Article:

Tirachini, A, Hurtubia, R, Dekker, T orcid.org/0000-0003-2313-8419 et al. (1 more author) (2017) Estimation of crowding discomfort in public transport: results from Santiago de Chile. Transportation Research Part A: Policy and Practice, 103. pp. 311-326. ISSN 0965-8564

https://doi.org/10.1016/j.tra.2017.06.008

Estimation of crowding discomfort in public transport: results from Santiago de Chile

Alejandro Tirachini[a], Ricardo Hurtubia[b],
Thijs Dekker[c] and Ricardo A. Daziano[d]

June 20, 2017

[a] Transport Engineering Division, Civil Engineering Department, Universidad de Chile, Santiago, Chile. Email: alejandro.tirachini@ing.uchile.cl

[b] School of Architecture and Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago, Chile. Email: rhg@ing.puc.cl

[c] Institute for Transport Studies (ITS), University of Leeds, UK. Email: t.dekker@leeds.ac.uk

[d] School of Civil and Environmental Engineering, Cornell University, Ithaca NY 14853, USA. Email: daziano@cornell.edu

Abstract

The relationship between train occupancy, comfort and perceived security is analysed, using data from a survey and stated choice (SC) study of users of Santiago's Metro (subway) system. Mode choice models where crowding is one of the main explanatory variables are estimated and crowding multipliers to measure its relevance on travel time disutility for sitting and standing are computed. An international comparison with previous studies from London, Paris, Singapore and Sweden is presented. The type of estimated models include Multinomial Logit, Mixed Logit, and Latent Class models. Results show that there is significant heterogeneity in crowding perception across the population. Users classes with low and high crowding multipliers are identified, in which gender, age and income play a role. In the SC survey, occupancy levels were shown with three alternative forms of representation (text, 2D diagram or photo), however we did not find relevant influences of the different forms of representation on crowding perception.

1 Introduction

In public transport, crowding refers to a subjective perception of the physical phenomenon represented by a high density of passengers in vehicles and at stations, stops and access-ways. In-vehicle crowding is, after price and travel time, one of the most important explanatory variables of mode choice. This is particularly true for public transport modes where high levels of crowding can result in physical discomfort, psychological burden and perceived risk and insecurity (Cox et al., 2006; Cheng, 2010; Mahudin et al., 2012). Moreover, crowding externalities (e.g. slower boarding and alighting from vehicles, increasing waiting times) have an important effect on the overall level of service and optimal fare of public transport systems (Tirachini et al., 2014).

Crowding in public transport is a common phenomenon in Santiago, Chile. Its city-wide integrated public transport system launched in February 2007, also known as the Transantiago system (Muñoz et al., 2014; Munizaga and Palma, 2012), deploys full fare integration between buses and Metro through the use of a single (smartcard) payment method. The implementation of Transantiago heavily loaded the metro network, making it the main artery of the system (Gómez-Lobo, 2012; Muñoz et al., 2014). The total number of daily passengers served by metro duplicated overnight and crowding conditions in the trains became extreme, reaching 6 passengers per square meter or more during peak hours[1]. This triggered many behavioural responses from the users ranging from selecting different modes of transport (there has been an increase in car and bicycle use) to route choices that, in regular crowding conditions, would be classified as being counter-intuitive or irrational (Raveau et al., 2014). For example, it may happen that users opt for longer routes in order to increase the chance of obtaining a seat in the train, or prefer not to board a train or bus because it is considered *too full* (although not reaching yet its full capacity). These behavioural responses reveal the extent to which users dislike crowding in public transport. A further case in point is provided by a user survey revealing that the attribute *comfort*, related to overcrowding, was the worst evaluated attribute of Transantiago (Yanez et al., 2010), a critical issue if we consider that comfort has been reported as a factor that reduce stress of public transport commuters (Legrain et al., 2015).

Despite the large impact of crowding on quality-of-service, the optimization model to design the Transantiago network (Fernandez et al., 2008) did not consider quality-of-service factors such as passenger density and service reliability valuation by users in the design of routes, optimal frequencies and vehicle sizes[2]. Instead the optimization model minimized the summation of users and operator costs. In other words, one minute travelling with five passengers per square meter was assumed to have the same weight in the users' cost function as one minute travelling with one passenger per square meter, thereby ignoring the discomfort of crowding on users.

Understanding and measuring the willingness to trade an increase in travel time for improved travel conditions in terms of reduced crowding levels, and vice-versa, is not only relevant for the planning of new public transport services, but also for the management of currently operating routes and services and cost-benefit analysis of policy interventions aimed at reducing crowding levels, either as a primary or secondary goal. Crowding multipliers (Wardman and Whelan, 2011; Tirachini et al., 2013) can be used for this objective. Crowding multipliers can be interpreted as a measure of how the disutility of travel time under different crowding levels relate to each other. Subsequently, they can be used to amplify the (monetary) value of in-vehicle time savings in order to account for the fact that reductions of travel time in crowded conditions are worth more than reducing travel time on a similar but less crowded trip.

The literature on crowding valuation has progressed quickly during the past ten years, and today we are aware of studies estimating the sensitivity of the value of travel time savings (VTTS) to different vehicle or station crowding conditions in Great Britain (Whelan and Crockett, 2009; Wardman and Whelan, 2011), the Paris region (Kroes et al., 2014; Haywood and Koning, 2015), Sydney (Hensher et al., 2011), Mumbai (Basu and Hunt, 2012), Los Angeles (Vovsha et al., 2013), Singapore (Tirachini et al., 2016), Hong Kong (Lam et al., 1999; Hörcher et al., 2017) and Santiago (Batarce et al., 2015, 2016), amongst other cities. Even in cycling research it was recently found that

[1]There are three reasons for this sudden increase in Metro usage: an integrated fare system in which users pay a very low fee for a bus-metro transfer; the redesign of parts of the bus network to serve as feeders of the metro network; and the noticeable reduction of bus service quality in terms of longer waiting and in-vehicle times, especially at the beginning of Transantiago

[2]In the design model, high occupancy of vehicles does not influence the perception of time but may increase the extension of waiting time through limited capacity considerations (Fernandez et al., 2008)

crowding (with other bicyclists) significantly influence route choice for bicyclists in Copenhagen (Vedel et al., 2017).

This paper makes a number of contributions to the crowding valuation literature. First, we test the impact of the crowding representation format on the perceived level of crowding, resulting travel behaviour and corresponding crowding valuation measures. To this end a stated choice survey is designed in which occupancy levels are presented to respondents either in the form of text, 2D diagrams or photos. Other studies have also used images (2D diagrams and photos) to describe crowding levels. Use of images has shown to influence the perception of attributes of the alternatives on stated preferences surveys (Rizzi et al., 2012) and facilitates the description of complex choice scenarios, where an exhaustive text-based description of the attributes would over-complicate the choice task (Motoaki and Daziano, 2015; Hurtubia et al., 2015). However, some evidence indicates that the form of representation used to describe single attributes has no effect on the perception of the respondent (Arentze et al., 2003).

Second, in this study the usual way to determine crowding externalities by means of a stated choice model is complemented by questions on the relationship between train occupancy and perceived levels of comfort and security, providing a link between subjective user perceptions and observable train occupancies.

Third, this paper follows the recommendations of Basu and Hunt (2012) who argue that significant care is required when establishing crowding multipliers based on Mixed Multinomial Logit (ML) models. In previous crowding valuation studies, user preferences have been estimated using Multinomial Logit (MNL) and ML models. In the realm of MNL models, Wardman and Whelan (2011) develops a meta-analysis of crowding multipliers using MNL values from 17 studies in Great Britain. Ease of application in optimal public transport supply models is one argument that has been used to support the use of MNL models in crowding valuation (Tirachini et al., 2014). Most studies, however, highlight that (unobserved) heterogeneity in crowding and time sensitivities is important to take into account.

Whelan and Crockett (2009)'s ML model assumes a normal distribution to introduce unobserved heterogeneity in user preferences towards crowding levels in trains, and find that around 25% of respondents have 'wrong signed' taste parameters. The authors, however, discard the use of the lognormal distribution as a solution, given that it may shift the mean of the (crowding sensitive) VTTS parameter. The referred study of Basu and Hunt (2012) for crowding valuation in Mumbai, compares MNL and ML models using a triangular distribution for travel time parameters for different crowding levels, as a way to avoid the issue of large spreads in unconstrained distributions. In this study, we acknowledge the limitations of the lognormal density, but prefer its use as the resulting densities for the crowding multipliers are analytically tractable and much better behaved when looking at the median values. Additionally, we contrast the MNL and ML models to a Latent Class (LC) specification. Results show that significant heterogeneity in crowding perception exist across the population, as exposed by estimated ML and LC models. Gender, income and age are significant variables in explaining heterogeneity in crowding disutility. MNL, mean LC and median ML models produce similar sitting and standing crowding multipliers for a given occupancy level, unlike mean ML values which produce crowding multipliers that are unreasonably high.

Finally, an international comparison of crowding multipliers with values found in other cities is performed. We find that the Santiago Metro crowding multipliers are close to those previously found in the Paris Metro system (Kroes et al., 2014) and in Hong Kong's Mass Transit Railway (MTR) network (Hörcher et al., 2017) On a more local level, this is the first article in which the value of sitting and standing are separately estimated in Santiago. The sitting and standing taste parameters can, in turn, be used to estimate the value of having a seat when travelling, through the computation of standing multipliers, as done in Section 6.

The paper is organized as follows: Section 2 describes the survey and its main results, while Section 3 focuses on the analysis of the relationship between crowding, comfort and perceived security.

Section 4 describes the methodology for the estimation of the proposed models. Section 5 shows and discusses results while Section 6 compares Santiago's crowding multipliers with those from other cities and countries. Finally, Section 7 concludes the paper.

2 Data Collection

A survey to measure the relevance of crowding in route choice was designed and executed. In order to simplify the choice task, only metro-based alternatives were considered and fare was excluded as an attribute (because in Santiago, within a time period, the metro fare is fixed regardless of trip distance).

The main survey included seven sections:

1. Background and socio-economic characteristics: e.g. gender, age, income, occupation and access to car.
2. Metro usage: average numer of times the respondent travels by metro each week and characteristics of latest trip (origin, destination, travel time, crowding level).
3. Smartphone availability and use: if the respondent has a smartphone, and if so, what (s)he uses it for while traveling by metro, and how frequently the smartphone is used.
4. Stated choice (SC) component: six binary choice tasks in which the respondent needs to choose between two alternatives for metro trips (see details below).
5. Crowding perception: the respondent is asked about how secure and how comfortable (s)he feels for three different crowding levels (low, medium and high).
6. Crowding description: the respondent is asked which phrase most accurately describes a specific crowding level shown on either a 2D diagram or a photo.
7. Trip perception and time use: the respondent is asked how (s)he feels about particular situations like having to share a reduced space with strangers, if (s)he likes to use a smartphone, read, listen to music, talk to people, etc., while travelling by metro.

In the SC component three attributes were used to characterize an alternative: i) travel time, ii) occupancy level, and iii) whether the passenger has to stand or can sit down during the trip. Travel time is pivoted around the travel time of the respondent's latest metro trip (question set 2). Five attribute levels are specified around this base travel time (-25%, -12.5%, 0, +12.5% and +25%). The crowding attribute was presented by means of six levels. The levels go from 1 (almost empty train) to 6 (completely full train). The way in which the crowding attribute was presented to respondents varied across versions of the survey. We used three alternative representation formats: i) text, ii) 2D diagrams (bird's-eye view), and iii) photos taken inside a metro car (edited with a photo edition software, if necessary, to match with the exact number of persons required for a particular passenger density level). In Figure 1 we show an example of a choice task, as shown to respondents, in which train occupancy level is depicted by means of a 2D diagram. The representation of all six occupancy levels and representation formats are shown in Figures 7 and 8 and Table 8 the Appendix. In total, a design of 12 choice tasks was constructed, grouped in two blocks of 6 tasks. Each respondent was presented with a single block of choice tasks and a single representation format.

The survey was programmed on the online survey platform Qualtrics. After a pilot carried out in September 2014, the final survey was conducted in October 2014 by a private consultant. In the pilot, the SC survey was designed using an orthogonal design; whereas for the final survey a

Figure 1: Example of stated choice task

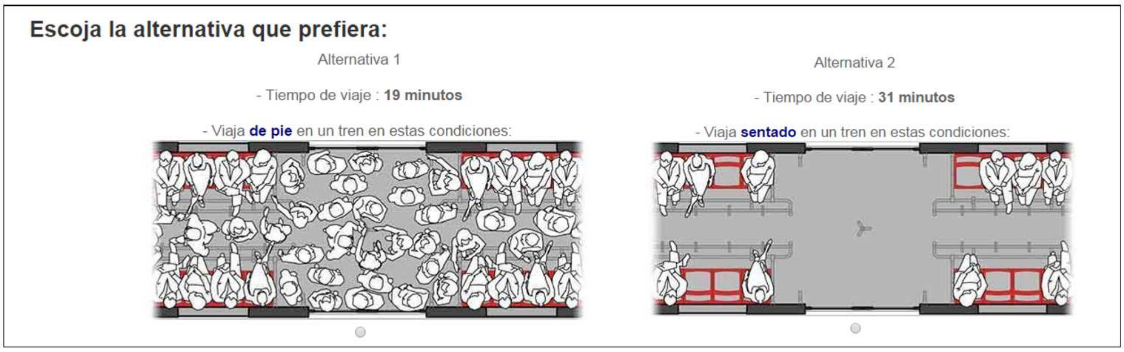


Table 1: Income profile network versus survey

Total Metro			Survey		
Household income (Euro/month)	Percent.	Accum.	Personal income (Euro/month)	Percent.	Accum.
0 - 448	18%	18%	0 - 299	15%	15%
448 - 1,194	46%	65%	299 - 597	20%	35%
1,194 - 2,090	22%	86%	597 - 896	19%	54%
2,090 – 2,836	8%	94%	896 – 1,493	16%	70%
2,836 – 3,731	3%	97%	1,493 – 2,239	11%	81%
3,731 or higher	3%	100%	2,239 or higher	19%	100%

D-efficient design was constructed using the SC experimental design software NGene (Rose et al., 2008). Priors for the parameters were obtained from the pilot study.

Two survey application methods were used: (a) online, in which the survey is distributed by email to a panel of respondents from the consultant, (b) face-to-face, in which surveyors with tablets interview metro users outside selected stations. The total number of correct complete surveys is 413 (210 online surveys, 203 face-to-face surveys). The sampling strategy attempted to resemble the income profile of Santiago´s metro users, as described by a network-wide origin-destination survey performed by the Metro company in 2013. Accordingly, Metro stations with different user income profile were chosen. The percentage of users by income range in both the total network survey and our survey is shown in Table 1.

From Table 1, likely there is a slight over-representation of higher-income users in our sample, as 70% and 81% of our respondents have personal incomes lower than 1,493 and 2,239 Euros, whilst on the network 86% of users report a household income lower than 2,090 Euros. However, there is no indication of large differences in income between the two samples. Regarding gender and age representation, 55% of metro users in Santiago are women (47% female respondents in our survey) and 48% of metro users are 29 years old or younger (30% of respondents in our survey are in that age range). The fact that our survey was applied only to adults partially explains the under-representation of young users in our sample.

3 The relationship between occupancy level, perceived comfort and security

In this section we focus on the relationship between occupancy levels in Metro trains, as shown to survey respondents in section five of the survey, and their perceived level of comfort and security.

Out of the six levels for the crowding attribute (see Appendix) three were shown to the respondents [3]. This was done after the SP part of the survey in order to not influence response patterns. For each of the three levels the following questions were asked:

- How secure do you feel to travel in these conditions? (security with respect to theft, or physical and psychological threat)
- How comfortable do you feel to travel in these conditions?

Respondents had to rate each level of occupancy on a 1 to 7 Likert scale, where 1 meant very insecure (very uncomfortable), and 7 meant very secure (very comfortable). The 1 to 7 scale has the advantage of been highly intuitive in Chile since it is the scale of marks in the Chilean education system (where 7 is the maximum possible mark, 1 is the lowest mark and 4 is the minimum mark to pass). Results of the average score for the six occupancy levels are shown in Fig. 2 where, to ease understanding, all six levels are shown with their respective 2D representation.

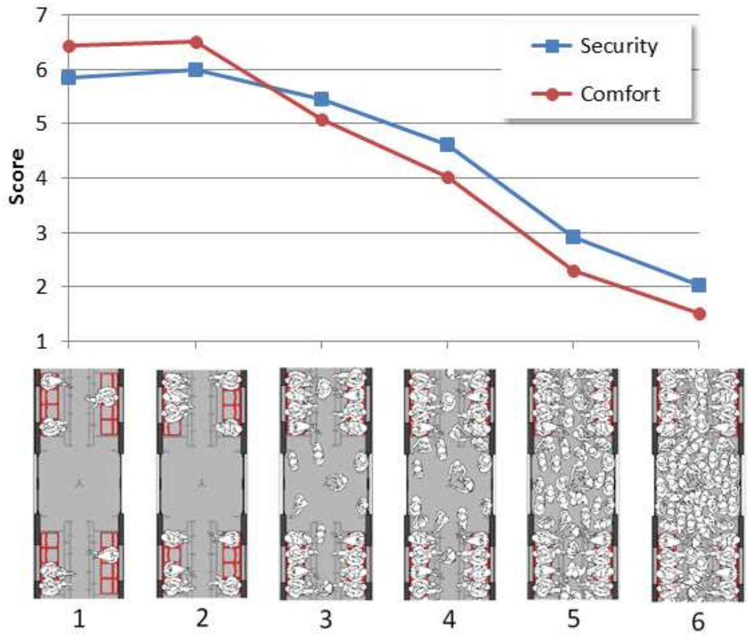
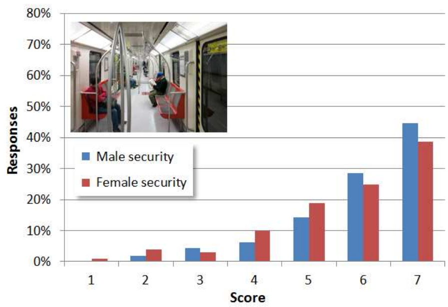


Figure 2: Average security and comfort levels for different occupancy levels

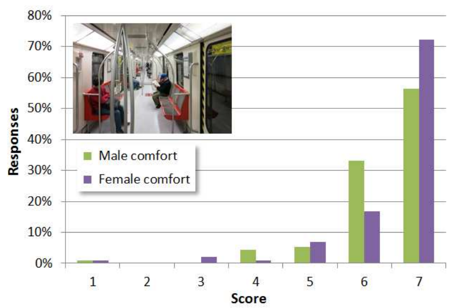
On average, users do not perceive a difference in comfort or security between levels 1 and 2 or occupancy, in which all passengers are sitting, and therefore it can be suggested that the main variable affecting both security and comfort is the presence of standees (in fact, both scores are 0.1 points higher in level 2, but the difference is not statistically significant at the 5% level). Due to the presence of standees the level of comfort drops quicker than the level of safety between levels 2 and 3. From level 3 and above, the perceived security has a higher average mark than perceived comfort. Notably, between levels 4 to 6 perceived comfort and security are dropping at a similar pace.

A more detailed analysis can be presented by moving beyond average scores. To ease understanding, we only present histograms of answers for occupancy levels 1 (the lowest), 3 (medium) and 6 (the highest), for all forms of crowding representation shown to respondents (see Fig. 3). It is interesting to note that there is more variation in the answers to the security question than in the

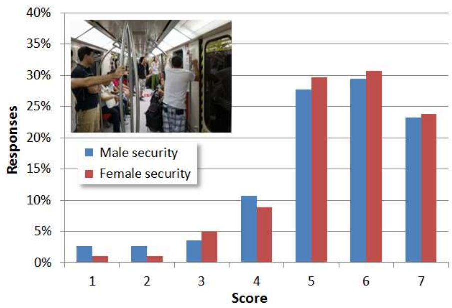
[3] the three levels were randomly chosen between crowding levels 1 and 2, 3 and 4, and 5 and 6



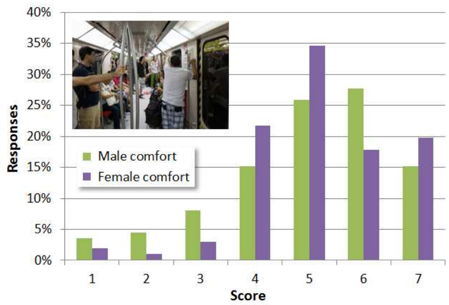
(a) Occupancy level 1, security



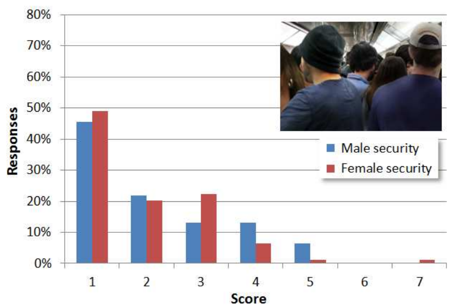
(b) Occupancy level 1, comfort



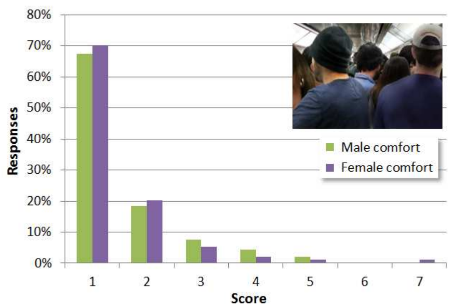
(c) Occupancy level 3, security



(d) Occupancy level 3, comfort



(e) Occupancy level 6, security

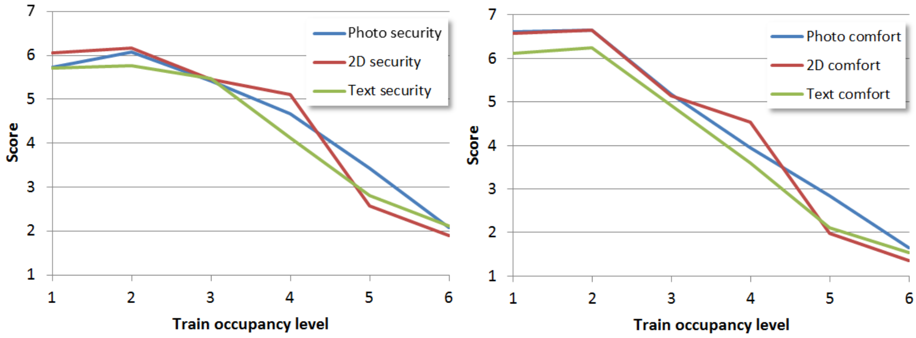


(f) Occupancy level 6, comfort

Figure 3: Perceptions of security and comfort, share of responses per score level for three ocupancy levels

answers to perceived comfort. For instance, respondents clearly relate an almost empty train with a high level of comfort (Fig. 3b), however less than 50% of respondents feel that situation as "very secure" (Score 7 in Fig. 3a). This finding is in line with the hypothesis of Cox et al. (2006), who state that the relationship between security and train occupancy varies by crime type, as muggings are more likely to happen in crowded trains but assaults are more likely to happen in empty trains. A similar outcome is observed with the histograms of occupancy level 6 (Fig. 3e and 3f), which 68-70% of respondent perceive as "very uncomfortable", but less than 50% of respondents perceive it as "very insecure". Therefore, there exists a more straightforward relationship between occupancy and the perception of comfort, than between occupancy and the perception of (in)security. Regarding gender differences, it is observed than men tend to feel more secure but less comfortable in an almost empty train than women (Fig. 3a and 3b), however, when comparing mean scores there are no significant differences for gender.

With respect to differences in perception of security and comfort among the forms of representation for occupancy, Fig. 4 shows average scores for all occupancy levels. No discernible tendency is observed in the perception of security. In the case of comfort perception, it is found that for low and medium occupancy levels the text representation has a lower average score than 2D and



(a) Security (b) Comfort

Figure 4: Average scores security and comfort per form of representation

Photo, which may point towards a misrepresentation of the actual comfort conditions of a text explanation compared against graphical forms.

Overall, we find that feelings of insecurity and discomfort increase with density and number of passengers standing in a metro carriage.

4 Choice modelling: methodology

In this section we introduce the discrete choice models used to estimate crowding multipliers for Santiago's Metro system. Our survey included a binary stated choice (SC) component, in which each choice task presented two alternative metro routes to the respondent, as previously depicted by Fig. 1. The choice between scenarios 1 and 2 in choice task situation $t = 1, \dots, T$ for individual $n = 1, \dots, N$ is modelled using the following random utility maximization (RUM) specification:

$$\begin{aligned} U_{1nt} &= \beta_{TT} TT_{1nt} + \beta_{TTdens}[TT_{1nt} \times dens_{1nt}] + \beta_{TTdensST}[TT_{1nt} \times dens_{1nt} \times 1_{stdg_{1nt}}] + \varepsilon_{1nt} \\ U_{2nt} &= \beta_{TT} TT_{2nt} + \beta_{TTdens}[TT_{2nt} \times dens] + \beta_{TTdensST}[TT_{2nt} \times dens_{2nt} \times 1_{stdg_{2nt}}] + \beta_0 + \varepsilon_{2nt} \end{aligned} \quad (1)$$

where TT_{int} is travel time in alternative i (min), $density_{int}$ is passenger density (pax/m^2) and $1_{stdg_{int}}$ is a binary variable indicating whether the passenger is standing or not. $\boldsymbol{\beta} = (\beta_0, \beta_{TT}, \beta_{TTdens}, \beta_{TTdensST})$ then represents a vector of corresponding preference parameters, and ε_{int} denotes the error term. The latter is assumed to follow a Type-I extreme value distribution such that logit type models can be estimated (Train, 2009) using the well-known MNL choice probabilities:

$$P_n^{MNL} = \prod_{t=1}^{T} \left[\frac{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta})}{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta}) + \exp(\mathbf{x}'_{2nt}\boldsymbol{\beta})} \right]^{y_{1nt}} \left[\frac{\exp(\mathbf{x}'_{2nt}\boldsymbol{\beta})}{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta}) + \exp(\mathbf{x}'_{2nt}\boldsymbol{\beta})} \right]^{y_{2nt}} \quad (2)$$

where $\mathbf{x}_{1nt} = (0, TT_{1nt}, TT_{1nt} \times dens_{1nt}, TT_{1nt} \times dens_{1nt} \times 1_{standing_{1nt}})$, $\mathbf{x}_{2nt} = (1, TT_{2nt}, TT_{2nt} \times dens_{2nt}, TT_{2nt} \times dens_{2nt} \times 1_{standing_{2nt}})$, and where $y_{int} = 1$ if alternative i was chosen in choice situation t.

The above model specification is motivated by previous model specifications used to derive crowding multipliers (e.g. Whelan and Crockett, 2009; Wardman and Whelan, 2011; Tirachini et al., 2013). Passenger load (i.e. density measures) are interacted with travel time to represent a higher dis-utility of crowding for longer trips; and if the passenger is standing then there is empirical evidence that crowding is even more bothersome (Wardman and Whelan, 2011). These hypotheses are in line with our results in Section 3. The crowding multipliers can accordingly be derived as the marginal utility of travel time under crowding conditions over marginal utility of travel time under non-crowded conditions:

$$\begin{aligned} CM^{\text{sitting}} &= \frac{\beta_{TT} + \beta_{TTdens}\text{dens}}{\beta_{TT}} = 1 + \lambda_1 \cdot \text{dens} \\ CM^{\text{standing}} &= \frac{\beta_{TT} + \beta_{TTdens}\text{dens} + \beta_{TTdensST}\text{dens}}{\beta_{TT}} = 1 + (\lambda_1 + \lambda_2) \cdot \text{dens} \end{aligned} \tag{3}$$

β_{TT} is the travel time parameter, whereas β_{TTdens} and $\beta_{TTdensST}$ are the parameters associated with the product between travel time and density for sitting and standing, respectively. Moreover $\lambda_1 = \frac{\beta_{TTdens}}{\beta_{TT}}$ and $\lambda_2 = \frac{\beta_{TTdensST}}{\beta_{TT}}$. Therefore, CM^{sitting} represents the crowding multiplier for a passenger who is seated, and CM^{standing} is the respective multiplier for a standing passenger. Standard errors for the crowding multipliers will be calculated using the Delta method (Daly et al., 2012). We will specifically test for differences in the crowding multipliers across the alternative representation formats of the crowding attribute.

The second specification is the mixed logit model where we assume there is unobserved heterogeneity in β_n across respondents. The heterogeneity is captured by a mixing density of the form $f(\beta_n|\theta)$, where θ represents the hyper parameters characterising the mixing density, such as the mean and standard deviation. As a result the expected choice probability for observing the sequence of choices by individual n is now given by:

$$P_n^{ML} = \int_{\beta_n} \prod_{t=1}^{T} \left[\frac{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta_n})}{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta_n}) + \exp(\mathbf{x}'_{2nt}\boldsymbol{\beta_n})}\right]^{y_{1nt}} \left[\frac{\exp(\mathbf{x}'_{2nt}\boldsymbol{\beta_n})}{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta_n}) + \exp(\mathbf{x}'_{2nt}\boldsymbol{\beta_n})}\right]^{y_{2nt}} f(\beta_n|\theta)d\beta_n \tag{4}$$

We explicitly account for the fact that β_n^{TT},β_n^{TTdens} and $\beta_n^{TTdensST}$ are expected to be negative by specifying a lognormal mixing density. The benefit of using a lognormal density is that the λ parameters in Eq. 3 have finite moments (and therefore also the crowding multipliers) (e.g. Daly et al., 2011) and are analytically tractable. Hence, there is no need for simulation exercises after estimation.

The third specification is derived according to the Latent Class model (Greene and Hensher, 2003). If we assume that the parameters β_n are random with a discrete instead of a continuous heterogeneity distribution, then for class q utility becomes:

$$\begin{aligned} U_{1nt}^{(q)} &= \beta_{TT}^{(q)} TT_{1nt} + \beta_{TTdens}^{(q)}[TT_{1nt} \times dens_{1nt}] + \beta_{TTdensST}^{(q)}[TT_{1nt} \times dens_{1nt} \times 1_{stdg_{1nt}}] + \varepsilon_{1nt}^{(q)} \\ U_{2nt}^{(q)} &= \beta_{TT}^{(q)} TT_{2nt} + \beta_{TTdens}^{(q)}[TT_{2nt} \times dens_{2nt}] + \beta_{TTdensST}^{(q)}[TT_{2nt} \times dens_{2nt} \times 1_{stdg_{2nt}}] + \beta_0^{(q)} + \varepsilon_{2nt}^{(q)} \end{aligned} \tag{5}$$

where $\boldsymbol{\beta} = \boldsymbol{\beta}^{(q)}$ with probability $w_n^{(q)} = \exp(\mathbf{z}'_n\gamma^{(q)})/\sum_{q=1}^{Q}\exp(\mathbf{z}'_n\gamma^{(q)})$, with $\mathbf{z}_n$ denoting sociodemographic characteristics of the individual and where the class-specific constant $\gamma^{(1)} = \mathbf{0}$ is normalised. We assume assignment to class is influenced by gender, age and income.

$$\begin{aligned} w_n^1 &= \frac{1}{1 + \exp(\gamma^{(2)} + \gamma_{\text{male}} 1_{\text{male}_n} + \gamma_{\text{age}}\text{age}_n + \gamma_{\text{inc}}\text{inc}_n} \\ w_n^2 &= \frac{\exp(\gamma^{(2)} + \gamma_{\text{male}} 1_{\text{male}_n} + \gamma_{\text{age}}\text{age}_n + \gamma_{\text{inc}}\text{inc}_n)}{1 + \exp(\gamma^{(2)} + \gamma_{\text{male}} 1_{\text{male}_n} + \gamma_{\text{age}}\text{age}_n + \gamma_{\text{inc}}\text{inc}_n)} \end{aligned} \tag{6}$$

where age_n, inc_n and 1_{male_n} stand for age in years, personal income range and whether individual n is a male, respectively.

Table 2: Basic MNL

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)	
intercept $(i = 2)$	0.130	0.042	3.136	0.002	**
TT	-0.101	0.010	-10.306	< 2.2e-16	***
TTdens	-0.010	0.001	-10.086	< 2.2e-16	***
TTdensST	-0.007	0.001	-6.950	0.000	***
Log-Likelihood:	-1628.8				
McFadden R^2:	0.043777				

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

In this latent class model, the probability of the observed sequence of choices for an individual is given by:

$$P_n^{LC} = \sum_{q=1}^{2} \left\{ w_n^{(q)} \prod_{t=1}^{T} \left[\frac{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta}^{(q)})}{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta}^{(q)}) + \exp(\mathbf{x}'_{2nt}\boldsymbol{\beta}^{(q)})} \right]^{y_{1nt}} \left[\frac{\exp(\mathbf{x}'_{2nt}\boldsymbol{\beta}^{(q)})}{\exp(\mathbf{x}'_{1nt}\boldsymbol{\beta}^{(q)}) + \exp(\mathbf{x}'_{2nt}\boldsymbol{\beta}^{(q)})} \right]^{y_{2nt}} \right\}. \tag{7}$$

The maximum likelihood estimator of the full vector of parameters θ can be derived by plugging the correct P_n into $\arg\max_{\boldsymbol{\theta}} \ell(\mathbf{y}|\mathbf{X};\boldsymbol{\theta}) = \sum_{n=1}^{N} \ln(P_n(\boldsymbol{\theta}))$. In the case of the Mixed Logit model, the likelihood needs to be simulated by considering a Monte Carlo approximation of P_n for which we use 1,500 halton draws.

5 Choice modelling: results

5.1 Estimation results

Results for the basic multinomial logit model (MNL) are presented in Table 2.[4] All parameter estimates are significant and have the expected sign. The intercept for alternative 2 is significant and indicating a potential bias towards choosing the alternative presented on the right hand side. Left-right bias is not uncommon in the stated choice literature. Such effects, however, often become less pronounced when moving towards more sophisticated model structures. Note that we also tested whether there was a penalty for standing during the length of the trip irrespective of the occupancy level, but this parameter turned out to be insignificant and was therefore not presented.

The second MNL model (Table 3) examines the impact of the crowding representation format on occupancy perceptions and, accordingly, behavioural responses. During the analysis, the 2D diagrams were considered as the referential crowding representation format. Table 3 reveals that perception bias is not present in our dataset. On the one hand, this is reassuring as the alternative representation formats were carefully developed. On the other hand, this is a remarkable result considering the amount of cognitive effort required from the respondent when being presented with a text description of crowding levels (see Table 8 in the Appendix). Since the representation format has no impact on the model results, the respective control variables are excluded in the remaining analyses.

Results for the ML and LC model are presented in Tables 4 and 5. Both models reveal a significant improvement in model fit over the MNL base model, highlighting there is substantial heterogeneity in sensitivities to travel time and crowding levels across respondents. In the ML model there is still a tendency to prefer the right alternative, but this effect is no longer significant in the LC

[4] All models are estimated using the R package *gmnl* (Sarrias and Daziano, 2015)

Table 3: Basic MNL accounting for type of crowding representation

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)	
2:(intercept)	0.131	0.042	3.137	0.002	**
TT	-0.101	0.010	-10.303	$<$ 2.2e-16	***
TTdens	-0.011	0.001	-8.107	0.000	***
TTdensST	-0.006	0.002	-4.013	0.000	***
TTdens (photo)	0.002	0.001	1.124	0.261	
TTdens (text)	0.000	0.002	0.240	0.810	
TTdensST (photo)	-0.001	0.002	-0.350	0.727	
TTdensST (text)	-0.002	0.002	-0.835	0.404	
Log-Likelihood:	-1627.6				

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 4: ML model using lognormal mixing densities

Coefficients	Estimate	t-stat	p-value	Sign.
Intercept ($i = 2$)	0.175	2.963	0.003	**
TT - μ	-1.596	-12.753	0.000	***
TTdens - μ	-3.791	-27.105	0.000	***
TTdensST - μ	-4.561	-18.583	0.000	***
TT - σ	1.012	7.931	0.000	***
TTdens - σ	1.498	11.078	0.000	***
TTdensST - σ	1.813	10.415	0.000	***
Log-Likelihood:	-1404.9			
obs	2467			
n	413			
draws	1500			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

model. After transformation of the lognormal parameters, the median (and mean) travel time and crowding level sensitivities are all higher compared to the basic MNL model, but these values are surrounded by significant heterogeneity. We provide a more detailed discussion when looking at the crowding multipliers, which are the model outcomes of interest. The LC model indicates that an individual is more likely to belong to Class 2 if (s)he is male, young and has higher income.[5] Travellers belonging to Class 2 are very sensitive to travel time, but much less sensitive to crowding levels than members of Class 1. This is a reasonable result regarding the role of age and gender in the class membership equation, but not necessarily regarding income as wealthier passengers might be more negatively affected by a large passenger density than lower income travellers, as found by Haywood et al. (2017) in Paris. We explain the observed effect as a result of higher income people being more adverse to long travel times, i.e. having higher values of time. This is, however, an inconclusive interpretation given that in our survey the trade-off between travel time, level of train occupancy and trip fare was not present, as fare was not an attribute in the SP experiment. We now turn to the crowding multipliers derived from the above models.

[5] We experimented with models having more than two classes and allowing for unobserved preference heterogeneity within classes. However, this respectively resulted in counter-intuitive parameter estimates and signs of model over-specification. Also ML and LC models in 'time space', i.e. directly estimating the crowding multipliers, were estimated. These did not offer additional insights

Table 5: Latent Class Model

Coefficients	Estimate	t-stat	p-value	Sign.
Class 1: Intercept ($i = 2$)	0.115	1.472	0.141	
Class 1: TT	-0.090	-4.127	0.000	***
Class 1: TTdens	-0.029	-11.563	0.000	***
Class 1: TTdensST	-0.021	-6.511	0.000	***
Class 2: Intercept ($i = 2$)	0.173	1.898	0.058	.
Class 2: TT	-0.223	-13.258	0.000	***
Class 2: TTdens	-0.005	-3.539	0.000	***
Class 2: TTdensST	-0.010	-6.964	0.000	***
Class Membership (class 2)				
Intercept	0.268	1.868	0.062	.
Gender	0.394	3.877	0.000	***
Age	-0.022	-6.555	0.000	***
Income	0.446	4.192	0.000	***
Log-Likelihood:	-1456.8			
obs	2467			
n	413			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.2 Crowding Multipliers

Table 6 presents the crowding multipliers when metro users are able to sit whilst travelling. When there are no other passengers standing, i.e. $pax/m2 = 0$, the 'regular' value of travel time savings applies, irrespective of the preferred model specification. As expected, the crowding multipliers are increasing with passenger density, showing that increased crowding levels increase the disutility of travel time. Metro users are therefore willing to accept longer travel times in return for less crowded conditions. Subsequently assuming metro users are also willing to pay for reductions in travel time, allows us to infer they are willing to pay more for reductions in travel time under crowded conditions. This willingness-to-pay increases with crowding density.

The multiplicative relation between λ and density in equation (3), however, also causes the standard error of the crowding multiplier to go up with density ($pax/m2$). This is consistent across the three model specifications. Standard errors increase further when moving from the MNL model to the more complex ML and LC models. The latter increase in the standard error is caused by introducing a more flexible model specification. Standard errors are notably higher for the mean crowding multipliers of the ML model and for Class 1 of the LC model than for the Median ML model and Class 2 of the LC model. In the ML model, the fat upper tail of the lognormal distribution causes both the mean and the standard error of the crowding multipliers to go up. People with a high crowding sensitivity have less of an impact on the median crowding multiplier. In ML models it is not uncommon to find that the median of the mixing density, or its WTP-like transformation, is most comparable to the MNL estimates. This is a direct result of the density's tails having a smaller impact on the median than on the mean (e.g. Borjesson et al., 2012). The tail of the distribution also has an impact on the Class 1 crowding multipliers of the LC model, but this effect is less pronounced due to the estimation of only a discrete number of classes rather than a continuous distribution as done by the ML model.

For the ML model, the median crowding multipliers closely correspond to those for the MNL model. As discussed above, the fat-tail of the lognormal density spurs the mean of the ML crowding multipliers up to an unreasonably high level relative to the values usually found in the extant

Table 6: Crowding multipliers: Sitting conditions

	MNL		ML				LC			
	Mean		Mean		Median		Class 1		Class 2	
pax/m2	Est.	St. Err	Est.	St. Err	Est.	St. Err	Est.	St. Err	Est.	St. Err
0	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
1	1.10	0.01	1.57	0.16	1.11	0.01	1.33	0.06	1.02	0.01
2	1.20	0.01	2.14	0.32	1.22	0.03	1.65	0.12	1.04	0.01
3	1.30	0.02	2.71	0.49	1.33	0.04	1.98	0.18	1.06	0.02
4	1.40	0.03	3.28	0.65	1.45	0.06	2.30	0.24	1.08	0.02
5	1.50	0.03	3.85	0.81	1.56	0.07	2.63	0.29	1.10	0.03
6	1.60	0.04	4.42	0.97	1.67	0.08	2.95	0.35	1.13	0.03

literature (sitting multipliers not larger than 2, and standing standing usually not larger than 3, even under very crowded conditions). On the other hand, median estimations (up to 1.7 for sitting and 2.0 for standing) are very similar to those of the MNL model and within the range of values found in e.g., Great Britain (Wardman and Whelan, 2011). Regarding our mean ML values, it is often not recommended to use such high values for policy evaluations and in many national value of time savings studies (e.g. Borjesson et al., 2012) censoring approaches are applied accordingly. The LC model, however, provides a more reasonable alternative where part of the sample has a high crowding multiplier, which is somewhat tempered by a second latent class of travellers experiencing only a limited disutility of crowding.

A very similar story emerges from Table 7. The crowding multipliers of MNL model are highly comparable to median ML value. The mean ML crowding multipliers and associated standard errors are again unreasonably high for which the LC model provides a more acceptable alternative.

In the Latent Class model we observe quite different crowding multipliers when comparing Classes 1 and 2, as shown in Tables 6 and 7: Class 1 (more likely higher income younger males) has very large crowding multipliers with mean values 2.95 for sitting and 4.33 for standing with 6 pax/m2, whilst Class 2 have lower multipliers of 1.13 and 1.39 for the same density of standees. When computing average multipliers for both classes combined, taking into account the probability of class membership for all respondent in the sample, we obtain an average multiplier that go up to 2.1 for sitting and 3.0 for standing. These values are larger than the crowding multipliers implied by the MNL and (median) ML values, as shown in Fig. 5, and also seem to be too large when compared to most of the existent international literature. We conclude that even though there is quite a substantial amount of heterogeneity in users aversion to crowding, a good indication of crowding multipliers for the population would be values up to 1.5-1.6 for sitting, and up to 2.0-2.3 for standing, for a density of standees of 6 pax/m2.

The actual levels of the crowding multipliers will be contrasted against other national and international measures in Section 6.

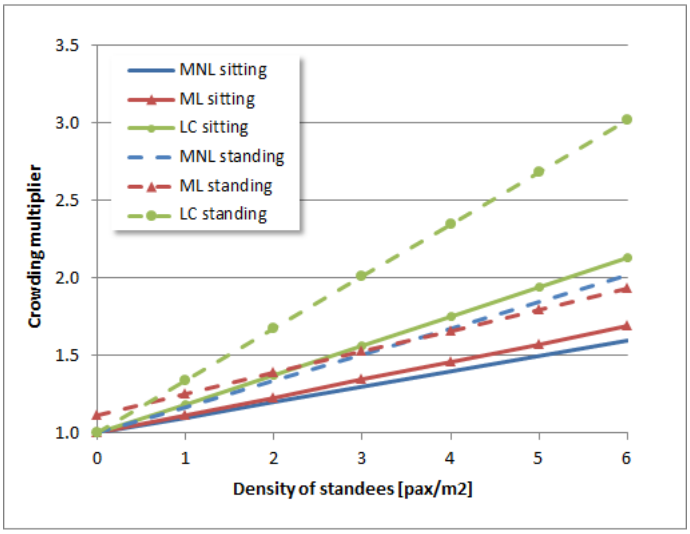
6 International Comparisons

We compare our median ML multipliers with those of London and South East (SE) England (Whelan and Crockett, 2009), the Paris region (Kroes et al., 2014), Singapore (Tirachini et al., 2016), Hong Kong (Hörcher et al., 2017) and Swedish cities (Bjorklund and Swardh, 2015). Crowding multipliers for sitting and standing are shown in Fig. 6a. Sitting multipliers in Santiago are almost equal to those recently estimated in Hong Kong and not far from those in Paris and London SE, whereas Sweden has lower sitting multipliers (up to 1.15 for 4 pax/m2). For standing, the estimated multipliers in Santiago are similar to those in Paris, slightly lower to those in Hong Kong and

Table 7: Crowding multipliers: Standing conditions

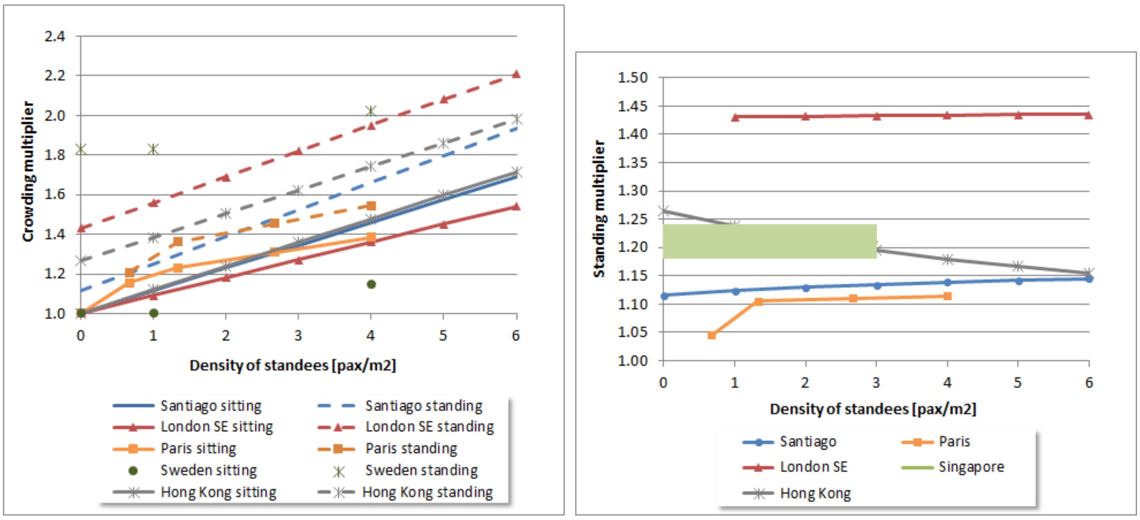
	MNL		ML				LC			
	Mean		Mean		Median		Class 1		Class 2	
pax/m2	Est.	St. Err	Est.	St. Err	Est.	St. Err	Est.	St. Err	Est.	St. Err
0	1.00	NA	1.00	NA	1.00	NA	1.00	NA	1.00	NA
1	1.17	0.01	2.02	0.29	1.16	0.02	1.56	0.09	1.07	0.01
2	1.33	0.02	3.03	0.59	1.33	0.04	2.11	0.18	1.13	0.01
3	1.50	0.03	4.05	0.88	1.49	0.06	2.67	0.27	1.20	0.02
4	1.67	0.03	5.06	1.17	1.65	0.08	3.22	0.36	1.26	0.03
5	1.84	0.04	6.08	1.47	1.81	0.10	3.78	0.45	1.33	0.03
6	2.00	0.05	7.10	1.76	1.98	0.11	4.33	0.54	1.39	0.04

Figure 5: Comparison of implied crowding multipliers: MNL, ML, LC models



clearly lower to those estimated in Sweden and London SE.

On the other hand, Fig. 6b depicts the value of having a seat, that is the ratio between the standing and sitting multipliers. We find the Santiago values closer to those of the Paris Metro system for sitting and standing. The value of having a seat in Santiago and Paris are for the most part between 1.10 and 1.15, which means that travel time is valued between 10 and 15 percent more when standing than when sitting. The value of having a seat in Hong Kong is estimated between 1.15 and 1.27, and is a decreasing function of the density by construction of the model (Hörcher et al., 2017). The London SE value of a seat is much higher at 1.44, which is possibly explained by a longer trip distance in the British study (it includes interurban travel) and having trains with more seats. As shown in the diagrams, in Santiago metro trains have very few seats and the probability of getting a seat is close to zero in peak hours (except for users that board trains at the first station of a line), and therefore people may not give a great value to having a seat since they are used to stand. The value of having a seat in Singapore's MRT was estimated between 1.18 and 1.24, a value that lies between those in Santiago and London. Therefore, we conclude that with evidence from four urban heavy rail systems, value of travel time savings when travelling standing should be around 1.10-1.26 larger than the value of travel time savings when sitting, a value that likely increases for suburban or interurban longer trips.



(a) Crowding multipliers

(b) Standing multipliers

Figure 6: International comparison for crowding and standing multipliers. Own elaboration based on Whelan and Crockett (2009), Kroes et al. (2014), Tirachini et al. (2016), Bjorklund and Swardh (2015) and Hörcher et al. (2017)

7 Conclusions

Mode choice models where crowding is one of the main explanatory variables were estimated. A basic Multinomial Logit (MNL) model, a Latent Class (LC) model and a Mixed Logit (ML) model were estimated and crowding multipliers were computed for each of them. Additionally, the relevance of the type of representation of the crowding level was tested, showing it has no significant effect.

Results show that crowding is relevant to explain user behaviour in Santiago, and that different travel time multipliers for sitting and standing could be estimated. The quantification of the crowding effect and the value of having a seat has the potential to influence project appraisal, allowing to consider different benefits for users under different crowding conditions. This would have been of use, for example, in the public transport design model for Santiago, where it was assumed that, while travelling, one minute is worth the same regardless of crowding conditions in trains or buses. Our results can be used to estimate the value of increasing service frequency, increasing train size or increasing the number of seats as measures to improve the service quality. We found that the sitting multiplier is up to 1.5-1.6 for a density of standees of 6 pax/m2, whereas the standing multiplier goes up to a value between 1.9 and 2.2 for the same density. The MNL and the median ML were not far from each other, which in the case of Santiago allows us to infer that for policy evaluation the use of crowding multipliers from a simple MNL model is enough to model the crowding sensitivity of the population as a whole. However, significant heterogeneity is present in our sample, which could be picked up by both ML and LC models. We used a latent class model to differentiate between groups of users that have different preferences. The group with low crowding sensitivity is more likely to be populated by younger people, males and users with higher income, whereas the group that is more sensitive to crowding is more likely to have females, older people and lower income travellers.

Regarding policy implications, the estimated crowding multipliers should be tried in the evaluation of changes to the existing metro network and service in, for example, the number of seats per train or increasing/reducing the service frequency in peak and off-peak periods (as analysed by Tirachini et al. (2014) for buses and de Palma et al. (2015) for trains). Without a crowding disutility, increasing train frecuency only has a value on reducing waiting time. The approach presented here can be used to estimate the effect of that intervention on the comfort of travel time, for a real metro line in Santiago. The bias that arises when ignoring crowding for the estimation of public transport demand has been analytically and numerically assessed by Tirachini et al. (2014) and Batarce et al. (2016); this issue must be taken into account in the economic assessment of public transport projects as already done by a few countries, such as Sweden, France, England and Australia (for a review see OECD/ITF (2014)).

Finally, when comparing the results obtained in this article with the extant literature, it is interesting to analyse the similarities of the Santiago results in particular to those of Paris and Hong Kong, taking into account the fact that the research methods used by the authors and the contexts are different: in Santiago and Paris, stated preferences have been used while in Hong Kong revealed preferences have been inferred using large automatic fare collection (AFC) and automatic vehicle location (AVL) databases. Importantly, we cannot confirm that crowding multipliers obtained from stated preferences might be larger than those from revealed preferences, as suggested by Kroes et al. (2014) and Hörcher et al. (2017), because we found mixed results when comparing different cities and research methods. The advent of large AFC and AVL databases for the estimation of crowding and standing externalities (as recently advanced by Tirachini et al. (2016) and Hörcher et al. (2017), with the implementation of route choice methods) paves the way for the extended use of revealed preferences for the economic analysis of crowding discomfort and other quality-of-service attributes in the near future. It is expected that as more RP-based results arise, a clearer picture of potential stated preferences biases will be obtained.

Acknowledgments

This research was supported by CONICYT Chile: the FONDECYT project "Social effects and Quality of Service Valuation in Public Transport Services" (Grant 11130227, P.I. Alejandro Tirachini) funded the field application of the survey and the visit of Thijs Dekker to Universidad de Chile in September 2014; the FONDECYT project "Understanding perception of qualitative attributes in urban space and its effects on user behavior: a mathematical modeling framework" (Grant 11130637, P.I. Ricardo Hurtubia) funded the visit of Ricardo A. Daziano to Universidad Católica in March 2015. The authors also acknowledge support from the Complex Engineering

Systems Institute (ICM: P05-004-F, CONICYT: FB0816), the Center for Sustainable Urban Development (CEDEUS, CONICYT/FONDAP 15110020) and the University Transportation Research Center - Region 2 (UTRC 49198-29-28). Research assistance by Felipe Hernández in developing the 2D and photo train occupancy representations and helping setting-up the online survey is greatly appreciated.

References

Arentze, T., Borgers, A., Timmermans, H., and DelMistro, R. (2003). Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transportation Research Part E: Logistics and Transportation Review*, 39(3):229–244.

Basu, D. and Hunt, J. D. (2012). Valuing of attributes influencing the attractiveness of suburban train service in mumbai city: A stated preference approach. *Transportation Research Part A: Policy and Practice*, 46:1465–1476.

Batarce, M., Munoz, J. C., de Dios Ortuzar, J., Raveau, S., Mojica, C., and Rios, R. A. (2015). Use of mixed stated and revealed preference data for crowding valuation on public transport in santiago, chile. *Transportation Research Record: Journal of the Transportation Research Board*, 2535:73–78.

Batarce, M., Muñoz, J. C., and de Dios Ortúzar, J. (2016). Valuing crowding in public transport: Implications for cost-benefit analysis. *Transportation Research Part A: Policy and Practice*, 91:358 – 378.

Bjorklund, G. and Swardh, J. (2015). Valuing in-vehicle comfort and crowding reduction in public transport. *2015 hEART Conference, Copenhagen, September.*

Borjesson, M., Fosgerau, M., and Algers, S. (2012). Catching the tail: Empirical identification of the distribution of the value of travel time. *Transportation Research Part A: Policy and Practice*, 46(2):378 – 391.

Cheng, Y.-H. (2010). Exploring passenger anxiety associated with train travel. *Transportation*, 37(6):875–896.

Cox, T., Houdmont, J., and Griffiths, A. (2006). Rail passenger crowding, stress, health and safety in britain. *Transportation Research Part A: Policy and Practice*, 40(3):244–258.

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

Daly, A., Hess, S., and Train, K. (2011). Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, 39(1):19–31.

de Palma, A., Kilani, M., and Proost, S. (2015). Discomfort in mass transit and its implication for scheduling and pricing. *Transportation Research Part B: Methodological*, 71:1 – 18.

Fernandez, J. E., de Cea, J., and Malbran, R. H. (2008). Demand responsive urban public transport system design: Methodology and application. *Transportation Research Part A*, 42(7):951–972.

Gómez-Lobo, A. (2012). The ups and downs of a public transport reform: the case of transantiago. *Serie documentos de trabajo SDT354, Universidad de Chile, Departamento de Economía, Santiago, Chile.*

Greene, W. H. and Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8):681–698.

Haywood, L. and Koning, M. (2015). The distribution of crowding cost in public transport: new evidence from paris. *Transportation Research Part A: Policy and Practice*, 77:182–201.

Haywood, L., Koning, M., and Monchambert, G. (2017). Crowding in public transport: Who cares and why? *Transportation Research Part A: Policy and Practice*, 100:215 – 227.

Hensher, D., Rose, J., and Collins, A. (2011). Identifying commuter preferences for existing modes and a proposed metro in sydney, australia with special reference to crowding. *Public Transport*, 3(2):109–147.

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

Hurtubia, R., Guevara, A., and Donoso, P. (2015). Using images to measure qualitative attributes of public spaces through sp surveys. *Transportation Research Procedia*, 11:460–474.

Kroes, E., Kouwenhoven, M., Debrincat, L., and Pauget, N. (2014). Value of crowding on public transport in ï¿œle-de-france, france. *Transportation Research Record*, 2417:37–45.

Lam, W. H., Cheung, C.-Y., and Lam, C. (1999). A study of crowding effects at the hong kong light rail transit stations. *Transportation Research Part A: Policy and Practice*, 33(5):401 – 415.

Legrain, A., Eluru, N., and El-Geneidy, A. M. (2015). Am stressed, must travel: The relationship between mode choice and commuting stress. *Transportation Research Part F: Traffic Psychology and Behaviour*, 34:141 – 151.

Mahudin, N. D. M., Cox, T., and Griffiths, A. (2012). Measuring rail passenger crowding: Scale development and psychometric properties. *Transportation research part F: traffic psychology and behaviour*, 15(1):38–51.

Motoaki, Y. and Daziano, R. A. (2015). A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. *Transportation Research Part A: Policy and Practice*, 75:217–230.

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

Muñoz, J. C., Batarce, M., and Hidalgo, D. (2014). Transantiago, five years after its launch. *Research in Transportation Economics*, 48:184–193.

OECD/ITF (2014). Valuing convenience in public transport. *ITF Round Tables, No. 156. OECD Publishing, France.*

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

Rizzi, L. I., Limonado, J. P., and Steimetz, S. S. (2012). The impact of traffic images on travel time valuation in stated-preference choice experiments. *Transportmetrica*, 8(6):427–442.

Rose, J. M., Bliemer, M. C., Hensher, D. A., and Collins, A. T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological*, 42(4):395–406.

Sarrias, M. and Daziano, R. (2015). *gmnl: Multinomial Logit Models with Random Parameters*. R package version 1.0.

Tirachini, A., Hensher, D. A., and Rose, J. M. (2013). Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. *Transportation Research Part A: Policy and Practice*, 53(0):36 – 52.

Tirachini, A., Hensher, D. A., and Rose, J. M. (2014). Multimodal pricing and optimal design of urban public transport: The interplay between traffic congestion and bus crowding. *Transportation Research Part B: Methodological*, 61(0):33 – 54.

Tirachini, A., Sun, L., Erath, A., and Chakirov, A. (2016). Valuation of sitting and standing in metro trains using revealed preference. *Transport Policy*, 47:94–104.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Vedel, S. E., Jacobsen, J. B., and Skov-Petersen, H. (2017). Bicyclists' preferences for route characteristics and crowding in copenhagen: A choice experiment study of commuters. *Transportation Research Part A: Policy and Practice*, 100:53 – 64.

Vovsha, P., Simas Olivera, M., Davidson, W., Chu, C., Farley, R., Mitchell, M., and Vyas, G. (2013). Statistical analysis of transit user preferences including in-vehicle crowding and service reliability. *92nd Annual Meeting of the Transportation Research Board (TRB)*.

Wardman, M. and Whelan, G. (2011). Twenty years of rail crowding valuation studies: Evidence and lessons from british experience. *Transport Reviews*, 31(3):379–398.

Whelan, G. and Crockett, J. (2009). An investigation of the willingness to pay to reduce rail overcrowding. In *Proceedings of the first International Choice Modelling Conference*, Harrogate, England.

Yanez, M. F., Mansilla, P., and Ortuzar, J. d. D. (2010). The santiago panel: measuring the effects of implementing transantiago. *Transportation*, 37(1):125–149.

Appendix: Crowding representations in the SC experiments

Three different types of representation of the crowding level were used in the SC experiments: 2d diagrams, photos and text descriptions. Because it offers the possibility of depicting standing passenger density in a very accurate way, the 2D diagram was built as the referential way to represent crowding. Figure 7 shows the 6 crowding levels and their corresponding representation with 2D diagrams while Figure 8 shows the corresponding photos used for each level. Table 8 shows the text used to represent each of level.

Figure 7: Crowding levels using 2D diagrams

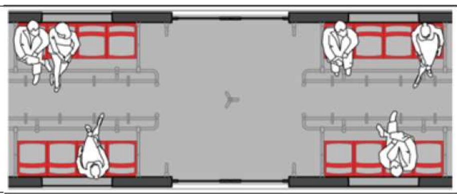
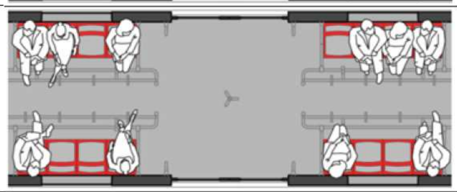
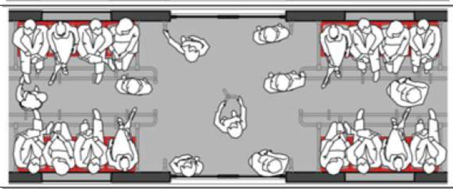
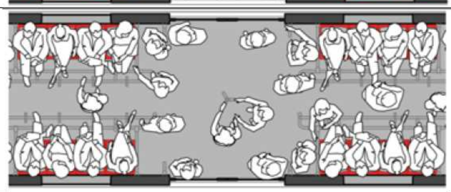
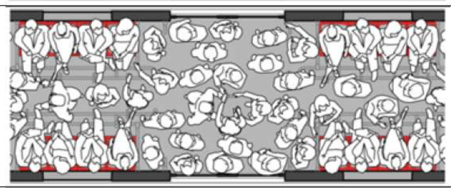
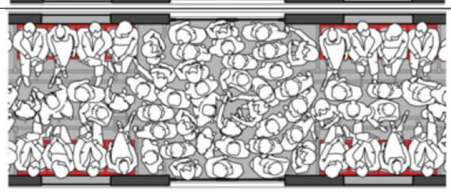
Crowding level	Diagram (shown to respondents)	Description (not shown to respondents)
1		35% seats occupied, 0 standees
2		69% seats occupied, 0 standees
3		100% seats occupied, 1 pax/m2 standing
4		100% seats occupied, 2 pax/m2 standing
5		100% seats occupied, 4 pax/m2 standing
6		100% seats occupied, 6 pax/m2 standing

Figure 8: Crowding levels using photos

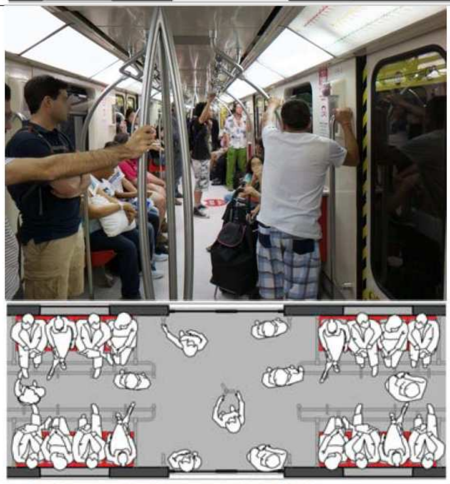
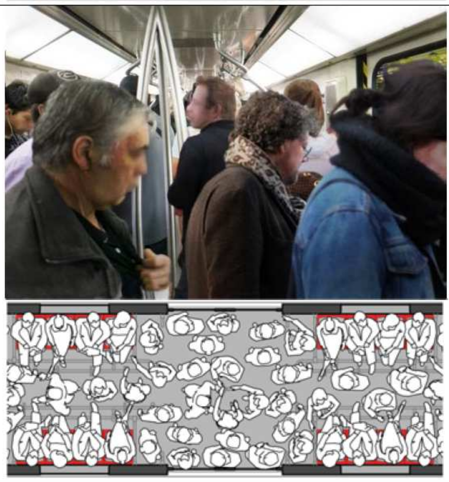
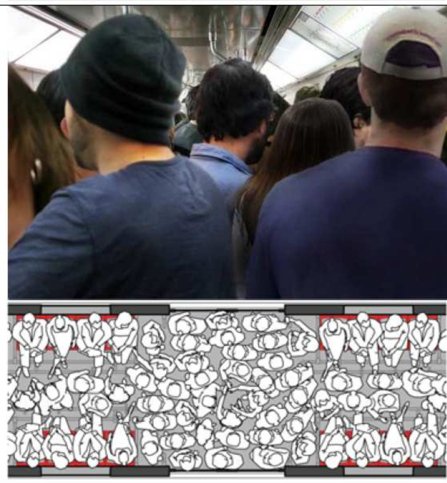
level	Diagram and photo representation	level	Diagram and photo representation
1		2	
3		4	
5		6	

Table 8: Crowding levels using text

level	Description
1	Less than half of seats are occupied. No one is standing.
2	More than half of seats are occupied. No one is standing.
3	All seats are occupied. Few people standing, there is no difficulty moving.
4	All seats are occupied. People standing, minor difficulty moving.
5	All seats are occupied. Many people standing, it is difficult to move.
6	All seats are occupied. Maximum number of people standing, maximum difficulty to move.