MODELLING PERCEIVED QUALITY FOR URBAN TRANSPORT SYSTEMS USING WEIGHTED VARIABLES AND RANDOM PARAMETERS

Eneko Echaniz Beneitez
Researcher, University of Cantabria, Spain
Luigi dell'Olio
Associate Professor, University of Cantabria, Spain
Ángel Ibeas Portilla
Associate Professor, University of Cantabria, Spain

ABSTRACT

In this article, an Ordered Logit model is proposed considering systematic and random variations in tastes.

The methodology followed for the creation of this model consisted, in first place, in obtaining data using a revelled preferences survey. In the survey, each user had to evaluate, following a qualitative scale, each one of the attributes of the analysed transport system. The variables evaluated in the survey had been grouped into six groups, and for each group, users had to order the attributes belonging to the group, using a ranking based method, from the most important to de least important, and, in the same way, with the groups itself. Once the database is formed, a generic model have been created, establishing this model as a comparative base for the rest. Next, two more models have been estimated one considering systematic users variations and the other one combining the systematic variations with weighted variables. Additionally, three new models have been calculated as an evolution of the previous ones using random variables as representation of systematic and random variations in user's tastes.

The results shows that as model's complexity increase, an improvement in model fit is achieved.

Keywords

Perceived Quality, Ordered Probit, Urban Public Transport, Revealed Preference Survey, Focus Group.

1. INTRODUCTION

Understanding the level of satisfaction of transport public users is essential in order to developed public transport systems based marketing plans in urban areas.

A global definition of the quality has not been achieved yet. Reeves and Bednar (1994) concluded that a global definition of "quality" does not exist, and, depending the circumstances or the market where the product or service can be located, different definitions

of the concept quality exists.

Even if the quality of tangible products has been widely described and easily measured, the quality of a given service is still undefined. Parasuraman, Zeithaml y Berry have extensively worked in the understanding of the quality of services. In 1985, they developed a model, in which, is concluded that the customers or users use similar criteria for evaluating the level of quality of a service. The parameters were grouped in 10 categories, which later end up becoming 5 because a notable similarities between them, with 22 sub-items, creating the scale known as SERVQUAL in 1988. This system have been used for a long time in several researches, and the authors themselves developed in 1996 the model explaining the impact of the quality in user behaviour.

Metri (2006) presents a total quality context (Total Quality Management: TQM) and the guidelines to use by transportation agencies. He studied the total quality of transport (Total Quality Transportation: TQT), its organization structure and the 14 points model `proposed by Deming, which is the bases for the management and the efficiency of a business process. The quality can be studied by using two perspectives: the Perceived Quality, based on users experience (dell´Olio, 2010), and the Expected Quality, the one that use expect to have if they tend to use the service (dell´Olio, 2011). The way to analyse each of one is different. Both concepts are complementary and necessary to understand the global quality of the system.

Hu and Jen (2006) made a study based on perceptions and expectations of the users in Taipei bus services, using a 4 dimension and 20 variables scale, similar to the SERVQUAL scale.

Filipović et al. (2009) researched the expected and the perceived quality of the public transport in the city of Belgrade, based on system characteristics, an importance ranking of the properties ant the level of satisfaction of users with the existing service.

Furthermore, Eboli and Mazzulla (2009) proposed a new index to evaluate the quality on transport "Heterogeneous Customer Satisfaction Index" (HCSI), inspired in "Customer Satisfaction Index" (CSI) but considering the heterogeneity present in user's evaluations.

Macário (2010) worked in the definition and the measurement of the service quality in transportation. The aim of this work was to develop tools in order to evaluate the Level of Service (LoS) and the quality of the service in different transport markets. The tools are divided in user's perception (service quality) and manager perception (Level of Service).

Other existing methodologies use the quality as a global concept studying, in most cases, only the users perceive quality of the public transport system. Garrido y Ortúzar (1994) developed a ranking to obtain the perceived importance of some attributes of the service. This ranking was used to determine the relative weights of the attributes inside the level of service. For the determination of the relative weights the methodology proposed by Cook and Fress (1988) was used. In this study, multinomial Logit models were used to calculate the elasticities of the demand regarding the evaluated variables. The main conclusion was that the results obtained following both methods differs in a notorious way.

Friman (2004) carried out a study approaching two questions: first, the effect that the improvement in quality has in the overall satisfaction of the transport service system; Second, the effect in perceived quality that negative incidents have. The principal conclusion

of this study was that the positive effects that the improvement in service quality has in the overall satisfaction of the service has a limited scope.

Benjamin (2006) made a research to study the perception of improvements focused in growth people with reduced mobility. For that, two surveys were made, one before and one after the improvements. A regression model relating the increase of travellers between both surveys was estimated, based on the evaluation given by users. The results of the study show that the adopted measures had a positive influence in perceived quality.

Hensher et al., (2010) had studied the influence that experience / usage frequency has in perception of the service quality by users. To that end, Ordered Probit models have been used, both standard and random, comparing the results obtained in both types. The authors support the fact the improvement in quality are a way to encourage client loyalty. A survey was made asking each user about the general perception if the quality and the importance and satisfaction regarding several defining attributes of the service with the objective of obtaining the most influencing factors.

Recently, different methodologies have been applied in order to define, in a more accurate way, service quality in public transport. Actually there is an open discussion about the different methodologies that can be applied. As recent studies have proof, neuronal networks, classifications trees or structural equations can be also used for defining service quality in public transport (de Oña et al. 2012, 2013, Garrido et al. 2014). Still there is not clear yet, which methodology is the best for representing perceived quality, having all of them their positive and negative aspects.

In this research a detailed study of the users perceived quality have been carried out, considering user heterogeneity, systematic and random variations in user's tastes and weighted variables. The estimated quality have been determined for the whole transport system.

The validity and utility of the proposed methodology have been tested in a real application: Public Transport System in the city of Santander, where the perceived quality have been determined for 15 lines composing the city public transport service. With all the econometric models calculated, a model fit comparison have been made, with the aim of stablishing the benefits of using more complex models against more simple ones.

This article is divided in other 6 sections. In section 2, a breve review of the state of the art regarding service quality in public transport service is done. In section 3, the methodology applied is explained, this section is also divided in two, the first part is focused in explaining the data gathering methodology and, the second one, focused on the models required for the practical application. In section 4, the results obtained in the practical application are shown. Next, a comparative of the estimated models is presented. Finally, the conclusions of this research are summarized, ending up with the list of the references consulted.

2. METODOLOGY

The methodology used in this study can be separated in two parts. In one hand, the data gathering process is explained. In the other hand, the models that are used in the practical

application are bravely explained.

2.1. Data Gathering

The variables that represent the quality in the public transport have been defined using an analysis of the existing international bibliography, and also, making a series of focus groups involving public transport users of the city of Santander. In the following table (Table 1) the list of the determined variables can be seen, which, have been grouped due to similar specifications.

Level of Service	Walking time until the bus stop
	Waiting Time
	Travel Time
	Time from the stop to de real destination
	Ticket price
Offered Service	Transfer easiness
	Offered Service (Timetable, frequencies)
	Service reliability
	Special Lines for football, concerts, etc.
	Night / Weekends Service
	Lines Coverage
Information	Information in the bus stops
	Information in informatics platforms
	Information inside buses
Comfort	Occupation grade
	Air conditioning / calefaction system
	Space for people with reduced mobility
	Comfort of the buses
	Buses cleanness
	Possibility for carrying objects: surfboards, luggage, animals, etc.
Drivers behaviour	Driving style
	Driver kindness
Sustainability	Hybrid / Biofuel buses implantation
	Acoustic pollution

Table 1 – Transport Service attributes

In addition, each one of the interviewed have been asked for some characterization information, which ones are listed in the table below (Table 2):

Characterisation Variables
Sex
Age
Employment Status
Driving license possession
Car possession
Reason of the trip
Number of trips done per week
Usual payment form
Monthly income

Table 2 – Characterisation Variables

2.2. Models used

Once the data is collected, the perceived quality have been estimated using Ordered Probit models. Hereafter, a breve explication of this type of models is done.

2.2.1. Ordered Models

For this section, the book Modelling Ordered Choices: A Primer (Greene, H. y Hensher, A. 2010) have been taken as a basis.

The Ordered Probit model, in its newest form, based on a regression, was proposed by McKelvey y Zavoina (1971, 1975) to analyse of ordered, categorized and no quantitate elections.

The ordered models are based on dividing a continuous utility space in discrete bands using a thresholds based system.

$$
\begin{aligned}
q_i^* &= \theta' v_i + \varepsilon_i, i = 1, \dots, n, \\
q_i &= 1 \ si \ \mu_{-1} < q_i^* \le \mu_{i1} \\
&= 2 \ si \ \mu_0 < q_i^* \le \mu_{i2} \\
&= 3 \ si \ \mu_1 < q_i^* \le \mu_{i3} \\
&= \cdots \\
&= J \ si \ \mu_{J-1} < q_i^* \le \mu_J.
\end{aligned}
\tag{1}
$$

In a first approach, the assumption of constant coefficients and threshold parameters for all users is made. The key idea of the model resides in that the observations made are not just an accumulation of a number of discrete observation that might be ordered somehow, but the transformation of an only continuous variable that has to be ordered.

The models contains the marginal utilities θ, plus, J+2 threshold parameters, μ_J, all of them to be estimated using n observations. The collected data consist on the v_i independent variables of each observation and the results q_i. The model is completed by the random error parameter ε_i. It is assumed that the error ε_i s distributed according to a known CDF distribution and defined in the whole real domain. This assumption also includes the independence and hexogeneity from v_i.

Focusing in the problem of this research. Let's suppose a band of five possibilities for each evaluation users have to do, where the options are:

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

Each evaluation q_i^* user does, does not provide the real value of the evaluation q_i^*, but a restricted version of it, a version limited to 5 options, one of which is the nearest to the real and exact quality perceived by the user.

The probabilities of the observed answers are:

$$Prob[q_i = j|v_i] = Prob[\varepsilon_i \leq \mu_j - \theta'v_i] - Prob[\mu_{j-1} - \theta'v_i], J = 0,1,...,J \qquad (2)$$

So that, the model describe the probabilities of obtaining the different values. There is not an obvious relation between the dependant variable q_i and the independent ones q_i, because q_i is just a label that represents a probability band.

For estimating the model parameters, it is necessary to set up a number of normalizations, which are discussed as follows.

In the first place, in order to maintain positive signs of all the probabilities, it is necessary that $\mu_j > \mu_{j-1}$. Secondly, the space comprehend by the model should be the whole real space, so, $\mu_{-1} = -\infty$ and $\mu_j = +\infty$. Given the fact that the data usually do not have information about the scale of the dependant variable (in which case if the scale of q_i^* is modified by a positive value, modifying the scale of the unknown μ_j and θ values by the same value, all the properties of the observations remains the same), so, it make nor possible to estimate the free variance $Var[\varepsilon_i] = \sigma_\varepsilon^2$. It is advisable to use a restriction based on $\sigma_\varepsilon = cte, \bar{\sigma}$. Is typical to use a variance equal to one for the Probit models and equal to $\pi^2/3$ for the Logit ones. Finally, taking the assumption of the existence of a constant term in the model, it is necessary to establish $\mu_0 = 0$.

The calculation of the parameters of the model is made by applying a maximum likelihood estimation problem Pratt (1984) y Greene (2007, 2008). The log likelihood function is:

$$\log L = \sum_{i=0}^n \sum_{j=0}^J m_{ij} \log[F(\mu_j - \theta'v_i) - F(\mu_{j-1} - \theta'v_i)] \qquad (3)$$

Where $m_{ij} = 1$ if $y_i = j$ and 0 if not. The maximization is done with summited to restricion of, $\mu_{-1} = -\infty$, $\mu_0 = 0$, and $\mu_J = +\infty$. The remaining restrictions, $\mu_{j-1} < \mu_j$, can, in a first instance, be imposed by using structural parametres.

$$\mu_j = \mu_{j-1} + e^{\alpha_j}$$
$$= \sum_{m=1}^j e^{\alpha_m} \qquad (4)$$

However, usually it should not be necessary.

The estimation of the model parameters, in practice, is estimated using specific software, such as NLOGIT, which is the one selected for this very research.

2.3.2. Models developed for this study

As was told in the previous part, 6 ordered models have been defined.

The first model, or Base Model (BM), is based on a lineal and direct relation between the partial evaluation of the service attributes v_{ik} and the global valoration of the transport service q_i, the mathematical expression is represent as follows:

$$q_i = \theta_0 + \sum_{k=1}^{N} \theta_k \cdot v_{ik} + \varepsilon_{ik} \tag{5}$$

with:
$k \in [1,2, \dots, N]$

where:
q_i:General evaluation of user i
θ_0: Model Constant
N:Number of evaluated attributes
θ_k:Coefficient of the variable k
v_{ik}:Evaluation done for each user i of each variable k

The second model, have been taken as an evolution of the BM model, considering systematic variations in user tastes (SVM). To achieve that, n-1 dummy variables z_{ic} have been created for each of the categories describes in table 2, with n as the number of options within each category. The mathematical expression of this model is:

$$q_i = \theta_0 + \sum_{k=1}^{N} \theta_k \cdot v_{ik} + \sum_{k=1}^{N} \theta_{ck} \cdot z_{ic} \cdot v_{ik} + \varepsilon_{ik} \tag{6}$$

with:
$k \in [1,2, \dots, N]$
$z_{ic} = 0 \ o \ 1$

where:
θ_{ck}: Coefficient of the variable k interacting with variable c
z_{ic}:Characterisation coefficient related to a user i and a characterisation variable c

The third model, considers, in addition to the systematic variations, the weighting of the variables (WSVM). To stablish the weighting of each variable, two new weighting variables have been created $(\delta_{nk}, \delta_{mk})$ with values between 0 and 1. This variables represent the importance of a variable against the others, obtaining the value 1 when a variable is the most

important. The weight have been determined in two stages, in one hand, la weight of a variable k inside each group δ_{nk}, aand in the other hand, the importance of the group m the variable k belongs against the other groups δ_{mk}. The rseluntan expression remains like:

$$q_i = \theta_0 + \sum_{k=1}^{N} \theta_k \cdot v_{ik} + \sum_{k=1}^{N} \theta_{ck} \cdot z_{ic} \cdot v_{ik} + \sum_{k=1}^{N} \theta_{nk} \cdot \delta_{nk} \cdot v_{ik}$$
$$+ \sum_{k=1}^{N} \theta_{nmk} \cdot \delta_{nk} \cdot \delta_{mk} \cdot v_{ik} + \sum_{k=1}^{N} \theta_{nck} \cdot \delta_{nk} \cdot z_{ic} \cdot v_{ik} \qquad (7)$$
$$+ \sum_{k=1}^{N} \theta_{nmck} \cdot \delta_{nk} \cdot \delta_{mk} \cdot z_{ic} \cdot v_{ik} + \varepsilon_{ik}$$

with:
$k \in [1,2, \dots, N]$
$z_{ic} = 0 \; o \; 1$
$\delta_{nk} \in [0,1]$
$\delta_{mk} \in [0,1]$

Where:
θ_{nk}: Coefficient of the variable k weighted by δ_{nk}
δ_{nk}: Weighting coefficcient for every n variable k inside a group m
δ_{mk}: Weighting coefficcient for every group m related to a variable k
θ_{nmk}: Coefficient of the variable k weighted by δ_{nk} and δ_{mk}
θ_{nck}: Coefficient of the simply weighted variable k interacting with variable c
θ_{nmck}: Coefficient of the double weighted variable k interacting with variable c

The remaining three models are the extension of the previous ones with the addition of considering systematic and random variation in user's tastes, for that, variables have been considered randomly distributed. The acronyms of this models are RBM, RSVM and RWSVM. For the sake of brevity the mathematical expressions of this models will be omitted, as they are coincident with the previous three.

3. PRACTICAL APLICATION

The principal aim of this study is to define the quality perceived from the users of public transport services, knowledge that will be useful as a start point for the design of efficient marketing policies. This research have been made in the city of Santander, in the north of Spain, a coastal small-medium size city. To this end, a number of surveys have been carried out in order to obtain the necessary data for estimating the models described in section 2.3.2.

3.1. Data collection
As is told, the aim of this research is to obtain the key variables that defines the quality perceived by users. For that, several surveys have been done to public transport users, in every line of the city of Santander managed by the public entity TUS. The number of

analysed lines ascend to 15.

In the next table a list of the numbers of surveys carried out by line is shown.

Line	Number of surveys done
L1	94
L2	93
L3	84
L4	63
L5C1	40
L5C2	40
L6C1	38
L6C2	37
L7C1	40
L7C2	39
L11	29
L12	38
L13	39
L14	38
L16	35
Total	747

Table 3 – Number of surveys

To stablish the minimum number of surveys needed for each line the expression (8) have been used. A usual expression for interception surveys as is shown in previous studies (dell´Olio, 2010, 2011).

$$n \geq \frac{p(1-p)}{\left(\frac{e}{z}\right)^2 + \frac{p(1-p)}{N}} \tag{8}$$

For that, the most conservative values have been taken, where: P=0,5; e=10%; z=1,69 and N=number of users of each line in rush hour.

The surveys have been made during all day, inside the buses as well as in bus stops, making the survey with relation to the ongoing trip or the finished one. The methodology to carry out the surveys and obtain the data have been based on a specifically created tablet app (Fig. 1), in which, each interviewer has downloaded and installed in a tablet. For each variable, the users had to make a qualitative evaluation choosing between Very Bad, Bad, Normal, Good or Very Good. Once the evaluation was done, the users had ordered from most to least important the line specific variables. In the same way, once the evaluation of the attributes and the ordination was done, the global evaluation of the transport service have been asked, for which, the same qualitative scale was followed. To finish the survey, the users had to order from most to least important all the 6 groups defined in Table 1.
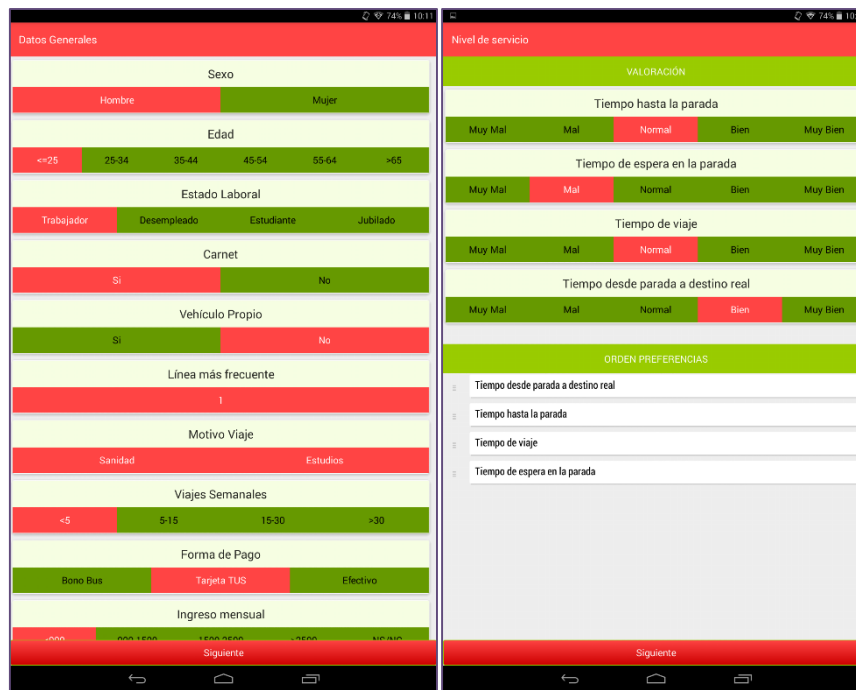
Fig. 1 – Survey App

3.2. Estimated models

This research is focused on comparing different perceived quality models considering the whole public transport system as a whole. The research have been made taking into account 15 lines of Santander with 747 observations sample.

For the modelling, Ordered Probit models have been used, described in part 3.3.2. Subsequently, random variations in tastes have been analysed.

Models are based on a Global Evaluation of the service q_i made by each one of the interviewed user i.

Given the large list of variables referring service attributes, a large number of them have ended being not statistically significant, therefore, the models that are shown as follows omit al not significant variables. Same with the interactions that explain systematic variations and the weighted variables.

3.2.1. Ordinary Ordered Models

Down below a three ordinary Ordered Probit models are shown.

3.2.1.1. Base Model (BM)

BM		
F.Log-Likelihood	-497,064	
Variable	Coef.	Test t
Constant	-4,307	-11,53
Walking time until the bus stop	0,248	4,01
Waiting Time	0,136	2,25

Travel Time	0,202	2,87
Ticket prices	0,078	1,54
Transfer easiness	0,263	4,21
Offered Service	0,064	0,98
Service reliability	0,178	2,52
Night / Weekends Service	0,132	2,22
Lines Coverage	0,250	3,91
Information *	0,270	2,86
Comfort **	0,579	6,22
Driving style	0,285	4,38
Threshold parameters		
μ_1	1,745	17,35
μ_2	4,606	33,67

*Average of the three variables within the group.

**Average between Comfort of the busses and Air conditioning / calefaction system.

Table 4 – Base Model (BM)

In this first model (Table 4), which one have been stabilised as a comparative base for the rest of the models, it can be seen that the vast majority of attributes evaluated in the survey v are not significant when the overall perceived quality q_i is being explained.

Regarding model consistency, it can be seen that the signs of the independent variables v_{ik} are correct, this means that all off them are positive, whereas the constant is negative. This statement confirms the specifications of the model, because an improvement of one of the attributes need to improve the overall satisfaction. In addition, as was told in the chapter 3.2.1, about the threshold parameters of the model, the positive probability of obtaining an outcome with value 0 (which is the combination of Very Bad and Bad evaluations as a result of not obtaining a Very Bad evaluation in any observation) the value obtained in the model has to be a value comprehended between -∞ and 0, this means that, if all the variables have a positive coefficient, the constant has to be negative. This fact is also repeated in the rest of the models.

As far as variables significance concern, all the variables excepting the Ticket Prices and Offered Service turn out significant. In both cases, it have been decided to maintain them because of the correct sign of the coefficient. The most influent variables that affects the Global Evaluation is Comfort, which duplicate the next most influent ones.

3.2.1.2. Systematic Variations Model (SVM)

The calculated model considering systematic variations in user tastes (Table 5) shows a substantial improvement in the model fit. In this model, the independent variable present in BM have been maintained, with the exception of the Ticket Prices, which only appears in the interactions. Two dummie variables have been added (Reason: Work and Trips>30) because in both cases the variable appear to be significant, but also, in both cases the sign is

negative so the perceived quality is reduced, such effect turn to think that regular customers and people that need the transport for work are more critic with the service than the other type of users.

In respect of interactions, it can be seen that only two of them shows a negative sign, what is related with the diminution of the effect of that variable within the global evaluation, in both cases, the affected variable is Comfort, whose contribution to the overall quality is being reduced for young people and oldest people. Is important to highlight the interactions regarding trip reason, due to the fact that all the trip reasons appears in at least one interaction.

SVM		
F.Log-Likelihood	-467,011	
Variable	Coef.	Test t
Constant	-4,141	-10,38
Walking time until the bus stop	0,203	3,05
Waiting Time	0,129	2,05
Travel Time	0,160	2,12
Transfer easiness	0,211	3,05
Offered Service	0,060	0,88
Service reliability	0,177	2,41
Lines Coverage	0,162	2,21
Information *	0,324	3,27
Comfort **	0,663	6,49
Driving style	0,283	4,18
Reason: Work	-0,818	-2,18
Trips>30	-1,972	-1,96
Interactions		
Age_25_35 * Travel Time	0,510	3,05
Age_25_35 * Comfort **	-0,514	-3,06
Employed * Night / Weekends Service	0,133	2,45
Student * Transfer easiness	0,129	2,25
Retired * Transfer easiness	0,569	2,47
Retired * Comfort **	-0,474	-1,98
Reason: Home * Ticket price	0,177	2,99
Reason: Work * Lines Coverage	0,427	3,08
Reason: Study * Waiting Time	0,141	1,89
Reason: Health * Ticket price	0,377	3,19
Reason: Shopping * Walking time until the bus stop	0,131	1,89
Reason: Leisure * Walking time until the bus stop	0,147	2,32
Trips_5_15* Transfer easiness	0,071	1,95
Trips>30* Driving style	0,986	2,56
Income< 900* Walking time until the bus stop	0,080	2,21
Threshold parameters		
μ_1	1,894	16,88

μ_2	4,921	31,74

*Average of the three variables within the group.

**Average between Comfort of the busses and Air conditioning / calefaction system.

Table 5 – Systematic Variations Model (SVM)

3.2.1.3. Weighted and Systematic Variations Model (WSVM)

In accordance with the stabilised in Table 6, the prediction model improves when the weighting is applied. As is shown, not all the significant variables maintain their significances when user's importance criteria is applied.

The effect of each weighted variable changes from one to other. For example, in the case of the Walking time until the bus stop or the Travel Time, people that consider this variables as very important affects that the contributions of this variables to the global evaluation is lower than before. While, the interactions involving weighted variables generate the opposite effect, generating a bigger contribution of the interaction as importance increase.

WSYM		
F.Log-Likelihood	-463,186	
Variable	Coef.	Test t
Constant	-4,325	-10,79
Walking time until the bus stop	0,289	4,20
Waiting Time	0,135	2,13
Travel Time	0,221	2,75
Transfer easiness	0,207	2,98
Offered Service	0,065	0,95
Service reliability	0,200	2,72
Lines Coverage	0,173	2,35
Information [*]	0,326	3,27
Comfort [**]	0,697	6,79
Driving style	0,310	4,56
Reason: Work	-0,822	-2,20
Trips>30	-0,972	-1,36
Interactions		
Age_25_35* Travel Time	0,480	2,86
Age_25_35* Comfort [**]	-0,483	-2,87
Employed * Night / Weekends Service	0,100	1,87
Student * Transfer easiness	0,111	1,93
Retired * Transfer easiness	0,592	2,58
Retired * Comfort [**]	-0,531	-2,22
Reason: Work * Lines Coverage	0,444	3,19
Reason: Study * Waiting Time	0,152	2,08
Reason: Leisure * Walking time until the bus stop	0,153	2,47
Trips_5_15* Transfer easiness	0,077	2,14
Weighted Variables		

δ_n* Walking time until the bus stop	-0,133	-2,28
δ_n*δ_m* Travel Time	-0,158	-2,58
δ_n* Driving style *Trips>30	0,709	2,46
δ_n*δ_m* Ticket price * Reason: Home	0,219	3,31
δ_n*δ_m* Ticket price * Reason: Health	0,453	3,52
δ_n*δ_m* Walking time until the bus stop * Reason: Shopping	0,426	2,92
Threshold parameters		
μ_1	1,903	16,82
μ_2	4,955	31,49

*Average of the three variables within the group.

**Average between Comfort of the busses and Air conditioning / calefaction system.

Table 6 – Weighted Systematic Variations Model (WSVM)

3.2.2. Random Variables Models

The models considering systematic and random variations in user tastes will be shown below.

3.2.2.2. Random Base Model (BM)

The first calculated random model is an extension of the BM. As it can be seen in the table below (Table 7), the variables used for this RBM model at the same as in BM, although, some of the variables have been considered as randomly distributed, this variables are the Ticket Price, Transfer easiness, Information and Comfort. For all these variable, the distribution assumed have been the Normal Distribution.

Regarding the deviations of this parameters, it can be seen that, apart from Comfort, the variances are quite important, so that, the need of establishing a random distribution for these variables have been demonstrated.

RBM		
F.Log-Likelihood	-493,477	
Variable	Coef.	Test t
Constant	-5,589	-12,23
Walking time until the bus stop	0,319	4,48
Waiting Time	0,192	2,78
Travel Time	0,241	3,18
Offered Service	0,093	1,38
Service reliability	0,231	3,12
Night / Weekends Service	0,182	2,68
Lines Coverage	0,310	4,53
Driving style	0,370	5,36
Means for random parameters		
Ticket price	0,111	2,02
Transfer easiness	0,349	5,03

Information *	0,321	2,98
Comfort **	0,762	7,38
Scale parameters for distributions of random parameters		
Ticket price	0,217	8,60
Transfer easiness	0,216	10,20
Information *	0,109	5,64
Comfort **	0,034	1,71
Threshold parameters		
μ_1	2,223	13,52
μ_2	5,992	22,09

*Average of the three variables within the group.

**Average between Comfort of the busses and Air conditioning / calefaction system.

Table 7 – Random Base Model (RBM)

3.2.2.2. Random Systematic Variations Model (RSVM)

In the case of Random and Systematic Variations model (Table 8), the differences with the not random model resides in the fact that the interactions involving random variables are also randomly distributed. As in the RBM, all random variables are considered distributed following a Normal Distribution Function.

Also, in this model, the randomly distributed variables are the same as in the RBM and the interactions are consistent with SVM.

In the same way, the sings of the random variables and interactions are also the same, being the only negative ones the ones related with comfort. Moreover, all the random parameters, the scale parameters and the interactions turn out to be significant.

RSVM		
F.Log-Likelihood	-464,444	
Variable	Coef.	Test t
Constant	-5,734	-11,20
Walking time until the bus stop	0,279	3,62
Waiting Time	0,219	2,89
Travel Time	0,208	2,47
Service reliability	0,239	2,92
Lines Coverage	0,232	2,81
Driving style	0,393	5,12
Reason: Work	-1,182	-2,80
Trips>30	-3,037	-1,97
Interactions		
Age_25_35* Travel Time	0,754	3,70
Employed* Night / Weekends Service	0,198	3,14
Reason: Work * Lines Coverage	0,588	3,66
Reason: Study * Waiting Time	0,171	1,73
Reason: Shopping*Walking time until the bus stop	0,183	2,26

Reason: Leisure * Walking time until the bus stop	0,194	2,50
Trips>30* Driving style	1,519	2,70
Income< 900* Walking time until the bus stop	0,112	2,56
Means for random parameters		
Ticket price	0,010	0,15
Transfer easiness	0,287	3,47
Information [*]	0,454	3,85
Comfort [**]	0,969	8,02
Scale parameters for distributions of random parameters		
Ticket price	0,187	7,09
Transfer easiness	0,259	10,96
Information [*]	0,166	7,68
Comfort [**]	0,112	5,09
Heterogeneity in the means of random parameters		
Reason: Home*Ticket price	0,218	2,94
Reason: Health*Ticket price	0,516	3,42
Student * Transfer easiness	0,199	2,81
Retired * Transfer easiness	0,845	3,12
Trips_5_15* Transfer easiness	0,103	2,34
Retired * Comfort [**]	-0,729	-2,57
Age_25_35* Comfort [**]	-0,768	-3,66
Threshold parameters		
μ_1	2,592	12,88
μ_2	6,895	20,42

*Average of the three variables within the group.

**Average between Comfort of the busses and Air conditioning / calefaction system.

Table 8 – Random Systematic Variations Model (RSVM)

3.2.2.3. Random, Weighted and Systematic Variations Model (RWSVM)

The las model, the one considering systematic and random variations in user tastes (Table 9), shows the best fit of all of them.

As is sown in the table below, in this case, the radon variables, all of them normally distributed, differs from the previous RBM and RSVM. Because of the effect of the weighting, the ticket price has become a not random variable.

Despite this fact, a weighted variable, δ_n* Walking time until the bus stop, have become a random one.

Regarding the sings and the significance level of the variables taken into account, it can be seen that, all of they are statistically significant and have the correct sign.

RWSYM		
F.Log-Likelihood	-461,725	
Variable	Coef.	Test t
Constant	-7,530	-12,26
Walking time until the bus stop	0,454	5,14

Waiting Time	0,217	2,57
Travel Time	0,324	3,23
Offered Service	0,180	2,14
Service reliability	0,356	3,83
Lines Coverage	0,270	2,86
Driving style	0,547	6,39
Reason: Work	-1,243	-2,72
Interactions		
Age_25_35* Travel Time	0,852	3,83
Employed* Night / Weekends Service	0,167	2,48
Reason: Work * Lines Coverage	0,167	2,48
Reason: Study* Waiting Time	0,260	2,39
Reason: Leisure * Walking time until the bus stop	0,263	3,12
Weighted Variables		
$\delta_n*\delta_m*$ Travel Time	-0,283	-3,38
δ_n* Driving style *Trips>30	0,626	3,37
$\delta_n*\delta_m*$ Ticket price * Reason: Home	0,359	4,25
$\delta_n*\delta_m*$ Ticket price * Reason: Health	0,841	4,66
$\delta_n*\delta_m*$ Walking time until the bus stop * Reason: Shopping	0,721	3,47
Means for random parameters		
Transfer easiness	0,371	3,95
Information [*]	0,539	4,20
Comfort [**]	1,234	8,63
δ_n* Walking time until the bus stop	-0,190	-2,39
Scale parameters for distributions of random parameters		
Transfer easiness	0,219	8,65
Information [*]	0,284	10,44
Comfort [**]	0,344	11,26
δ_n* Walking time until the bus stop	0,309	7,52
Heterogeneity in the means of random parameters		
Student* Transfer easiness	0,206	2,56
Retired * Transfer easiness	1,013	3,32
Trips_5_15* Transfer easiness	0,140	2,93
Retired* Comfort [**]	-0,911	-2,82
Age_25_35* Comfort [**]	-0,846	-3,71
Threshold parameters		
μ_1	3,109	12,92
μ_2	8,526	18,46

[*]Average of the three variables within the group.

[**]Average between Comfort of the busses and Air conditioning / calefaction system.

Table 9 – Random, Weighted and Systematic Variations Model (RWSVM)

4. DISCUSSION

Once the models are calculated, in Table 10, a comparison of all the models is shown, taking into account the Log-Likelihood value and the Degrees of Freedom of each one.

Model	Log-Likelihood	Degrees of freedom
BM	-497,064	15
SVM	-467,011	30
WSYM	-463,186	31
RBM	-493,477	19
RSVM	-464,444	34
RWSYM	-461,725	34

Table 10 – Model Comparison

The models prediction rate improves in a progressive way as models become more complex. Making a comparison between the models, with and without considering random variables, it can be seen that in all cases the random variable model fits better. Although, this improvement is not as substantial as the inclusion of systematic variations for the first time, both in models SVM and RSVM, from what, we can consider that the randomization of the independent variables is widely explained with the systematic variations.

Same happens with the weighting, though the model improvement is evident, the model fit does not improve in a substantial way. Nevertheless, the best model obtained was the las calculated one, in which one, all systematic variations, weighting and randomization, just as it have been demonstrated in previous researches (Bordagaray et al. 2012).

As a result of the interesting result obtained in this study, it has been noticed the need of carry out a further researches considering line heterogeneity, for what, the same methodology will be applied for each one of the 15 lines analysed as one in this study.

5. CONCLUSIONS

In this research, a methodology to study the actual users perceived quality for urban public transport have been explained. The obtained results show the point of view of public transport passengers but the principal porpoise it to stablish useful guidelines for improving the quality of the service.

It has been demonstrated that the Ordered Probit Models are capable of representing user's perceived quality in public transport in a realistic way.

From the estimated models make it clear the necessity of stablishing taste variations in users, for what a user characterization is required. Moreover, when a variable importance is included in the survey, it allows to implement a weighting in the models by using specific δ variables, which improves the models fit. Also, asking for evaluating the importance of each variable enable the collection of a raking of the most important variables in a revealed way. The randomization of the variables have been studied using random variables which

represents the variation present in user's tastes. Only a few of all the available variables turn out to be random, from that fact, it can be said that the systematic variation explained by interaction explain largely the variation in the rest of not random variables.

Finally, it has to be said the application of the weighting system in models estimation improves existing methodologies, therefore, a better guidelines for quality improvement policies can be achieve.

REFERENCES

Benjamin, J. M., & Price, G. N. (2006). A study of the impact of APTS on service quality perceptions of elderly and disabled riders. *Journal of Public Transportation*, 9(1), 3.

Bordagaray, M., dell'Olio, L., Ibeas, A., & Cecín, P. (2014). Modelling user perception of bus transit quality considering user and service heterogeneity. *Transportmetrica A: Transport Science*, 10(8), 705-721.

Cook, W. D., & Kress, M. (1988). Deriving weights from pairwise comparison ratio matrices: An axiomatic approach. *European Journal of Operational Research*, 37(3), 355-362.

de Oña, J., de Oña, R., & Calvo, F. J. (2012). A classification tree approach to identify key factors of transit service quality. *Expert Systems with Applications*, 39(12), 11164-11171.

de Oña, J., de Oña, R., Eboli, L., & Mazzulla, G. (2013). Perceived service quality in bus transit service: a structural equation approach. *Transport Policy*, 29, 219-226.

dell'Olio, L., Ibeas, A., & Cecín, P. (2010). Modelling user perception of bus transit quality. *Transport Policy*, 17(6), 388-397.

dell'Olio, L., Ibeas, A., & Cecin, P. (2011). The quality of service desired by public transport users. *Transport Policy*, 18(1), 217-227.

Eboli, L., & Mazzulla, G. (2009). A new customer satisfaction index for evaluating transit service quality. *Journal of Public transportation*, 12(3), 2.

Filipović, S., Tica, S., Živanović, P., & Milovanović, B. (2009). Comparative analysis of the basic features of the expected and perceived quality of mass passenger public transport service in Belgrade. *Transport*, 24(4), 265-273.

Friman, M. (2004). Implementing quality improvements in public transport. *Journal of Public Transportation*, 7(4), 3.

Garrido, C., de Oña, R., & De Oña, J. (2014). Neural networks for analyzing service quality in public transportation. *Expert Systems with Applications*, 41(15), 6830-6838.

Garrido, R. A., & Ortuzar, J. D. D. (1994). Deriving public transport level of service weights from a multiple comparison of latent and observable variables. *Journal of the Operational Research Society*, 1099-1107.

Greene, W. H. & Hensher, D. A. (2010). *Modelling ordered choices: a primer*. Cambridge University Press.

Greene, W., (2007). *LIMDEP Version 9.0: Reference Guide*. Plainview, NY: Econometric Software.

Greene, W., (2008). *Econometric Analysis, 6th edn*. Englewood Cliffs, NJ: Prentice Hall.

Hensher, D. A., Mulley, C., & Yahya, N. (2010). Passenger experience with quality-

enhanced bus service: the tyne and wear 'superoute'services. *Transportation*, 37(2), 239-256.

Hu, K. C., & Jen, W. (2006). Passengers' perceived service quality of city buses in Taipei: scale development and measurement. *Transport Reviews*, 26(5), 645-662.

McKelvey, R. & Zavoina, W. (1971). An IBM Fortram IV Program to Perform N-Chotomus Multivariate Probit Analysis. *Behavioural Science*, 16(2), March, 186-7.

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of mathematical sociology*, 4(1), 103-120.

Metri, B. A. (2006). Total quality transportation through Deming's 14 points. *Journal of Public Transportation*, 9(4), 3.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *the Journal of Marketing*, 41-50.

Parasuraman, A., Zeithaml, V. A. & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of retailing*, 64(1), 12-40.

Pratt, J. W. (1981). Concavity of the log likelihood. *Journal of the American Statistical Association*, 76(373), 103-106.

Reeves, C. A., & Bednar, D. A. (1994). Defining quality: alternatives and implications. *Academy of management Review*, 19(3), 419-445.