CROWDING VALUATION IN PUBLIC TRANSPORT BASED ON STATED PREFERENCE DATA FROM SANTIAGO DE CHILE

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

M.S.

by

Yiming Gong

August 2021

ABSTRACT

Based on stated preference (SP) survey data from Santiago metro system, this thesis analyzes the influence of congestion on passengers' perceived travel time. The survey data includes 413 respondents' socioeconomic characteristics, as well as their decisions in each choice scenario (2,461 valid in total), where both options may have a varied journey duration, crowding level, and whether or not a seat is available. Discrete choice models such as multinomial logit (MNL), mixed multinomial logit (MMNL), and latent class logit (LCL) are estimated in both preference space and crowding multiplier space (CM-space). The findings reveal that passengers are willing to accept longer travel times in exchange for less crowded conditions and there is significant heterogeneity in the amount of willingness throughout the population. Specifically, crowding is viewed as a negative factor by metro users, represented as a perceptional weight on travel time that may reach a value of 1.5-1.7 for sitting and 1.9-2.2 for standing at a density of 6 standees per square meter. Moreover, crowding levels in the SP survey are represented using three distinct forms (text, 2D diagram, or photo), but no significant impact of the varied forms of representation on crowding perception is discovered. The results also show that utilizing a smartphone has a significant impact on decreasing passengers' perceived congestion levels when they have to stand throughout the journey. There is also an international comparison with prior research from Seoul, New York, Sweden, Île-de-France, London and the South East of England, and Netherlands.

Keywords: Choice modeling, Discrete choice model, Random parameters, Parametric heterogeneity, Crowding multiplier, Standing multiplier

BIOGRAPHICAL SKETCH

Yiming Gong is in his second year of studies in Civil and Environmental Engineering at Cornell University. He or she will get his Master of Science in Transportation Systems Engineering degree in August 2021.

Yiming graduated in 2019 from Central South University in Changsha, China. Yiming opted to pursue his master's degree in Transportation Systems Engineering at Cornell University because he or she had a great interest in investigating intelligent transportation challenges and an urgent desire to make existing transportation systems more efficient.

Yiming spent his memorable years at Cornell University learning various skills that can be applied to solving intelligent transportation problems from lectures provided by a variety of programs and departments such as Civil & Environmental Engineering, Operation Research & Information Science, Computer Science, and System Engineering.

Aside from academics, sports like snowboarding bring excitement to Yiming's stay in Ithaca. While traveling through the United States, he or she is immensely inspired by the beauty of nature and the welcoming society.

To my beloved parents.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr Ricardo A. Daziano of the Department of Civil & Environmental Engineering for his assistance and advice as my committee chair. His accomplishments in the field of engineering decision making research inspired me to go into the unknown and pursue my interests throughout my academic education. Dr Daziano assisted me whenever I needed it and provided valuable feedback on my thesis. Without his assistance, I would not have been able to effectively implement the choice models in my thesis.

As the committee minor member, I would appreciate the assistance of Prof. H. Oliver Gao of the Department of Civil & Environmental Engineering. His energizing demeanor inspired me to persevere despite the rigors of this research.

I would also want to express my gratitude to the lecturers who have been my mentors through my two-year graduate program. I am obliged to Prof. Linda K. Nozick for giving the opportunity to work as an assistant, to Prof. David P. Williamson, Drs Samitha Samaranayake, Faisal Alkaabneh, and Francis Vanek for constantly encouraging me to think critically and inspiring me to pursue my interests further after graduation.

Data collection was supported by the CONICYT Chile grant "FONDECYT: Social effects and Quality of Service Valuation in Public Transport Services" (Grant 11130227, P.I. Alejandro Tirachini).

Finally, I want to convey my heartfelt appreciation to my parents and friends for always being there for me whenever I needed them. I would not be who I am now if it weren't for them.

It is because of all of you that I am pleased to be a Cornellian. Thank you very much.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Literature Review

The complicated character of the crowding issue drew considerable attention in the area of psychology as early as the 1970s [24][34]. Daniel Stokols [48], a psychologist, believes, the word ”crowding” includes both objective and subjective components. The number of individuals per unit of space is used to quantify objective crowding, while subjective crowding refers to a person’s felt state of mind that may develop when there is a large discrepancy between anticipated and actual interpersonal distance [6]. As a result, when physical space becomes too restricted, one of the most important aspects of subjective crowding is the “felt loss of behavioral flexibility and privacy” [44]. Furthermore, crowding externalities have a detrimental impact on public transportation performance. Crowding, for example, may result in sluggish boarding, which increases travel time and decreases travel time reliability, or security problems such as trampling, which endangers passengers’ safety [49]. As a result, Strategic consulting company Sinclair Knight Merz (SKM) [47] has identified four major crowding effects: in-vehicle crowding, platform congestion, excessive waiting time, and extended dwell time. The primary concern of this thesis, however, is in-vehicle crowding impact. [1]

In-vehicle crowding is a situation in which a large number of passengers occupy the majority of the space in a vehicle, causing passengers to be uncomfortable. In-vehicle crowding discomfort, in the context of public transportation,

[1] Unless otherwise stated, crowding in the following discussion refers to in-vehicle crowding.

is an individual's subjective feeling of discomfort when confronted with a high density of people on public vehicles such as buses and metros. It may not only cause bodily and emotional pain for the passenger, but it may also lead to a sense of uneasiness over life and property [16][18]. As a result, congestion has a substantial impact on passengers' travel route and mode selection. Given the subjective impact on passengers' feelings, crowding level is therefore the third explanatory variable in determining passenger mode choice after trip price and duration. How much more travel time is a passenger ready to accept in exchange for less crowding? What is the marginal rate of substitution between trip time and crowding level, in other words?

Crowding multipliers [55] may be utilized to answer the questions. The idea of crowding multipliers is straightforward: the value of travel time for passengers in crowded and uncrowded situations differs. In other words, crowding multipliers reflect the perceived in-vehicle times of passengers. Because passengers' perceived in-vehicle time in a packed state should be greater than in a less packed condition given the same actual travel time, travel time reductions in crowded situations have greater value than travel time reductions in less crowded situations. As a result, they may reflect the real time-based value of journey time in congested situations.

For more than a decade, crowding valuation has been a hotly debated academic issue. Crowding valuation studies have been conducted in the UK [55][56], Bogota [26], Paris [27], Île-de-France [35], Sydney [28], Mumbai [7], Los Angeles [54], New York [5], Singapore [51], Sweden [13], Seoul[46], Netherlands [58], Hong Kong [30], Santiago [10][9][50], etc. The majority of studies estimated crowding multipliers using a multinomial logit (MNL) model, while some used a mixed multinomial logit (MMNL[2]) model to account for unobserved preference heterogeneity. Normal distribution [56], triangular distribution [7], and lognormal distribution [50] were employed in the mixed logit models. Recent research has also argued against employing a parametric distribution in a mixed logit model. Bansal et al. [5] utilized the semi-nonparametric logit model, which includes the logit-mixed logit (LML) model and the mixture-of-normals MNL (MON-MNL) model, to quantify the preference difference in passengers in a more flexible manner.

1.2 Background information

In Santiago, Chile, crowding on public transportation is a frequent occurrence. Its city-wide integrated public transportation system, known as Red Metropolitana de Movilidad [43][8] (English: Metropolitan Mobility Network; named Transantiago until March 2019), was inaugurated in February 2007. Local (feeder) bus routes, major bus lines, and the Metro (subway) network are all part of the system. It has an integrated pricing system that enables customers to move from bus to bus, metro to bus or metro to train on a two-hour time limit from the first trip (maximum of two changes) for the price of one ticket by using a single contactless smart card called "Bip! card". Bus-to-metro transfers costs 0.03 USD during Horario Valle (English: low-use hours) and 0.12 USD during Horario Punta (English: rush hour). [3]

Transantiago's implementation has been difficult, since the reduced bus fleet and new routes have been insufficient to effectively service the population. The

[2]The acronym ML is also often used.
[3]Source: Red Metropolitana de Movilidad Official site

most common criticisms include a shortage of buses and irregular frequencies, a lack of or inadequate infrastructure (such as segregated lanes, prepaid areas, and bus stops), the network's coverage, and the number of transfers required for longer journeys. As a consequence, the Metro has become congested [25][8], reaching 6 standees per square meter or more during peak hours. This elicited a wide range of behavioral reactions from users, ranging from changing means of transportation (there has been an increase in automobile and bicycle usage) to route choices that, under normal crowding circumstances, would be characterized as counter-intuitive or illogical [45]. For example, customers may choose longer routes in order to improve their chances of getting a seat on the train, or they may choose not to join a train or bus because it is deemed overcrowded [50]. Therefore, understanding the relationship between crowding and passenger's travel choice is essential. It is important not only for the planning of new public transportation services, but also for the management of existing routes and services and the cost-benefit analysis of policy interventions aimed at reducing crowding levels, either directly or indirectly [50].

1.3 Contributions

This thesis makes three contributions. First, based on stated preference (SP) survey data, this thesis constructs a multinomial logit model in preference space to examine passengers' choice of public transportation, with crowding level being one of the most relevant explanatory characteristics. The SP survey data originates from a survey created on the Qualtrics online survey platform. The poll, which began in September 2014 and finished in October 2014, was conducted in two ways: online and in person. There were 413 total respondents that completed the poll (210 online surveys and 203 face-to-face surveys). The survey's specifics are covered in Chapter 2. In addition to the fundamental multinomial logit model, the thesis investigates the impact of various crowding level display forms on passengers' perceived crowding level. There are three presenting formats: 2D schematics of public transportation carriages (bird's eye view), images shot inside public transportation carriages, and text. Whelan and Crockett [56], Batarce, Munoz, and Ortuzar [9] depicted the amount of crowding in the form of photographs. Crowding is defined as the likelihood of occurrence by Lu, Fowkes, and Wardman [39], while Li, Gao, and Tu [37] explain and illustrate various crowding levels in the form of colors. Tirachini et al. [50] utilized three distinct forms to describe congestion levels and stated that the varied presentation styles had no meaningful influence on passengers' route and mode choice. Moreover, this thesis analyzes the effect of various smartphone use on respondents' perceived crowding level.

Second, this thesis employs mixed multinomial logit and latent class logit models in preference space to account for unobserved preference heterogeneity. The MMNL model makes the assumption that the parameters have a lognormal distribution. Even though the lognormal distribution may misspecify the underlying mixing density, it has the virtue of being analytically tractable and produce a decent estimate of crowding multipliers when the median value is used. Sociodemographic factors like gender, wealth, age, car ownership, occupation and number of metro trips per week are taken into account in the LCL model.

Finally, crowding multipliers (CM) and standing multipliers (SM) of MNL and MMNL models when passengers are sitting or standing are calculated.

These models are re-estimated in CM-space, which directly offer the crowding multipliers (similar as WTP-space). Furthermore, a worldwide comparison of CM and SM is carried out across seven cities or regions, including Santiago, Seoul, New York, Sweden, London and the South East (SE) of England, Île-de-France, and Netherlands.

CHAPTER 2

DATA ANALYSIS

This chapter provides a thorough description of the survey data. The survey included respondents' background and socioeconomic characteristics (e.g. race, gender, age, income, occupation and car ownership), metro usage, smartphone availability and use, crowding perception, crowding description, and Stated preference (SP) data. The details are illustrated in Section 2.1, 2.2, 2.3, 2.4, 2.5, 2.5, respectively.

2.1 Background and socioeconomic characteristics

The survey includes information on respondents' background and socioeconomic variables such as gender, age, education level, automobile and smartphone ownership, income, etc. As indicated in Figure D.1, 47.2 percent of respondents are female, while 52.8 percent are male. In terms of age distribution, Figure D.2 demonstrates that the majority of respondents are under the age of 50. Figure D.3 shows that 35.8 percent of respondents have a middle school education and 30.3 percent have a bachelor's degree. Figure D.4 and Figure D.5 demonstrate that 62.2% of respondents own vehicles and 69.2% possess smartphones, respectively. Figure D.6 and Table C.1 depict the income distribution. 64.0% respondents have an income lower than 800 $, however, the respondent with an income larger than 1500 $ also takes up to 18.9%.

2.2 Metro Usage

This section discusses the respondents' use of public transportation. Figure D.7 shows the association between automobile ownership and the number of weekly metro journeys. It may be inferred that if a respondent has a vehicle, he or she is more likely to make less metro journeys, compared with other respondent who does not own a vehicle.

The link between average travel time and auto ownership is shown in Figure D.8. Unfortunately, there is no discernible trend, i.e., journey time is unrelated to auto ownership.

Another factor to consider is the form of transportation, i.e., is the journey completed solely by metro or combined with bus? Figure D.9 indicates that 53.3 percent of journeys include just the metro, whereas 37 percent use both the metro and the bus. The association between travel time and trip type is shown in Figure D.10.

2.3 Smartphone availability and use

In this part, the population consists only of respondents who own a smartphone. Figure D.11 reveals that 35% of respondents use their cellphones during the route and 24% use their cellphones for the bulk of the travel. Only 18.5 percent of respondents never use their smartphone when traveling. Respondents are then presented with a range of scenarios, and their reactions are recorded in Figure D.12. According to the survey, most respondents prefer to use their smartphone even on a short journey; however, no significant difference is discovered in the other situations: around half of respondents will use their smartphone while the remainder will not.

2.4 Crowding perception

Respondents are asked how safe and comfortable they feel at three distinct crowding levels: low (level 1 and 2), medium (level 3 and 4) and high (level 5 and 6). Note that the three levels are selected at random from congestion levels 1 and 2, 3 and 4, and 5 and 6 to ensure that each respondent will evaluate all three crowding levels (low, medium and high). Respondents are asked to evaluate each degree of occupancy on a scale of 1 to 7, with 1 indicating extremely insecure (or very uncomfortable) and 7 indicating very secure (or very comfortable). Because it is the scale of marks in the Chilean school system (where 7 is the highest possible score, 1 is the lowest value, and 4 is the minimum mark to pass), the 1 to 7 scale has the benefit of being very intuitive in Chile [50]. The average score for the six occupancy levels is given in Figure 2.1[1], along with their corresponding 2D representations for ease of comprehension. Users do not perceive a difference in comfort or security between levels 1 and 2 or occupancy, in which all passengers sit, and thus it can be suggested that the main variable affecting both security and comfort is the presence of standees (the difference is not statistically significant at the 5% level). The presence of standees causes the degree of comfort to fall faster than the level of safety between levels 2 and 3. From level 3 and above, felt security outperforms perceived comfort on average. Notably, perceived comfort and security are declining at a comparable rate across levels 4 and 6.

[1]The 2D diagram in the figure is credit to A. Tirachini et al.

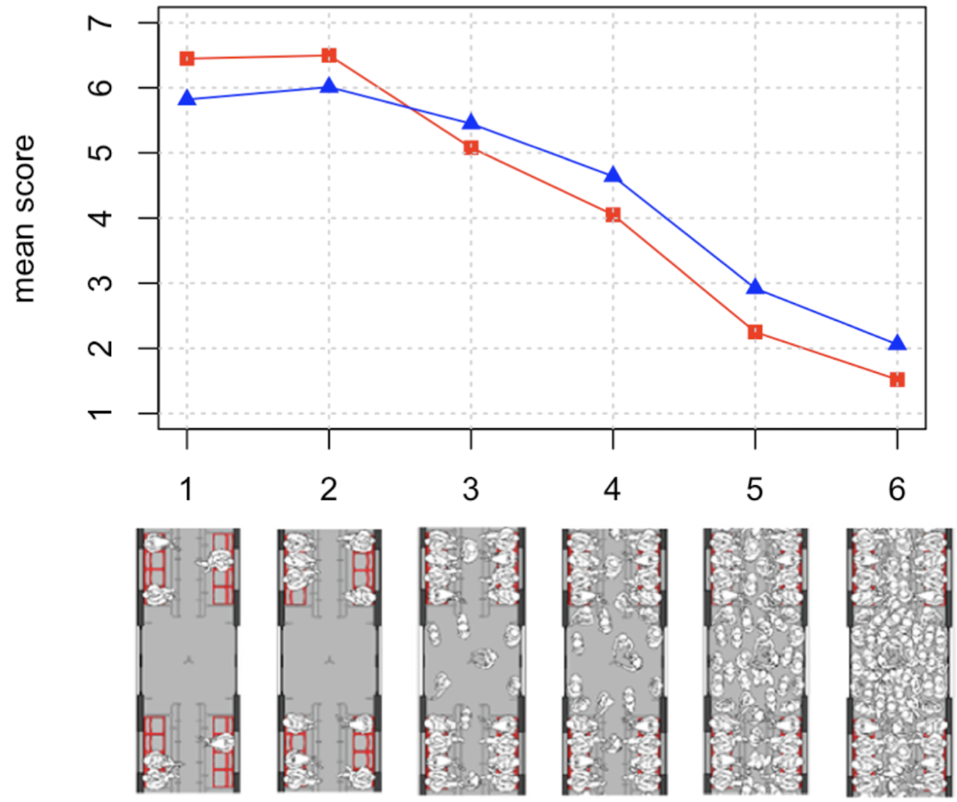


Figure 2.1: Mean security and comfort score for different crowding levels.

According to Figure D.13, when the occupancy level is low, most respondents feel comfortable and safe; when the occupancy level is medium, most respondents feel neither comfortable nor uncomfortable and slightly safe; and when the occupancy level is high, the majority of respondents feel uncomfortable and unsafe. It's also worth noting that the variation in the responses to the security questions is greater than the variation in the responses to the comfort questions. This makes sense since various crimes are associated with various levels of occupancy. Muggings, for example, are more likely to occur when there are a large number of individuals present, while attacks are more likely to occur when there are a small number of people present. In other words, the link between sense of comfort and occupancy level is more likely to be linear, but the association between impression of security and occupancy level is more akin to a bell-shaped curve. When comparing the average score, the survey does not disclose a significant difference between men and women; nonetheless, it can be inferred that when the occupancy level is low, males feel more comfortable but less secure than women.

2.5 Crowding description

The survey includes three types of crowding descriptions: 2D graphics, photographs, and text. An example of 2D graphics is shown in Figure D.14.[2] and an example of crowding levels using photos is shown in Figure D.15.[3] The standees density varies from 0 (no passenger is standing) to 6 pax/m^2 (technical capacity). It should be noted that a single crowding description format is utilized for each participant, i.e., one respondent cannot be provided with two or more alternative crowding description forms.

2.6 Stated preference (SP) survey data

Each responder is presented with six binary choice problems and asked to choose one of the two alternatives in each instance. Each option includes three distinct characteristics, including travel duration, crowding level (expressed in one of the three aforementioned forms), and whether he or she has a seat or must stand throughout the journey. Figure 2.2[4] shows a sample of a choice scenario as presented in the survey. Table C.3 is an example of one of the six choice

[2]The figure is credit to A. Tirachini et al.
[3]The figure is credit to A. Tirachini et al.
[4]The diagram is credit to P. Bansal, et al.

tasks presented to one respondent. The first column *ID* relates to a respondent's unique id, and each respondent is requested to make a decision in every separate choice task, numbered 1 to 6. Note that the crowding level is shown in text format here, and the significance of each number is shown in Table C.4. The value 1 in column *Stand or not* indicates that the responder has to stand during the trip, whereas 0 indicates that the responder has a seat. In each choice situation, the respondent must choose either alternative 1 or alternative 2, with no other options available.

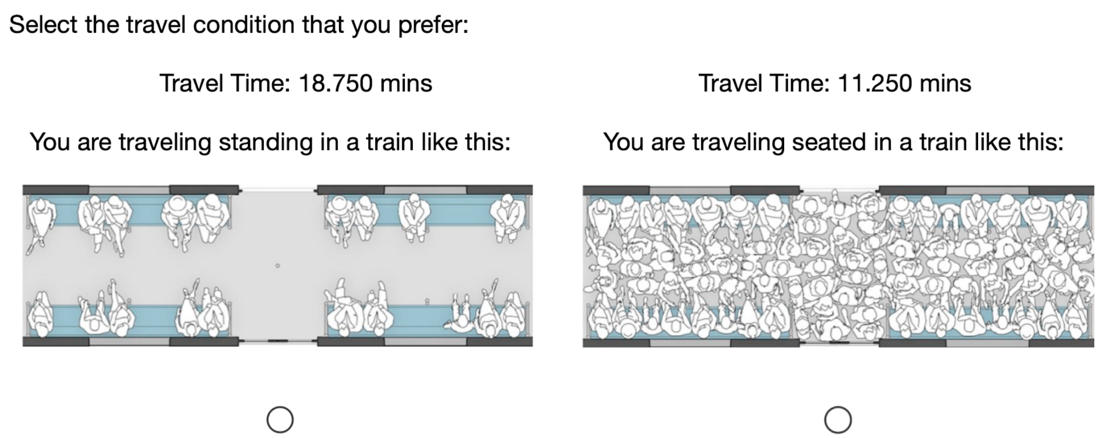


Figure 2.2: Stated preference survey sample.

CHAPTER 3

DISCRETE CHOICE MODELS

In this chapter, three different discrete choice models are built to estimate the crowding multipliers for Santiago's metro users. According to Section 2.6, each respondent is faced with six different binary choice tasks, and he or she needs to choose one of the two alternatives (there is no opt out choice) in each choice task. Note that in the following sections, the choice task scenario is denoted by t, $t = 1, \ldots, T$, whereas each responder is denoted by n, $n = 1, \ldots, N$. Appendix A contains a list of the major notations used in the models.

The multinomial logit (MNL) model [40] may describe observable choice heterogeneity by interacting attributes with individual characteristics, unobserved choice heterogeneity necessitates additional assumptions and a separate model. To account for unobserved preference heterogeneity, Boyd and Mellman [14] developed the mixed multinomial logit (MMNL) model, which adds random parameters that follow a prespecified parametric continuous mixing distribution to the MNL model. Following the landmark work by McFadden and Train [41], which demonstrated that any random utility maximization model can be approximated by MMNL provided the mixing distributions of the random parameters are set properly, MMNL quickly became common practice in choice modeling research. In addition to MMNL, others assumed that the population can be divided into finite classes or clusters, i.e., the number of preference parameters is discrete, as in the latent class logit (LCL) specification: Kamakura and Russell [31], DeSarbo et al. [21], and Bhat [11]. Latent class logit model can be viewed as a special case of MMNL because any mixing distribution can be approximated by increasing the number of classes to infinity.

Most research in the MMNL literature have employed normal heterogeneity distributions. However, Louviere and Eagle [38], Fosgerau et al. [23], and Adamowicz et al. [1] have argued that the normal mixing distributions may introduce misspecification problems if the assumed distribution is not appropriate for the data. For example, if a normal distribution is assumed, researchers may obtain a positive marginal utility of cost. Therefore, lognormal distribution has been used in this case when some parameters are believed to have a fixed sign for everyone, see examples in Tirachini et al. [50].

However, in reality, the theoretical generality of MMNL model is limited by the difficulties of defining and estimating parameter distributions that are both sufficiently flexible and computationally practical [4]. A model that is more flexible than MMNL model is called mixed-mixed logit, or mixture-of-normals multinomial logit (MON-MNL) model. The MON-MNL model is essentially a logit model with random parameters represented by a discrete mixture of continuous (Gaussian) heterogeneity distributions. The flexible representation of random heterogeneity in MON-MNL comes from the notion that any continuous distribution may be approximated to a chosen degree of accuracy by a discrete mixture of normal distributions [22]. Keane and Nada [32] used simulated and expressed preference datasets to compare MON-MNL with MMNL and discovered that MON-MNL outperformed parametric models in terms of the Bayes Information Criterion (BIC).

Train [52] proposed logit-mixed logit (LML) model for defining the distribution of random parameters that is mathematically simple but allows for a great degree of flexibility. This specification can estimate any mixing distribution to any degree of precision. Using flexible forms such as polynomials, splines, and step functions, the researcher defines variables to characterize the shape of the mixing distribution. The gradient of the loglikelihood is simple to compute, making estimation easier.

While LML model assumes all utility parameters to be random, Bansal [4] argued that that fixed parameters are usually needed in practice. For instance, (a) to account for alternative-specific fixed effects, the utility should contain fixed alternative-specific constants (ASCs); (b) because preferences for a particular covariate may not differ among people, the calculated parameter may simply be non-random; (c) by allowing the mean of a random parameter (e.g., marginal utility of price) to vary by sociodemographics, a parsimonious specification of heterogeneity combines various means (deterministic taste variation) with the same variance of the unobserved component (e.g., gender). This taste variation specification necessitates the addition of a fixed parameter on the covariate's interaction with the demographic dummy (e.g., price $\times$ gender). Bansal developed the LML-FR model, extending the LML model to include both random and fixed parameters.

In this chapter, different discrete choice models including basic MNL, MMNL (or ML) and LC are implemented, see in Section 3.1, 3.2 and 3.3. MOM-MNL and LML are covered in Appendix B.

3.1 Multinomial logit model

First, a multinomial logit model is constructed, with just three characteristics of the alternatives taken into account: travel time, crowding level, and a binary indicator indicating whether the passenger is standing or sitting throughout the ride. It is expected that when journey duration varies, the coefficient of crowding level will differ, since the impact of crowding on a short trip vs a long trip is obviously not the same. Similarly, it is assumed that the binary indicator's coefficient is linked to both trip time and crowding level. As a result, alternative metro routes 1 and 2 have the following utility functions:

$$\begin{aligned} u_{1nt} &= \beta_{TT} TT_{1nt} + \beta_{TTd}(TT_{1nt} \times d_{1nt}) + \beta_{TTdSt}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) + \varepsilon_{1nt} \\ &= x'_{1nt}\beta + \varepsilon_{1nt} \\ u_{2nt} &= \beta_{TT} TT_{2nt} + \beta_{TTd}(TT_{2nt} \times d_{2nt}) + \beta_{TTdSt}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) + \beta_0 + \varepsilon_{2nt} \\ &= x'_{2nt}\beta + \varepsilon_{2nt} \end{aligned} \tag{3.1}$$

where $\beta = (\beta_{TT}, \beta_{TTd}, \beta_{TTdSt}, \beta_0)$ is the vector of corresponding estimated parameters, $x_{1nt} = (TT_{1nt}, TT_{1nt} \times d_{1nt}, TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}, 0)$ and $x_{2nt} = (TT_{2nt}, TT_{2nt} \times d_{2nt}, TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}, 1)$. Note that the constant for u_{1nt} is set to 0 to ensure identification. TT_{jnt} denotes the travel time for alternative j, d_{jnt} denotes the crowding level for alternative j and $\mathbf{1}_{Stjnt}$ is the binary indicator of whether passenger n is stand or not in choice task t for alternative j. ε_{jnt} is the error term and it is assumed that ε_{1nt} and ε_{2nt} are iid and follow Gumbel distribution, or type-I extreme value (EV1) distribution [53], i.e., $\varepsilon_{jnt} \overset{iid}{\sim} EV1(\mu, \lambda)$, where μ is the location and λ is the scale. The logit type models is based on the assumption, and the following part derives the multinomial logistic choice probabilities. In this case, $J = 2$ since there are two alternatives.

The utility function may be used to rank options based on preferences. For example, if alternative 1 is selected, it has a higher utility value than the other alternatives. Therefore, the probability that alternative 1 is chosen can be written as follows:

$$\begin{aligned} P_{1nt} &= P(u_{1nt} > u_{2nt}) \\ &= P(x'_{int}\beta + \varepsilon_{1nt} > x'_{2nt}\beta + \varepsilon_{2nt}) \\ &= P(\varepsilon_{2nt} - \varepsilon_{1nt} < x'_{1nt}\beta - x'_{2nt}\beta) \\ &= F(x'_{1nt}\beta - x'_{2nt}\beta) \end{aligned} \tag{3.2}$$

where the last equation follows from the fact that the difference of two independent random variables ε_{i1} and ε_{i2} that follows EV1 distribution is logistically distributed [19][3], and F stands for the cumulative distribution function (CDF). Since $\epsilon_{i2}, \epsilon_{i1} \overset{iid}{\sim} EV1(\mu, \lambda)$, then:

$$\varepsilon_{2nt} - \varepsilon_{1nt} \sim \Lambda(\frac{\varepsilon_{2nt} - \varepsilon_{1nt} + \mu - \mu}{\lambda}) = \Lambda(\varepsilon_{2nt} - \varepsilon_{1nt}) \tag{3.3}$$

where Λ denotes the CDF of logistic regression. Hence, the probability of choosing alternative 1 becomes:

$$\begin{aligned} P_{1nt} &= \Lambda(x'_{1nt}\beta - x'_{2nt}\beta) \\ &= \frac{\exp(x'_{1nt}\beta - x'_{2nt}\beta)}{1 + \exp(x'_{1nt}\beta - x'_{2nt}\beta)} \\ &= \frac{\exp(x'_{1nt}\beta)}{\exp(x'_{1nt}\beta) + \exp(x'_{2nt}\beta)} \end{aligned} \tag{3.4}$$

Similarly, the probability of choosing alternative 2 is:

$$P_{2nt} = \frac{\exp(x'_{2nt}\beta)}{\exp(x'_{1nt}\beta) + \exp(x'_{2nt}\beta)} \tag{3.5}$$

The probability for respondent n making the sequence of choices in all choice tasks may be expressed as follows:

$$P_n = \prod_{t=1}^{T} \left[\frac{\exp(x'_{1nt}\beta)}{\exp(x'_{1nt}\beta) + \exp(x'_{2nt}\beta)}\right]^{y_{1nt}} \left[\frac{\exp(x'_{2nt}\beta)}{\exp(x'_{1nt}\beta) + \exp(x'_{2nt}\beta)}\right]^{y_{2nt}} \tag{3.6}$$

where y_{jnt} equals to 1 if and only if alternative j is chosen by respondent n in choice task t.

Several different statistic values are utilized to assess the model's goodness of fit. Rho-square, or ρ^2 value is the percentage increase in the loglikelihood of the model compared to the equally likely model, and it takes the following form [42]:

$$\rho^2 = 1 - \frac{\mathcal{L}(\hat{\beta}_{MLE})}{\mathcal{L}(0)} \tag{3.7}$$

where $\mathcal{L}(0)$ is the loglikelihood of a equally likely model or a null model, i.e., when all the parameters are set to 0 and hence all choice are chosen with the same probability:

$$\mathcal{L}(0) = \sum_{n=1}^{N} \ln\left(\frac{1}{J_n}\right) \tag{3.8}$$

where J_n is the total number of alternatives respondent n has. The loglikelihood value for a logistic regression model is always negative (because the likelihood contribution from each observation is a probability between 0 and 1). If the model does not outperform the null model in predicting the result, then $\mathcal{L}(\hat{\beta}_{MLE})$ will not be much larger than $\mathcal{L}(0)$, and in this case $\rho^2 \approx 0$; if the model was exceptionally excellent, those with a success (1) outcome would have a fitted probability close to one, and those with a failure (0) outcome would have a fitted probability near to zero. In this case $\mathcal{L}(\hat{\beta}_{MLE})$ is close to 0 and hence $\rho^2 \approx 1$. However, usually the value is not big; in fact a Rho-square between 0.2 to 0.4 indicate excellent model fit [7]. But since ρ^2 value is monotonic in the number of parameters K, it has the risk of overfitting. In order to penalize for the number of parameters, adjusted Rho-squared $\bar{\rho}^2$ is preferred [42]:

$$\bar{\rho}^2 = 1 - \frac{\mathcal{L}(\hat{\beta}_{MLE}) - K}{\mathcal{L}(0)} \tag{3.9}$$

In addition to ρ^2 and $\bar{\rho}^2$, Akaike information criterion (AIC) and Bayesian information criterion (BIC) are also useful. Let N be the number of observations, then AIC and BIC are defined as follows [2][15][57]:

$$\begin{aligned} AIC &= 2K - 2\mathcal{L}(\hat{\beta}_{MLE}) \\ BIC &= K\ln(N) - 2\mathcal{L}(\hat{\beta}_{MLE}) \end{aligned} \tag{3.10}$$

The preferred model, given a collection of candidate models fitting the data, is the one with the lowest AIC value [2]. Thus, AIC rewards quality of fit (as measured by the likelihood function), but it also contains a penalty that increases as the number of estimated parameters increases. Overfitting is discouraged by the penalty, which is desirable since increasing the number of parameters in the model nearly always enhances the quality of fit. The model with the lowest BIC value is recommended [57], as is the model with the lowest AIC value. A lower BIC indicates that there are fewer explanatory factors, a better fit, or both. The BIC penalizes free parameters more severely than the AIC, albeit this depends on the size of n and the relative magnitude of n and k.

Let us now return to the MNL estimate. The estimation of β is shown in Table 3.1. Note that all models in this thesis are estimated using the R package **Apollo** [29]. The result shows that these parameters are all significant. The

Table 3.1: Basic MNL.

Coefficients	Estimate	Std. Error	t-value	p-value
β_0	0.145	0.041	3.550	0.000
β_{TT}	-0.098	0.015	-6.456	0.000
β_{TTd}	-0.012	0.002	-6.451	0.000
β_{TTdSt}	-0.008	0.001	-5.516	0.000
Log-likelihood:	-1680.618			
Rho-square:	0.0172			
Adj.Rho-square:	0.0148			
AIC:	3369.24			
BIC:	3392.48			

positive constant β_0 indicates that there is a potential bias towards choosing alternative 2, the one that is shown on the right side of each choice task. Such bias, or so-called left-right bias [17] is common in discrete choice experiment and its influence can be ignored if the model is sophisticated. β_{TT} is negative and it is reasonable since increasing the travel time of a alternative makes it less attractive, i.e., the probability of choosing this alternative decreases. As for the other two coefficients, they are all negative which meets with the expectation. Similarly, when the crowding level increase, or when the passenger has to stand, the probability of choosing such alternative decreases.

Recall that the crowding multipliers is time-based (not money-based), hence they can be treated as the marginal utility of travel time in crowding conditions over that of travel time in uncrowded conditions:

$$
\begin{aligned}
CM_{sitting} &= \frac{\beta_{TT} + \beta_{TTd} \times d}{\beta_{TT}} = 1 + \lambda_1 \times d \\
CM_{standing} &= \frac{\beta_{TT} + \beta_{TTd} \times d + \beta_{TTdSt} \times d}{\beta_{TT}} = 1 + (\lambda_1 + \lambda_2) \times d
\end{aligned} \tag{3.11}
$$

where $CM_{sitting}$ and $CM_{standing}$ stand for the crowding multiplier for a passenger who is sitting and standing correspondingly. Moreover, $\lambda_1 = \beta_{TTd}/\beta_{TT}$ and $\lambda_2 = \beta_{TTdSt}/\beta_{TT}$. Section 4.2 contains a comprehensive explanation of the crowding multiplier.

The second MNL model tests the influence of different crowding presentation formats on respondent's perceived crowding level and route choice. In this case, 2D diagram is set to be the reference, and hence there are only two dummy variables indicating whether the crowding is presented in photo or in text. Then, the utility function takes the following form:

$$
\begin{aligned}
u_{1nt} =& \beta_{TT} TT_{1nt} + \beta_{TTd}(TT_{1nt} \times d_{1nt}) + \beta_{TTdSt}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) + \\
& \beta_{TTd\_text}(TT_{1nt} \times d_{1nt}) Text + \beta_{TTdSt\_text}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) Text + \\
& \beta_{TTd\_photo}(TT_{1nt} \times d_{1nt}) Photo + \beta_{TTdSt\_photo}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) Photo + \varepsilon_{1nt} \\
u_{2nt} =& \beta_{TT} TT_{2nt} + \beta_{TTd}(TT_{2nt} \times d_{2nt}) + \beta_{TTdSt}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) + \\
& \beta_{TTd\_text}(TT_{2nt} \times d_{2nt}) Text + \beta_{TTdSt\_text}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) Text + \\
& \beta_{TTd\_photo}(TT_{2nt} \times d_{2nt}) Photo + \beta_{TTdSt\_photo}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) Photo + \beta_0 + \varepsilon_{2nt}
\end{aligned}
\tag{3.12}
$$

where $Photo$ and $Text$ are dummies. β_{TTd_photo} and β_{TTdSt_photo} are the parameters for passenger density and whether the passenger is stand or not when the crowding presentation format is photo. Similarly, β_{TTd_text} and β_{TTdSt_text} denotes the parameters for passenger density and whether the passenger is stand or not when the format is text. Results of the estimation is shown in Table 3.2. The

Table 3.2: MNL considering different crowding representation formats.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.145	0.041	3.557	0.000
β_{TT}	-0.098	0.0152	-6.462	0.000
β_{TTd}	-0.012	0.002	-5.634	0.000
β_{TTdSt}	-0.007	0.002	-4.028	0.000
β_{TTd_photo}	0.001	0.002	0.494	0.622
β_{TTd_text}	-0.001	0.002	-0.574	0.566
β_{TTdSt_photo}	-0.001	0.002	-0.399	0.690
β_{TTdSt_text}	-0.002	0.002	-0.949	0.343
Log-likelihood:	-1679.444			
Rho-square:	0.0179			
Adj.Rho-square:	0.0132			
AIC:	3374.89			
BIC:	3421.37			

p-value of β_{TTd_photo}, β_{TTd_text}, β_{TTdSt_photo} and β_{TTdSt_text} are larger than 0.05, which indicates that there is no significant perception bias at 95% confidence level. Because the crowding presentation format has no effect on the perceived crowding level of respondents, the following models all omit the format control variables.

However, new specifications of Equation 3.13 where the interactions with photo and text only happen for the highest/the two highest/the three/the four levels of crowding differs from the results that all levels of crowding are considered in Equation 3.12. The new specification is:

$$\begin{aligned}
u_{1nt} =&\beta_{TT}TT_{1nt} + \beta_{TTd}(TT_{1nt} \times d_{1nt})+ \\
&\beta_{TTdSt}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt})+ \\
&\beta_{TTd_text}(TT_{1nt} \times d_{1nt})Text \times \mathbf{1}_{CLnt})+ \\
&\beta_{TTdSt_text}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt})Text \times \mathbf{1}_{CLnt})+ \\
&\beta_{TTd_photo}(TT_{1nt} \times d_{1nt})Photo \times \mathbf{1}_{CLnt})+ \\
&\beta_{TTdSt_photo}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt})Photo \times \mathbf{1}_{CLnt}) + \varepsilon_{1nt} \\
u_{2nt} =&\beta_{TT}TT_{2nt} + \beta_{TTd}(TT_{2nt} \times d_{2nt})+ \\
&\beta_{TTdSt}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt})+ \\
&\beta_{TTd_text}(TT_{2nt} \times d_{2nt})Text \times \mathbf{1}_{CLnt})+ \\
&\beta_{TTdSt_text}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt})Text \times \mathbf{1}_{CLnt})+ \\
&\beta_{TTd_photo}(TT_{2nt} \times d_{2nt})Photo \times \mathbf{1}_{CLnt})+ \\
&\beta_{TTdSt_photo}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt})Photo \times \mathbf{1}_{CLnt}) + \beta_0 + \varepsilon_{2nt}
\end{aligned} \tag{3.13}$$

A new indicator variable $\mathbf{1}_{CLnt}$ is added into the utility, which $\mathbf{1}_{CLnt} = 1$ if and only if the crowding level of at least one of the two alternatives in choice situation n for respondent t is larger than some certain value, otherwise $\mathbf{1}_{CLnt} = 0$. For instance, if only the highest level of crowding is considered, then $\mathbf{1}_{CLnt} = 1$ means that at least one alternative has crowding level equals to 6. The results are shown in Table C.5, C.6, C.7, C.8 correspondingly. Note that when considering the four highest levels of crowding, the estimation results are the same as Table 3.2, which considers all the six levels of crowding. This is due to the fact that no choice situation in the survey has both choices with crowding levels less than 3. Unlike when there is no significant perception bias when all levels of crowding are considered, when only the highest/two highest/three highest levels of crowding are considered, the photo format tends to make respondent feel more uncomfortable than 2D diagram, as it describes the crowding situation in a more intuitive way and thus makes respondent feel as immersive. The main distinction between text and picture is that when text interacts with just the greatest degree of crowding, it is no longer relevant. One potential reason is that text, to some degree, has a similar impact to photos in that it gives the responder with a good imagination of the actual scene (which is exactly what a photo shows). Text, on the other hand, cannot give the same impression as a picture when discussing an extreme instance.

The third MNL model tests the influence of different smartphone usage on respondent's perceived crowding level. Three questions emerge from this specification:

- Is there an effect of smartphone possession on respondents' crowding valuation?
- Does the frequency with which respondents use their smartphones influence their crowding valuation?
- Will a respondent's perceived crowding level change if he or she chooses not to use his smartphone while the train is overcrowded?
- Will respondents' crowding valuation be affected by their use of smart-

phones for various purposes?

The following specification of utility function is used, taking the first question (Is there a relationship between smartphone ownership and respondents' crowding valuation?) for example:

$$\begin{aligned} u_{1nt} =& \beta_{TT}TT_{1nt} + \beta_{TTd}(TT_{1nt} \times d_{1nt}) + \beta_{TTdSt}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) + \\ & \beta_{TTd\_phone}(TT_{1nt} \times d_{1nt})Phone + \beta_{TTdSt\_phone}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt})Phone \\ u_{2nt} =& \beta_{TT}TT_{2nt} + \beta_{TTd}(TT_{2nt} \times d_{2nt}) + \beta_{TTdSt}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) + \\ & \beta_{TTd\_phone}(TT_{2nt} \times d_{2nt})Phone + \beta_{TTdSt\_phone}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt})Phone \end{aligned} \tag{3.14}$$

where *Phone* is the dummy variable. If $Phone = 1$, the responder has a smartphone; otherwise, $Phone = 0$. β_{TTd_phone} and β_{TTdSt_phone} are the parameters for passenger density and whether the passenger is standing or sitting when the responder possesses a smartphone. The estimate results are given in Table 3.3. The findings indicate that, although smartphone ownership has no impact on

Table 3.3: MNL considering phone ownership.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.144	0.041	3.536	0.000
β_{TT}	-0.098	0.015	-6.444	0.000
β_{TTd}	-0.012	0.002	-5.472	0.000
β_{TTdSt}	-0.010	0.002	-5.607	0.000
β_{TTd_phone}	-0.001	0.002	-0.395	0.693
β_{TTdSt_phone}	0.003	0.002	1.986	0.047
Log-likelihood:	-1678.632			
Rho-square:	0.0183			
Adj.Rho-square:	0.0148			
AIC:	3369.26			
BIC:	3404.13			

felt congestion while the respondent is sitting, it has a substantial affect (at 95% confidence level) on perceived crowding when the respondent has to stand throughout the journey. The presence of a smartphone reduces the respondent's perceived crowding levels, as shown by the positive sign of β_{TTdSt_phone}. In other words, while standing throughout the journey, smartphone users are less sensitive to congestion than non-smartphone owners. This may be explained by the distractions provided by smartphones, such as viewing movies, listening to music, and using social media, among other things, which diverted passengers' attention away from the discomfort and insecurity created by congestion. If the passenger has a seat, this distraction is little; however, if the passenger must stand throughout the trip, it becomes substantial.

It should be noted that owning a smartphone does not guarantee that the passenger would use it, although according to the survey data, just 18.5 % of passengers have smartphones but choose not to use them during the journey (see in Figure D.11). As a result, it is reasonable to assume that smartphone owners are also smartphone users. To validate this assumption, a new specification where $Phone\&Use$ is the new dummy and it equals to 1 if and only if the passenger has a smartphone and will use it during the trip is considered:

$$
\begin{aligned}
u_{1nt} =& \beta_{TT} TT_{1nt} + \beta_{TTd}(TT_{1nt} \times d_{1nt}) + \beta_{TTdSt}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) + \\
& \beta_{TTd\_phone\&use}(TT_{1nt} \times d_{1nt}) Phone\&Use + \\
& \beta_{TTdSt\_phone\&use}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) Phone\&Use \\
u_{2nt} =& \beta_{TT} TT_{2nt} + \beta_{TTd}(TT_{2nt} \times d_{2nt}) + \beta_{TTdSt}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) + \\
& \beta_{TTd\_phone\&use}(TT_{2nt} \times d_{2nt}) Phone\&Use + \\
& \beta_{TTdSt\_phone\&use}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) Phone\&Use
\end{aligned}
\tag{3.15}
$$

The results shown in Table 3.4 is close to Table 3.3. This supports the assumption that smartphone owners are also smartphone users.

Table 3.4: MNL considering phone ownership and use.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.144	0.041	3.540	0.000
β_{TT}	-0.098	0.015	-6.443	0.000
β_{TTd}	-0.012	0.002	-5.718	0.000
β_{TTdSt}	-0.010	0.002	-5.750	0.000
$\beta_{TTd\_phone\&use}$	-0.001	0.001	-0.461	0.645
$\beta_{TTdSt\_phone\&use}$	0.003	0.002	1.998	0.046
Log-likelihood:	-1678.606			
Rho-square:	0.0184			
Adj.Rho-square:	0.0148			
AIC:	3369.21			
BIC:	3404.08			

The next MNL model is estimated and the results are shown in Table C.9. $\beta_{TTd\_all}, \beta_{TTd\_maj}, \beta_{TTd\_half}, \beta_{TTd\_lhalf}$ are the parameters for passenger density when passengers use their phones for the whole journey, the majority of the journey, half of the journey, and less than half of the journey, respectively. Similarly, $\beta_{TTdSt\_all}, \beta_{TTdSt\_maj}, \beta_{TTdSt\_half}, \beta_{TTdSt\_lhalf}$ are the parameters for passenger density when the passenger is standing under different frequencies of smartphone use. To guarantee identification, the frequency that a smartphone is never utilized by the passenger is assigned as a reference. The estimate results show that there is no significant crowding perception bias across different frequencies of smartphone usage.

Another MNL model investigates if various purposes for smartphone usage affect perceived crowding levels. According to the findings in Table C.10, when passengers use their smartphone for telephone conversations and work (with apps, work emails, etc.), their perceived crowding level increases, whereas when they use their smartphone for messaging (WhatsAPP, etc.), watching videos, and playing games installed on their phone or online, their perceived crowding level decreases. It can be concluded that perceived crowding rises when people use their smartphones for work and reduces when they use them for leisure. The other purpose of smartphone use, however, had no effect on perceived crowding levels at 95% confidence level. All of this becomes insignificant when it comes to perceived crowding levels when passengers are standing, which may be linked to the fact that many opt not to use cellphones or use them less while standing.

3.2 Mixed multinomial logit model

In the previous section, a multinomial logit model is built, which assumes that the parameters are the same for every respondent. However, respondents might differ in their valuations for each attributes, due to some reasons like individual preferences. Thus, in this section, a mixed multinomial logit (MMNL) model is built to test the assumption that there is unobserved heterogeneity in β. This means that there is a unique $\beta_n = (\beta_{nTT}, \beta_{nTTd}, \beta_{nTTdSt}, \beta_{n0})$ for each respondent n, but each respondent will have the same parameter β_n for different choice situations t. In MMNL model, it is assumed that $\beta_n \sim f(\beta_n|\theta)$, where f is some known distribution and θ are the unknown parameters of the distribution, such as mean and variance. If the values of β_n for each respondent n is known, the choice probability would be:

$$P(y_{jnt}|\beta_n) = \frac{\exp(x'_{jnt}\beta_n)}{\sum_j \exp(x'_{jnt}\beta_n)} \tag{3.16}$$

For our model, there are only two alternatives, and the probability conditional on β_n that respondent n makes the sequence of choices is the same as

MNL model in Section 3.1:

$$P(y_n|\beta_n) = \prod_{t=1}^{T} \left[\frac{\exp(x'_{1nt}\beta_n)}{\exp(x'_{1nt}\beta_n) + \exp(x'_{2nt}\beta_n)} \right]^{y_{1nt}} \left[\frac{\exp(x'_{2nt}\beta_n)}{\exp(x'_{1nt}\beta_n) + \exp(x'_{2nt}\beta_n)} \right]^{y_{2nt}} \tag{3.17}$$

The choice probability should be the integral of the conditional probability over the density of β_n. Therefore, the choice probability for respondent n is:

$$\begin{aligned} P_n^{MMNL} &= \int_{\beta^n} P(y_n|\beta_n) f(\beta_n|\theta) d\beta_n \\ &= \int_{\beta_n} \prod_{t=1}^{T} \left[\frac{\exp(x'_{1nt}\beta_n)}{\exp(x'_{1nt}\beta_n) + \exp(x'_{2nt}\beta_n)} \right]^{y_{1nt}} \left[\frac{\exp(x'_{2nt}\beta_n)}{\exp(x'_{1nt}\beta_n) + \exp(x'_{2nt}\beta_n)} \right]^{y_{2nt}} f(\beta_n|\theta) d\beta_n \end{aligned} \tag{3.18}$$

Then the loglikelihood function would be the log sum of the above equation:

$$\begin{aligned} \mathcal{L}^{MMNL} &= \sum_{n=1}^{N} \ln\left(\int_{\beta_n} P(y_n|\beta_n) f(\beta_n|\theta) d\beta_n \right) \\ &\approx \sum_{n=1}^{N} \ln\left(\frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T} P(y_{jnt}|\beta_n^{(r)}) \right) \end{aligned} \tag{3.19}$$

The last equation is what known as maximum simulated likelihood estimation [36][12]. $r = 1, \ldots, R$ is the number of draws and $\beta_n^{(r)}$ is the parameter that is drawn from the distribution $f(\beta_n|\theta)$ that needs to be estimated. Repeat R times and the average results will be the estimation of the choice probability P_n^{MMNL}. By the law of large number [20], this is an unbiased estimator of P_n^{MMNL} [12]. As $R \to \infty$, it is consistent and smooth with respect to the unknown parameters.

In practice, $f(\beta_n|\theta)$ can be lognormal, uniform, or any other distributions. When all parameters are assumed to have the same sign for each responder, the lognormal distribution works well. In our case, parameter β_{nTT}, β_{nTTd} and β_{nTTdSt} are believed to be negative for all respondents. Therefore, a lognormal density is used, i.e., $\ln \beta_n \sim N(\mu, \sigma^2)$ and hence μ and σ are the parameters to be estimated.

Another reason for choosing lognormal is that in this case both β_{nTTd} and β_{nTTdSt} would have finite moments, and according to Equation 3.11, this would make sure that the crowding multipliers also have finite moments.

The number of Halton draws is 2000 and the estimation results are shown in Table 3.5. Results show that there is significant heterogeneity across respondents. Another MMNL model where parameters are assumed to be normally distributed is reported in Table C.11. All describing statistic values indicate that the first MMNL model performs better than MNL model.

Table 3.5: Mixed multinomial logit model using lognormal mixing densities.

Coefficients	Estimate	Std.Error	t-value	p-value
β_0	0.192	0.052	3.675	0.000
β_{TT} - μ	-2.019	0.193	-10.449	0.000
β_{TT} - σ	1.087	0.163	6.682	0.000
β_{TTd} - μ	-4.109	0.231	-17.825	0.000
β_{TTd} - σ	1.627	0.225	7.246	0.000
β_{TTdSt} - μ	-4.516	0.239	-18.873	0.000
β_{TTdSt} - σ	1.110	0.220	5.051	0.000
Log-likelihood:	-1494.816			
Number of observations:	2467			
Number of individuals:	413			
Number of draws:	2000			
Rho-square:	0.1258			
Adj.Rho-square:	0.1217			
AIC:	3003.63			
BIC:	3044.31			

3.3 Latent class logit model

Similar to the MMNL model that considers the heterogeneity across respondents, latent class logit (LCL) model assumes that β_n follows a discrete instead of a continuous heterogeneity distribution. Different from continuous distribution, if β_n follows a discrete distribution, then the number of β_n will be finite and the utility function for each class q will be:

$$\begin{aligned} u_{1nt}^{(q)} &= \beta_{TT}^{(q)} TT_{1nt} + \beta_{TTd}^{(q)}(TT_{1nt} \times d_{1nt}) + \beta_{TTdSt}^{(q)}(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt}) + \varepsilon_{1nt}^{(q)} \\ &= x'_{1nt}\beta^{(q)} + \varepsilon_{1nt}^{(q)} \\ u_{2nt}^{(q)} &= \beta_{TT}^{(q)} TT_{2nt} + \beta_{TTd}^{(q)}(TT_{2nt} \times d_{2nt}) + \beta_{TTdSt}^{(q)}(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt}) + \beta_0^{(q)} + \varepsilon_{2nt}^{(q)} \\ &= x'_{2nt}\beta^{(q)} + \varepsilon_{2nt}^{(q)} \end{aligned} \tag{3.20}$$

where $\beta^{(q)} = (\beta_{TT}^{(q)}, \beta_{TTd}^{(q)}, \beta_{TTdSt}^{(q)}, \beta_0^{(q)})$. For each possible $\beta^{(q)}$, it has an associated probability $w_n^{(q)}$ for respondent n. Similar as MMNL model, if the values of $\beta^{(q)}$ for each class q is known, the choice probability will be:

$$P(y_{jnt}|\beta^{(q)}) = \frac{\exp(x'_{jnt}\beta^{(q)})}{\sum_j \exp(x'_{jnt}\beta^{(q)})} \tag{3.21}$$

The unconditional probability that respondent n makes the sequence of choices is:

$$P(y_n) = \sum_{q=1}^{Q} w_n^{(q)} \prod_{t=1}^{T} p(y_{jnt}|\beta^{(q)}) \tag{3.22}$$

Therefore, the likelihood function is:

$$\ell^{LCL} = \prod_{n=1}^{N} \sum_{q=1}^{Q} w_n^{(q)} \prod_{t=1}^{T} P(y_{jnt}|\beta^{(q)}) \tag{3.23}$$

There are different ways to assign $w_n^{(q)}$. To ensure that $\sum_q w_n^{(q)} = 1, \forall n$ and $w_n^{(q)} > 0$, a multinomial logit is specified:

$$w_n^{(q)} = \frac{\exp(\gamma^{(q)})}{\sum_{q=1}^{Q} \exp(\gamma^{(q)})}, q = 1, \ldots, Q \tag{3.24}$$

where $\gamma^{(q)}$ is a constant and $\gamma^{(1)} = 0$ ensures identification. It is also possible to add sociodemographics to the probabilities:

$$w_n^{(q)} = \frac{\exp(h'_n\gamma^{(q)})}{\sum_{q=1}^{Q} \exp(h'_n\gamma^{(q)})}, q = 1, \ldots, Q \tag{3.25}$$

where h_n is a vector of sociodemographics of respondent n. In our model, h_n contains income, car ownership, age, gender, occupation and number of metro trips per week (return = 2 trips):

$$
\begin{aligned}
w_n^{(1)} &= \frac{1}{1+\exp(\gamma^{(2)}+\gamma_{inc}inc_n+\gamma_{auto}\mathbf{1}_{auto_n}+\gamma_{age}age_n+\gamma_{male}\mathbf{1}_{male_n}+\gamma_{ocup}ocup_n+\gamma_{tp}tp_n)} \\
w_n^{(2)} &= \frac{\exp(\gamma^{(2)}+\gamma_{inc}inc_n+\gamma_{auto}\mathbf{1}_{auto_n}+\gamma_{age}age_n+\gamma_{male}\mathbf{1}_{male_n}+\gamma_{ocup}ocup_n+\gamma_{tp}tp_n)}{1+\exp(\gamma^{(2)}+\gamma_{inc}inc_n+\gamma_{auto}\mathbf{1}_{auto_n}+\gamma_{age}age_n+\gamma_{male}\mathbf{1}_{male_n}+\gamma_{ocup}ocup_n+\gamma_{tp}tp_n)}
\end{aligned}
\tag{3.26}
$$

where inc_n, age_n, $ocup_n$ and tp_n represent the income, age, occupation and number of metro trips per week of respondent n. $\mathbf{1}_{auto_n}$ and $\mathbf{1}_{male_n}$ are the binary indicators of whether respondent n has a car or is a male respectively. Then, the loglikelihood function is:

$$
\mathcal{L}^{LCL} = \sum_{n=1}^{N} \ln\left\{\sum_{q=1}^{2} w_n^{(q)} \prod_{t=1}^{T} \left[\frac{\exp(x'_{1nt}\beta^{(q)})}{\exp(x'_{1nt}\beta^{(q)})+\exp(x'_{2nt}\beta^{(q)})}\right]^{y_{1nt}} \left[\frac{\exp(x'_{2nt}\beta^{(q)})}{\exp(x'_{1nt}\beta^{(q)})+\exp(x'_{2nt}\beta^{(q)})}\right]^{y_{2nt}}\right\}
\tag{3.27}
$$

The estimation result is shown in Table 3.6. The absolute value of β_{TT} of class 2 is larger than that of class 1, indicating that class 2 has a higher value of travel time. Moreover, class 1 has larger crowding multipliers for both standing and sitting. These all indicate that class 1 is more sensitive to crowding than class 2 is. Regarding intercept parameters, β_0 for class 1 is no longer significant, which means that the potential bias towards choosing alternative 2 decreases.

However, β_{TT} for class 1 is not significant, with a p-value larger than 0.05. This implies that the value of travel time is split within classes 1. Class 2, on the other hand, has a significant β_{TT} under 95 % confidence level. Males, those with a higher income, those who are younger, those who possess a car, and those who take more metro trips per week are more likely to be classified as class 2. In terms of age and occupation (significant under 95% confidence level), this is a fair outcome. This makes sense since younger individuals are less likely to be bothered by congestion, even if they must stand for the whole of the journey. They are busier and stronger than the elderly, therefore journey time is more important than congestion. It is also reasonable regarding occupation, since students are more concerned with travel time since they utilize the subway to go to school and the penalty for being late is severe.

Table 3.6: Latent class logit model.

Coefficients	Estimate	Std.err.	t-value	p-value
Class 1: β_0	0.087	0.084	1.042	0.297
Class 1: β_{TT}	-0.028	0.044	-0.647	0.518
Class 1: β_{TTd}	-0.035	0.006	-6.254	0.000
Class 1: β_{TTdSt}	-0.017	0.005	-3.514	0.000
Class 2: β_0	0.214	0.069	3.114	0.002
Class 2: β_{TT}	-0.220	0.027	-8.258	0.000
Class 2: β_{TTd}	-0.011	0.002	-4.338	0.000
Class 2: β_{TTdSt}	-0.012	0.002	-5.350	0.000
Class Membership (Class 2)				
$\gamma^{(2)}$	0.838	0.550	1.523	0.128
γ_{male}	0.375	0.254	1.475	0.140
γ_{inc}	0.095	0.067	1.415	0.157
γ_{age}	-0.019	0.009	-2.115	0.034
γ_{auto}	0.477	0.264	1.808	0.071
γ_{ocup}	-0.282	0.117	-2.417	0.016
γ_{tp}	0.016	0.028	0.575	0.565
Log-likelihood:	-1514.984			
Rho-square:	0.1119			
Adj.Rho-square:	0.1031			
AIC:	3059.97			
BIC:	3147.09			

Other latent class logit models were considered as well. Instead of 2 classes, 3 classes are used and the estimation result is shown in Table C.12. Class 3 individuals are the least susceptible to crowding circumstances, and they are more likely to be male, have a higher income, own a car, and take more metro trips per week. However, elderly individuals no longer fall into this category, indicating that the connection between age and crowding sensitivity may not be as simple as it seems. Another latent class logit model with continuous random parameters (travel time are assumed to be lognormally distributed) was built and the result is shown in Table C.13. However, $\beta_{TT} - \mu$ of class 1 is not significant.

CHAPTER 4

RESULTS ANALYSIS

4.1 Choice model results comparison

Table 3.1 shows the result for the base multinomial logit model and Table 3.5 presents the result for the mixed multinomial logit model. To compare two nested models, likelihood-ratio test can be used. The likelihood-ratio test compares the quality of fit of two competing statistical models based on the ratio of their likelihoods, especially one discovered via maximizing over the whole parameter space and another discovered after imposing a restriction. If the constraint (i.e., the null hypothesis) is supported by the observed data, the difference between the two likelihoods should be less than sampling error [33]. As a result, the likelihood-ratio test determines if this ratio is statistically different from one, or, more precisely, if its natural logarithm is substantially distinct from zero. A simple hypothesis test is as follows:

$$\begin{aligned} H_0 &: \theta = \theta_U \\ H_1 &: \theta = \theta_R \end{aligned} \tag{4.1}$$

where θ_U are the parameters of unrestricted model and θ_R are the parameters of restricted model. Suppose restricted model imposes M restrictions, then the limiting distribution of the likelihood ratio test is:

$$-2\left(\mathcal{L}(\hat{\theta}_R; y) - \mathcal{L}(\hat{\theta}_U; y)\right) \sim \chi^2_M \tag{4.2}$$

$\mathcal{L}(\hat{\theta}_R; y)$ and $\mathcal{L}(\hat{\theta}_U; y)$ represent the loglikelihood achieved by the restricted model and unrestricted model correspondingly. For a 95% confidence level, reject H_0 if test statistics is larger than $\chi^2_{95\%,M}$.

Another test for model comparison is the Ben-Akiva & Swait test. Ben-Akiva & Swait test is based on adjusted ρ^2 and can be used for non-nested models:

$$P(\bar{\rho}_1^2 - \bar{\rho}_1^2 \geq z) \leq \Phi\left(-\sqrt{2z\mathcal{L}(0) + df_1 - df_2}\right) \tag{4.3}$$

where z is the observed difference in adjusted ρ^2, Φ is the cumulative standard normal distribution, and df_1 and df_2 are the degree of freedom for model 1 and model 2.

The MMNL model is significantly better than MNL model, which is supported by both the likelihood-ratio test and Ben-Akiva & Swait test. The test result is shown in Table 4.1 and 4.2, and the p-value for likelihood-ratio test and Ben-Akiva & Swait test are both below 0.05 for 95% confidence level. This again supports the assumption that there is significant heterogeneity across individuals. The simulated distribution for each parameters are shown in Figure 4.1, 4.2 and 4.3.

Table 4.1: Likelihood-ratio test for MNL and MMNL.

	Log-likelihood	df
MNL	-1680.62	4
MMNL	-1494.82	7
Difference	185.80	3
Likelihood ratio test-value	371.6	
Degrees of freedom	3	
Likelihood-ratio test p-value	0.000	

df is the number of parameters for the model.

Table 4.2: Ben-Akiva & Swait test for MNL and MMNL.

	LL (0)	LL	df	Adj.Rho-square
MNL	-1709.99	-1680.62	4	0.0148
MMNL	-1709.99	-1494.82	7	0.1217
Difference	0.00	185.80	3	0.1069
Ben-Akiva & Swait test p-value	0.000			

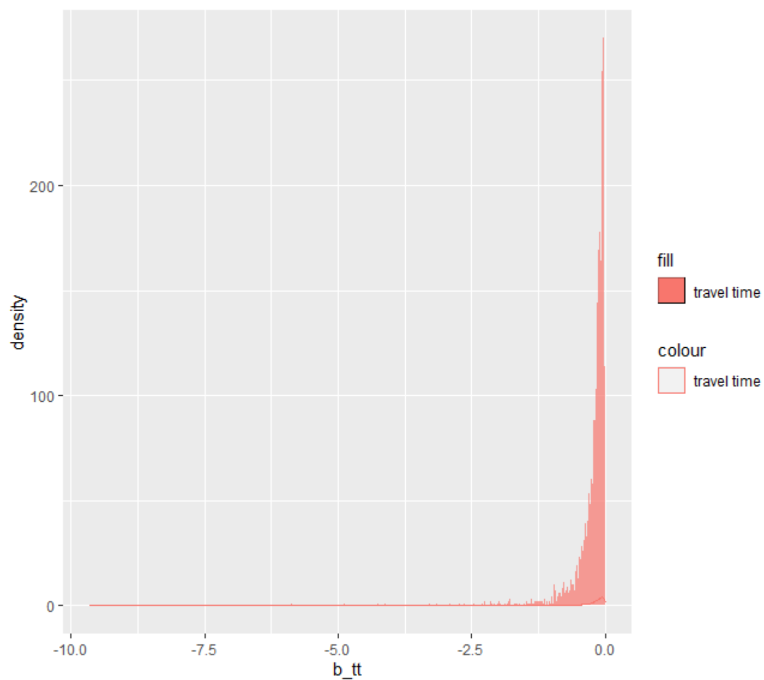


Figure 4.1: β_{TT} simulated distribution histogram, sample = 5,000.

4.2 Crowding multipliers

4.2.1 CM-space

In the valuation of crowding, the distributional assumptions on random heterogeneity can be imposed in two ways: (a) by specifying the distribution of marginal utilities (preference space) and then deriving the distribution of the implied crowding multiplier, or (b) similar to willingness-to-pay space (WTP-

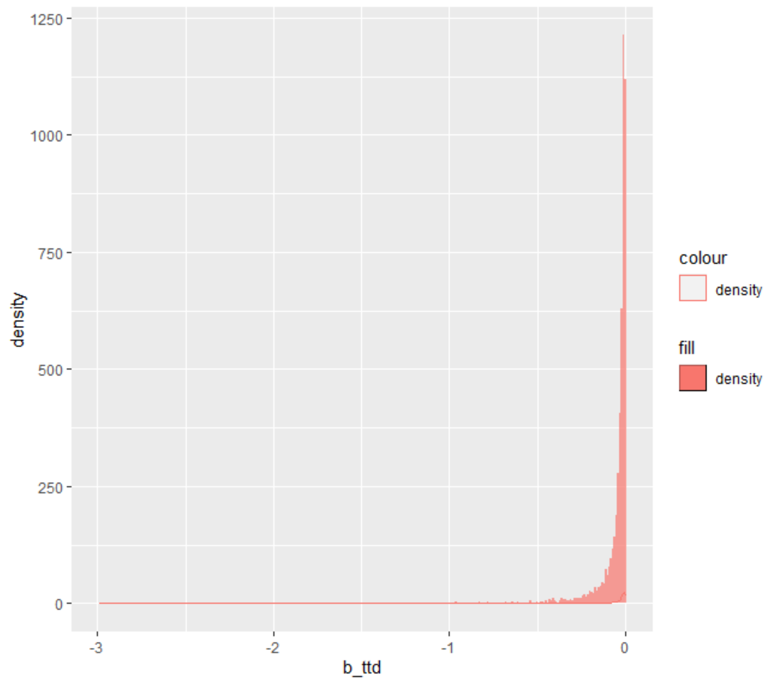


Figure 4.2: β_{TTd} simulated distribution histogram, sample = 5,000.

space), by specifying the distribution of the crowding multiplier directly (CM-space). All the aforementioned model can be estimated in crowding multipliers space (CM-space). This new specification can provide an insight on crowding multipliers directly. Equation 4.4 specifies the utility in CM-space:

$$\begin{aligned} u_{1nt} &= \beta_{TT}[TT_{1nt} + \lambda_1(TT_{1nt} \times d_{1nt}) + \lambda_2(TT_{1nt} \times d_{1nt} \times \mathbf{1}_{St1nt})] + \varepsilon_{1nt} \\ &= x'_{1nt}\beta + \varepsilon_{1nt} \\ u_{2nt} &= \beta_{TT}[TT_{2nt} + \lambda_1(TT_{2nt} \times d_{2nt}) + \lambda_2(TT_{2nt} \times d_{2nt} \times \mathbf{1}_{St2nt})] + \beta_0 + \varepsilon_{2nt} \\ &= x'_{2nt}\beta + \varepsilon_{2nt} \end{aligned} \tag{4.4}$$

Recall from Section 3.1, $\lambda_1 = \beta_{TTd}/\beta_{TT}$ and $\lambda_2 = \beta_{TTdSt}/\beta_{TT}$. Therefore, $CM_{sitting}$ and $CM_{standing}$ can be written as:

$$\begin{aligned} CM_{sitting} &= \frac{\beta_{TT} + \beta_{TTd} \times d}{\beta_{TT}} = 1 + \lambda_1 \times d \\ CM_{standing} &= \frac{\beta_{TT} + \beta_{TTd} \times d + \beta_{TTdSt} \times d}{\beta_{TT}} = 1 + (\lambda_1 + \lambda_2) \times d \end{aligned} \tag{4.5}$$

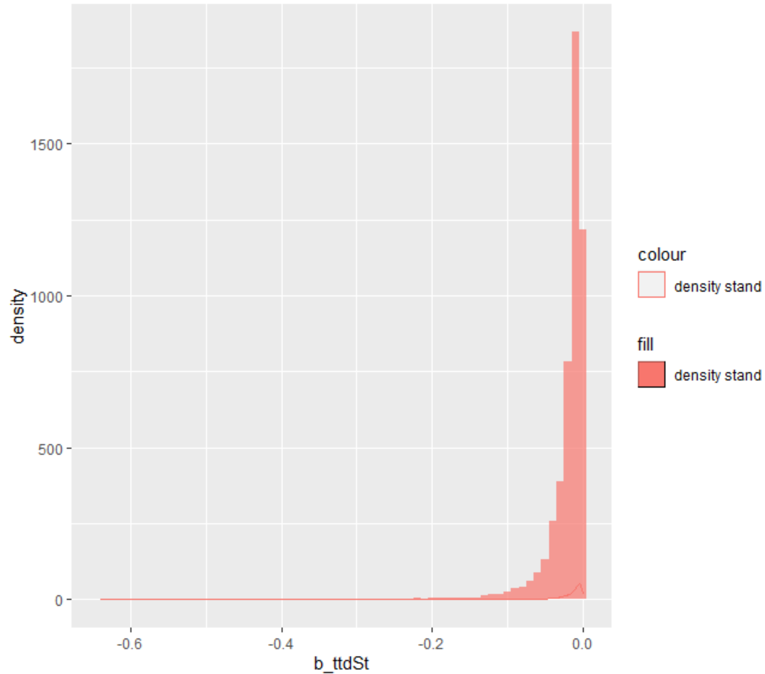


Figure 4.3: β_{TTdSt} simulated distribution histogram, sample = 5,000.

Standing multiplier calculates the increase in the value of travel time saved while standing over sitting and it can be written as the following:

$$SM = \frac{CM_{standing}}{CM_{sitting}} \tag{4.6}$$

It is worth noting that the CM-space specification not only allows for direct interval inference for the crowding multiplier, but also allows for the direct assumption of heterogeneity distributions. Working with ratios is difficult from an econometric standpoint, since it may lead to odd forms of the mixing distribution and unlikely values of the crowding multipliers.

The estimation results for MNL and MMNL model in CM-space are shown in Table 4.3 and Table 4.4. For LCL model in CM-space, the estimation result is shown in Table C.14.

Table 4.5 shows the crowding multipliers when metro passengers may sit

Table 4.3: MNL model in CM-space.

Coefficients	Estimate	Std. Error	t-value	p-value
β_0	0.145	0.041	3.550	0.000
β_{TT}	-0.098	0.015	-6.459	0.000
λ_1	0.121	0.008	16.025	0.000
λ_2	0.082	0.008	10.916	0.000
Log-likelihood:	-1680.618			
Rho-square:	0.0172			
Adj.Rho-square:	0.0148			
AIC:	3369.24			
BIC:	3392.48			

Table 4.4: MMNL model in CM-space.

Coefficients	Estimate	Std.Error	t-value	p-value
β_0	0.189	0.053	3.528	0.000
β_{TT} - μ	-1.869	0.188	-9.921	0.000
β_{TT} - σ	1.475	0.315	4.683	0.000
λ_1 - μ	-2.330	0.105	-22.266	0.000
λ_1 - σ	1.886	0.238	7.926	0.000
λ_2 - μ	-3.107	0.208	-14.965	0.000
λ_2 - σ	1.628	0.232	7.032	0.000
Log-likelihood:	-1504.034			
Number of observations:	2467			
Number of individuals:	413			
Number of draws:	2000			
Rho-square:	0.1204			
Adj.Rho-square:	0.1164			
AIC:	3022.07			
BIC:	3062.74			

while traveling. When pax/m^2 = 0, i.e., no passenger is standing, the crowding multipliers is only related to the value of travel time which is equal to 1 in this case. The crowding multipliers increase with passenger density, as expected, indicating that increased crowding levels increase the disutility of travel time. As a result, metro users are willing to accept longer travel times in exchange for less crowded conditions. Assuming metro users are also willing to pay for reduced travel time, it can be inferred that they are willing to pay more for reduced travel time under crowded conditions. This willingness to pay rises as crowding density rises.

The standard error of the crowding multipliers grows with density, according to Equation 3.11, and this is consistent with both models. It is worth noting that the mean crowding multipliers of the MMNL model are substantially bigger than those of the MNL model, which is due to the fat upper tail of the lognormal density function. This impact might be mitigated by utilizing the median rather than the mean. The median crowding multipliers will not be affected by the minority of respondents who are more sensitive to crowding situations. As a result, the median crowding multipliers of the MMNL model are similar to those of the MNL model.

When it comes to standing crowding multipliers, the situation is fairly close. Even though the mean crowding multipliers of MMNL model are significantly larger, the crowding multipliers of the MNL model are quite similar to the median MMNL value. The comparison of different crowding multipliers is shown in Figure 4.4. It can be concluded that, despite considerable variety in user aversion to crowding, values up to 1.5–1.7 for sitting and up to 1.9–2.2 for standing at a density of standees of 6 pax/m^2, values up to 1.2-1.3 for sitting and values up to 1.4-1.6 for standing at a density of standees of 3 pax/m^2 would be a fair indicator of crowding multipliers for Santiago metro passengers.

The outcome of the standing multiplier, according to Equation 4.6, is displayed in Table 4.7. In Santiago, the value of a seat ranges between 1.04 and 1.29, implying that standing travel time is valued between 4% and 29% more than seated transit time.

Table 4.5: Crowding multipliers: sitting conditions.

	MNL		ML			
pax/m^2	Mean Est.	St. Err	Mean Est.	St. Err	Median Est.	St. Err
0	1.000		1.000		1.000	
1	1.121	0.011	1.556	0.195	1.098	0.006
2	1.242	0.022	2.113	0.390	1.196	0.013
3	1.363	0.032	2.669	0.586	1.293	0.019
4	1.484	0.043	3.226	0.781	1.391	0.026
5	1.606	0.054	3.782	0.976	1.489	0.032
6	1.727	0.065	4.338	1.171	1.587	0.038

Table 4.6: Crowding multipliers: standing conditions.

	MNL		ML			
pax/m^2	Mean Est.	St. Err	Mean Est.	St. Err	Median Est.	St. Err
0	1.000		1.000		1.000	
1	1.203	0.015	1.733	0.231	1.142	0.013
2	1.407	0.030	2.466	0.463	1.284	0.026
3	1.610	0.044	3.199	0.694	1.426	0.039
4	1.814	0.059	3.932	0.926	1.568	0.052
5	2.017	0.074	4.665	1.157	1.709	0.066
6	2.220	0.089	5.397	1.388	1.851	0.079

4.2.2 International comparison of crowding multiplier

The comparison of crowding multipliers and standing multipliers of this thesis with those of London and the South East (SE) of England[56], Seoul [46], New York [5], Sweden [13], Île-de-France [35] and Netherlands [58] is shown in Table 4.9. Crowding multipliers sitting is shown in Figure 4.5 and crowding multipli-

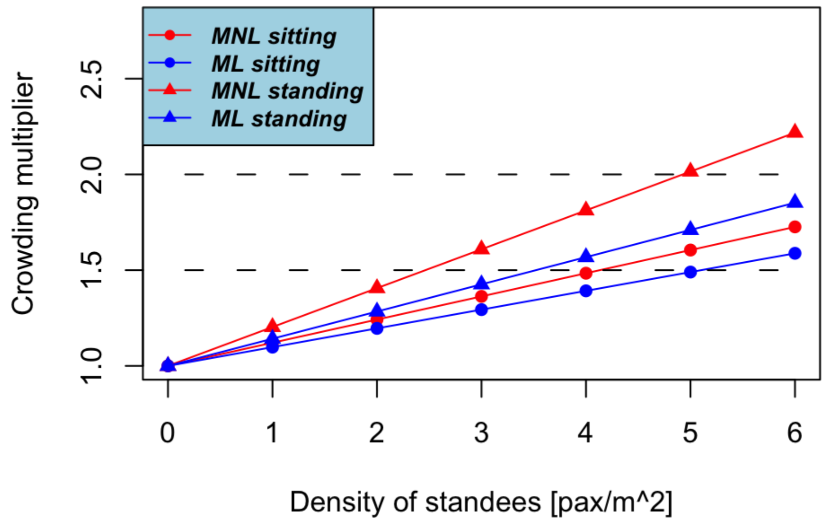


Figure 4.4: Comparison of crowding multipliers: MNL and ML models.

Table 4.7: Standing multipliers.

	MNL	ML	
pax/m^2	Mean Est.	Mean Est.	Median Est.
1	1.073	1.114	1.040
2	1.133	1.167	1.074
3	1.181	1.199	1.103
4	1.222	1.219	1.127
5	1.256	1.233	1.148
6	1.285	1.244	1.166

ers standing is shown in Figure 4.6.[1] The crowding multiplier sitting at 6 pax/m^2 in this thesis is close to that of all cities except Seoul and New York, which both surpass 2. As a result, these two cities are among the top 50 most populous in the world, with Seoul ranked 33st and New York ranking 44th.[2] All cities'

[1]Note that in order to plot CM on the same x axis, the technical capacity for Seoul [46] and Île-de-France [35] is both set to 5 pax/m^2

[2]World Population Prospects (United Nations, 2021)

sitting crowding multipliers in the MMNL model are greater than those in the MNL model. Santiago lies in the middle, with New York having the highest and Seoul having the lowest.

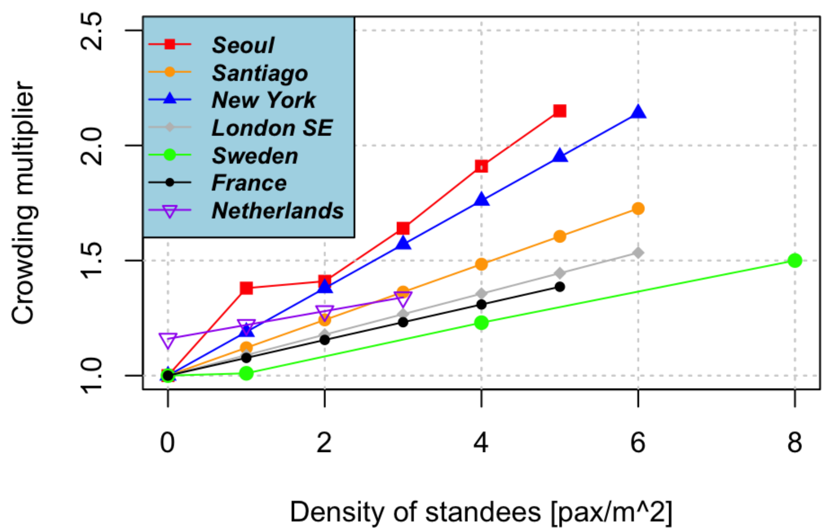


Figure 4.5: International comparison for crowding multiplier sitting.

For the MNL standing CM, only Île-de-France was under the 2.0 CM threshold and Seoul exceeds the 3.0 CM threshold. Santiago is about equivalent to London and the South East of England. Santiago had the lowest standing CM peak for the mixed logit standing CM, at 5.40, while Seoul had the greatest. According to Table 4.8[3] and Figure 4.7[4] (For detailed means of transportation to work/school in the three cities, see in Figure D.16[5]), One possible explanation is that Santiago inhabitants spend a higher proportion of their income on transportation than residents of the other two cities (Santiago is 11.8% , while Seoul is 8.3% and New York is 6.5%). Therefore, it is possible to argue that metro

[3]Data source: NUMBEO
[4]Original figure comes from NUMBEO
[5]Original figure comes from NUMBEO

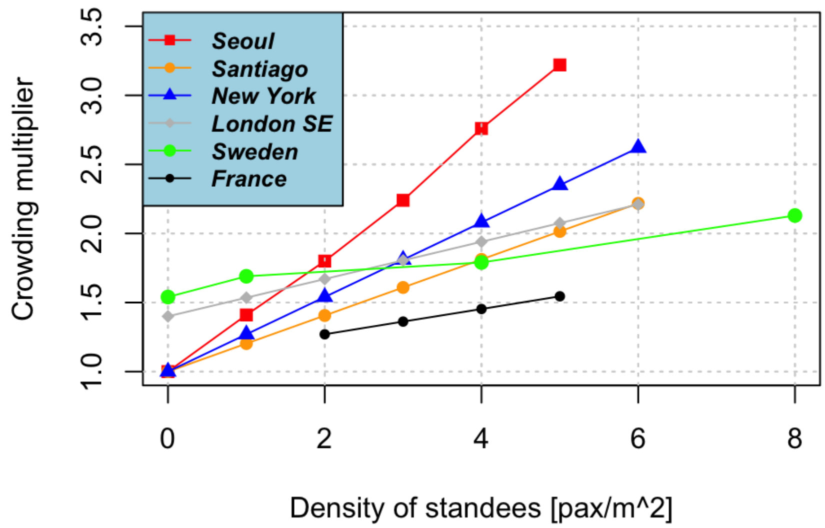


Figure 4.6: International comparison for crowding multiplier standing.

passengers in Santiago are less sensitive to congestion in the metro system than passengers in Seoul and New York.

Table 4.8: Lives comparison between Santiago, Seoul and New York.

City	Santiago	Seoul	New York
Population	6,811,595	9,967,677	8,230,290
Average Monthly Net Salary (After Tax)	641.78 $	2,596.08 $	6,526.66 $
A single person estimated monthly costs without rent	686.60 $	1,079.14 $	1,345.05 $
Average transportation cost (one-way ticket)	1.07 $	1.10 $	2.75 $
Average transportation cost (monthly pass)	54.67 $	48.55 $	130.00 $

Note: All data is from 2021.

Passengers' appraisal of crowding when they have a seat does not change much among cities. However, when passengers are forced to stand throughout their journey, the value of crowding varies dramatically from city to city. This

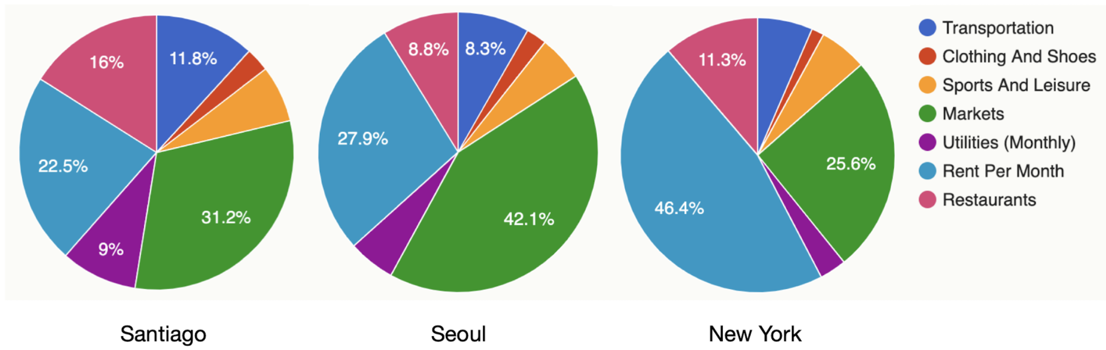


Figure 4.7: Distribution of living expenses in Santiago, Seoul and New York.

may be explained by the fact that various metro systems have varied designs and routes. A metro train, for example, with fewer seats implies that customers are unlikely to acquire a seat during peak hours. Meanwhile, travelers are more sensitive to congestion if they must stand for extended interurban journeys. As a result, it is possible to deduce that, when compared to sitting circumstances, crowding multipliers standing are also influenced by the design of metro trains and the metro system.

Table 4.9: International Comparison of Crowding Multiplier (CM) and Standing Multiplier (SM) Results.

							CM at highest CL		
Author	Year	City	Method	Transit mode	CL	Standing pax/m^2 or Number of passengers (% of total number of seats) at the highest CL	Sitting	Standing	SM
Shin et al. [46]	2021	Seoul	MNL	Metro	6	200%	2.15	3.22	1.50
			MMNL				3.41	7.19	2.11
Bansal et al. [5]	2019	New York	MNL	Subway	5	5	2.13	2.65	1.24
			MMNL				5.95	6.92	1.16
Björklund and Swärdh [13]	2017	Sweden	MNL	PT	4	8	1.50	2.13	1.42
Kroes et.al [35]	2014	Île-de-France	MNL	Metro	8	250%	1.39	1.55	1.12
Whelan and Crocket [56]	2009	London SE	MNL	Rail	7	6	1.63	2.04	1.25
Yap et al. [58]	2020	Netherlands	MNL	Bus	4	3	1.34	-	-
This thesis	2021	Santiago de Chile	MNL	Metro	6	6	1.73	2.22	1.29
			MMNL				4.34	5.40	1.24

Note: PT stands for *Public Transportation*, CL stands for *Crowding level*, CM stands for *Crowding multiplier*, SM stands for *Standing multiplier*.

Note: Two different ways of quantifying crowding levels are included, either measuring the number of standing passengers per m^2, or the number of passengers over the number of seats.

Note: CM for MMNL model is the mean value, not the median value.

Note: All researches use stated preference data.

CHAPTER 5

CONCLUSION

Crowding is included as one of the key explanatory factors in choice models, and it is an important factor in explaining user behavior in Santiago. Crowding multipliers and standing multipliers are obtained after estimating basic multinomial Logit models, mixed multinomial logit models, and latent class logit models. In addition, the significance of the form of representation of the crowding level was investigated, and it was shown to have no significant influence. Furthermore, results show that (a) while smartphone ownership has no effect on felt congestion while the respondent is sitting, it has a significant effect on perceived crowding when the respondent has to stand the entire journey; (b) there is no significant crowding perception bias across different frequencies of smartphone usage; and (c) perceived crowding increases when people use their smartphones for work and reduces when they use them for entertainment.

The assessment of crowding effect and the value of having a seat has the potential to influence project appraisal by allowing different benefits for users to be evaluated at different crowding levels. This might have been helpful in the Santiago public transportation planning model, since it was assumed that one minute is worth the same whether on trains or buses. The results of this thesis may be used to assess the value of increasing service frequency, train size, or seat capacity as methods for improving service quality. The crowding multiplier sitting is determined to be 1.5–1.7 at a density of standees of 6 pax/m2, but the crowding multiplier standing is between 1.9 and 2.2 for the same density. Crowding multipliers for the MNL and the median MMNL are not far apart, suggesting that in the instance of Santiago, using crowding multipliers from a basic MNL model is sufficient to predict the crowding sensitivity of the population as a whole. However, substantial variability exists in our dataset, which both MMNL and LCL models can detect. Using the LCL model, this thesis attempts to discern between groups of consumers with varied preferences. However, distinct groups cannot be readily distinguished based on their backgrounds and socioeconomic status.

The models are also estimated in CM-space in this thesis. In the crowded valuation literature, preference space models have previously been calculated. However, only a few articles specifically mention random crowding multipliers. It is worth noting that the CM-space specification not only allows for direct interval inference for the crowding multiplier, but also allows for the direct assumption of heterogeneity distributions (whereas preference space imposes a multiplier with a 'ratio of a mixture of normals'). Working with ratios is difficult from an econometric standpoint, since it may lead to odd forms of the mixing distribution and high (unlikely) values of the crowding multipliers.

Moreover, results obtained in this thesis are compared with results from Seoul, New York, Sweden, Île-de-France, London and the South East of England, and Netherlands. Because all of these studies use stated preference data, there is no reason to be concerned about the suggestion made by Kroes et al. [35] and Hörcher et al. [30] that crowding multipliers obtained from stated preferences data may be larger than those obtained from revealed preferences data. The results indicate that Santiago's crowding multiplier is around the average of these cities, suggesting that Santiago metro passengers are neither greatly nor seldom affected by congestion.

In terms of policy implications, the anticipated crowding multipliers should be tested in the assessment of modifications to the current metro network and service, such as increasing or decreasing service frequency in peak and off-peak hours. Without a crowding disutility, increasing train frequency just reduces waiting time. The method given here may be used to quantify the impact of such intervention on travel time comfort for an actual metro line in Santiago.

BIBLIOGRAPHY

[1] Wiktor Adamowicz, David Bunch, Trudy Ann Cameron, Benedict GC Dellaert, Michael Hanneman, Michael Keane, Jordan Louviere, Robert Meyer, Thomas Steenburgh, and Joffre Swait. Behavioral frontiers in choice modeling. *Marketing Letters*, 19(3):215–228, 2008.

[2] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[3] Narayanaswamy Balakrishnan. *Handbook of the logistic distribution*. CRC Press, 1991.

[4] Prateek Bansal. *Improved Estimation of Flexible Logit Models and an Extension to a Model with a t-distributed Error Kernel*. PhD thesis, Cornell University, 2019.

[5] Prateek Bansal, Ricardo Hurtubia, Alejandro Tirachini, and Ricardo A. Daziano. Flexible estimates of heterogeneity in crowding valuation in the new york city subway. *Journal of Choice Modelling*, 31:124–140, 2019.

[6] Robert A. Baron and Deborah R. Richardson. *Human aggression*. Springer Science & Business Media, 2004.

[7] Debasis Basu and John Douglas Hunt. 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(9):1465–1476, 2012.

[8] Marco Batarce, Juan Carlos Muñoz, Juan de Dios Ortúzar, Sebastián Raveau, Carlos Mojica, and Ramiro Alberto Ríos. Use of mixed stated and revealed preference data for crowding valuation on public transport in santiago, chile. *Transportation Research Record*, 2535(1):73–78, 2015.

[9] Marco Batarce, Juan Carlos Muñoz, and Juan de Dios Ortúzar. Valuing crowding in public transport: Implications for cost-benefit analysis. *Transportation Research Part A: Policy and Practice*, 91:358–378, 2016.

[10] Marco Batarce, Juan Carlos Muñoz, Juan de Dios Ortúzar, Sebastián Raveau, Carlos Mojica, and Ramiro Alberto Ríos. Use of mixed stated and revealed preference data for crowding valuation on public transport in santiago, chile. *Transportation Research Record*, 2535(1):73–78, 2015.

[11] Chandra R. Bhat. An endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science*, 31(1):34–48, 1997.

[12] Chandra R. Bhat. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological*, 35(7):677–693, 2001.

[13] Gunilla Björklund, Jan-Erik Swärdh, et al. Valuing in-vehicle comfort and crowding reduction in public transport. Technical report, CTS-Centre for Transport Studies Stockholm (KTH and VTI), 2015.

[14] J.Hayden Boyd and Robert E. Mellman. The effect of fuel economy standards on the u.s. automotive market: An hedonic demand analysis. *Transportation Research Part A: General*, 14(5):367–378, 1980.

[15] Kenneth P. Burnham, David R. Anderson, and Kathryn P Huyvaert. Aic model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral ecology and sociobiology*, 65(1):23–35, 2011.

[16] Yung-Hsiang Cheng. Exploring passenger anxiety associated with train travel. *Transportation*, 37(6):875–896, 2010.

[17] Keith Chrzan. Three kinds of order effects in choice-based conjoint analysis. *Marketing Letters*, 5(2):165–172, 1994.

[18] Tom Cox, Jonathan Houdmont, and Amanda Griffiths. Rail passenger crowding, stress, health and safety in britain. *Transportation Research Part A: Policy and Practice*, 40(3):244–258, 2006.

[19] John S. Decani and Robert A. Stine. A note on deriving the information matrix for a logistic distribution. *The American Statistician*, 40(3):220–222, 1986.

[20] Frederik Michel Dekking, Cornelis Kraaikamp, Hendrik Paul Lopuhaä, and Ludolf Erwin Meester. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media, 2005.

[21] Wayne S. Desarbo, Venkatram Ramaswamy, and Steven H. Cohen. Market segmentation with choice-based conjoint analysis. *Marketing Letters*, 6(2):137–147, March 1995. Copyright: Copyright 2007 Elsevier B.V., All rights reserved.

[22] Thomas S. Ferguson. A bayesian analysis of some nonparametric problems. *The annals of statistics*, pages 209–230, 1973.

[23] Mogens Fosgerau and Stephane Hess. A comparison of methods for representing random taste heterogeneity in discrete choice models. *European Transport Trasporti Europei*, (42):1–25, 2009.

[24] Matthew L. Fried and Victor J. DeFazio. Territoriality and boundary conflicts in the subway. *Psychiatry*, 37(1):47–59, 1974.

[25] Andrés Gómez-Lobo et al. 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*, 2012.

[26] Gonzalo Guerra and J. Bocarejo. Congestion cost in mass transit systems; pricing and investment policy implications—case study: Bogota's brt system. In *13th world conference on transportation research, Rio de Janeiro*, 2013.

[27] Luke Haywood and Martin Koning. The distribution of crowding costs in public transport: New evidence from paris. *Transportation Research Part A: Policy and Practice*, 77:182–201, 2015.

[28] David A. Hensher, John M. Rose, and Andrew T. Collins. Identifying commuter preferences for existing modes and a proposed metro in sydney, australia with special reference to crowding. *Public Transport*, 3(2):109–147, 2011.

[29] Stephane Hess and David Palma. Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of choice modelling*, 32:100170, 2019.

[30] Daniel Hörcher, Daniel J. Graham, and Richard J. Anderson. Crowding cost estimation with large scale smart card and vehicle location data. *Transportation Research Part B: Methodological*, 95:105–125, 2017.

[31] Wagner A. Kamakura and Gary J. Russell. A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, 26(4):379–390, 1989.

[32] Michael Keane and Nada Wasi. Comparing alternative models of heterogeneity in consumer choice behavior. *Journal of Applied Econometrics*, 28(6):1018–1045, 2013.

[33] Gary King. *Unifying political methodology: The likelihood theory of statistical inference*. Cambridge University Press, 1989.

[34] Eric S. Knowles. The gravity of crowding: Application of social physics to the effects of others. *Human response to crowding*, pages 183–218, 1978.

[35] Eric Kroes, Marco Kouwenhoven, Laurence Debrincat, and Nicolas Pauget. Value of crowding on public transport in île-de-france, france. *Transportation Research Record*, 2417(1):37–45, 2014.

[36] Lung-Fei Lee. On efficiency of methods of simulated moments and maximum simulated likelihood estimation of discrete response models. *Econometric Theory*, pages 518–552, 1992.

[37] Hao Li, Kun Gao, and Huizhao Tu. Variations in mode-specific valuations of travel time reliability and in-vehicle crowding: Implications for demand estimation. *Transportation Research Part A: Policy and Practice*, 103:250–263, 2017.

[38] Jordan J. Louviere and Thomas Eagle. Confound it! that pesky little scale constant messes up our convenient assumptions. In *Sawtooth software conference*. Sawtooth Software Inc, 2006.

[39] Hui Lu, Tony Fowkes, and Mark Wardman. Amending the incentive for strategic bias in stated preference studies: Case study in users' valuation of rolling stock. *Transportation Research Record*, 2049(1):128–135, 2008.

[40] Daniel McFadden et al. Conditional logit analysis of qualitative choice behavior. 1973.

[41] Daniel McFadden and Kenneth Train. Mixed mnl models for discrete response. *Journal of Applied Econometrics*, 15(5):447–470, 2000.

[42] Daniel L. McFadden and Tom Domencich. Urban travel demand: A behavioral analysis. 1975.

[43] Marcela A. Munizaga and Carolina Palma. 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, 2012.

[44] Feifei Qin. Investigating the in-vehicle crowding cost functions for public transit modes. *Mathematical Problems in Engineering*, 2014:502708, Feb 2014.

[45] Sebastián Raveau, Zhan Guo, Juan Carlos Muñoz, and Nigel H.M. Wilson. 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, 2014.

[46] Hosuk Shin, Dong-Kyu Kim, Seung-Young Kho, and Shin-Hyung Cho. Valuation of metro crowding considering heterogeneity of route choice behaviors. *Transportation Research Record*, 2675(2):162–173, 2021.

[47] Sinclair Knight Merz (SKM). Critical review of transport modelling tools. 2011.

[48] Daniel Stokols. On the distinction between density and crowding: some implications for future research. *Psychological review*, 79(3):275, 1972.

[49] Alejandro Tirachini, David A. Hensher, and John M. Rose. Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. *Transportation Research Part A: Policy and Practice*, 53:36–52, 2013.

[50] Alejandro Tirachini, Ricardo Hurtubia, Thijs Dekker, and Ricardo A. Daziano. Estimation of crowding discomfort in public transport: Results from santiago de chile. *Transportation Research Part A: Policy and Practice*, 103:311–326, 2017.

[51] Alejandro Tirachini, Lijun Sun, Alexander Erath, and Artem Chakirov. Valuation of sitting and standing in metro trains using revealed preferences. *Transport Policy*, 47:94–104, 2016.

[52] Kenneth Train. Mixed logit with a flexible mixing distribution. *Journal of choice modelling*, 19:40–53, 2016.

[53] Kenneth E. Train. *Discrete choice methods with simulation*. Cambridge university press, 2009.

[54] Peter Vovsha, Marcelo G. Simas Oliveira, William Davidson, Chaushie Chu, Robert Farley, Michael Mitchell, and Gaurav Vyas. Statistical analysis of transit user preferences including in-vehicle crowding and service reliability. Technical report, 2014.

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

[56] Gerard Whelan and Jon Crockett. An investigation of the willingness to pay to reduce rail overcrowding. In *Proceedings of the first International Conference on Choice Modelling, Harrogate, England*, volume 30. Citeseer, 2009.

[57] Ernst Wit, Edwin van den Heuvel, and Jan-Willem Romeijn. 'all models are wrong...': an introduction to model uncertainty. *Statistica Neerlandica*, 66(3):217–236, 2012.

[58] Menno Yap, Oded Cats, and Bart van Arem. Crowding valuation in urban tram and bus transportation based on smart card data. *Transportmetrica A: Transport Science*, 16(1):23–42, 2020.

APPENDIX A

NOTATIONS

- $n = 1, \ldots, N$: respondent or passenger
- $t = 1, \ldots, T$: choice task situation
- $j = 1, \ldots, J$: alternative
- u_{jnt}: utility of alternative j for respondent n in choice task t
- TT_{jnt}: travel time (min) of alternative j for respondent n in choice task t
- d_{jnt}: passenger density (pax/m^2) of alternative j for respondent n in choice task t
- $\mathbf{1}_{jnt}$: a binary indicator of whether the passenger n is standing or not in alternative j of choice task t
- $\beta = (\beta_{TT}, \beta_{TTd}, \beta_{TTdSt}, \beta_0)$: a vector of corresponding preference parameters in MNL model
- $\beta_n = (\beta_{nTT}, \beta_{nTTd}, \beta_{nTTdSt}, \beta_{n0})$: a vector of corresponding preference parameters of respondent n in MMNL model
- $\beta_{TT} - \mu, \beta_{TT} - \sigma, \beta_{TTd} - \mu, \beta_{TTd} - \sigma, \beta_{TTdSt} - \mu, \beta_{TTdSt} - \sigma, \beta_0$: parameters of interest of MMNL model
- $\beta^{(q)} = (\beta^{(q)}_{TT}, \beta^{(q)}_{TTd}, \beta^{(q)}_{TTdSt}, \beta^{(q)}_0)$: a vector of corresponding preference parameters of class q in LCL model
- $w^{(q)}_n$: probability for $\beta^{(q)}$ of respondent n
- h_n: a vector of sociodemographics for respondent n
- γ_q: a vector of parameters for h_n in class q
- λ_1: the ratio of β_{TTd} and β_{TT}

- λ_2: the ratio of β_{TTdSt} and β_{TT}
- K: the total number of parameters of interest
- $CM_{standing}$: crowding multiplier in standing conditions
- $CM_{sitting}$: crowding multiplier in sitting conditions
- SM: standing multiplier

APPENDIX B

APPENDIX OF CHAPTER 3

B.1 Mixture-of-normals logit (MON-MNL) model

The name of MOM-MNL model or mixed-mixed logit model is kind of self-explained. The random parameters in MMNL model are described by a discrete mixture of continuous (Gaussian) heterogeneity distributions in MOM-MNL model. With C components in the mixture, utility derived for individual n in component c from making choice j in choice situation t is:

$$U^c_{jnt} = x'_{jnt}\alpha_c + z'_{jnt}\beta^c_n + \varepsilon_{jnt} \tag{B.1}$$

where x_{jnt} are the alternative-specific characteristics and z_{jnt} are random parameters, with a fixed marginal utility α_c and a random marginal utility β^c_n of class c respectively. Note that $\beta^c_n \sim \mathcal{N}(\gamma_c, \Delta_c)$ and $\varepsilon_{jnt} \overset{iid}{\sim} EV1(0, 1)$. The probability of choosing alternative j by individual n of class c in choice situation t, conditional on β^c_n is:

$$P_{jnt}(y_{jnt}|\alpha_c, \beta^c_n) = \frac{\exp(x'_{jnt}\alpha_c + z'_{jnt}\beta^c_n)}{\sum_{j=1}^{J} \exp(x'_{jnt}\alpha_c + z'_{jnt}\beta^c_n)} \tag{B.2}$$

Then the conditional likelihood function given that respondent n makes the sequence of choices is:

$$\ell^{MON-MNL}_n = \prod_{t=1}^{T} \prod_{j=1}^{J} \left[P_{jnt}(y_{jnt}|\alpha_c, \beta^c_i)\right]^{y_{jnt}} \tag{B.3}$$

Define a membership component $w = (w_1, \ldots, w_N)$. W_n is a random value with probability $P(w_n = c) = s_c$, where $0 \leq s_c \leq 1$ and $\sum_c s_c = 1$. Class membership probability s_c may be given as logit-type expressions, with link functions characterizing classes based on sociodemographics. Therefore, the unconditional probability is:

$$P_i(\theta) = \sum_{c=1}^{C} \left\{ s_c \left[\int_{\beta} \ell_n^{MON-MNL} f(\beta|\gamma_c, \Delta_c) d\beta \right] \right\} \tag{B.4}$$

where θ is the parameters of interests and $\theta = \{\alpha_1, s_1, \gamma_1, \Delta_1, \ldots, \alpha_C, s_C, \gamma_C, \Delta_C\}$. The loglikelihood of MON-MNL model is:

$$\mathcal{L}^{MON-MNL} = \sum_{n=1}^{N} \ln \left(\sum_{c=1}^{C} \left\{ s_c \left[\int_{\beta} \ell_n^{MON-MNL} f(\beta|\gamma_c, \Delta_c) d\beta \right] \right\} \right) \tag{B.5}$$

The loglikelihood function does not have a close-form solution. Again, maximum simulated likelihood estimation can be used.

B.2 Logit-mixed logit (LML) model

Logit-mixed logit model uses a discrete mixing distribution over a finite support set S. The probability of a random β_n belongs to a specific value β_{nr} is:

$$w_n(\beta_n = \beta_{nr}|\gamma) = \frac{\exp(z(\beta_{nr})'\gamma)}{\sum_s \exp(z(\beta_{ns})'\gamma)} \tag{B.6}$$

where γ is a vector of parameters and $z(\beta_{nr})$ is vector-valued function (e.g. spline, step, or polynomial function) that captures the shape of the mixing distribution. For instance, if $z(\beta_{nr})$ is step function. Let $T_m, m \in \{1, \ldots, M\}$ be a partition of S, then the probability $w_n(\beta_n = \beta_{nr}|\gamma)$ becomes:

$$w_n(\beta_n = \beta_{nr}|\gamma) = \frac{\exp(\sum_{m=1}^{M} \mathbf{1}(\beta_{nr} \in T_m)\gamma_m)}{\sum_s \exp(\sum_{m=1}^{M} \mathbf{1}(\beta_{nr} \in T_m)\gamma_m)} \tag{B.7}$$

where $\mathbf{1}(\beta_{nr} \in T_m)$ is a vector of M indicators that identify the subset containing β_{nr}. The conditional choice probability is:

$$P_{jnt}(y_{jnt}|\alpha, \beta_{nr}) = \frac{\exp(x'_{jnt}\alpha + z'_{jnt}\beta_{nr})}{\sum_{j=1}^{J} \exp(x'_{jnt}\alpha + z'_{jnt}\beta_{nr})} \tag{B.8}$$

Therefore, the unconditional likelihood function is:

$$\ell_n^{LML} = \sum_{r \in S} \prod_{t=1}^{T} \prod_{j=1}^{J} \left[P_{jnt}(y_{jnt}|\alpha, \beta_{nr}) w_n(\beta_n = \beta_{nr}|\gamma) \right]^{y_{jnt}} \tag{B.9}$$

and the loglikelihood is:

$$\mathcal{L}^{LML} = \sum_{n=1}^{N} y_{jnt} \sum_{t=1}^{T} \sum_{j=1}^{J} \ln \left[\sum_{r \in S} P_{jnt}(y_{jnt}|\alpha, \beta_{nr}) w_n(\beta_n = \beta_{nr}|\gamma) \right] \tag{B.10}$$

Since the support set S is too large, direct maximization of the sample loglikelihood is computationally intractable in practice. Maximum simulated loglikelihood estimation can be applied by using a randomly generated subset $\mathcal{S} \in S$:

$$\hat{\theta}_{MSLE} = \arg\max_{\theta,\gamma} \mathcal{L}^{LML} = \sum_{n=1}^{N} y_{jnt} \sum_{t=1}^{T} \sum_{j=1}^{J} \ln \left[\sum_{r \in \mathcal{S}} P_{jnt}(y_{jnt}|\alpha, \beta_{nr}) w_n(\beta_n = \beta_{nr}|\gamma) \right] \tag{B.11}$$

APPENDIX C

APPENDIX OF TABLES

Table C.1: Income profile.

Household income ($/month)	Percent.	Accum.
0-200	14.8%	14.8%
200-400	19.9%	34.7%
400-600	19.1%	53.8%
600-800	10.2%	64.0%
800-1000	5.8%	69.8%
1000-1500	11.1%	80.9%
1500 or higher	18.9%	100%

Table C.2: Descriptive sample statistics (N = 413).

Sociodemographics variables	Min	Max	Median	Mean	Std. Dev.
Male indicator	0	1	0	0.460543	0.498594
Age (years)	17	73	33	35.03083	11.93868
Education levels	2	6	5	4.477805	1.187198
Occupation	1	6	3	3.379778	1.208151
Car ownership indicator	0	1	1	0.676326	0.468022
Personal income as share of household income	1	5	3	2.69852	1.431465

Table C.3: Example of six stated choice components for a respondent.

		Alternative 1			Alternative 2		
ID	Choice task	Travel time (mins)	Crowding level	Stand or not	Travel time (mins)	Crowding level	Stand or not
1	1	15.000	6	1	16.875	4	1
1	2	11.250	4	1	16.875	2	0
1	3	15.000	6	0	13.125	5	1
1	4	18.750	1	0	11.250	6	0
1	5	16.875	1	0	15.000	3	0
1	6	11.250	3	0	13.125	1	0

Table C.4: Crowding level description in text format.

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

Table C.5: MNL considering different crowding representation formats, where the interactions with Photo and Text only happen for the highest levels of crowding.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.187	0.043	4.334	0.000
β_{TT}	-0.084	0.016	-5.167	0.000
β_{TTd}	-0.009	0.002	-4.413	0.000
β_{TTdSt}	-0.006	0.002	-3.038	0.002
β_{TTd_Photo}	-0.005	0.002	-2.523	0.012
β_{TTd_Text}	-0.002	0.002	-1.164	0.244
β_{TTdSt_Photo}	-0.002	0.002	-1.018	0.309
β_{TTdSt_Text}	-0.004	0.002	-2.076	0.038
Log-likelihood:	-1674.455			
Rho-square:	0.0208			
Adj.Rho-square:	0.0161			
AIC:	3364.91			
BIC:	3411.4			

Table C.6: MNL considering different crowding representation formats, where the interactions with Photo and Text only happen for the two highest levels of crowding.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.175	0.042	4.167	0.000
β_{TT}	-0.101	0.016	-6.395	0.000
β_{TTd}	-0.009	0.002	-4.705	0.000
β_{TTdSt}	-0.007	0.002	-3.699	0.000
β_{TTd_Photo}	-0.006	0.002	-3.329	0.000
β_{TTd_Text}	-0.008	0.002	-3.958	0.000
β_{TTdSt_Photo}	-0.002	0.002	-0.860	0.390
β_{TTdSt_Text}	-0.003	0.002	-1.457	0.145
Log-likelihood:	-1667.391			
Rho-square:	0.0249			
Adj.Rho-square:	0.0202			
AIC:	3350.78			
BIC:	3397.27			

Table C.7: MNL considering different crowding representation formats, where the interactions with Photo and Text only happen for the three highest levels of crowding.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.142	0.041	3.470	0.000
β_{TT}	-0.108	0.016	-6.913	0.000
β_{TTd}	-0.010	0.002	-5.025	0.000
β_{TTdSt}	-0.009	0.002	-4.687	0.000
β_{TTd_Photo}	-0.005	0.002	-2.752	0.000
β_{TTd_Text}	-0.006	0.002	-3.453	0.001
β_{TTdSt_Photo}	0.000	0.002	0.145	0.885
β_{TTdSt_Text}	-0.001	0.002	-0.471	0.638
Log-likelihood:	-1672.98			
Rho-square:	0.0216			
Adj.Rho-square:	0.017			
AIC:	3361.96			
BIC:	3408.45			

Table C.8: MNL considering different crowding representation formats, where the interactions with Photo and Text only happen for the four highest levels of crowding.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.145	0.041	3.557	0.000
β_{TT}	-0.098	0.0152	-6.462	0.000
β_{TTd}	-0.012	0.002	-5.634	0.000
β_{TTdSt}	-0.007	0.002	-4.028	0.000
β_{TTd_Photo}	0.001	0.002	0.494	0.622
β_{TTd_Text}	-0.001	0.002	-0.574	0.566
β_{TTdSt_Photo}	-0.001	0.002	-0.399	0.690
β_{TTdSt_Text}	-0.002	0.002	-0.949	0.343
Log-likelihood:	-1679.444			
Rho-square:	0.0179			
Adj.Rho-square:	0.0132			
AIC:	3374.89			
BIC:	3421.37			

Table C.9: MNL considering phone usage frequency.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.134	0.049	2.709	0.007
β_{TT}	-0.115	0.019	-6.093	0.000
β_{TTd}	-0.015	0.003	-4.986	0.000
β_{TTdSt}	-0.008	0.003	-2.811	0.005
β_{TTd_all}	0.004	0.003	1.499	0.134
β_{TTdSt_all}	-0.002	0.003	-0.672	0.502
β_{TTd_maj}	0.003	0.003	0.945	0.345
β_{TTdSt_maj}	0.000	0.003	-0.142	0.887
β_{TTd_half}	-0.003	0.004	-0.970	0.332
β_{TTdSt_half}	0.004	0.004	0.998	0.318
β_{TTd_lhalf}	-0.006	0.003	-1.659	0.097
β_{TTdSt_lhalf}	-0.004	0.003	-1.143	0.253
Log-likelihood:	-1149.137			
Number of Observations:	1706			
Rho-square:	0.0282			
Adj.Rho-square:	0.0181			
AIC:	2322.27			
BIC:	3404.13			

Table C.10: MNL considering different purpose of phone use.

Coefficients	Estimate	Std.Error.	t-value	p-value
β_0	0.137	0.051	2.698	0.007
β_{TT}	-0.122	0.020	-6.239	0.000
β_{TTd}	-0.015	0.003	-4.917	0.000
β_{TTdSt}	-0.010	0.003	-3.731	0.000
β_{TTd_social}	-0.002	0.002	-0.922	0.357
β_{TTdSt_social}	-0.002	0.002	-0.824	0.410
$\beta_{TTd\_message}$	0.005	0.002	2.196	0.028
$\beta_{TTdSt\_message}$	0.001	0.002	0.250	0.803
β_{TTd_tele}	-0.005	0.002	-2.129	0.033
β_{TTdSt_tele}	0.002	0.002	0.865	0.387
β_{TTd_news}	-0.001	0.002	-0.354	0.723
β_{TTdSt_news}	-0.001	0.002	-0.361	0.718
β_{TTd_ebook}	-0.002	0.003	-0.581	0.561
β_{TTdSt_ebook}	0.001	0.003	0.379	0.705
β_{TTd_music}	-0.002	0.002	-0.977	0.328
β_{TTdSt_music}	0.002	0.002	1.037	0.300
β_{TTd_video}	0.009	0.004	2.186	0.029
β_{TTdSt_video}	0.004	0.004	0.954	0.340
β_{TTd_game}	0.006	0.002	2.566	0.010
β_{TTdSt_game}	0.003	0.002	1.377	0.169
β_{TTd_work}	-0.007	0.003	-2.713	0.007
β_{TTdSt_work}	-0.001	0.003	-0.335	0.738
Log-likelihood:	-1096.058			
Number of Observations:	1646			
Rho-square:	0.0393			
Adj.Rho-square:	0.02			
AIC:	2236.12			
BIC:	2355.05			

Table C.11: Mixed multinomial logit model using normal mixing densities.

Coefficients	Estimate	Std.Error	t-value	p-value
β_0	0.193	0.051	3.757	0.000
β_{TT} - μ	-0.167	0.027	-6.171	0.000
β_{TT} - σ	0.248	0.024	10.320	0.000
β_{TTd} - μ	-0.026	0.003	-8.337	0.000
β_{TTd} - σ	0.021	0.003	7.156	0.000
β_{TTdSt} - μ	-0.016	0.002	-7.274	0.000
β_{TTdSt} - σ	-0.007	0.005	-1.472	0.141
Log-likelihood:	-1512.314			
Number of observations:	2467			
Number of individuals:	413			
Number of draws:	2000			
Rho-square:	0.1156			
Adj.Rho-square:	0.1115			
AIC:	3038.63			
BIC:	3079.3			

Table C.12: Latent class logit model, with 3 classes.

Coefficients	Estimate	Std.err.	t-value	p-value
Class 1: β_0	8.465	14.616	0.579	0.562
Class 1: β_{TT}	-7.277	36.767	-0.198	0.843
Class 1: β_{TTd}	-0.402	3.010	-0.133	0.894
Class 1: β_{TTdSt}	0.433	0.817	0.530	0.596
Class 2: β_0	0.082	0.085	0.961	0.337
Class 2: β_{TT}	-0.020	0.044	-0.461	0.645
Class 2: β_{TTd}	-0.035	0.006	-6.065	0.000
Class 2: β_{TTdSt}	-0.017	0.005	-3.337	0.001
Class 3: β_0	0.210	0.070	3.000	0.003
Class 3: β_{TT}	-0.211	0.028	-7.479	0.000
Class 3: β_{TTd}	-0.011	0.003	-4.143	0.000
Class 3: β_{TTdSt}	-0.013	0.002	-5.415	0.000
Class Membership (Class 2)				
$\gamma^{(2)}$	-1.023	1.744	-0.587	0.557
$\gamma^{(2)}_{male}$	-0.026	0.783	-0.034	0.973
$\gamma^{(2)}_{inc}$	0.014	0.257	0.054	0.957
$\gamma^{(2)}_{age}$	0.080	0.080	1.001	0.317
$\gamma^{(2)}_{auto}$	0.376	0.933	0.403	0.687
$\gamma^{(2)}_{ocup}$	0.209	0.451	0.464	0.643
$\gamma^{(2)}_{tp}$	-0.001	0.072	-0.019	0.985
Class Membership (Class 3)				
$\gamma^{(3)}$	-0.538	1.763	-0.305	0.760
$\gamma^{(3)}_{male}$	0.384	0.813	0.473	0.637
$\gamma^{(3)}_{inc}$	0.122	0.268	0.456	0.648
$\gamma^{(3)}_{age}$	0.065	0.081	0.804	0.421
$\gamma^{(3)}_{auto}$	0.931	0.968	0.962	0.336
$\gamma^{(3)}_{ocup}$	-0.079	0.465	-0.170	0.865
$\gamma^{(3)}_{tp}$	0.021	0.074	0.287	0.774
Log-likelihood:	-1507.149			
Rho-square:	0.1165			
Adj.Rho-square:	0.1012			
AIC:	3066.3			
BIC:	3217.31			

Table C.13: Latent class logit model, with continuous random parameters.

Coefficients	Estimate	Std.err.	t-value	p-value
Class 1: β_0	0.459	0.321	1.431	0.152
Class 1: $\beta_{TT} - \mu$	-5.119	3.946	-1.297	0.195
Class 1: $\beta_{TT} - \sigma$	-0.799			
Class 1: β_{TTd}	-0.224	0.059	-3.804	0.000
Class 1: β_{TTdSt}	-0.095	0.028	-3.382	0.001
Class 2: β_0	0.173	0.052	3.355	0.001
Class 2: $\beta_{TT} - \mu$	-1.842	0.150	-12.251	0.000
Class 2: $\beta_{TT} - \sigma$	0.940	0.121	7.778	0.000
Class 2: β_{TTd}	-0.015	0.002	-6.741	0.000
Class 2: β_{TTdSt}	-0.015	0.002	-7.816	0.000
Class Membership (Class 2)				
$\gamma^{(2)}$	2.285	0.494	4.622	0.000
γ_{male}	-0.073			
γ_{inc}	0.012			
γ_{age}	-0.007	0.008	-0.859	0.390
$\gamma^{(2)}_{auto}$	0.110			
$\gamma^{(2)}_{ocup}$	-0.218	0.101	-2.155	0.031
$\gamma^{(2)}_{tp}$	0.009	0.022	0.421	0.673
Log-likelihood:	-1479.159			
Rho-square:	0.1329			
Adj.Rho-square:	0.1229			
AIC:	2992.32			
BIC:	3091.06			

Table C.14: Latent class logit model in CM-space.

Coefficients	Estimate	Std.err.	t-value	p-value
Class 1: β_0	0.087	0.084	1.042	0.297
Class 1: β_{TT}	-0.028	0.043	-0.654	0.513
Class 1: β_{TTd}	1.229	1.734	0.709	0.478
Class 1: β_{TTdSt}	0.604	0.784	0.771	0.441
Class 2: β_0	0.214	0.069	3.114	0.002
Class 2: β_{TT}	-0.220	0.027	-8.258	0.000
Class 2: β_{TTd}	0.049	0.007	6.554	0.000
Class 2: β_{TTdSt}	0.056	0.006	9.590	0.000
Class Membership (Class 2)				
$\gamma^{(2)}$	0.838	0.550	1.523	0.128
γ_{male}	0.375	0.254	1.475	0.140
γ_{inc}	0.095	0.067	1.415	0.157
γ_{age}	-0.019	0.009	-2.116	0.034
γ_{auto}	0.477	0.264	1.808	0.071
γ_{ocup}	-0.282	0.117	-2.417	0.016
γ_{tp}	0.016	0.028	0.575	0.566
Log-likelihood:	-1514.984			
Rho-square:	0.1119			
Adj.Rho-square:	0.1031			
AIC:	3059.97			
BIC:	3147.09			

APPENDIX D

APPENDIX OF FIGURES

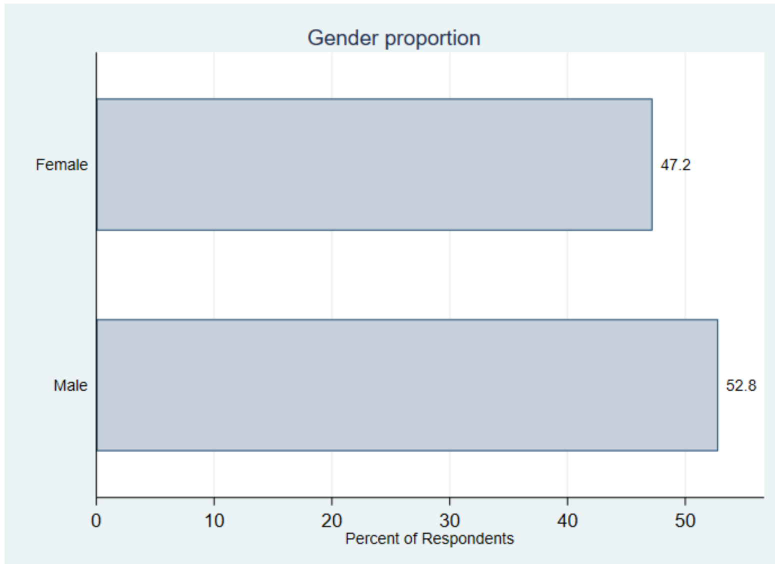


Figure D.1: Gender proportion.

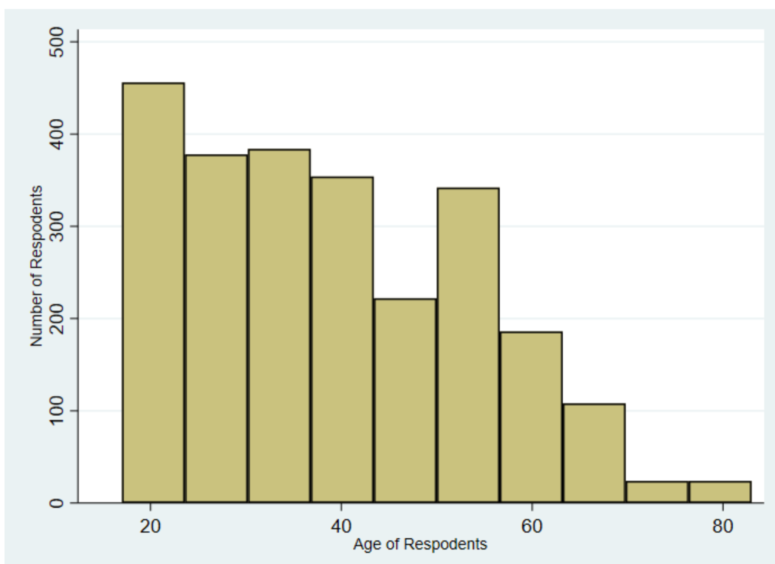


Figure D.2: Age proportion.

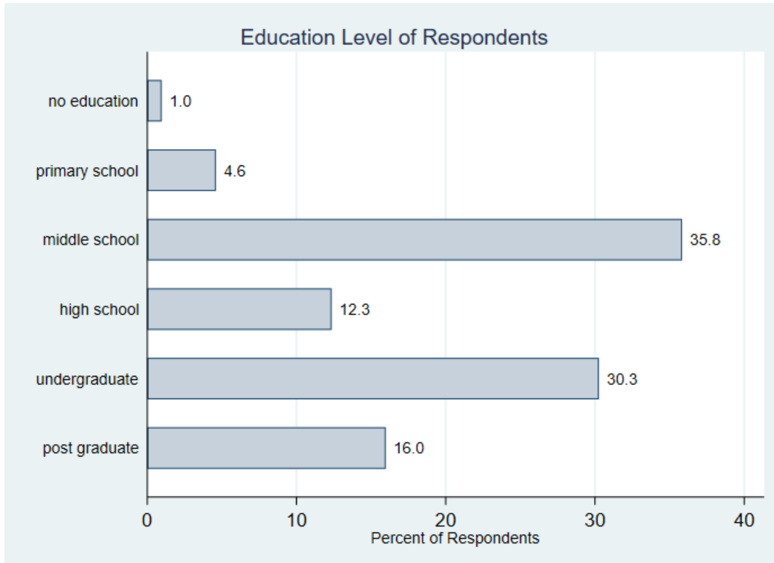


Figure D.3: Education level.

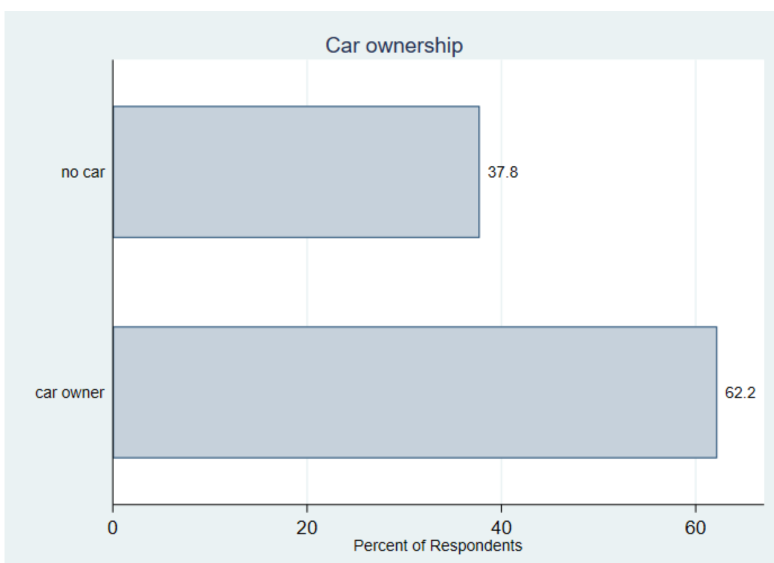


Figure D.4: Car ownership.

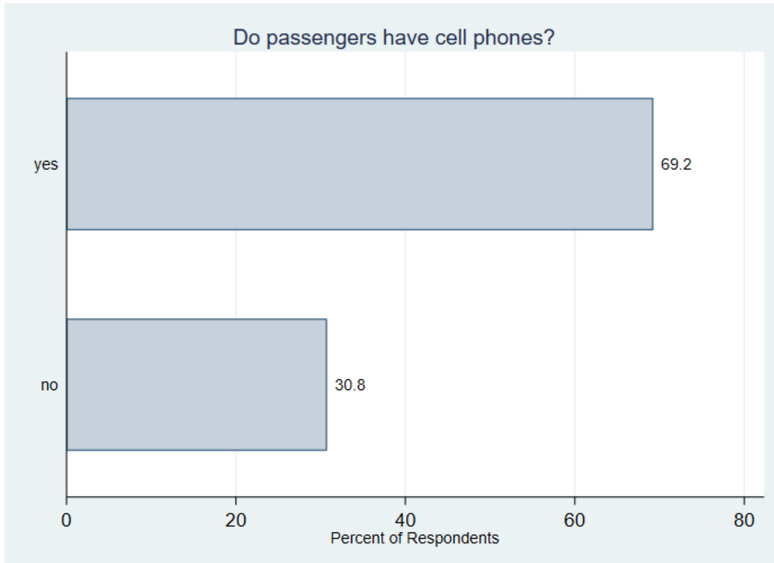


Figure D.5: Smartphones percentage.

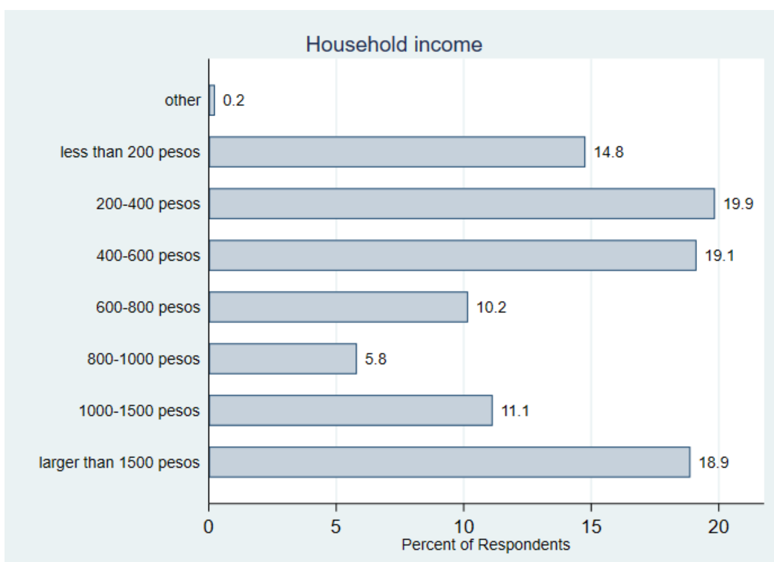


Figure D.6: Income.

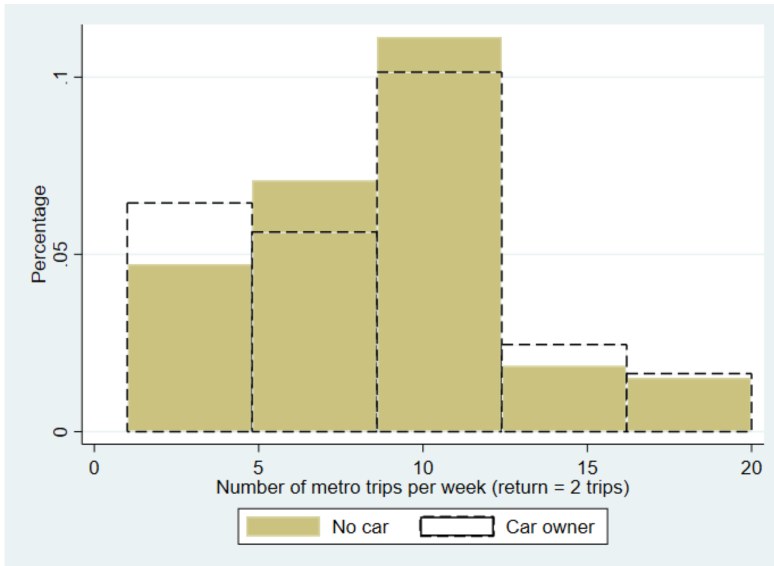


Figure D.7: The relationship between car ownership and number of metro trips.

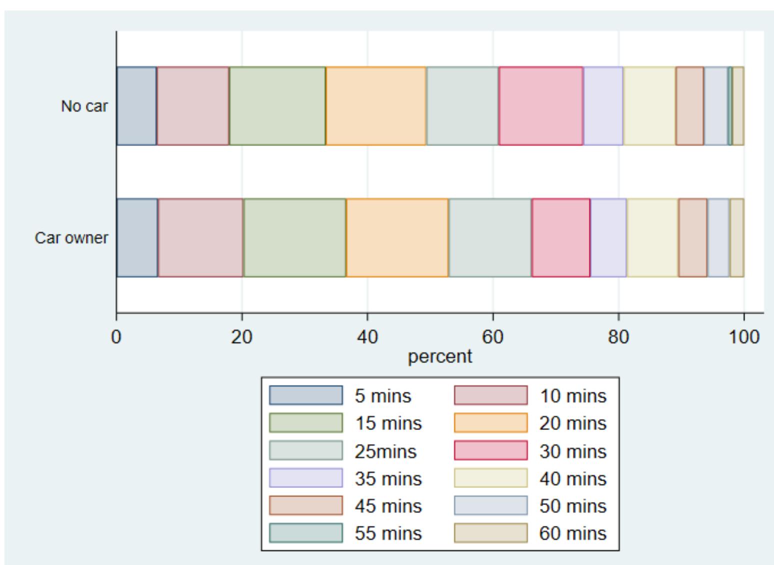


Figure D.8: The relationship between car ownership and average travel time.

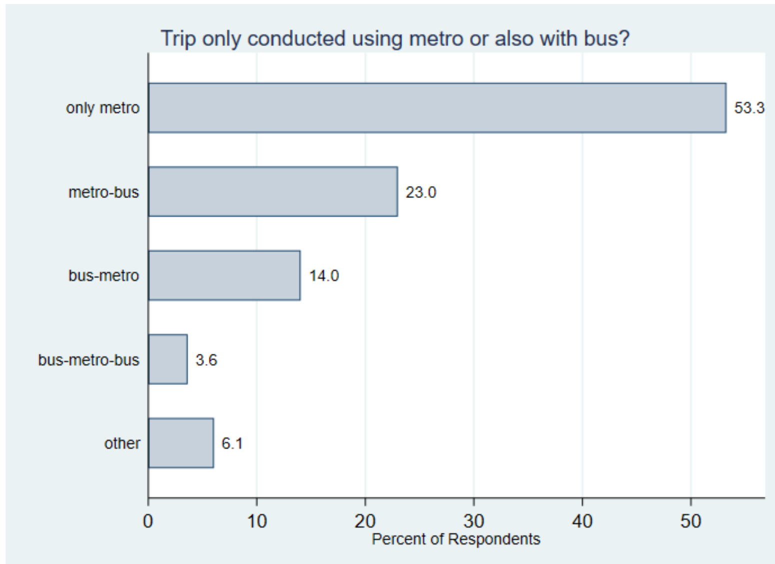


Figure D.9: Trip type.

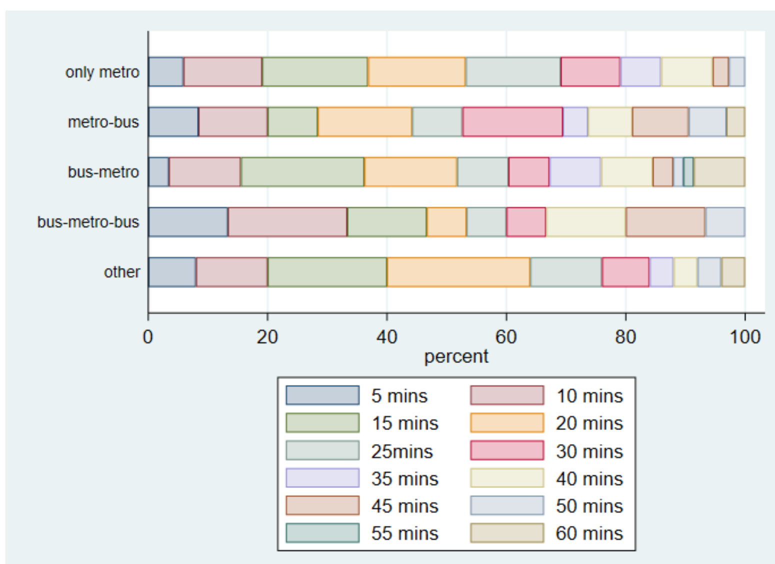


Figure D.10: The relationship between trip type and travel time.

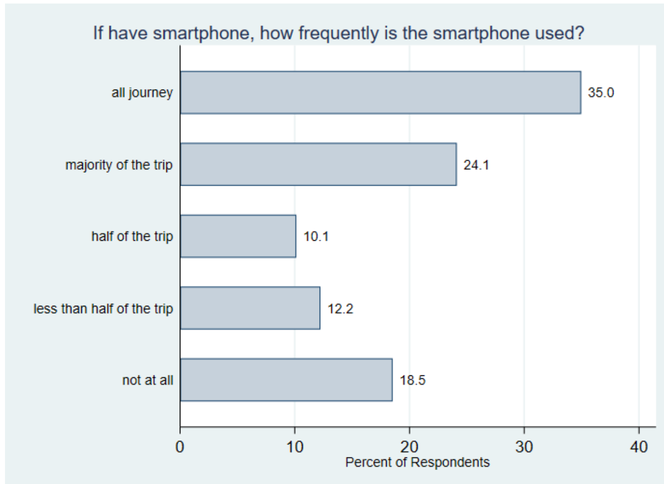


Figure D.11: Smartphone usage frequency.

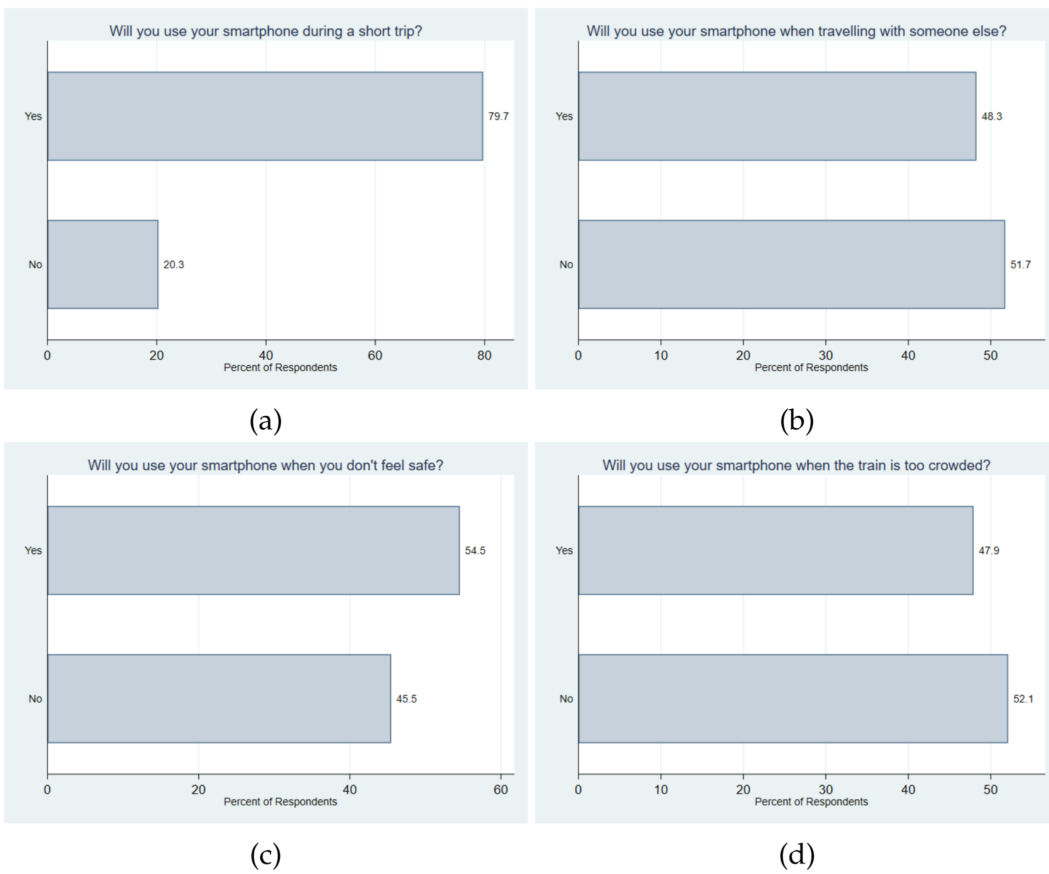
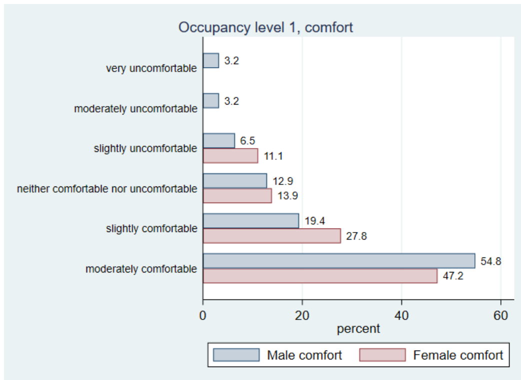
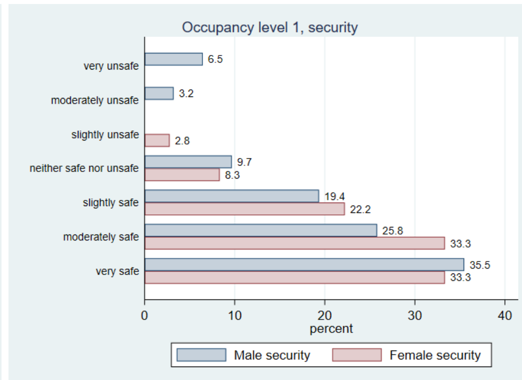


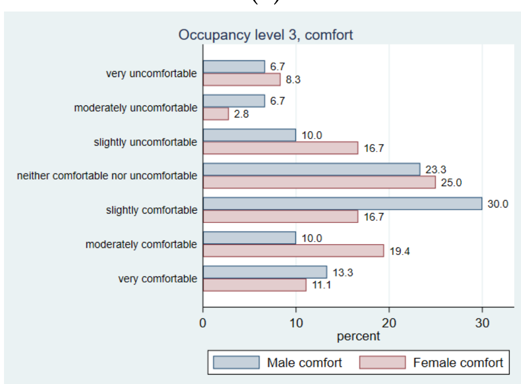
Figure D.12: Situations which the respondents will use smartphone. (a) during a short trip; (b) when travelling with others; (c) when he or she feels unsafe; (d) when the train is too crowded.



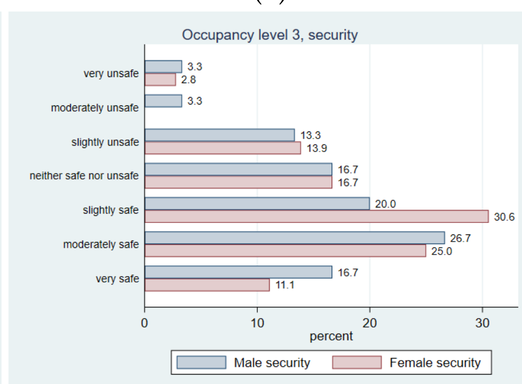
(a)



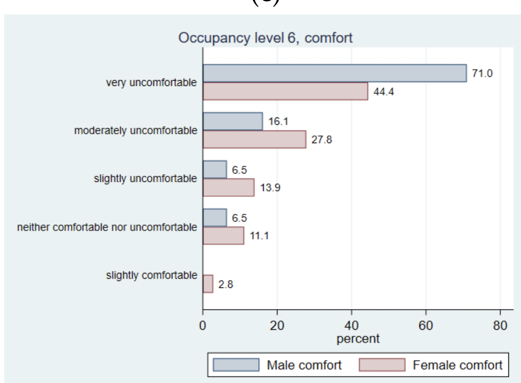
(b)



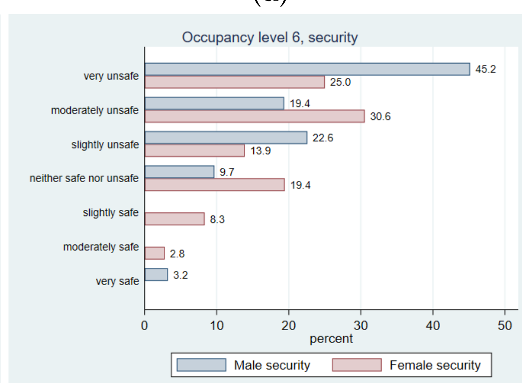
(c)



(d)



(e)



(f)

Figure D.13: Crowding perception: how secure and how comfortable the respondent feels for three different crowding levels. (a) occupancy level low, comfort; (b) occupancy level low, security; (c) occupancy level medium, comfort; (d) occupancy level medium, security; (e) occupancy level high, comfort; (f) occupancy level high, security.

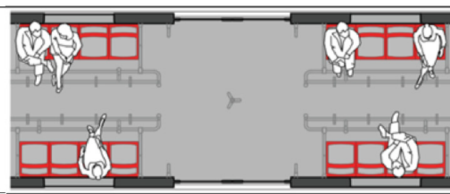
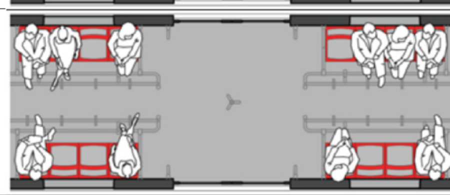
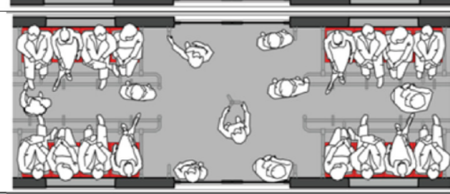
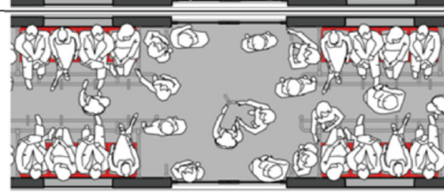
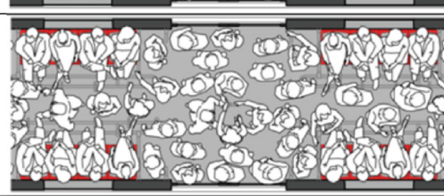
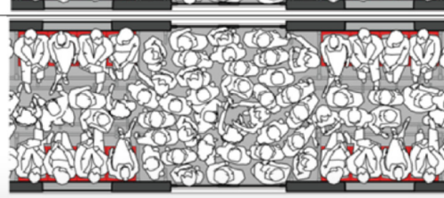
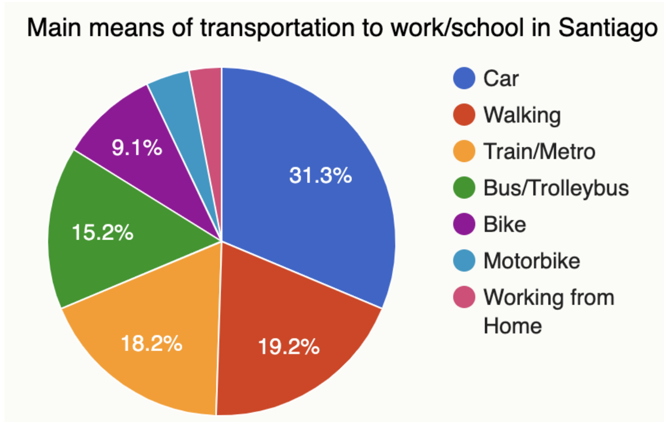
Crowding level	Diagram (shown to respondents)	Description (not shown to respondents)
1		35% seats occupied, 0 standees
2		69% seats occupied, 0 standees
3		100% seats occupied, 1 pax/m2 standing
4		100% seats occupied, 2 pax/m2 standing
5		100% seats occupied, 4 pax/m2 standing
6		100% seats occupied,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

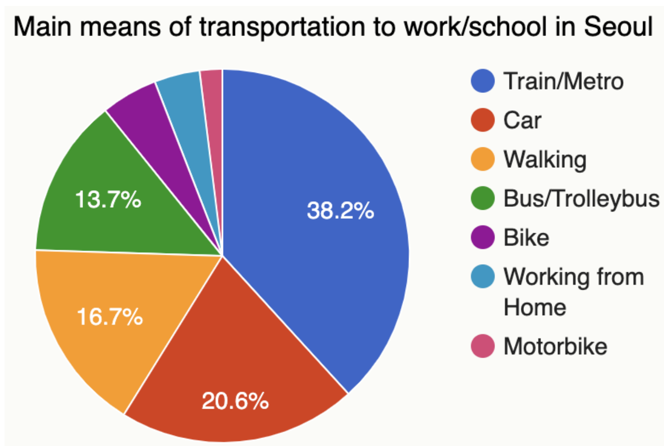
Figure D.14: Crowding levels using 2D diagrams.

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

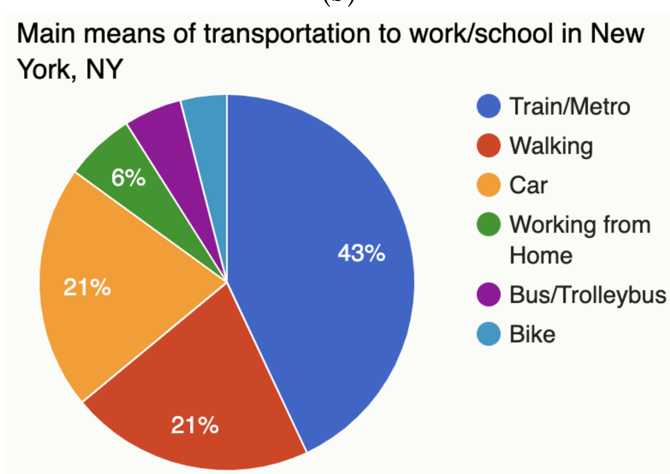
Figure D.15: Crowding levels using photos.



(a)



(b)



(c)

Figure D.16: Main means of transportation to work/school in (a) Santiago, (b) Seoul and (c) New York.