MODELLING USER SATISFACTION IN PUBLIC TRANSPORT SYSTEMS CONSIDERING MISSING INFORMATION

Eneko Echaniz
University of Cantabria, Transport Systems Research Group, E.T.S. Ingenieros de Caminos, Canales y Puertos, Avda. de los castros S/N 39005 Santander, Cantabria (Spain)
email: echanize@unican.es
tlf: +34 615007261

Chinh Ho
The University of Sydney, Institute of Transport and Logistics Studies, The University of Sydney Business School, 378 Abercrombie St, Darlington NSW 2008 (Australia)

Andres Rodriguez
University of Cantabria, Transport Systems Research Group, E.T.S. Ingenieros de Caminos, Canales y Puertos, Avda. de los castros S/N 39005 Santander, Cantabria (Spain)

Luigi dell'Olio
University of Cantabria, Transport Systems Research Group, E.T.S. Ingenieros de Caminos, Canales y Puertos, Avda. de los castros S/N 39005 Santander, Cantabria (Spain)

ABSTRACT

Collecting data to obtain insights into customer satisfaction with public transport services is very time-consuming and costly. Many factors such as service frequency, reliability and comfort during the trip have been found important drivers of customer satisfaction. Consequently, customer satisfaction surveys are quite lengthy, resulting in many interviews not being completed within the aboard time of the passengers/respondents. This paper questions as to whether it is possible to reduce the amount of information collected without a compromise on insights. To address this research question, we conduct a comparative analysis of different Ordered Probit models: one with a full list of attributes vs. one with partial set of attributes. For the latter, missing information was imputed using three different methods that are based on *modes*, single imputations using predictive models and multiple imputation. Estimation results show that the partial model using the Multiple Imputation method behaves in a similar way to the model that is based on the full survey. This finding opens an opportunity to reduce interview time which is critical for most customer satisfaction surveys.

Keywords: Missing information, multiple imputation, user satisfaction, ordered probit, perceived quality.

1. INTRODUCTION

Research on the perceived quality or the satisfaction of the users usually relies on customer satisfaction surveys conducted using a revealed preference survey method. Data collection is usually the most time-consuming and costly part, especially when a face-to-face survey method is used. While this survey method undoubtedly delivers the high data quality, its completion/response rate depends heavily on the interview duration with lengthy questionnaire resulting in a lower response/completion rate. Thus, finding a way to shorten the survey length would improve the effectiveness of customer satisfaction studies. This article proposes a way to do so through a comparative analysis of different models. These models are based on data from customer satisfaction surveys with full and partial list of attributes. Partial dataset is obtained after randomly deleting half of the information available in the original survey. No statistical difference between the two methods will mean that it is possible to reduce the amount of data collected in customer satisfaction surveys. To this end, missing data are imputed using three different methods in order to identify the most adequate method for imputing non-collected information. The first method uses the "mode" of each attribute to fill out the data for respondents who were not shown these attributes. The second uses Ordered Probit models for each attribute and the final method uses Multiple Imputation Process. Different Ordered Probit models are then estimated for the different databases and results compared to check if the models obtained with the partial information databases are correlated with the model based on the complete database.

The remaining of this paper includes 7 sections. The next section summarises a general view of the state of the art regarding the study of satisfaction in public transport systems, reviewing the most relevant studies. Methodology is described in section 3 with analysis results presented in section 4. Section 5 discusses the important findings and identify areas for future research.

2. LITERATURE REVIEW

Satisfaction surveys have proved to be a reliable and robust method to measure the users' perceived quality of public transport systems. Many studies contributed to this fast-growing literature, from a generic analysis of perceived quality (Parasuraman et al. 1985) to the more advanced method that focuses on the provision of public transport services (dell'Olio et al., 2010; dell'Olio et al., 2011; Fellesson and Friman, 2008; Rojo et al., 2013; Wongwiriya et al., 2017).

Most of these studies have focused on identifying key drivers/attributes of the transport system that best describe public transport services. Examples are the Quattro project (EC 1999) which used eight sets of attributes, or the work of Hensher et al. (2003) that employed Service Quality Index (SQI). Another line of research in this literature focused on improving the method used for modelling the data collected. A variety of modelling methods have been

used such as basic statistics (Eboli and Mazzulla 2011), Ordered Data Models (Bordagaray et al. 2014; dell'Olio et al. 2010; Echaniz et al. 2017), structural equations model (Das et al. 2017; de Oña et al. 2013; Rahman et al. 2016) and decision tree (Hernandez et al. 2016; Machado-León et al. 2017; de Oña et al. 2016; Tsami and Nathanail 2017). Recently, we see some exceptions that aim to optimize the data collection. Typical examples are Rose and Bliemer (Rose and Bliemer 2014; Rose and Bliemer 2013; Rose and Bliemer 2009) where efficient stated preference S-design is used to minimize the sample size; however, similar efforts in optimising surveys are not observed in the revealed preference domain.

The data collection process is essential part of any customer satisfaction study that usually use on-board intercept followed by face to face interviews (Bordagaray et al. 2014; dell'Olio et al. 2010; Echaniz et al. 2017) or self-administered questionnaire accessible via QR codes or URL links provided at intercept at public transport stops/stations where the passengers board or alight (Guirao et al. 2015). As for customer satisfaction surveys, the survey duration is a key factor for obtaining valid and quality responses. A long questionnaire generates rich data for the subsequent analysis but this significantly reduces the response/completion rate, resulting in fewer samples for a given budget. Conversely, short surveys can improve sample size at the expense of less data being collected such that statistical model results are not very reliable and robust, since the model cannot control for some important factors that were not collected in the survey. Trade-offs between data richness and budget depends directly on the target sample size, the duration of the survey and the survey method (face to face, online, or app). The aim of this study is to obtain robust models, not by reducing the number of observations, but reducing the amount of data required from each respondent, in other words, reducing the time required to complete each survey. The benefit of reducing surveying time increases as the sample size becomes larger.

The literature shows that user satisfaction studies usually require a long list of factors that requires respondent's feedback, with the data collection lasting from several weeks to even a few months. In Rahman et al., (2016) for example, surveys were conducted to 2008 public transport users during the months of June and July 2015, with a survey consisting of two sections, one for obtaining socio-economic data and another for obtaining the satisfaction of 21 attributes of the system. In Rissel et al., (2016) a total of 512 online surveys were conducted during the months of September and October to obtain information on the mode of transport used by users and the level of satisfaction they had with it. In Guirao et al., (2016) 850 face-to-face surveys were carried out, of which 813 were valid complete answers. The length of the respondent period was 2 weeks. In Tyrinopoulos and Antoniou (2008) 1474 survey where set as minimum to have a sample representative enough for their study, where a total of 5 companies where evaluated. In St-Louis et al. (2014) an online approach was taken. The invitation to participate in the survey was sent via email targeting university staff and students. The response rate was 31.7%, 3377 complete responses from the 20,851 invitation sent. Participation was incentivised by different prizes and all respondents received a reminder 2 weeks after the first email was sent. At the end, the survey was kept

active for 35 days during March and April 2013. In Abenoza et al., (2017) unlike the previous ones, there was a very extensive database obtained by the Swedish Public Transport Barometer for the years between 2001 and 2014 with about 450,000 useful telephone surveys. The aforementioned studies are only a small example of satisfaction studies carried out in the last few years. Hence, it can be seen that public transport satisfaction studies require the completion of a large number of surveys. Therefore, an improvement in the efficiency of this process would considerably improve the total cost of the entire process, as long as the quality of the data allows subsequent analyses.

Regarding the modelling methodology, several studies have shown (Bordagaray et al. 2014; dell'Olio et al. 2010; Dell'Olio et al. 2011; Echaniz et al. 2017; Rojo et al. 2013) that ordered data models are very adequate to model customer satisfaction with public transport services. These models requires a series of very specific data, which are composed by a dependent variable, the overall satisfaction of the service, and independent variables, the attributes of the service. Each respondent must evaluate all the variables, which means that for a survey in which 24 attributes are used to define the system, the respondent must answer at least 25 questions (24 attributes and the overall satisfaction), in addition to background questions relating to individual characteristics such as age and gender. These need of complete data observations have been the main reasons for choosing this modelling method for this study.

To analyse a missing database as if all the data were available, it is necessary to establish a methodology to fill in the missing information. The statistical processes relating to the missing information have evolved considerably in recent years. It has been a subject of many sociology and psychology studies, of which the work of Schafer and Graham (2002) stands out. The authors provide an extensive review of the state of the art regarding the types of missing data and the different imputation methods available. There are several types of missing data according to the nature of the reason why they are missing, according to the classification rooted to Rubin (1976).

In order to classify the current study within this classification it is necessary to understand that the missing information has not been due to a decision of the respondent not to answer a question, but due to the design. That is, part of the available information has been deliberately eliminated to form a reduced survey with fewer questions. Since the elimination of the data has followed a random criterion, the nature of the missingness does not depend on any of the variables belonging to the survey, being a case defined as "missing completely at random" MCAR. In (Graham et al. 1996) this type of scenarios is defined as Planned Missing Value patterns, in other words, survey was intentionally planned to have missing information. When the type of missing information responds to an MCAR nature, simple methods can be applied (Donders et al. 2006). However, in this study we find ourselves in a special MCAR situation, where all the observations have missing data, and therefore there are techniques like listwise deletion (deletion of observations with missing data) that cannot be used. However, within these simple methods (Donders et al. 2006), or also called older

methods (Schafer and Graham 2002), there are several suitable methods for the case that concerns us. On the one hand, there is the possibility of replacing the missing data with the means of the available observations of each variable, in our particular case, as we are dealing with discrete qualitative responses it has been considered more accurate to replace the missing data with the most common response (mode) of that variable. On the other hand, it is possible to apply the single imputation method, based on imputing a single value for each missing data, filling it with a plausible value. This imputation is made by inferring a value for the missing data based on the information that we have available. Although this simple methods can give acceptable results, several studies recommend using more sophisticated methods based on Maximum likelihood (Graham et al. 1996) or Multiple imputation (Donders et al. 2006; Graham et al. 1996). More explicitly, in Donders et al. (Donders et al. 2006) it is demonstrated that the use of the multiple imputation (MI) approach leads to results with correct standard errors, especially in situations where missing data is MCAR. In the same way, Graham and Schafer (1999) showed that MI performs very well in small samples even with as much as 50% missing data.

The MI method was initially developed by Donald Rubin (Rubin 1977) and has proven to be a very effective method to obtain missing data in non-responses. Indeed, MI method is very popular in social sciences and medical studies. Examples in the field of medicine include (Burton et al. 2007; van Buuren et al. 1999; Newgard et al. 2018; Pettersson et al. 2018; Sterne et al. 2009; Troyanskaya et al. 2001) and (Alegria et al. 2004; Allison 2000; König et al. 2018; Love et al. 2018; Phan et al. 2016; Roth 1994; Zou 2015) for social sciences. Very few applications have been found in transport research with exceptions being (Chiou et al. 2014; Henrickson et al. 2015; Li et al. 2015; Tang et al. 2015) that uses MI method to fill the missing data in traffic flows or loop detectors.

Thus, to the best of authors' knowledge, previous studies have mainly focused on filling missing information caused by problems during the surveying process i.e.: non-responses. In this specific case, we propose the possibility of using the same methodology to verify that it is possible to obtain similar results considering a partial sample, similar to what was proposed in (Graham et al. 1996). Moreover, this technique has not been yet applied for the specific case of users' satisfaction in transit services.

3. MODEL SPECIFICATIONS

3.1. Ordered Probit Modelling

The Ordered Probit model was first proposed by McKelvey and Zavoina, 1975, 1971 for the analysis of choices and ordered, categorized or non-quantitative responses.

The ordered data models are based on dividing a continuous utility space (users' satisfaction in this case) in discrete bands through a system of limitations (Greene y Hensher, 2010).

$$q_i^* = \beta' x_i + \varepsilon_i, i = 1, \dots, n,$$
$$q_i = 1 \text{ if } \mu_{-1} < q_i^* \leq \mu_{i1}$$
$$= 2 \text{ if } \mu_0 < q_i^* \leq \mu_{i2}$$
$$= 3 \text{ if } \mu_1 < q_i^* \leq \mu_{i3}$$
$$= \dots$$
$$= J \text{ if } \mu_{J-1} < q_i^* \leq \mu_J.$$

(1)

The key idea of the model is that the observations made are not a simple accumulation of discrete results that can be ordered in a certain way, but consist of a transformation of a single continuous variable that must be ordered.

The model contains the unknown marginal utilities, β, in addition to J + 2 threshold parameters μ_j, all of them to be estimated by n observations. The data consists on the variables x_i of each observation and the resulting observation q_i of each one of them. The random variable εi completes the model. It is assumed that the random variable εi is distributed according to a known distribution function and defined throughout the real domain. Focusing the models on the problem raised in this study. Let's suppose a series of answers available for each of the respondents, where the options are the following:

- 0 Very Bad
- 1 Bad
- 2 Normal
- 3 Good
- 4 Very Good

The regression model shows an underlying and at the same time not observable preference on the evaluated question, q_i^*. Each individual surveyed does not provide the value of q_i^*, but a limited version of it divided into five possible options, one of which is closest to his exact preference. The probabilities associated with the observed responses are:

$$Prob[q_i = j|x_i] = Prob[\varepsilon_i \leq \mu_j - \beta' x_i] - Prob[\mu_{j-1} - \beta' x_i], J = 0,1, \dots, J \quad (2)$$

The established model describes the probability of occurrence of the values of the results. It does not describe a direct relationship between the evaluation q_i and the parameters xi, because there is no obvious regression relationship between both parameters, since q_i is mere a label.

For the estimation of the parameters it is necessary to establish a series of normalizations. First, to keep the positive signs for all probabilities, it is necessary that $\mu_j > \mu_{j-1}$. Second, if the model must exist in the complete real domain, then $\mu_{-1} = -\infty$ y $\mu_j = +\infty$. Since the

data does not contain unconditional information about the scale of the dependent variable (in case of modifying the scale of q_i^* with any positive value, modifying the scale of the unknown values μ_j y β with the same value the characteristics of the observations will remain the same) it is not possible to estimate the free variance parameter $Var[\varepsilon_i] = \sigma_\varepsilon^2$. It is advisable to make a restriction based on $\sigma_\varepsilon = const.$, $\bar{\sigma}$. It is usual to assume variance equal to one in the case of a Probit model and variance equal to $\pi^2/3$ in the case of Logit. Finally, assuming that $\beta'x_i$ has a constant term, it is necessary to set $\mu_0 = 0$. The calculation of the parameters of the models is done by a maximum likelihood estimation (Greene 2007; Greene 2008; Pratt 1981), which equation to maximize is:

$$\log L = \sum_{i=0}^{n} \sum_{j=0}^{J} m_{ij}\log[\, F(\mu_j - \beta'x_i) - F(\mu_{j-1} - \beta'x_i)] \tag{3}$$

3.2. Multiple Imputation for missing information

The goal of the multiple imputation is to complete the missing data, in a way that the resulting data can be statistically analysed and modelled in a similar way to the complete database. The theoretical foundation on which the multiple imputation is based on is the repetitive imputation (Rubin 2004; Rubin 1996; Rubin 1977). This means that for each missing data value m values (as opposed to 1) are imputed. Considering the fact that the missing data have been eliminated randomly, it can be said that the missing data corresponds to a MCAR type, so the use of this method is appropriate (Donders et al. 2006) .

The methodology used to perform the multiple imputation is called the Fully Conditional Specification (FCS), which uses an iterative Monte Carlo method with Markov chains (van Buuren 2007). The FCS approach is based on variable-by-variable imputation of data, specifying an estimation model for each one of the variables with missing data. The FCS tries to define $P(X, C, R|\theta)$ by specifying a conditional density $P(X_i|C, X_{-i}, R, \theta_i)$ for each X_i , this density is used to impute X_i^{mis} given some C, X_{-i} and R. An imputation consists of a complete cycle through all X_i (van Buuren 2007). Where X represents the evaluation of the attributes, C the characterization variables, θ the parameters of the imputation model and R an indicator that show if X is a missing or observed value. The imputation is made by using the Gibbs sampling methodology (Casella et al. 2016; Gilks et al. 1996) assuming that the conditional density distribution exists. This methodology has been used in a large number of simulation studies (Brand 1999; Brand et al. 2003; Van Buuren et al. 2006; Horton et al. 2016; Raghunathan et al. 2001) that have provided sufficient evidence that the results obtained through the FCS are generally unbiased and have adequate coverage.

In order to optimize the imputation process it has been assumed that the satisfaction data is a scale type variable with values between 0 and 4, so the imputation model follows a linear regression methodology rounding to the nearest whole value. This has enabled the imputed values to match the actual values of the data. It has been proven that the predictive mean

matching procedure, a variant of the linear regression that equals the imputed values with the closest observed value, generates worse results. In Graham et al. (2007) they recommend a high number of imputations for these types of cases. However it has been empirically verified that, for this specific practical case, with 5 imputations the results are acceptable enough. So it has been decided to maintain this number of imputations (5) mainly for efficiency reasons. For the regression model, Y_j corresponds with the attributes with missing evaluations and X with all the socioeconomic variables (Table 1) plus the overall satisfaction of the service.

3.3. Comparison

According to the final objective of this study, in which it is intended to analyse if it is possible to obtain similar results based on a partial information database, 3 methodologies are proposed to perform the modelling.

The starting point is the model that it will be called BASE, which is estimated considering the complete database. For the rest of the models, half of the satisfaction data have been randomly eliminated, that is, creating a hypothetical scenario where only 12 of the 24 attributes would have been answered by the respondents. For modelling, missing information need to be fulfilled so 3 different methods have been used to achieve that.

The first method is based on using the "mode" of the answers to complete the missing information of each attribute. That is, to use the most common value among the respondents for each attribute. In other words, the satisfaction of a user who does not make the evaluation of an attribute will be equal to the most commonly chosen value by those that did evaluate it. This model will be called MODE throughout the rest of the article.

The second method consists of estimating J Ordered Probit models, one for each of the attributes. This way, a missing satisfaction value is imputed from a model estimated with the existing responses for that attribute, based on the socioeconomic characteristics of the people who have evaluated it. Being explained as:

$$y_{ji}^* = \delta_{ji}\, \beta' x_i + \varepsilon_i, i = 1, \dots, n, y\ j = 1, \dots, 24, \dots \qquad (4)$$

Where, y_{ji}^* represents each one of the 24 evaluated attributes and x_i the different socioeconomic variables; δ_{ji} obtains the value 1 if the attribute j is evaluated by the respondent i and 0 otherwise, up to a maximum of $\sum_{j=1}^{24} \delta_{ij} = 12$ per respondent, since it has been assumed that in the restricted version of the survey the respondents would only perform half of the evaluation exercises . This model will be called ATTRIBUTE throughout the article.

Finally, the last method used to complete the missing data have been through the use of the multiple imputation procedure (section 3.2). As indicators to infer the missing data, both the socioeconomic variables and the evaluations made to all the attributes have been used, as well as the overall satisfaction of the service. A total of 5 imputations have been carried out with 100 interactions each. The results of the Multiple Imputation consist in the generation of 5 new databases, 1 for each imputation. In order to obtain a single model, an OP model have been estimated for each one of these databases and then the average of the parameters have been use for the comparison. This model will be referred by the acronym MI.

4. CASE STUDY

4.1. Satisfaction survey

The data used in this study was obtained from a satisfaction survey carried out in 2015 in the city of Santander, a small-medium size coastal city located in the north of Spain. At the time when the survey was conducted, the city had around 173,000 inhabitants with the metropolitan area reaching 240,000 residents. Buses are the only public transport of the city. The bus network has 22 lines, of which 16 were surveyed in this study.

Field surveys were conducted over 15 working days in the months of April and May 2015. The surveys were carried out on board using a face to face method. In case the survey could not be completed during the respondent's journey, the interviewer had two options: either leaving the bus with the respondent and finishing the survey at the stop or discarding the survey and find another respondent on board. If the respondent chose the former option, they then wait for the next bus to come, and then continue on-board recruitments and interviews. In both cases the efficiency of the survey process was affected. The minimum sample size n was set in 700 completed surveys, being calculated by using equation (5) (Bordagaray et al. 2014; dell'Olio et al. 2010; Echaniz et al. 2018). For which, the most conservative value was taken: p =0.5. In the end, a total of 747 complete observations were obtained with a ratio of approximately 4 complete surveys per hour per interviewer.

$$n \geq \frac{p(1-p)}{\left(\frac{e}{z}\right)^2 + \frac{p(1-p)}{N}}$$

$Where$:
e: $standard\ error\ of\ the\ parameter\ estimate$
z: $critical\ z-statistic\ value\ ofr\ the\ desired\ level\ of\ confience\ (eg\ 95\%)$
N: $Number\ of\ passengers\ at\ rush\ hour$
p: $probability\ associated\ with\ the\ choice$

(5)

The survey included two main parts with the collected information summarised in Table 1. The first part seeks the respondent's socioeconomic characteristics and usage of public

transport services. The second part focuses on user overall satisfaction with public transport service (OS) and on a subset of attributes that represent different aspects of the service. Level of satisfaction was measured using a 5 point Likert scale.

**Table 1 - Variables included in the survey**

Characterization	PT Service attributes
Gender	Access time to the stop (AT)
Age	Waiting time at the stop (WT)
Work status	On board travel time (TT)
Driving license ownership	Egress time, last stop to final destination (TD)
Car ownership	price/fare (PR)
Trip purpose	Ease of transfer (TR)
Frequency of use	Service frequencies (SE)
Usual payment system	Service reliability/punctuality (SR)
Monthly income	Special lines (EL)
	Night services / weekend services (NS)
	Line coverage (LC)
	Information at stops (IS)
	Information on webpage and mobile platforms (IWM)
	Information on board (IB)
	Occupancy level (OC)
	Air conditioned vehicles (CA)
	Priority seats for people with disability (RM)
	Comfort of the buses (CM)
	Cleanliness (CL)
	Possibility to carry large objects (OB)
	Driving style (DS)
	Driver kindness (DK)
	Hybrid buses (HY)
	Noise (NO)

The sample was made up of 71% women, who are over-represented. Two thirds of the respondents are under 44 years old and nearly half working (49%) with a further quarter studying part-time or full-time. About six in ten respondents having a driving license (59%) but only four in ten own a car. Regarding the use of public transport service, regular users, defined as those using bus services between 5 and 15 times per week, accounts for nearly half of the sample where the main reason for bus travelling is commuting (work or study). The vast majority of the respondents use contactless card with cash payment accounting for only 5% of the sample. Regarding personal income level, a majority of the respondents have low to medium income levels, with high income respondents accounting for only 8%. A little more than the third part of the respondents (38%) preferred not to answer this question, a usual result since it is a very sensitive question. Table 2 provides a summary of the respondents

Table 2 - Descriptive analysis

Gender	Male	29%
	Female	71%
Age	<25	28%
	25-34	16%
	35-44	22%
	45-54	16%
	55-64	11%
	>65	7%
Work status	Worker	49%
	Unemployed	17%
	Student	25%
	Retired	9%
Driving license	Yes	59%
	No	41%
Car ownership	Yes	40%
	No	60%
Frequency of use	< 5 trips/week	29%
	5 - 15 trips/week	50%
	15 - 30 trips/week	18%
	> 30 trips/week	3%
Trip purpose	Home	32%
	Work	22%
	Study	13%
	Health	4%
	Shopping	7%
	Leisure	11%
	Other	11%
Payment system	Contactless transport card	95%
	Cash	5%
Monthly income	Low (< 900€)	31%
	Medium (900 - 1500€)	23%
	High (1500 - 2500€)	7%
	Very high (> 2500€)	1%
	Unknown (not reported)	38%

Regarding user satisfaction with the public transport service, **Figure** 1 shows the level of satisfaction of respondents. For brevity, the 5-point Likert scales are coded from 0 to 4, with 0 being "Very Bad" and 4 being "Very good". Additionally, since the aim of the study is to analyse the possibility of obtaining similar results based on a reduced data base. A comparison was made comparing, on one hand, the average satisfactions obtained for the different attributes through the complete database. And on the other hand, the average value of the satisfaction obtained after eliminating half of the available information (partial database), just as it has been done for the modelling process (section 3.3).
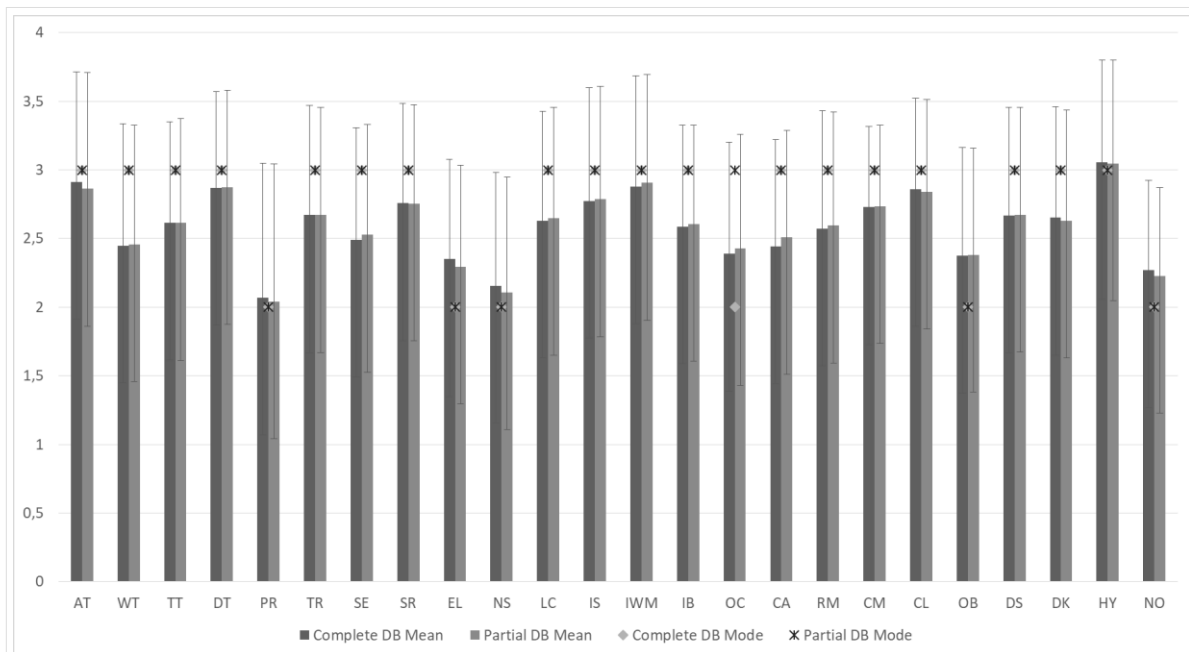
Figure 1 – Users' satisfaction

The results show that users are generally satisfied with the service and with all aspects that describe the services they use. The attribute that is considered as worst is the fare. This fact can be understood as a strategic response, since the users do not tend to evaluate this attribute well for fear of a possible increase of the service fares, still, the average value shows that it is not considered as an unsatisfactory factor for users. On the contrary, an attribute that valued the most is the use of hybrid buses. Any action associated with an environmental improvement of the service is generally considered good by the users.

The comparison made between the two databases shows that, even having half of the information, the average difference in satisfaction level between the two datasets is small, with differences in means being less than 3% in all cases. The biggest difference in the mode is found in the occupation attribute, where the mode changes from a "normal" evaluation (value 2) to a "Good" evaluation (value 3). One possible reason for this would be a random elimination of the attributes, the worse attributes may have been eliminated. However, this difference is only shown in one variable of the whole set of attributes so it can be considered an outlier. The standard deviations also show small differences, usually less than 3% with the exception of the Access time to the stops (AT) which shows a variation close to 6%. We can safely conclude that based on the results shown in Figure 1 the average level of satisfaction are very similar between the two datasets.

4.2. Modelling results

Four Ordered Probit models were developed. One model was estimated using the complete dataset and this model is referred to as BASE. The remaining three models are developed from the partial dataset obtained after randomly deleting 50% of the evaluations made in the

original survey, which resembles a hypothetical survey where only half of the attributes would have been answered by each respondent. Missing information was imputed using the three methods described in section 3.4.

The attributes included in each model have been selected following a step-by-step process until the resulting parameters have the correct sign (positive sign except for the constant that must be negative (Echaniz et al. 2017)) and are statistically significant. This can be seen in where t-values are included in parentheses. Significant parameters are shown in bold (at least at a 10% level) so that similarities and differences between models can be spotted easier. The significant parameters are largely similar between different models. The most similar model to BASE in terms of the significance level of the parameters is the MODE model, where 79% of the 24 parameters have the same level of significance compared to BASE. The model calculated using the database fulfilled by attribute specific models (ATTRIBUTE) shows a lower correlation, with 71% of the parameters showing a similar significance level. Finally, the MI model, derived from Multiple Imputation, lie in the middle with a coincidence of 75% of the parameters.

Only two threshold parameters are shown in the models because there was no "Very Bad" evaluation observed in the survey for the dependent variable (OS). Instead, the value 0 now represents the grouping of the "Very Bad" and "Bad" responses.

Table 3 - Ordered probit models: parameter estimates and t-values

Parameters	Base	Mode	Attribute	MI[1]
Constant	**-4.37** **(-11.22)**	**-7.81** **(-11.89)**	**-8.47** **(-15.25)**	**-3.81** **(-10.82)**
Access time to the stop	**0.23** **(3.71)**	**0.26** **(3.35)**	**0.35** **(4.37)**	**0.27** **(5.05)**
Waiting time at the stop	**0.13** **(2.14)**	**0.18** **(2.59)**	**0.36** **(4.73)**	**0.22** **(4.37)**
On board travel time	**0.18** **(2.47)**	**0.25** **(3.11)**	**0.39** **(4.52)**	**0.16** **(3.37)**
Egress time	**0.12** **(1.65)**	**0.24** **(2.56)**	-	-
Price/fare	0.07 (1.37)	0.09 (1.34)	0.07 (1.14)	-
Ease of transfer	**0.26** **(4.2)**	**0.31** **(3.93)**	**0.36** **(4.23)**	**0.24** **(4.07)**
Service frequencies	0.09 (1.29)	**0.29** **(3.88)**	0.13 (1.53)	**0.19** **(3.21)**
Service reliability	**0.19** **(2.69)**	**0.30** **(3.49)**	**0.30** **(3.26)**	**0.02** **(1.7)**
Special lines	-	-	-	**0.1**

				(2.86)
Night/weekend services	0.14 (2.38)	-	0.22 (2.78)	0.12 (2.35)
Line coverage	0.25 (3.92)	0.32 (4.2)	0.31 (3.91)	0.14 (2.94)
Information at stops	0.11 (1.78)	0.22 (2.82)	0.19 (2.27)	0.02 (2.44)
Information on webpage and Mobile platforms	-	0.19 (2.32)	0.27 (3.28)	0.08 (1.64)
Information on board	-	-	0.31 (3.56)	-
Occupancy level	-	-	0.18 (2.41)	0.03 (2.85)
Air conditioned vehicles	0.26 (3.99)	0.16 (2.14)	0.37 (4.52)	0.19 (3.58)
Priority seats for people with disability	-	-	0.24 (3.09)	0.04 (4.2)
Comfort of the buses	0.37 (4.29)	0.30 (2.99)	-	0.42 (5.57)
Cleanliness	-	-	-	-
Possibility to carry large objects	-	0.16 (2.24)	-	-
Driving style	0.30 (4.55)	0.23 (2.96)	0.25 (3.04)	0.25 (4.28)
Driver kindness	-	0.14 (1.82)	-	-
Hybrid buses	-	-	-	-
Noise	-	-	0.21 (2.24)	-
Mu(01)	1.73 (17.21)	1.52 (18.02)	2.31 (17.01)	1.73 (17.21)
Mu(02)	4.60 (33.49)	4.09 (37.05)	5.55 (30.2)	4.6 (33.37)
Log-likelihood	-494.86	-675.79	-603.9	-510.26
Psuedo-R^2	0.73	0.63	0.70	0.71
Degrees of freedom	17	19	20	15

Both the evaluation of general satisfaction (OS) and the satisfaction of the attributes has been measured following the same Likert scale. Therefore a comparison between the parameters of the same model can be made, understanding that a parameter of greater value will give greater importance to its corresponding attribute. The most influencing parameter is comfort (CM), which shows the highest parameter value for BASE, ATTRIBUTE and MI models. The comfort on board the bus is followed by the driving style (DS), which also represents how comfortable the ride is. Without considering the variables that turn out to be highly statistically insignificant, the ticket price (PR) show the lower parameter value, which means that the price to get to the service is not really important to define users' satisfaction. Service

related attributes (travel time, waiting time…) show medium level values and those attributes that are clearly additional to the basic service, such as, special lines, cleanliness or noise turn out not statistically significant. The trend is similar in all models. Thereupon, it can be said that user satisfaction is highly defined by the comfort during the trip. Users could be used to the actual level of service and see it as acceptable. In such a way that users feeling more satisfied would come from attributes related with how comfortable the trip is.

When comparing different models, even taking data from similar sources and based on the same scale, because constant values and threshold parameters are different, a direct comparison is not possible. For this reason, all model parameters have been standardized before comparing them. In Figure 1 it can be seen that the correlation between the normalized models is considerable. In this case, without going into detail of each parameter individually, it can be seen that the MI model shows a similar trend to the BASE model. That is, the normalized parameters vary jointly, obtaining high values in the MI model when the values are high in the BASE model and vice versa.
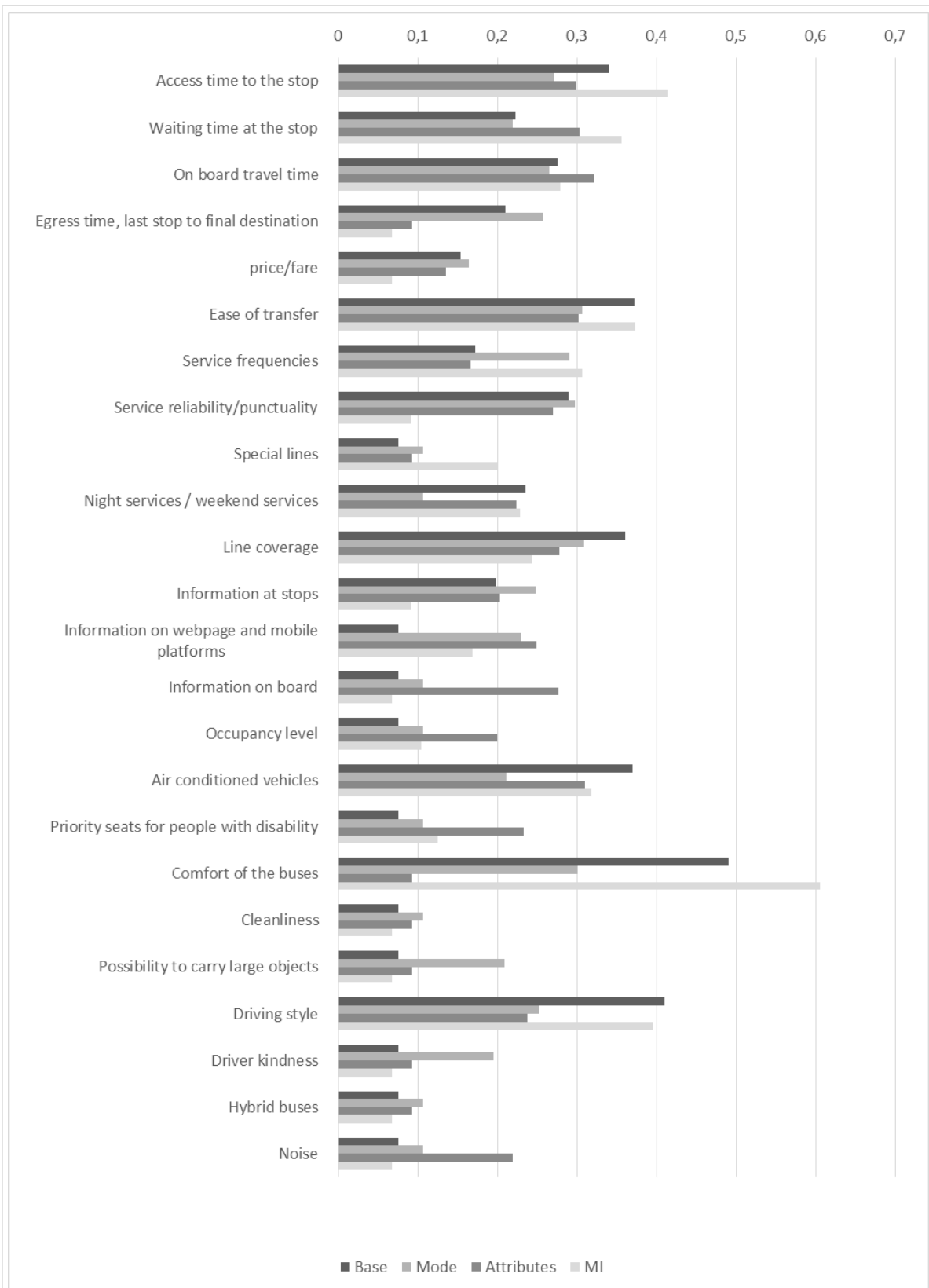
Figure 2 - Normalized model parameters: a comparison of models with full vs. partial data

This is not the case for all the variables. There are some cases where the correlation between these models is not weak. For example: egress time (DT) and bus fare (PR) are not significant

in the MI model but they are in the BASE model. The Pearson correlation coefficients calculated between models show a very high correlation, coefficient values compared to BASE model are 0.95 for MODE, 0.97 for ATTRIBUTE and 0.99 for MI.

Finally, it is essential to compare the explanatory ability that have been achieved with the different models based on goodness of fit indicators, such as Log-Likelihood value or count R^2 value (Echaniz et al. 2017; Greene and Hensher 2010).   shows that the prediction capability of the BASE model is the best one, which shows a Log-likelihood value of -494.86 and count R2 value of 0.73, that is, the model is able to explain 73% of the variation observed in the data. None of the models with missing data can match the prediction capacity that is obtained with the BASE model; however, the differences are relatively small, with the MI model having1% less in predictive power. The MODE model has the worst predictive power, where the loss of accuracy is up to 10%.

An additional to predictive power, we carry a Vuong test (Vuong 1989) to compare the models with results shown in table 4. Z values close to 0 mean that the two models behave similarly. Absolute values greater than $Z = 1.96$ consider that the two models exhibit different behaviour at the 95% level of confidence. As can be seen, the BASE model based on fill information outperforms alternative models fitted with partial information. The only model that behaves statistically similar to the Base model the MI model, with a z-value of 0.39. Therefore, MI model can be considered a slightly worse model than the BASE model but not different in a statistical sense. The other two models show values considerably larger than 1.96, and thus they are expected to behave differently from the Base model.

Table 4 - Vuong test for non-nested models

Z (Model 1 vs Model 2)		Model 2			
		Base	Mode	Attribute	MI
Model 1	Base	0.00	7.77	4.23	0.39
	Mode	-	0.00	-2.29	-7.21
	ATT	-	-	0.00	-3.83
	MI	-	-	-	0.00

5. CONCLUSIONS

This paper has shown a method to analyse public transport users' satisfaction based on partial information data. In addition, the empirical evidence included in this paper has shown that Ordered Probit Models, widely used in the analysis of users' satisfaction, can be estimated from a partial database with a minimum loss of information.

It has been observed that, even considering half of the available data, the descriptive analysis of attribute evaluations suffers a very small variation. Therefore, it can be said that it is not necessary to collect all the data if what is wanted is to simply study the average satisfaction of the users. This is a common practice among public transport operators in order to obtain

a clear picture of users' satisfaction towards their services. Thus, the lessons learned in this study could provide a considerable economic advantage for companies by requiring less time and resources to conduct the surveys. Having said that, it is necessary to emphasize that the results obtained in this study belong to a medium-sized coastal city, with a single public transport system (and operator) based on bus services, where the averages and deviations of the evaluations of the attributes and overall satisfaction of transport system are those that have been shown throughout the article. In consequence, caveats should be taken when extrapolating this study to cities or other public transport modes without prior analysis.

Regarding the modelling, the best methodology to fulfil the missing data turn out to be the Multiple Imputation (MI), which has allowed to obtain similar results to the ones obtained with the complete data. Vuong test carried out has shown that both models (the one obtained with the complete dataset and the one obtained after applying MI to impute the missing information) behave similarly.

The main output of this study is that the comparison between the models has shown that there is the possibility of obtaining very similar results with very similar fits to reality even starting from a partial information datasets. This allows to optimize the resources so that the time and the cost of the surveys can be reduced to a great extent, reducing the loss of information caused from the modelling of the data. Future studies will be focussed on reducing even more the need of data by applying different methodologies in order to obtain similar results as the ones obtained with ordered models.

Another way to optimize the surveying process is to study the possibility of applying different methodologies to obtain similar results obtained with the models used in this study, an issue that will be addressed in future studies.

6. ACKNOWLEDGEMENTS

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

AUTHORS' CONTRIBUTION
E Echaniz: Modelling and manuscript writing C Ho: Manuscript writing and reviewing A Rodriguez: Data gathering and modelling L dell'Olio: Manuscript reviewing and editing.

7. REFERENCES

Abenoza, R.F., Cats, O., Susilo, Y.O.: Travel satisfaction with public transport: Determinants, user classes, regional disparities and their evolution. Transportation Research Part A: Policy and Practice. 95, 64–84 (2017). doi:10.1016/j.tra.2016.11.011

Alegria, M., Takeuchi, D., Canino, G., Duan, N., Shrout, P., Meng, X.-L., Vega, W., Zane, N., Vila, D., Woo, M., Vera, M., Guarnaccia, P., Aguilar-Gaxiola, S., Sue, S., Escobar, J., Lin, K.-M., Gong, F.: Considering Context, Place, and Culture: The National Latino and Asian American Study. Int J Methods Psychiatr Res. Int J Methods Psychiatr Res. 13, 208–220 (2004)

Allison, P.D.: Multiple imputation for missing data: A cautionary tale. Sociological Methods and Research. 28, 301–309 (2000). doi:10.1177/0049124100028003003

Bordagaray, M., Dell'Olio, L., Ibeas, A., Cecín, P.: Modelling user perception of bus transit quality considering user and service heterogeneity. Transportmetrica A: Transport Science. 10, 705–721 (2014)(a). doi:10.1080/23249935.2013.823579

Brand, J.: Development , implementation and evaluation of multiple imputation strategies for the statistical analysis of incomplete data sets. (1999)

Brand, J.P.L., Van Buuren, S., Groothuis-Oudshoorn, K., Gelsema, E.S.: A toolkit in SAS for the evaluation of multiple imputation methods. Statistica Neerlandica. 57, 36–45 (2003). doi:10.1111/1467-9574.00219

Burton, A., Billingham, L.J., Bryan, S.: Cost-effectiveness in clinical trials: using multiple imputation to deal with incomplete cost data. Clinical Trials: Journal of the Society for Clinical Trials. 4, 154–161 (2007). doi:10.1177/1740774507076914

van Buuren, S.: Multiple imputation of discrete and continuous data by fully conditional specification. Statistical Methods in Medical Research. 16, 219–242 (2007). doi:10.1177/0962280206074463

van Buuren, S., Boshuizen, H.C., Knook, D.L.: Multiple imputation of missing blood pressure covariates in survival analysis. Statistic in Medicine. 18, 681–694 (1999). doi:10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R [pii]

Van Buuren, S., Brand, J.P.L., Groothuis-Oudshoorn, C.G.M., Rubin, D.B.: Fully conditional specification in multivariate imputation. 76, 1049–1064 (2006). doi:10.1080/10629360600810434

Casella, G., George, E.I., Casella, G., George, E.I.: Explaining the Gibbs Sampler Stable URL : http://www.jstor.org/stable/2685208 Linked references are available on JSTOR for this article : Explaining the Gibbs Sampler. 3, 167–174 (2016)

Chiou, J.-M., Zhang, Y.-C., Chen, W.-H., Chang, C.-W.: A functional data approach to missing value imputation and outlier detection for traffic flow data. Transportmetrica B: Transport Dynamics. 2, 106–129 (2014). doi:10.1080/21680566.2014.892847

Das, T., Apu, N., Hoque, M.S.S., Hadiuzzaman, M., Xu, W.: Parameters Affecting the Overall Performance of Bus Network System at Different Operating Conditions: A Structural Equation Approach. In: Transportation Research Procedia. pp. 5063–5075. Elsevier B.V. (2017)

Dell'Olio, L., Ibeas, A., Cecin, P.: The quality of service desired by public transport users. Transport Policy. 18, 217–227 (2011). doi:10.1016/j.tranpol.2010.08.005

dell'Olio, L., Ibeas, A., Cecín, P.: Modelling user perception of bus transit quality. Transport Policy. 17, 388–397 (2010). doi:10.1016/j.tranpol.2010.04.006

Donders, A.R.T., van der Heijden, G.J.M.G., Stijnen, T., Moons, K.G.M.: Review: A gentle introduction to imputation of missing values. Journal of Clinical Epidemiology. 59, 1087–1091 (2006). doi:10.1016/j.jclinepi.2006.01.014

Eboli, L., Mazzulla, G.: A methodology for evaluating transit service quality based on

subjective and objective measures from the passenger's point of view. Transport Policy. 18, 172–181 (2011). doi:10.1016/j.tranpol.2010.07.007

EC: Quattro Final Report. , Paris (1999)

Echaniz, E., Dell'Olio, L., Ibeas, Á.: Modelling perceived quality for urban public transport systems using weighted variables and random parameters. Transport Policy. 67, 31–39 (2018). doi:10.1016/j.tranpol.2017.05.006

Echaniz, E., Dell'Olio, L., Ibeas, Á., Dell 'olio, L., Ibeas, Á.: Modelling perceived quality for urban public transport systems using weighted variables and random parameters. Transport Policy. (2017). doi:10.1016/j.tranpol.2017.05.006

Fellesson, M., Friman, M.: Perceived satisfaction with public transport service in Nine European cities. The Transportation Research Forum. 47, 874770 (2008). doi:10.5399/osu/jtrf.47.3.2126

Gilks, W.R. (Wally R.., Richardson, S. (Sylvia), Spiegelhalter, D.J.: Markov chain Monte Carlo in practice. Chapman & Hall (1996)

Graham, J.W., Hofer, S.M., MacKinnon, D.P.: Maximizing the Usefulness of Data Obtained with Planned Missing Value Patterns: An Application of Maximum Likelihood Procedures. Multivariate Behavioral Research. 31, 197–218 (1996). doi:10.1207/s15327906mbr3102_3

Graham, J.W., Olchowski, A.E., Gilreath, T.D.: How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. Prevention Science. 8, 206–213 (2007). doi:10.1007/s11121-007-0070-9

Graham, J.W., Schafer, J.L.: On the performance of multiple imputation for multivariate data with small sample size. Statistical strategies for small sample research. 50, 1–27 (1999)

Greene, W.: Limdep computer program: Version 9. Plainview, NY: Econometric Software. (2007)

Greene, W.H.: Econometric Analysis, 6th edn. Prentice Hall (2008)

Greene, W.H., Hensher, D. a.: Modeling Ordered Choices: A Primer. (2010)

Guirao, B., Eugenia López, M., Comendador, J.: New QR Survey Methodologies to Analyze User Perception of Service Quality in Public Transport: The Experience of Madrid. Journal of Public Transportation. 18, (2015)

Guirao, B., García-Pastor, A., López-Lambas, M.E.M.E., Garc??a-Pastor, A., L??pez-Lambas, M.E., García-Pastor, A., López-Lambas, M.E.M.E.: The importance of service quality attributes in public transportation: Narrowing the gap between scientific research and practitioners' needs. Transport Policy. 49, 68–77 (2016). doi:10.1016/j.tranpol.2016.04.003

Henrickson, K., Zou, Y., Wang, Y.: Flexible and Robust Method for Missing Loop Detector Data Imputation. Transportation Research Record: Journal of the Transportation Research Board. 2527, 29–36 (2015). doi:10.3141/2527-04

Hensher, D.A., Stopher, P., Bullock, P.: Service quality - developing a service quality index in the provision of commercial bus contracts. Transportation Research Part A: Policy and Practice. 37, 499–517 (2003). doi:10.1016/S0965-8564(02)00075-7

Hernandez, S., Monzon, A., de Oña, R.: Urban transport interchanges: A methodology for evaluating perceived quality. Transportation Research Part A: Policy and Practice. 84, 31–43 (2016). doi:10.1016/j.tra.2015.08.008

Horton, N.J., Lipsitz, S.R., Horton, N.J., Lipsitz, S.R.: Multiple Imputation in Practice : Comparison of Software Packages for Regression Models with Missing Variables Statistical Computing Software Reviews Multiple Imputation in Practice : Comparison of Software Packages for Regression Models With Missing Vari. 55, 244–254 (2016)

König, H.-H., Bleibler, F., Friederich, H.-C., Herpertz, S., Lam, T., Mayr, A., Schmidt, F.,

Svaldi, J., Zipfel, S., Brettschneider, C., Hilbert, A., de Zwaan, M., Egger, N.: Economic evaluation of cognitive behavioral therapy and Internet-based guided self-help for binge-eating disorder. International Journal of Eating Disorders. 51, 155–164 (2018). doi:10.1002/eat.22822

Li, L., Su, X., Zhang, Y., Lin, Y., Li, Z.: Trend Modeling for Traffic Time Series Analysis: An Integrated Study. IEEE Transactions on Intelligent Transportation Systems. 16, 3430–3439 (2015). doi:10.1109/TITS.2015.2457240

Love, H.A., Durtschi, J.A., Ruhlmann, L.M., Nelson Goff, B.S.: Army Soldiers and Suicidal Thoughts: The Impact of Negative Relationship Dynamics Moderated by the Dissolution of Romantic Relationships. Journal of Marital and Family Therapy. 44, 265–276 (2018). doi:10.1111/jmft.12252

Machado-León, J.L., de Oña, R., Baouni, T., de Oña, J.: Railway transit services in Algiers: priority improvement actions based on users perceptions. Transport Policy. 53, 175–185 (2017). doi:10.1016/j.tranpol.2016.10.004

McKelvey, R.D., Zavoina, W.: Ibm Fortran-Iv Program To Perform N-Chotomous Multivariate Probit Analysis, (1971)

McKelvey, R.D., Zavoina, W.: A statistical model for the analysis of ordinal level dependent variables. The Journal of Mathematical Sociology. 4, 103–120 (1975). doi:10.1080/0022250X.1975.9989847

Newgard, C.D., Malveau, S., Zive, D., Lupton, J., Lin, A.: Building A Longitudinal Cohort From 9-1-1 to 1-Year Using Existing Data Sources, Probabilistic Linkage, and Multiple Imputation: A Validation Study. Academic Emergency Medicine. acem.13512 (2018). doi:10.1111/acem.13512

de Oña, J., de Oña, R., Eboli, L., Mazzulla, G.: Perceived service quality in bus transit service: A structural equation approach. Transport Policy. 29, 219–226 (2013). doi:10.1016/j.tranpol.2013.07.001

de Oña, J., de Oña, R., López, G.: Transit service quality analysis using cluster analysis and decision trees: a step forward to personalized marketing in public transportation. Transportation. 43, 725–747 (2016). doi:10.1007/s11116-015-9615-0

Parasuraman, A., Zeithaml, V.A., Berry, L.L., Parasuraman: A conceptual model of service quality and its implications for future research, http://www.jstor.org/stable/1251430?origin=crossref, (1985)

Pettersson, M.E., Koppelman, G.H., Flokstra-de Blok, B.M.J., Kollen, B.J., Dubois, A.E.J.: Prediction of the severity of allergic reactions to foods. Allergy. 73, 1532–1540 (2018). doi:10.1111/all.13423

Phan, M.H., Keebler, J.R., Chaparro, B.S.: The Development and Validation of the Game User Experience Satisfaction Scale (GUESS). Human Factors: The Journal of the Human Factors and Ergonomics Society. 58, 1217–1247 (2016). doi:10.1177/0018720816669646

Pratt, J.W.: Concavity of the Log Likelihood. Journal of the American Statistical Association. 76, 103–106 (1981). doi:10.1080/01621459.1981.10477613

Raghunathan, T.E., Lepkowski, J.M., Van Hoewyk, J., Solenberger, P.: A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology. 27, 85–95 (2001)

Rahman, F., Das, T., Hadiuzzaman, M., Hossain, S.: Perceived service quality of paratransit in developing countries: A structural equation approach. Transportation Research Part A: Policy and Practice. 93, 23–38 (2016). doi:10.1016/j.tra.2016.08.008

Rissel, C., Crane, M., Wen, L.M., Greaves, S., Standen, C.: Satisfaction with transport and enjoyment of the commute by commuting mode in inner Sydney. Health Promotion Journal of Australia. 27, 80–83 (2016). doi:10.1071/HE15044

Rojo, M., dell'Olio, L., Gonzalo-Orden, H., Ibeas, Á.: Interurban bus service quality from

the users' viewpoint. Transportation Planning and Technology. 36, 599–616 (2013). doi:10.1080/03081060.2013.845432

Rose, J.M., Bliemer, M.C.J.: Constructing Efficient Stated Choice Experimental Designs. Transport Reviews. 29, 587–617 (2009). doi:10.1080/01441640902827623

Rose, J.M., Bliemer, M.C.J.: Sample size requirements for stated choice experiments. Transportation. 40, 1021–1041 (2013). doi:10.1007/s11116-013-9451-z

Rose, J.M., Bliemer, M.C.J.: Stated choice experimental design theory : The who , the what and the why. Handbook of Choice Modelling. 152–177 (2014)

Roth, P.L.: MISSING DATA: A CONCEPTUAL REVIEW FOR APPLIED PSYCHOLOGISTS. Personnel Psychology. 47, 537–560 (1994). doi:10.1111/j.1744-6570.1994.tb01736.x

Rubin, D.B.: Inference and missing data. Biometrika. 63, 581–592 (1976). doi:10.1093/biomet/63.3.581

Rubin, D.B.: Formalizing Subjective Notions About the Effect of Nonrespondents in Sample Surveys Formalizing Sub jective Notions About the Effect of Nonrespondents in Sample Surveys. Source Journal of the American Statistical Association. 72144202, 538–543 (1977)

Rubin, D.B.: Multiple Imputation after 18+ Years. Journal of the American Statistical Association. 91, 473–489 (1996). doi:10.1080/01621459.1996.10476908

Rubin, D.B.: Multiple imputation for nonresponse in surveys. Wiley-Interscience (2004)

Schafer, J.L., Graham, J.W.: Missing data: Our view of the state of the art. Psychological Methods. 7, 147–177 (2002). doi:10.1037/1082-989X.7.2.147

St-Louis, E., Manaugh, K., van Lierop, D., El-Geneidy, A.: The happy commuter: A comparison of commuter satisfaction across modes. Transportation Research Part F: Traffic Psychology and Behaviour. 26, 160–170 (2014). doi:10.1016/j.trf.2014.07.004

Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G., Wood, A.M., Carpenter, J.R.: Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ (Clinical research ed.). 338, b2393 (2009). doi:10.1136/bmj.b2393

Tang, J., Wang, Y., Zhang, S., Wang, H., Liu, F., Yu, S.: On Missing Traffic Data Imputation Based on Fuzzy C -Means Method by Considering Spatial–Temporal Correlation. Transportation Research Record: Journal of the Transportation Research Board. 2528, 86–95 (2015). doi:10.3141/2528-10

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., Altman, R.B.: Missing value estimation methods for DNA microarrays. Bioinformatics. 17, 520–525 (2001). doi:10.1093/bioinformatics/17.6.520

Tsami, M., Nathanail, E.: Guidance Provision for Increasing Quality of Service of Public Transport. In: Procedia Engineering. pp. 551–557 (2017)

Tyrinopoulos, Y., Antoniou, C.: Public transit user satisfaction: Variability and policy implications. Transport Policy. 15, 260–272 (2008). doi:10.1016/j.tranpol.2008.06.002

Vuong, Q.H.: Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. Econometrica. 57, 307 (1989). doi:10.2307/1912557

Wongwiriya, P., Nakamura, F., Tanaka, S., Ariyoshi, R.: User Satisfaction of Songtaew in Thailand: Case Study of Khon Kaen City. In: Transportation Research Procedia. pp. 4946–4953. Elsevier B.V. (2017)

Zou, M.: Gender, work orientations and job satisfaction. Work, Employment and Society. 29, 3–22 (2015). doi:10.1177/0950017014559267