left behind is perceived 3.5 times as negative as one minute initial waiting time. Furthermore, there are studies which quantify the value passengers attribute to reliability of their journey, typically using the variance or standard deviation of travel time as metric for travel time variability for this (for example Fosgerau and Karlström, 2010, Fosgerau and Engelson 2011, Li et al. 2016, Mishra et al. 2018). Whilst these studies allow for estimating the value passengers attribute to variability around the scheduled or expected travel time, these metrics do not accurately capture the higher marginal costs passengers might perceive only after being left behind. Although many studies estimate how one minute regular (initial) waiting time is perceived compared to on-board time (e.g. Balcombe et al. 2004, Bovy and Hoogendoorn-Lanser 2005), to the best of our knowledge, there is lack of knowledge as to how one minute waiting time after being denied boarding compares to initial waiting time and in-vehicle time for PT systems. Our study aims to fill this research gap.

The contribution of this study can be summarised as the development and application of a method to infer how waiting time after being denied boarding in crowded PT systems is valued by passengers, based on observed passenger route choice behaviour. This allows us to understand how perception coefficients of initial waiting time and denied boarding waiting time compare to each other and to other travel time components. In addition, denied boarding valuation in this study is entirely estimated based on observed route choices made by passengers in reality. Estimating denied boarding valuation from RP route choices is a non-trivial task, as denied boarding typically occurs only at specific locations and within strict time bands, whilst its occurrence is notoriously uncertain. This requires extra constraints for an appropriate choice set generation for which observed route choices can be used to estimate denied boarding perceptions. We propose a generic approach for choice set generation, based on which we estimate a discrete choice model using maximum likelihood estimation to infer denied boarding valuation. The proposed approach is applied to the metro network of Washington, DC. Furthermore, we use London as example to illustrate the policy implications of our study results in appraisal studies.

The remainder of this paper is structured as follows. Section 2 discusses the model formulation and the methodology for choice set generation. In Section 3, the Washington DC case study network is introduced, followed by a discussion of results and policy implications in Section 4. We formulate conclusions and limitations of our work in Section 5.

2. Methodology

In this section, we discuss the required data inputs (Section 2.1), model specification and estimation (Section 2.2), and choice set generation criteria (Section 2.3).

2.1. Data input

Our study requires a) information on individual passenger journeys and b) vehicle occupancy data for individual PT trips and trip segments as input. This information is obtained using automated data processed from AFC and AVL systems. Processed data from AFC systems provides information on the journeys made by individual passengers in the PT network of consideration, including the chosen route in the network along with its relevant attributes such as in-vehicle time, waiting time and number of transfers. By fusing passenger movement data from AFC systems and train movement data from AVL systems, the occupancy and crowding level can be obtained for each individual PT trip and segment, which allows for the calculation of denied boarding probabilities (see Section 2.2).

The data processing and inference steps required depend on the AFC system in place for a particular case study PT network (see Fig. 1). For example, in case of an entry-only AFC system which requires passengers only to tap in with their smart card at the start of a journey (leg) (e.g. bus systems with flat fares), destination inference algorithms are needed to infer the most plausible journey (leg) destination. Most commonly accepted destination inference algorithms are based on the so-called trip chaining method, as proposed by Trépanier et al. (2007), Munizaga and Palma (2012) and Sánchez-Martinez (2017). Conversely, boarding and alighting locations are empirically available for entry-exit systems. In case passengers need to tap out and tap in again when interchanging, a transfer inference algorithm is necessary to determine whether to consider an alighting location as an intermediate transfer or final destination. Transfer inference algorithms as for example proposed by Gordon et al. (2013) and Yap et al. (2017) use a rule-based approach to link individual AFC transactions to complete passenger journeys. When PT vehicles have on-board devices for passengers to tap in (or out), one can directly connect a passenger movement to a specific PT vehicle and its trip characteristics (such as realised running time or delays). When passengers are required to tap in and out at the station or platform, as typically the case for metro systems worldwide, a passenger-to-train assignment algorithm is needed to infer the most plausible route and trip a passenger has taken given the empirical tap in and tap out times at station gates (see e.g. Hörcher et al. 2017, Zhu et al. 2017a).

Given the extensive number of studies focusing on the development of destination inference, transfer inference, and passenger-to-train assignment algorithms, we adopt existing techniques in our work. The raw AFC and AVL data were processed using the ODX algorithm proposed by Sánchez-Martinez (2017) and made available to us. Based on the passenger entry and exit time, walking times to/from the platform and the realised departure times of different trains, it assigns passengers to the most plausible train. In the passenger-to-train assignment algorithm, no *a priori* assumptions are made regarding passenger valuation of any of the journey time components. This algorithm results in inferred routes for individual passenger journeys through a PT network. We then calculate the relevant characteristics of each passenger journey in terms of in-vehicle time, out-of-vehicle time, number of transfers and crowding levels.

For a given dataset, we remove days where large planned maintenance works took place based on an incident log file. Our study

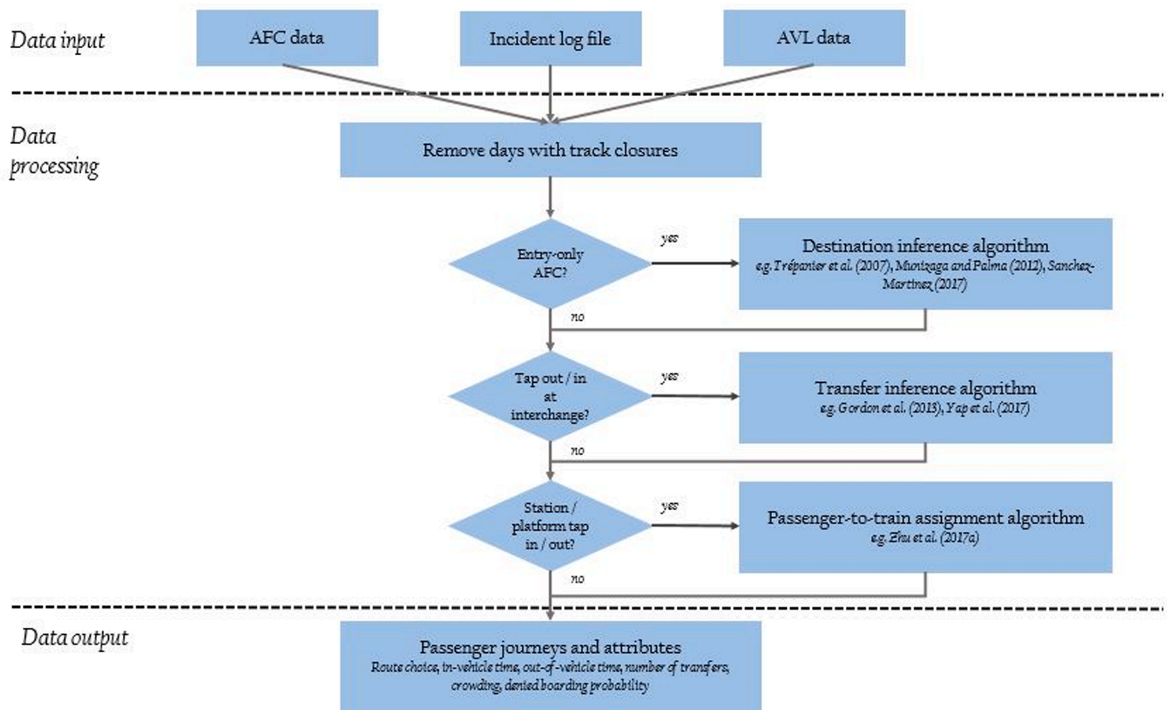

Fig. 1. Data processing.

uses observed route choices to estimate models explaining how passengers perceive denied boarding waiting time. Fundamentally, this requires anticipated crowding levels, based on which passengers make their route choice decision, to be aligned as much as reasonably possible with realised crowding levels inferred from AFC and AVL data. Large construction works can result in crowding levels where or when one might not expect crowding or denied boarding to occur. This can therefore create a discrepancy between (*a priori*) anticipated and (*a posteriori*) experienced denied boarding probabilities. By only including regular days of operation in our dataset, we aim to minimise the potential difference between expected and experienced journey attributes.

2.2. Model formulation

In this section, we discuss the attributes used for explaining passengers' route choice. First, we introduce some notations. We represent a PT network as a directed graph $G(S,E)$, where each node represents a PT stop $s \in S$ and each edge $e \in E$ reflects a directional and direct PT link between two stops. Each platform of a stop is represented by a separate node. This implies that transfer stations where different PT lines have separate platforms are composed of multiple nodes, despite sharing the same public stop name. We define PT trip k of a certain PT line with trip departure times $t_{s,k}^{dep}$. Processed AFC data contains journey data of individual passengers i between origin stop $o \subseteq S$ and destination stop $d \subseteq S$. Different observed paths between a certain OD pair are indicated by $a_{od} \in A_{od}$. Each path is defined as a unique sequence of stops$\{s_o, s_1, s_2 \cdots s_d\}$. All passenger journeys are categorised in time intervals t.

We calculate the expected attribute values between a certain origin–destination pair by averaging the realised attribute values of individual passenger journeys i (of set I) starting within time interval t. Choosing an appropriate time interval t is important, as a too large time interval (e.g. 60 or 120 min) potentially averages out unevenly distributed attribute values (such as crowding levels or denied boarding probabilities), thereby resulting in inaccurate attribute levels. This is especially relevant when studying denied boarding, as this might occur only during specific times. Notwithstanding, t should also not be chosen too small, as passengers will not have a clear expectation of (for example) denied boarding probabilities for each individual PT trip. In our study, we aim to strike a balance between these two ends by grouping journeys into 15-minute intervals. This allows us to capture differences in denied boarding probabilities per 15 min, for which it can be presumably expected that passengers are still able to distinguish denied boarding probabilities.

We incorporate the journey components *in-vehicle time* t^{ivt}, *walking time* t^{wkt}, *initial waiting time* $t^{wtt,i}$, *waiting time after being denied boarding* $t^{wtt,d}$, *number of transfers* n^{tf} and *load factor* lf as explanatory variables in the utility function of our route choice model. Fig. 2 illustrates the different journey time components for an example passenger journey. For an observed path $a_{od} \in A_{od}$, the in-vehicle time and number of transfers can be directly obtained from the processed AFC and AVL data for individual passenger journeys. The expected values for t^{ivt} are computed by simply taking the average over all individual passenger journeys i from the total set of observations I within a given time interval (Eq. (1)).

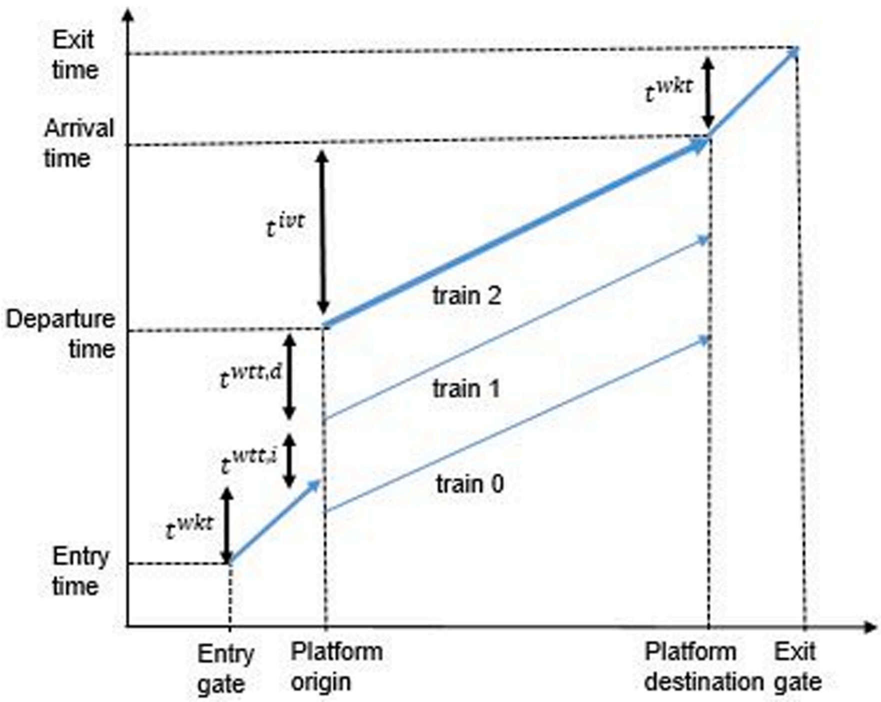

Fig. 2. Example passenger journey where a passenger boards the second departing train from the platform due to denied boarding for the first departing train (adapted from Zhu et al. 2017a).

$$t^{ivt}_{a_{od,t}} = \sum_{i \in I} t^{ivt}_{i_{a,od,t}} / |I_{a_{od,t}}| \quad (1)$$

For our study, walking times between each station entry/exit and platform and between different platforms are available from a model of the incumbent PT agency, which are directly used in our model for t^{wkt}. Alternatively, one could derive a passenger walking speed distribution in the passenger-to-train assignment to obtain t^{wkt} (see Zhu et al. 2017a, Hänseler et al. 2020). Before we calculate the expected values for t^{ivt} and t^{wkt}, we first remove outliers from the dataset, as these indicate unplanned disturbances or disruptions. Similar as for the exclusion of planned maintenance works (Section 2.1), data from disruptions is removed to align observed and expected attribute values as much as possible. For a given path a_{od}, an observation is considered an outlier if the value is larger than $q_3 + (1.5^*(q_3 - q_1))$ or smaller than $q_1 - (1.5^*(q_3 - q_1))$, where $q_3 - q_1$ reflects the inter-quartile range for that particular attribute for each a_{od}.

The total waiting time for each journey leg can be obtained by taking the time difference between the departure time of the boarded vehicle – using the passenger-to-train assignment model outcomes – and the estimated passenger arrival time at the platform $t^{arr}_{s,i}$. The latter can be determined based on the empirically available tap in time at the station gate (from AFC data) or the arrival time of the previous PT journey leg (from AVL data) and t^{wkt}. However, for estimating the valuation of denied boarding we need to distinguish between initial waiting time $t^{wtt,i}$ for the first arriving PT trip k_1, and waiting time after being denied boarding $t^{wtt,d}$until the departure time of the n^{th} PT trip k_na passenger actually boarded. To this end, we compare the total passenger wait time t^{wtt} with the realised headway between the boarded trip k_n and the leading trip k_{n-1} obtained from AVL data. If the total wait time is smaller than the leading headway, the passenger boarded the first departing PT trip k_1 after the platform arrival time (i.e. $t^{wtt,d}$ is assumed zero). In case t^{wtt} exceeds the realised headway, this may indicate denied boarding. To exclude other reasons causing a passenger not to board the first trip (e.g. buying food, making a phone call on the platform), we additionally confirm that the passenger load q of the previous PT trip k_{n-1} was at least 75% of the total crush capacity ϑ of the stock type operating this trip when departing from station s. We do not require q to be equal to ϑ, as passengers are typically not entirely uniformly distributed over the different carriages or parts of a vehicle. Depending on the waiting position on the platform, passengers might not be able to board a certain carriage, whilst some other carriages still have some residual capacity. As we do not have information available on loads per carriage (e.g. from load-weight data), nor on the exact waiting position of an individual passenger on the platform, we use 75% vehicle capacity utilisation as a proxy for overcrowding that may lead passengers to be left behind. This accounts for potential unevenness of the passenger load distribution, as for example empirically shown by Hänseler et al. (2020) who found a 13% difference in load between busiest and least busy train carriage. Eq. (2) formulates the calculation for waiting time for left behind passengers. The initial waiting time equals the difference between total waiting time and denied boarding waiting time (Eq. (3)). Again, outliers are removed to exclude potential unplanned disruptions from the dataset. Passenger journeys for which the realised headway with the leading trip exceeds 2.5 times the scheduled headway of that PT line during that period are considered outliers in this study. For a PT line with a scheduled 8-minutes headway, this implies that a 20-minutes gap between two subsequent trips is the threshold for being considered an outlier, suggesting a disruption. The expected initial waiting time and denied boarding waiting time per path and time period can then be calculated by averaging observed values for individual passengers per time interval, similar to Eq. (1) for t^{ivt}. The denied boarding probability $p^d_{a_{od}}$ for each path can be computed by calculating the fraction of journeys for which $t^{wtt,d}$ is larger than zero (Eq. (5)), using dummy variable x_1 as defined in Eq. (4). The expected denied boarding waiting time effectively equals the denied boarding probability, multiplied by the headway waiting time extension.

$$t^{wtt,d}_{s,i} = \begin{cases} t^{dep}_{s,k_n} - t^{dep}_{s,k_{n-1}} \text{ if } t^{wtt}_{s,i} > t^{dep}_{s,k_n} - t^{dep}_{s,k_{n-1}} \text{ and } q_{s,k_{n-1}} \geq 0.75^*\vartheta_{k_{n-1}} \\ 0 \text{ if } t^{wtt}_{s,i} < t^{dep}_{s,k_n} - t^{dep}_{s,k_{n-1}} \text{ or } (t^{wtt}_{s,i} > t^{dep}_{s,k_n} - t^{dep}_{s,k_{n-1}} \text{ and } q_{s,k_{n-1}} < 0.75^*\vartheta_{k_{n-1}}) \end{cases} \quad (2)$$

$$t^{wtt,i}_{s,i} = t^{dep}_{s,k_n} - t^{arr}_{s,i} - t^{wtt,d}_{s,i} \quad (3)$$

$$x_{1,i} = \begin{cases} 1 \text{ } if \text{ } t^{wtt,d}_{s,i} > 0 \\ 0 \text{ } if \text{ } t^{wtt,d}_{s,i} = 0 \end{cases} \quad (4)$$

$$p^d_{a_{od,t}} = \left(\sum_{i \in I} x_{1,i_{a,od,t}} \right) \Big/ |I_{a_{od,t}}| \quad (5)$$

We use the load factor lf_k to express passengers' in-vehicle time perception as a function of on-board crowding levels. This results in the estimation of an in-vehicle time crowding multiplier to reflect perceived on-board travel time. The load factor is the ratio between the on-board passenger volume q_k and the seat capacity θ_k, and thus equals one when $q_k = \theta_k$. Whilst it is typically preferred to use the standing density as a metric for crowding when $q_k > \theta_k$ to be able to capture differences in vehicle layouts (Wardman and Whelan 2011), we choose to use the load factor for the two following reasons. First, while the seat capacity θ and crush capacity ϑ are available for our case study, the maximum number of standing passengers per m^2 corresponding to the crush capacity is not provided. Second, in our case study all metro stock has the same ratio between seat and crush capacity, which implies that different vehicle layouts are irrelevant here. When q_k is smaller than 50% of the seat capacity, lf_k is set equal to zero. We did experiment with using an unconstrained load factor, a constrained load factor at 50% and 100% of the seat capacity, and estimating a piecewise linear crowding function below and above the seat capacity. However, estimating a linear in-vehicle time multiplier starting from 50% of the seat capacity resulted in the most plausible and significant results. From a theoretical perspective, setting this lower bound for the load factor can be justified, as passengers will on average start having to sit next to other passengers when 50% of the seats are occupied. One can expect this might increase the perceived in-vehicle time from this value onwards. For a given passenger journey i on path a_{od}, the load factor is calculated as the travel time weighted average value, based on the inferred load factor and realised PT running time for each link e_i which constitutes a part of this path E_i (Eq. (6)). E_i refers here to the set of links on path a_{od} for passenger journey i. The expected load factor per path and time period is computed by averaging values for individual passenger journeys $i$per time interval, similar to Eq. (1).

$$lf_i = \max\left(\frac{\sum_{e_i \in E_i} \frac{q_{k,e_i}}{\theta_k} \cdot t_{k,e_i}^{ivt}}{\sum_{e_i \in E_i} t_{k,e_i}^{ivt}} - 0.5, 0\right) \tag{6}$$

In the standard multinomial logit (MNL) model, there is no correction applied for potential overlap between two observed routes due to the Independence of Irrelevant Alternatives (IIA) assumption. To correct for this, we estimate a path size logit (PSL) model. By adding the natural logarithm of the path size factor $r_{a_{od}}$as correction term to our model, we can still estimate a closed form discrete choice model. The path size factor is calculated based on the share of a route between an OD pair which is commonly used by other routes $a_{od} \in A_{od}$ (Eq. (7)). The number of routes by which each link is used is indicated by j in Eq. (7) (ranging from 1 to $|A_{od}|$). As passengers will often not be aware of the detailed distance or running time between each segment, we expect that the fraction of common links might best approximate passenger's overlap perception, as this can be derived relatively easy from a map. From Eq. (7) can be seen that $r_{a_{od}}$ equals $\ln(1)$ in case of two routes having no overlap.

$$r_{a_{od}} = \ln\left(\sum_{j \in 1..|A_{od}|} \left[\left(\frac{|e_{a_{od}}|_j}{|e_{a_{od}}|}\right) * \left(\frac{1}{j}\right)\right]\right) \tag{7}$$

To obtain passengers' preferences related to denied boarding, we estimate the PSL model as abovementioned. In this study we assume a standard utility maximisation framework, which assumes that passengers choose the route which minimises their disutility. The expected disutility of a certain route $U(V,\varepsilon)$ consists of a structural utility component V, which is a vector of observable attributes with their corresponding weights, and a random utility component ε. The structural part of the utility function used in our model is presented in Eq. (8). Except for the lower bound constraint applied to the load factor attribute, we assume linear relations between each attribute and the corresponding disutility. Testing the estimation of a model using a piecewise linear function for waiting time valuation after denied boarding, instead of assuming a linear function, did not provide plausible results. The coefficients of the attributes are obtained by applying maximum likelihood estimation (MLE) using PythonBiogeme (Bierlaire 2016). In addition to the closed form PSL discrete choice model, we also estimate a mixed logit model to test whether significant taste heterogeneity between different respondents exists. For this purpose, coefficients for the standard deviation of each perception coefficient are added to the model, which is solved with simulated MLE using Halton draws. In this study, we do not estimate a panel effects model. It is possible that multiple journeys made by the same individual passenger are included in the dataset, which would ask for a correction for possible correlations between choices of the same individual. However, as no smart card ID was available for this study due to privacy regulations, it is not possible to identify records associated with the same individual and thus to account for panel effects. Instead, we therefore use robust standard errors and robust t-values to correct for a possible overestimation of values of coefficients.

$$V = \beta^{wkt} \cdot t^{wkt} + \beta^{wtt,i} \cdot t^{wtt,i} + \beta^{wtt,d} \cdot t^{wtt,d} + \left(\beta^{ivt} \cdot t^{ivt} \cdot \left(1 + \left(\beta^{lf} \cdot lf\right)\right)\right) + \beta^{tf} \cdot n^{tf} + \beta^{r} \cdot r \tag{8}$$

2.3. Choice set generation

As mentioned in Chapter 1, estimation of denied boarding perceptions adds additional constraints to the choice set generation process, in order to use a choice set which can capture trade-offs between expected waiting time induced by denied boarding and other journey attributes from revealed PT route choices. The criteria applied to arrive at a suitable choice set – pertaining to individual path alternatives as well as to the path set composition - are discussed below.

First, there should be at least *two different observed paths* taken between a certain OD pair to be included in the choice set. In our study, we only use observed route choices in the choice set, as this does not require making assumptions about feasible, non-observed routes and ensures that all paths included are considered attractive under some circumstances. We group stations which are within walking distance from each other by applying hierarchical clustering. We use a complete clustering which only groups stations together if the Euclidean distance between all of them is less than 500 m, as this can be assumed a reasonable maximum walking distance (see for example Munizaga and Palma 2012, Gordon et al. 2013). When performing a sensitivity analysis to this threshold value, the clustering results did not change when using a threshold of 400 (−20%) or 600 (+20%) Euclidean metres. After grouping origin and destination stations, there should be at least two different chosen paths in the dataset for a given time interval t. The average attribute values of each path starting within the same time interval are used as the expected attribute values when populating the chosen and non-chosen, observed route alternatives (see also Eq. (1)). We require the observed path probability to amount to at least 10% for all paths $a_{od,t} \in A_{od,t}$, as this makes sure that a certain path is not only chosen under very rare or unexpected circumstances, such as during an unplanned disruption.

Second, for each OD pair at least one path $a_{od,t} \in A_{od,t}$ should consist of a boarding station where *denied boarding can reasonably be* *expected*. This can be the first boarding station, and/or a transfer boarding location. We require this denied boarding probability $p^{d}_{a_{od,t}}$ (following the calculation in Eq. (5)) to be at least 10% for a specific path and 15-minute time interval t. We deem experiencing denied boarding at least once in every ten trips on average (i.e. once every two weeks for a daily traveller) the minimum for which one can assume that passenger might adjust their route choice in the anticipation thereof. We experimented using different thresholds for $p^{d}_{a_{od,t}}$ (10%, 25%, 40%) to test the sensitivity of our results to this assumption, and found that a threshold of 10% provided most plausible results.

Third, it is required that all paths are *physically distinctive* from each other and that these paths *do not share the same boarding stop(s) where denied boarding occurs* (using the second criteria above). Each path is thus composed of a unique stop sequence, in line with the definition in Section 2.2. Two routes using different PT lines on the same PT corridor, which both share the exact same infrastructure, do therefore not satisfy this criterion. The requirement for physically distinctive paths and denied boarding locations stems from the study objective to use anticipated route choice to derive denied boarding valuation. It is therefore important that passengers make their route choice decision based on the expected probability of being left behind and the additional waiting time induced thereby. If two routes share the same boarding location where denied boarding might occur, there is a risk that a passenger made a route choice decision for route option one, but was forced to board the alternative route option as the first arriving trip of route option one arrived at capacity. In this case, the empirical route choice does not align with the intended route choice.

The last three criteria are related to the choice set composition, ensuring sufficient *variance*, a clear *trade-off* and the absence of *dominance* between observed route alternatives. We check if there is sufficient variance in the attribute levels for in-vehicle time, initial and denied boarding wait time, and crowding between different paths $a_{od,t} \in A_{od,t}$ to be able to estimate a RP based route choice model. Furthermore, there must be a clear trade-off between either a higher expected denied boarding waiting time on the one hand and a shorter expected regular journey time or fewer transfers on the other hand, or the other way around. This enables us to only include choices which can expose trade-offs between denied boarding and other journey attributes. At last, we exclude OD pairs and time intervals from the choice set if there is a clear dominance where all attribute values of a certain path are more favourable than all attribute levels of another path.

3. Case study

We apply our method to the metro network of Washington, DC, as a case study to estimate denied boarding perceptions. At the time of consideration, the DC metro network consisted of 95 stations, which are served by six different lines (Red, Green, Yellow, Blue, Silver and Orange lines) (see Fig. 3). During the peak hours, the Red Line has a scheduled headway of 4 min with alternating services

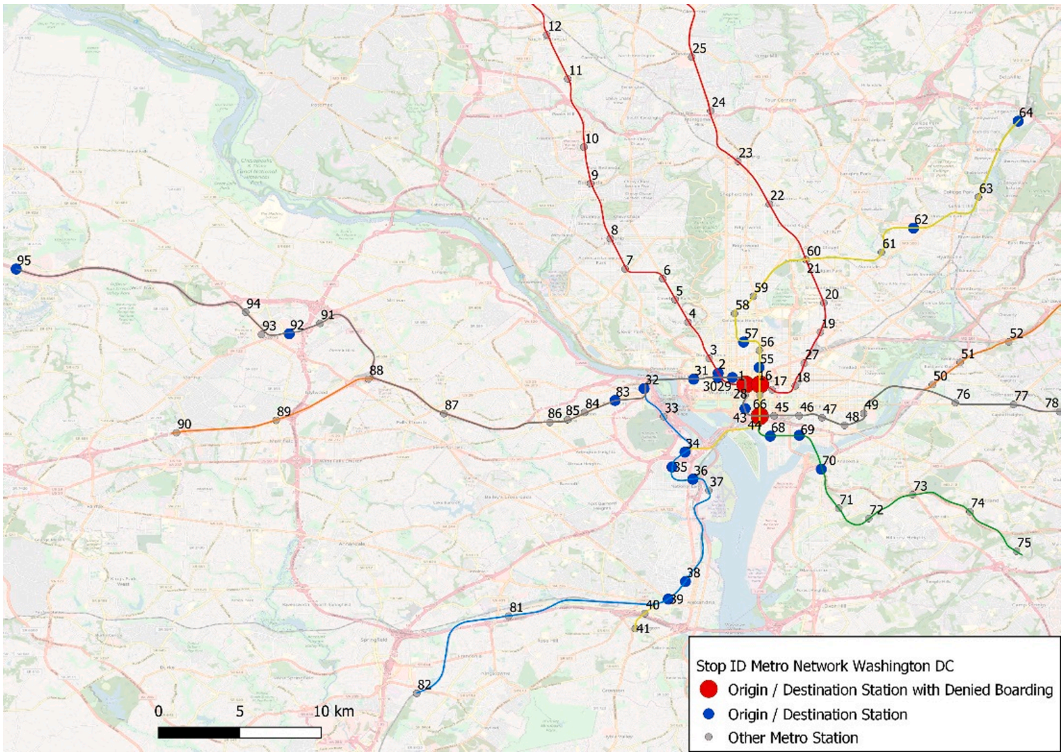

Fig. 3. Washington DC metro network with OD stops included in the choice set. *The different coloured lines indicate the metro lines (Red, Green, Yellow, Blue, Orange and Silver lines).*

terminating at Glenmont or Silver Spring at the north-eastern part of the line. All other lines have a scheduled peak headway of 8 min. As input for our study, we used processed AFC and AVL data of the entire month of September 2017. Based on the temporal distribution of denied boarding occurrences in the dataset, we only kept AFC and AVL data from Monday to Thursday (Labor Day as public holiday being excluded) in the AM and PM peak, as the vast majority of denied boarding instances occurred within these periods. In addition, the AM and PM peak typically contain the largest share of frequent, experienced passengers, who are most likely to have the ability to form expectations related to denied boarding probabilities based on previous travel experiences. No planned maintenance works took place during the AM and PM peak of this analysis period.

Using the millions of passenger metro journeys as initial input, we apply the choice set generation criteria as detailed in Section 2.3 to obtain our final choice set. From the dataset can be concluded that only three stations in the Washington DC metro network experience sufficient denied boarding to satisfy the minimum denied boarding probabilities (see the red-marked stations in Fig. 3):

- Metro Center (Red Line) eastbound to Gallery Place (PM: 17:30–17:45);
- Gallery Place (Red Line) westbound to Metro Center (AM: 08:15–09:00);
- L'Enfant Plaza (Yellow Line) southbound to Pentagon (PM: 17:15–18:00).

In Fig. 4, we report the average denied boarding probabilities for these three stations for all boarding passengers per 15 min interval within the analysis period. Notwithstanding, in the choice set denied boarding probabilities are calculated for each path and time interval separately. Given these results, we only include OD pairs in the choice set where passengers of at least one of the observed paths use one of these three stations (in the relevant direction). From all 8930 (95*94) theoretically possible OD pairs (excluding pairs with the same origin and destination), only 23 OD pairs satisfy all choice set criteria. The origin and destination stations of these pairs are shown on the map in Fig. 3. These 23 OD pairs are composed of 51 different paths in total, indicating that there are two different observed paths for most OD pairs. The maximum number of paths $|A_{od}|$ is three, given the limited number of cycles in the case study network. When also including the different 15-minute time intervals, the choice set consists of 37 OD pairs/time interval combinations (see Table 1). Please note that the time intervals in this table reflect the journey start time, which is not necessarily the same time interval as when a passenger boards at one of the stations with potential denied boarding. In total, 5336 observed route choices remain in the choice set for model estimation.

4. Results and discussion

4.1. Results

The maximum likelihood estimation performed in PythonBiogeme is solved using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm as iterative method to solve this nonlinear optimisation problem. The model converged after 37 iterations within 3 s on a regular i7 core Dell laptop. The null log likelihood of the initial model equals −15,252, whereas the final log likelihood for the estimated model is equal to −4161. Both the Rho-square and Rho-square-bar are 0.727. In the estimation process, the walking time coefficient β^{wkt} as specified in the utility function (Eq. (8)) was found to be statistically insignificant. This is probably due to insufficient variation in attribute levels of t^{wkt} in the choice set, as walking times within the metro network (between gates and platforms, or between platforms) are often very similar for different paths $a_{od} \in A_{od}$. We therefore conclude that there is not sufficient variation in our data in regard to walking times to allow estimating a significant walking time coefficient. Based on several previous studies (e.g. Balcombe et al. 2004, Bovy and Hoogendoorn-Lanser 2005), we can expect this coefficient to be significant if there would have been sufficient variation in the choice set, for example using a SP approach. Due to correlations between different attributes, we removed

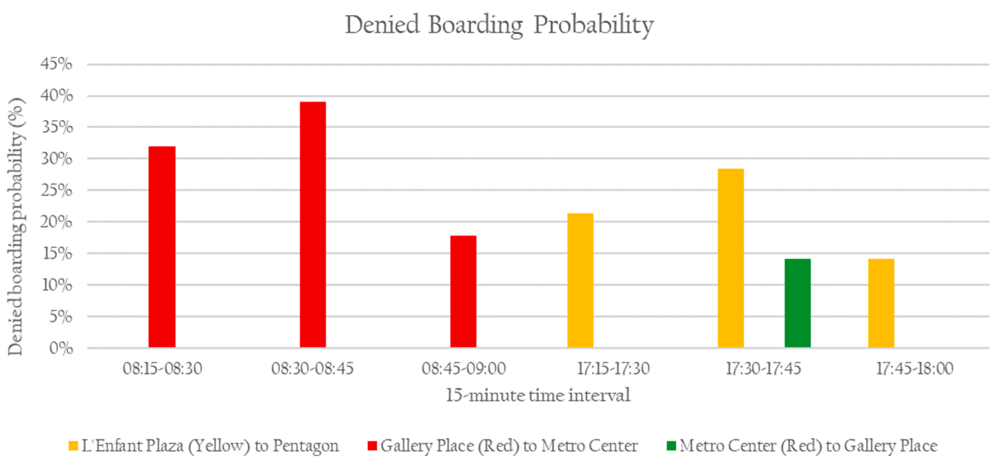

Fig. 4. Average denied boarding probability per station and per 15-minute time interval.

Table 1
OD pairs and time intervals included in the choice set.

OD Pair	From ID	Origin Stop Name	To ID	Destination Stop Name	Time Intervals
1	70	Anacostia	2/30	Farragut North/West	08:30–08:45
2	44/67	L'Enfant Plaza	32	Rosslyn	17:15–17:45
3	44/67	L'Enfant Plaza	83	Court House	17:00–17:15
4	44/67	L'Enfant Plaza	92	Tysons Corner	17:00–17:15
5	28	Metro Center	35	Pentagon City	17:45–18:00
6	28	Metro Center	36	Crystal City	16:45–17:00; 17:45–18:00
7	28	Metro Center	39	King St-Old Town	17:45–18:00
8	55	Mt Vernon Sq	31	Foggy Bottom-GWU	08:00–08:15; 08:30–08:45
9	55	Mt Vernon Sq	32	Rosslyn	08:30–08:45
10	69	Navy Yard-Ballpark	2/30	Farragut North/West	08:15–08:30
11	69	Navy Yard-Ballpark	95	Wiehle-Reston East	17:30–17:45
12	43	Smithsonian	34	Pentagon	17:00–17:30
13	43	Smithsonian	35	Pentagon City	17:00–17:45
14	43	Smithsonian	36	Crystal City	17:00–17:45
15	43	Smithsonian	38	Braddock Road	17:15–17:30
16	43	Smithsonian	39	King St-Old Town	17:00–17:45
17	68	Waterfront	2/30	Farragut North/West	08:45–09:00
18	64	Greenbelt	2/30	Farragut North/West	08:15–08:30
19	62	Prince Georges Plaza	2/30	Farragut North/West	08:15–08:45
20	62	Prince Georges Plaza	29	McPherson Square	08:15–08:30
21	32	Rosslyn	57	U Street	17:15–17:30
22	34	Pentagon	1/28	Metro Center	08:00–08:45
23	38	Braddock Road	1/28	Metro Center	08:00–08:30

walking time from the utility function and re-estimated a final model. Table 2 presents the estimation results of this final model. In this model, all remaining coefficients are statistically significant with all p-values <0.01. Furthermore, for all estimated coefficients the sign is plausible.

In this final model, we include six coefficients reflecting the standard deviation of the coefficients of these attributes to estimate a mixed logit model for assessing taste heterogeneity. All six coefficients are found insignificant, which means that no significant heterogeneity in valuation between passengers is observed. A possible explanation is that our choice set only consists of journeys made in the AM and PM between Mondays to Thursday, using routes with high crowding levels and denied boarding occurrences. Typically, these are routes most often used by frequent commuting or business travellers. We can therefore hypothesise there might not be significant heterogeneity in preferences between individuals of this particular segment. Due to the insignificant spread coefficients, we proceed with reporting results from the closed form path size logit model and discussing these model estimation results in the subsequent sub-section.

4.2. Discussion

In Fig. 5, we scale the estimated coefficients relative to the in-vehicle time coefficient. It can be seen that on average one minute of initial waiting time is perceived as 1.62 min (uncrowded) in-vehicle time. This ratio of 1.62 is very similar to the ratio between out-of-vehicle time (waiting and/or walking time) and in-vehicle time of 1.68 found in a recent RP based study which calibrates a route choice model based on smart card data in Seoul (Kim et al. 2020). Yap et al. (2020) found a ratio of 1.5 between out-of-vehicle and in-vehicle time based on a RP study applied in the Netherlands. This suggests that the found ratio of 1.62 is within a plausible range compared to other recent RP based studies. Earlier performed SP based studies often find a somewhat higher ratio between (initial) waiting and uncrowded on-board time, such as a ratio of 2.2 found by Bovy and Hoogendoorn-Lanser (2005).

Making a transfer is on average perceived as almost 8.5 min uncrowded in-vehicle time, which can be considered a fixed transfer penalty for transferring, above and beyond any additional travel times induced by transferring. This value lies between the transfer penalty of 5.2 min found in a RP study in the Netherlands when interchanging between urban trams and buses (Yap et al. 2020), and

Table 2
Estimation results.

Coefficient	Name	Value (robust t-value)
β^{ivt}	in-vehicle time	−0.0739** (−8.61)
$\beta^{wtt,i}$	initial waiting time	−0.120** (−4.63)
$\beta^{wtt,d}$	waiting time after denied boarding	−0.201** (−5.13)
β^{tf}	transfer penalty	−0.627** (−5.21)
β^{lf}	load factor	0.389** (3.16)
β^{r}	log-path size factor	−2.46** (−12.9)

Robust t-values in parentheses. * $p < 0.05$; ** $p < 0.01$.

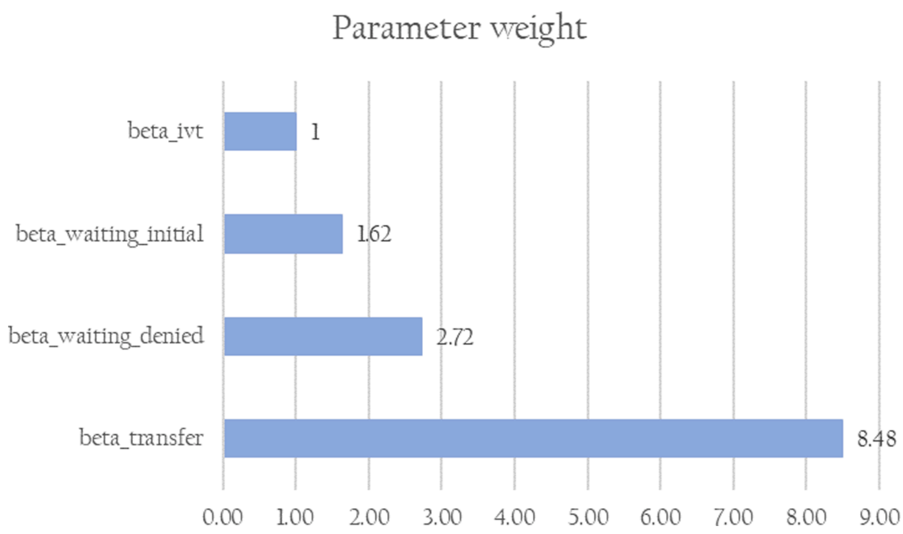

Fig. 5. Relative coefficient weights.

the average transfer penalty of 15–17 min found in a combined SP/RP study in Madrid for multimodal urban transfers (Garcia-Martinez et al. 2018). We can therefore conclude that the transfer penalty found in our study falls within the expected range of values. As we did not include walking time in our final model, it is possible that our transfer penalty is slightly inflated due to the implicit incorporation of walking time during a transfer, which might explain why our value is somewhat higher than for example found by Yap et al. (2020).

Our estimation results show that one minute waiting time after being left behind is on average perceived as 2.72 min uncrowded in-vehicle time. This means that one minute waiting time after denied boarding is on average perceived 68% more negatively, compared to one minute initial waiting time for the first arriving PT vehicle. Whilst there are no previous studies estimating this ratio, these results confirm the intuitive expectation that waiting time after being left behind is experienced more negatively than initial waiting time and is thus in line with our expectations. A possible explanation for this more negative waiting time valuation is the impact denied boarding has on passengers' final arrival time. Passengers travelling during the peaks - when denied boarding is most likely to occur - for commuting or other journey purposes with relatively inflexible schedules, may experience inconvenience from arriving later at their final destination than planned. For example, this higher perception coefficient could in that case reflect the higher marginal costs for passengers arriving too late at work or a business meeting. A second explanation for the found coefficient may be linked to uncertainty about the arrival time at the final destination when denied boarding, depending on the arrival time and denied boarding probability of the next train, which might be experienced more negatively by passengers.

In Fig. 6 (left), we report the estimated in-vehicle time crowding multiplier as function of the load factor (ratio between passenger volume and seat capacity). As explained in Section 2.2, a linear load factor coefficient starting when 50% of the seats are occupied provided the most plausible and significant results. In the left figure, one can see that one minute in-vehicle time is perceived as 1.19 min when all seats are occupied. This is very much in line with RP values of 1.265 found by Hörcher et al. (2017) for the metro network of Hong Kong, and the crowding multiplier of 1.16 found for the Netherlands by Yap et al. (2020). The crowding multiplier then linearly increases until the crush capacity is reached. As the crush capacity of the Washington DC metro stock equals a load factor of 2.65, the in-vehicle time multiplier at crush capacity is equal to 1.84. As mentioned in Section 2.2, we do not have information on which standing density corresponds to the crush capacity. However, if we assume that the crush capacity corresponds to typical manufacturer specifications of four standing passengers per m^2, we can convert the directly estimated figure of Fig. 6 (left) to a figure

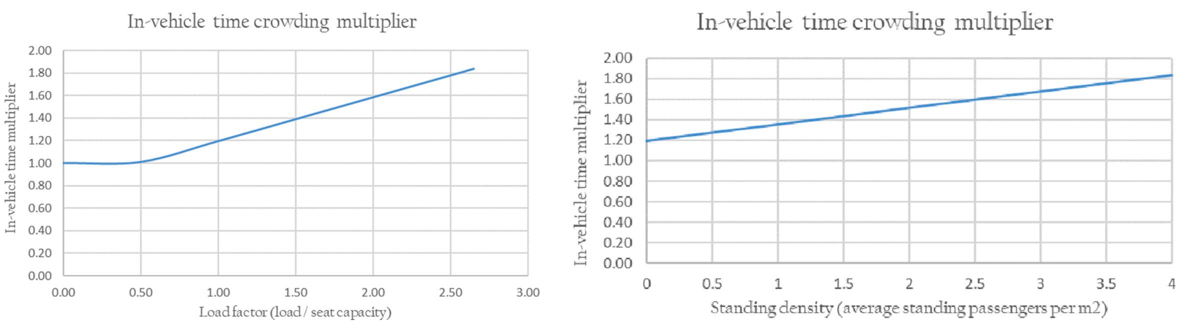

Fig. 6. In-vehicle time crowding multiplier as function of load factor (left) and standing density (right).

showing the crowding multiplier as a function of the standing density (Fig. 6, right). This would imply that the crowding multiplier increases with 0.16 for each increase in standing passengers per square metre. In comparison, Hörcher et al. (2017) found an increase of 0.119 for each increase in standing density in Hong Kong, whilst Tirachini et al. (2016) estimated a crowding multiplier increase of 0.18 in Singapore. Our crowding estimation results thus show to be consistent with those reported by other studies which estimated RP route choice models for metro networks worldwide, providing confidence in our estimation results including the denied boarding estimates.

4.3. Policy implications

We derive several policy implications from our model estimation results. First, based on our study it becomes possible to have a more accurate estimation of passenger and societal impacts of overcrowding in PT systems. When additional waiting time due to denied boarding is quantified and monetised using our higher waiting time coefficient, instead of the traditionally used initial waiting time coefficient, this leads to higher perceived journey times than previously assumed. This will result in larger impacts of overcrowding on PT attractiveness if quantified using our bespoke denied boarding waiting time coefficient. Second, our results can influence outcomes of business cases and societal cost-benefit analyses for schemes which aim at reducing PT crowding. When the benefits of measures which reduce overcrowding and denied boarding probabilities (such as providing higher train frequencies or a vehicle layout which offers increased capacity) are quantified by explicitly distinguishing between initial and denied boarding wait time perception, this can result in a higher benefit-cost ratio or larger societal benefits than previously assumed. Third, our study results have the potential to improve the accuracy of model forecasts of passengers flow distribution. When PT assignment models incorporate the higher denied boarding waiting time perception factor, as obtained from realised route choice behaviour, this would contribute to improved model validity. Moreover, it can also help improving route choice forecasts, as higher denied boarding probabilities can explain why some passengers prefer to use another route, even if this route might involve an additional transfer or longer journey times. Consequentially, an improved forecast accuracy can further improve the accuracy of modelled values which feed into business cases or cost-benefit analyses.

We illustrate the policy implications of our work using a simplified case study for the London Underground network. We consider the Waterloo & City line, which is a very short line running directly from Waterloo train station to Bank. The nominal running time equals only 4 min without any intermediate stops. Despite the short route, this line is heavily used by commuters who arrive by train at Waterloo station travelling into the city of London. As a result, the line suffers from high crowding levels and denied boarding during the AM peak (towards Bank) and PM peak (towards Waterloo station). Fig. 7 (left) shows the reference demand profile for the busiest AM hour between 08:00 and 09:00 from Waterloo to Bank, which is obtained from publicly available data from Transport for London (Transport for London 2018). The crush capacity per train is estimated based on specification of the number of carriages and seats per train. This figure illustrates that, even when assuming a uniform demand and supply pattern per 15-min interval, denied boarding occurs between 08:15 and 08:45. Within this hour, the total frequency equals 22 trains per hour per direction, of which 10 trains run between 08:15 and 08:45. As a hypothetical case, we test the impact of adding one extra train for each of these two busiest 15-min time intervals, thus effectively increasing the hourly frequency to 24 trains per hour. We compare the perceived passenger journey time hours between using the initial wait time coefficient for all waiting time, and applying the bespoke denied boarding wait time coefficient for the expected denied boarding wait time. Fig. 7 (right) shows these values for all passengers travelling on the Waterloo & City Line to Bank between 08:15 and 08:45. When uniform demand and supply is assumed per 15 min, adding these additional trains would resolve denied boarding issues. When calculating the passenger benefits from this measure as the difference in perceived passenger journey time between test and reference scenario, we can conclude that the default parameter set forecasts a reduction of 65 passenger hours. When applying our new estimated parameter set, the forecast crowding relief benefits equal a reduction of 75 passenger hours. This illustrates that not applying a higher wait time coefficient for waiting times induced by denied boarding results in an underestimation of the crowding relief benefits of schemes. For this case study, this benefit underestimation equals 14%, which can substantially affect the feasibility of the proposed scheme in practice.

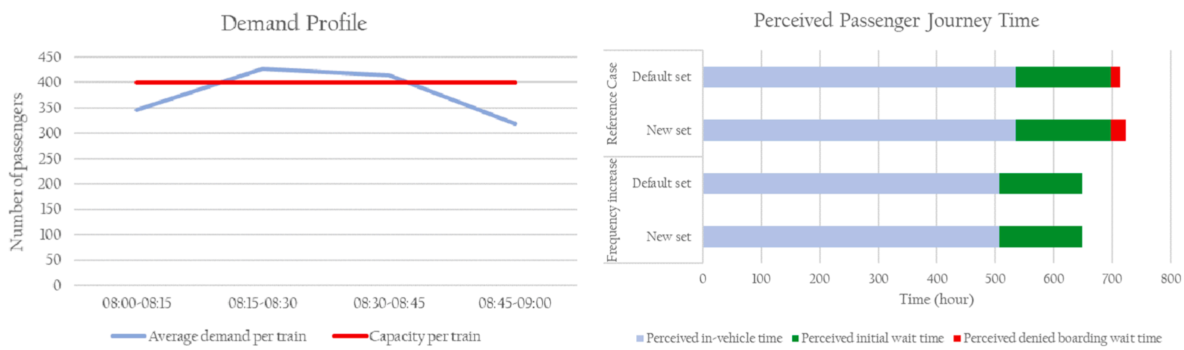

Fig. 7. Demand profile Waterloo & City Line (left) and journey time impacts of frequency increase between 08:15 and 08:45 (right).

5. Conclusions

We estimated a discrete choice model based on observed route choices to infer how passengers value waiting time after being denied boarding in crowded public transport networks. This gives us a quantitative indication of how waiting time is perceived by left behind passengers, compared to initial waiting time for the first arriving vehicle as well as compared to in-vehicle time. Instead of using passenger satisfaction surveys or other direct interrogation techniques, in our work the passenger perception after being denied boarding is inferred from realised route choice behaviour. Our work contributes to a better understanding of the impact of overcrowding in public transport on passengers' travel experiences and route choice decisions. The estimation results of our model confirm the intuitive expectation: waiting time after denied boarding is perceived more negatively compared to initial waiting time. We found that the expected waiting time after denied boarding is valued 68% more negatively compared to the expected initial waiting time. On average, one minute initial and denied boarding wait time are perceived as 1.62 and 2.72 min on-board an uncrowded vehicle, respectively. Not incorporating this more negative denied boarding wait time valuation can result in an underestimation of the passenger and societal impact of overcrowding in PT systems. Moreover, it can underestimate the benefits of potential crowding relief measures and as such underestimate the benefit-cost ratio of these measures and undermine the prioritisation of related efforts. Our findings emphasise the importance of explicitly distinguishing between initial and denied boarding wait time valuation for obtaining a more accurate quantification of impacts of measures. It also has the potential to improve the validation and forecasting performance of PT assignment models.

We identify three key limitations of our approach. First, we used an existing passenger-to-train assignment procedure to process raw AFC and AVL data. As our case study application entails a metro network where passengers need to tap in and tap out at the stations, passenger route choices and train crowding levels can only be inferred using such algorithm, instead of being directly available from empirical data. Although the adopted ODX algorithm by Sánchez-Martinez (2017) is well established and successfully applied and validated for many PT systems worldwide, inference processes are inevitably not error free, which can influence the accuracy of our input data. However, due to the relatively simple structure of the Washington DC metro network without extremely high service frequencies (22.5 trains per hour on the busiest corridor during the peaks), we expect this algorithm to perform well for our particular case study. Second, in our approach a passenger who does not board the first departing train is only considered as being denied boarding if the passenger load of this departed train exceeds 75% of the total capacity, thereby allowing for some uneven passenger distribution over the different parts of the train. When performing a sensitivity analysis to this assumption, the found ratio between initial and denied boarding waiting time coefficients shows to be robust regarding this assumption. However, the ratio between in and vehicle time and denied boarding wait time appears to be somewhat more sensitive to this assumption. We therefore recommend applying a load unevenness factor as accurate as possible for the particular case study of investigation, for example by calibrating this factor based on data from load-weight systems. Third, we did not have access to the smart card ID of individual passengers due to privacy regulation constraints. If such data would have been available, then it would have been possible to distinguish between frequent, experienced passengers and infrequent passengers, for example as done by Yap et al. (2020). As infrequent passengers will not have prior knowledge about denied boarding probabilities, one can assume that this passenger segment is not able to incorporate crowding and denied boarding into their route choice. The observed route attribute values for this group in particular will thus be less aligned with the expected route attribute values when making the route choice decision. By focusing solely on AM and PM peak journeys, we aimed to include as many frequent, experienced travellers in our choice set as possible, as the share of less experienced passengers is typically low during the peaks. However, it cannot be ruled out that a limited number of route choices made by infrequent passengers is still included in our dataset. As this group is not expected to incorporate crowding or denied boarding in their route choice decision due to unfamiliarity with these attribute values, it is possible that the estimated crowding and denied boarding coefficients in our work are on the lower end of the scale. For further research, testing the impact of segmentation between passenger groups on estimated coefficients is therefore recommended.

CRediT authorship contribution statement

Menno Yap: Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing. **Oded Cats:** Methodology, Writing - original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the Washington Metropolitan Area Transit Agency, WMATA, for the valuable cooperation and data provision.

References

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P., 2004. The demand for public transport: a practical guide. TRL Report TRL 593, Crowthorne, UK.
Batarce, M., Munoz, J.C., Ortuzar, J., 2016. Valuing crowding in public transport: Implications for cost-benefit analysis. Transp. Res. Part A 91, 358–378.
Batarce, M., Munoz, J.C., Ortuzar, J., Raveau, S., Mojica, C., Rios, R.A., 2015. Valuing crowding in public transport systems using mixed stated/revealed preference data: the case of Santiago. Transp. Res. Rec. 2535, 73–78.
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.
Börjesson, M., Eliasson, J., Franklin, J.P., 2012. Valuations of travel time variability in scheduling versus mean–variance models. Transp. Res. Part B 46, 855–873.
Bovy, P.H.L., Hoogendoorn-Lanser, S., 2005. Modelling route choice behaviour in multi-modal transport networks. Transportation 32, 341–368.
Cats, O., West, J., Eliasson, J., 2016. A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. Transp. Res. Part B 89, 43–57.
Fosgerau, M., Engelson, L., 2011. The value of travel time variance. Transp. Res. Part B 45, 1–8.
Fosgerau, M., Karlström, A., 2010. The value of reliability. Transp. Res. Part B 44, 38–49.
Garcia-Martinez, A., Cascajo, R., Jara-Diaz, S.R., Chowdhurry, S., Monzon, A., 2018. Transfer penalties in multimodal public transport networks. Transp. Res. Part A 114, 52–66.
Gordon, J.B., Koutsopoulos, H.N., Wilson, N.H.M., Attanucci, J.P., 2013. Automated inference of linked transit journeys in London using fare-transaction and vehicle location data. Transp. Res. Rec. 2343, 17–24.
Hamdouch, Y., Ho, H.W., Sumalee, A., Wang, G., 2011. Schedule-based transit assignment model with vehicle capacity and seat availability. Transp. Res. Part B 45, 1805–1830.
Hänseler, F.S., Van den Heuvel, J., Cats, O., Daamen, W., Hoogendoorn, S.P., 2020. A passenger-pedestrian model to assess platform and train usage from automated data. Transp. Res. Part A 132, 948–968.
Hörcher, D., Graham, D.J., Anderson, R.J., 2017. Crowding cost estimation with large scale smart card and vehicle location data. Transp. Res. Part B 95, 105–125.
Jenelius, E., Cats, O., 2015. The value of new public transport links for network robustness and redundancy. Transportmetrica A 11, 819–835.
Kim, K.M., Hong, S.-P., Ko, S.-J., Kim, D., 2015. Does crowding affect the path choice of metro passengers? Transp. Res. Part A 77, 292–304.
Kim, I., 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 47, 2179–2202.
Kroes, E., Kouwenhoven, M., Debrincat, L., Pauget, N., 2014. Value of crowding on public transport in Ile-de-France, France. Transp. Res. Rec. 2417, 37–45.
Li, H., Tu, H., Hensher, D.A., 2016. Integrating the mean-variance and scheduling approaches to allow for schedule delay and trip time variability under uncertainty. Transp. Res. Part A 89, 151–163.
Li, Z., Hensher, D., 2011. Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal. Transp. Policy 18, 880–887.
Miller, E., Sánchez-Martínez, G.E., Nassir, N., 2018. Estimation of passengers left behind by trains in high-frequency transit service operating near capacity. Transp. Res. Rec. 2672, 497–504.
Mishra, S., Tang, L., Ghader, S., Mahapatra, S., Zhang, L., 2018. Estimation and valuation of travel time reliability for transportation planning applications. Case Stud. Transport Policy 6, 51–62.
Munizaga, M.A., Palma, C., 2012. Estimation of a disaggregate multimodal public transport origin-destination matrix from passive smart card data from Santiago, Chile. Transp. Res. Part C 24, 9–18.
MVA Consultancy, 2008. Valuation of overcrowding on rail services. Report prepared for the UK Department of Transport.
Nuzzolo, A., Crisalli, U., Rosati, L., 2012. A schedule-based assignment model with explicit capacity constraints for congested transit networks. Transp. Res. Part C 20, 16–33.
Pel, A., Bel, N., Pieters, M., 2014. Including passengers' response to crowding in the Dutch national train passenger assignment model. Transp. Res. Part A 66, 111–126.
Raveau, S., Guo, Z., Munoz, J.C., Wilson, N.H.M., 2014. A behavioural comparison of route choice on metro networks: time, transfers, crowding, topology and socio-demographics. Transp. Res. Part A 66, 185–195.
Sánchez-Martinez, G.E., 2017. Inference of public transportation trip destinations by using fare transaction and vehicle location data: dynamic programming approach. Transp. Res. Rec. 2652, 1–7.
Tirachini, A., Sun, L., Erath, A., Chakirov, A., 2016. Valuation of sitting and standing in metro trains using revealed preferences. Transp. Policy 47, 94–104.
Tirachini, A., Hurtubia, R., Dekker, T., Daziano, R.A., 2017. Estimation of crowding discomfort in public transport: results from Santiago de Chile. Transp. Res. Part A 103, 311–326.
Transport for London, 2018. Crowding data TfL. Obtained from http://crowding.data.tfl.gov.uk/.
Trépanier, M., Tranchant, N., Chapleau, R., 2007. Individual trip destination estimation in a transit smart card automated fare collection system. J. Intell. Transp. Syst. 11, 1–14.
Wardman, M., Whelan, G., 2011. Twenty years of rail crowding valuation studies: evidence and lessons from British experience. Transp. Rev. 31, 379–398.
Yap, M.D., Cats, O., Van Arem, B., 2020. Crowding valuation in urban tram and bus transportation based on smart card data. Transportmetrica A 16, 23–42.
Yap, M.D., Cats, O., Van Oort, N., Hoogendoorn, S.P., 2017. A robust transfer inference algorithm for public transport journeys during disruptions. Transp. Res. Procedia 27, 1042–1049.
Zhu, Y., Koutsopoulos, H., Wilson, N.H.M., 2017a. A probabilistic Passenger-to-Train Assignment Model based on automated data. Transp. Res. Part B 104, 522–542.
Zhu, Y., Koutsopoulos, H., Wilson, N.H.M., 2017b. Inferring left behind passengers in congested metro systems from automated data. Transp. Res. Part C 23, 362–379.