Transport Policy xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Transport Policy

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

Route effect on the perception of public transport services quality

Jorge Chica-Olmo[a,*], Héctor Gachs-Sánchez[b], Carmen Lizarraga[c]

[a] Department of Quantitative Methods for Economy and Business, University of Granada, Spain
[b] Consultant, Managing director of Metropolitan Transport Consortium of Granada (2012-2016), Spain
[c] Department of Applied Economics, University of Granada, Spain

A R T I C L E I N F O

Keywords:
Quality
Transport service
Logit multilevel model
Nonlinear principal component analysis
Bus lines
Satisfaction

A B S T R A C T

User satisfaction is a key indicator of public service quality, especially for those services considered basic necessities. The conceptualization and measurement of transport service quality—a fundamental determinant of demand—poses challenges for conducting economic analyses and designing mobility policies.

Several operating companies are involved in the transport sector. In this paper, the bus network of the metropolitan area of Granada, Spain, is taken as a case study. The aim is to design a model of overall satisfaction based on level of satisfaction using a specific set of factors that take into account the individual characteristics of users and the differential effect of using different bus lines.

A combined method using nonlinear principal component analysis (NLPCA) and a logit multilevel model (LMLM) in two steps is applied to a satisfaction survey conducted by the Metropolitan Transport Consortium of Granada (Consorcio de Transporte Metropolitano del área de Granada) in 2013. The survey shows that even though customers within the metropolitan area of Granada are satisfied with the service received (67.26%), the level of satisfaction is not equal for all bus lines, with the perceived quality of some lines being above or below the average. This differential effect is due to different reasons, including the technical and functional performance of the operating companies, commercial speed and length and type of route, among others. Both the operators and the public administration need to focus their attention on these lines in order to design economic policy measures to improve bus lines with below standard compliance.

1. Introduction and literature review

In order to create a new model of mobility and travel in more sustainable ways, city dwellers must modify their usual behaviour, especially within metropolitan areas (Miralles Guasch, 2002; Lizárraga, 2006). As regards public transport, significant changes are needed which cannot be achieved only by improving the efficiency of vehicle design and traffic management. Changes must also be made in the way transport is considered and how solutions are identified and evaluated (Litman, 2003).

The aim of this study is to provide a design model to determine the perceived quality of the public transport service of the metropolitan area of Granada, Spain. To this end, a satisfaction survey of the bus network conducted in 2013 by the Metropolitan Transport Consortium of Granada was used as a case study. It is important to note that the interurban transport service of Granada is heterogeneous because it is managed by several operators with different lines or routes which provide service to fifty-two towns.

The measurement and conceptualization of public transport quality is one of the greatest challenges of economic analyses and mobility policies given the importance of such data for both the companies that provide these services and the public administration (Román et al., 2014). Quality of service, as well as transport fares, price of petrol, personal disposable income, unemployment rates and vehicle ownership, have been considered essential factors that determine public transport demand (Paulley et al., 2006; Holmgren, 2007; Cordera et al., 2015). Public policies aimed at promoting the use of public transport as a means to reduce traffic jams and pollution must create a more appealing image focused on markets in order to make the this type of transport more competitive than private vehicles (Beirão and Cabral, 2007: 478; dell'Olio et al., 2010: 388).

The study of public transport quality forms part of the field of service quality; an ambiguous concept at the crossroads of a wide range of attributes (Grönroos, 1984; Parasuraman et al., 1985; Hensher et al., 2003; Paulley et al., 2006; Beirão and Cabral, 2007; de Oña et al., 2016). For this reason, it opens up and interesting field of research with practical implications for transport suppliers and authorities.

The concept and method of measuring quality have evolved since

* Corresponding author.
E-mail address: jchica@ugr.es (J. Chica-Olmo).

http://dx.doi.org/10.1016/j.tranpol.2017.03.024
Received 25 July 2016; Received in revised form 1 March 2017; Accepted 30 March 2017

marketing began to study goods and services from a different approach. As regards the methods used to measure quality, Grönroos (1984) and Parasuraman et al. (1985) designed service quality models based on the correlation between expected performance and the subjective perception of the product. This approach permits identifying three characteristics of the service analysed here: the intangibility of some service components, the material and temporal heterogeneity in the delivery and reception of the service, and the inseparability of service production and consumption (Parasuraman et al., 1985: 42).

In public transport, tangibles coincide with the technical dimensions of Grönroos' service quality model (facilities, infrastructures, vehicles and staff training). However, intangible aspects are rather difficult to measure since they depend on subjective opinions such as safety, room temperature, drivers' attitude and cleanliness, among others. Unlike the evaluation of the quality of durable goods, the fact that the service (i.e. the displacement of a person between two points in a public transport vehicle) is provided, received and consumed immediately and simultaneously with other users (Grönroos, 1984: 39) implies that close attention should be given to the process of providing and receiving the service (Parasuraman et al., 1985: 42).

Public transport delivery and reception are not uniform; they vary from one day to another, from one operating company to another, from one route to another and even from one vehicle to another. This heterogeneity is due to the transport system organization and the diversity in the performance of each operating company due to material endowment (Grönroos, 1984) and to the attitude and competence of its staff. All these specific characteristics of the transport sector make the complete standardization of services impossible.

The functional quality dimension takes into account the process by which technical components are transferred to the public as a substantial feature. However, because services are produced in interaction with consumers, the technical dimension of quality does not on its own account for the perception of users. Public transport may be considered as a "high contact service" (Parasuraman et al., 1985: 43) in which the relationships between users and staff are very frequent and continuous. Both dimensions, "what is obtained by a service user" and "how it is obtained", are consumed and perceived simultaneously but differently depending on the individual. Perceived quality is determined by comparing the perceived service—a combination of functional and technical dimensions—with the expected service (Grönroos, 1984: 39).[1]

2. Methodology

Level of overall satisfaction is an aggregate measure of how to perceive satisfaction with different aspects of the transport system. Overall aggregate satisfaction, which is referred to here as "quality", is satisfaction with a specific set of features of the system. Satisfaction with specific features of the transportation system may be called "specific satisfactions" (del Castillo and Benitez, 2013). Specific satisfactions can be measured by ordinal categorical variables. The level of specific satisfaction is an example of a phenomenon which cannot be objectively measured but can be evaluated using ordinal variables (Ferrari et al., 2011).

Different methods have been used to study level of satisfaction with public services. A review of these methods can be found in Ferrari and Manzi (2014). According to the authors, the most widely used methods are logit, probit and lineal regression. Nevertheless, other methods such as principal component analysis (PCA) have been used to build synthetic measures of satisfaction for different services. The logit and probit models aim to explain service satisfaction, which is measured as a binary variable from a set of explanatory variables. Jilke and Van de Walle (2013), for instance, evaluated the existence of claims in different public services according to a set of socio-economic factors, while Fiorio et al. (2013) studied satisfaction with public transportation using four sets of explanatory variables (demographic, city-specific aggregate, travel and transportation).

Multilevel models (MLM) are another method that is used to study level of satisfaction. Borra and Chiavarini (2005) fit an ordered logit multilevel model with random intercept to explain the quality of public transport in Rome. The authors showed that quality does not depend only on a set of fixed factors, but also on contextual indicators concerning the demographic and environmental features of the municipality in which citizens live. Ji and Gao (2010) also used an ordered logit multilevel model to evaluate satisfaction with public transport in Beijing. They found that the number of bus stops, the access to the main places of the city, as well as people's socio-economic attributes have a significant effect on residents' satisfaction with public transportation.

In this study, we use a method that combines nonlinear principal component analysis (NLPCA) and a logit multilevel model (LMLM) in two steps. The aim is to explain the quality of service (binary variable) as a function of a set of explanatory variables. In a first step, NLPCA is used because the specific satisfaction variables are ordinal categorical. These variables should not be used directly as explanatory variables in the regression model because the marginal effect is not the same for all the values of these categorical variables. We are also interested in measuring users' satisfaction by reducing the observed multi-dimensional variables into a lower number of numerical variables. In a second step, LMLM is used to model the binary nature of the dependent variable, which depends on the effect that the synthetic variables obtained with NLPCA and other visible variables have on overall satisfaction with the interurban bus transport service in the metropolitan area of Granada. In addition, we attempt to analyse the differential effect of the bus route used by travellers on their perception of quality.

2.1. Nonlinear principal component analysis (NLPCA)

Ferrari et al. (2011) used a combination of two methods in two steps: NLPCA and MLM. In the first step, they used NLPCA to build a synthetic indicator (dependent variable) of overall satisfaction based on four relevant public services: landline telephone, electricity supply, postal service and rail service. In a second step, they used an MLM with random intercept to explain the synthetic indicator through a set of socio-economic variables (gender, age, income, etc.). The MLM included these sets of socio-economic variables, as well as the presence of random effects caused by variability among citizens and across countries.

NLPCA or categorical PCA is an optimal scaling method which belongs to non-linear multivariate analysis techniques. Although the aim of NLPCA is similar to that of the standard PCA, NLPCA allows scaling variables at different levels of measurement and identifying the nonlinear relationships among them. When information needs to be synthesized from a pack of numerical variables into a small set of components, standard PCA is a suitable method. However, when operating with mixed measurement levels (nominal, ordinal, and numerical variables), NLPCA is more appropriate (Ferrari and Manzi, 2014; Linting et al., 2007). Gifi (1990) provides a comprehensive explanation of nonlinear multivariate methods based on optimal scaling.[2]

Several methods can be used to select the number of factors, components and dimensions (Jackson, 1993). One of the most widely used methods consists in reducing an initially large number of dimensions using the Kaiser-Guttman criterion, which permits retain-

[1] Grönroos includes corporate image, that is, the perception consumers have of the company, as the third dimension of his quality model.

[2] An introduction to this method can be found in Linting et al. (2007).

ing those dimensions with an eigenvalue greater than the unit. This enables each component to explain a higher variance percentage than the percentage explained by each original variable by itself. Furthermore, Cronbach's alpha can be used as a measure of reliability such that the closer it is to its maximum value of 1, the greater the reliability of the scale.

In NLPCA, orthogonal rotation may be applied in the same way as standard PCA (Linting et al., 2007). The aim of orthogonal rotation is to find a simple structure with a similar but more easily discerned component loading pattern without a change in the variance percentage explained by each of the components. VARIMAX, for example, is a method to find an orthogonal rotation close to a simple structure (Bartholomew et al., 2008).

2.2. Logit multilevel model (LMLM)

MLM is also often referred to as a mixed-effects model. This type of model allows us to understand how nesting individuals within groups can explain the change in data variance. In MLM, it can be considered that there is a fixed effects part that can be observed which affects all individuals equally and some unobserved effects with a random component which can be modelled.

In our model, we assume that the respondents (level-1) taking different bus lines (level-2) are affected by idiosyncratic elements which are similar for each line due to different factors related to the technical management of the transport authority –i.e. route, number and location of stops, schedules–, to the actual performance of each operator –i.e. vehicle age, performance and maintenance, driver behaviour, room in the vehicle– and other intangible factors. Therefore, we consider that there is no independence between respondents who use the same bus line, even assuming that there is a base model or "baseline" for all individuals. In addition, we consider that there is a set of fixed effects affecting all individuals equally. As indicated, the mixed-effects model takes into account these fixed effects and assumes that there are random effects due to the idiosyncratic factors of each traveller.

LMLM is a particular case of a generalized linear mixed model (GLMM). The difference between MLM and LMLM is that the family distribution of LMLM is binomial rather than Gaussian. To account for the binary structure of the dependent variable, we specify a two-level binary logit model with random intercept (Snijders, 2011). Let us consider that Y_{ij} represents the individual response (1=good quality, 0=bad quality) within cluster j (bus line). In general, we consider a linear predictor, η, as a combination of fixed effects, $X\beta$, and random effects, $Z\gamma$:

$$\eta = X\beta + Z\gamma$$

For the binary outcome, we use a logistic link function, $g(\bullet)$, which relates the outcome Y to the linear predictor η:

$$g(\bullet) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \eta = X\beta + Z\gamma$$

where: $p_{ij} = \Pr(Y_{ij} = 1)$

In the case of a random intercept model $Z\gamma = \gamma_{0j}$ with $\gamma_{0j} \sim N(0, \sigma_0^2)$.

A measure of the importance of random effects is the variance partition coefficient (VPC) or intra-class correlation coefficient (ICC), which is the proportion of the variance attributed to variation among individuals. The ICC is interpreted as the correlation between two randomly selected individuals in one cluster (Ferrari et al., 2011). For a LMLM with random intercept, the ICC is obtained (Goldstein et al., 2002):

$$\rho = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_\varepsilon^2}$$

where: $\sigma_\varepsilon^2 = \pi^2/3 = 3.29$, which is the variance for the standard logistic distribution.

An alternative to ICC is the R-squared (Nakagawa and Schielzeth, 2010). We have used the marginal R-squared (R2m), which represents the variance explained by fixed factors and the conditional R-squared (R2c), which represents the variance explained by fixed and random factors (Nakagawa and Schielzeth, 2013).

3. Data

The data used in this study was drawn from the satisfaction survey conducted in March 2013 by the Metropolitan Transport Consortium of Granada. Of the 1720 surveys conducted, 1484 valid records were used.

The metropolitan area of Granada comprises the city of Granada and an additional 51 municipalities. The area spans 861 square kilometres and had nearly 600,000 inhabitants in 2013. The total length of the bus lines exceeds 805 km, which is serviced by 105 interurban vehicles. Transport services are provided by 54 bus lines, all of which have their origin and destination in the city of Granada. The bus lines are managed by sixteen operators, some of which from part of groups of firms.

Satisfaction surveys are the most widely used technique for analysing the quality of public transport services. However, qualitative techniques also have their place in academic research (Beirão and Cabral, 2007), primarily as a complement to quantitative studies (dell'Olio et al., 2010; Román et al., 2014.). These surveys are sometimes administered to both users and non-users of the service (Fellesson and Friman, 2008). Survey respondents can express their stated preferences (dell'Olio et al., 2011) from which it is possible to construct experimental prototypes founded on hypothetical behaviour that can be directly tested with the group of real and potential users (Kroes and Sheldon, 1988; Hensher et al., 2003; Paulley et al., 2006: 300; Asensio and Matas, 2008; Román et al., 2014).

In our research, a revealed preference approach has been used. As this approach takes into account respondents' current perception about the actual service they receive, information can be obtained regarding customers' opinions about service quality, the best and worst aspects of the services and any deficiencies in need of improvement (Morfoulaki et al., 2007).

The data was drawn from a satisfaction survey conducted in 2013 by a firm specializing in sampling techniques commissioned by the Metropolitan Transport Consortium. The survey was administered to a total of 1720 riders at all the bus network headlines. The sample was selected randomly for each of the 54 bus lines and sixteen operators.

The survey is structured into two sections. The first section contains information about the service (place and time of the interview, operator name, line number, origin and destination of the trip), personal characteristics of the respondent, including gender, age, driving license and vehicle ownership together with travel habits, reason for using public transport, motives for travelling, frequency of use, type of tickets or passes and type of access to bus stop (Table 1).

In the second section of the survey, users' are asked to give their opinion about the transportation service. Specifically, the respondents are asked to evaluate 13 items on an 11-point scale (0–10): information, punctuality, safety on board, driver behaviour, vehicle cleanliness, room inside the vehicle, temperature, easy access to get on/off the vehicle, fares, speed, daily service frequency, proximity to bus-stops and schedules.[3]

Moreover, five response options are used to measure the overall quality of the service (very poor, poor, satisfactory, good and very

[3] The survey takes as a reference the items that the European Metropolitan Transport Authorities (EMTA) uses in its annual Benchmarking of customer satisfaction with public transport in Europe (BEST) survey and includes four checkpoints that characterize in detail the comfort inside the vehicle: room, cleanliness, temperature and access.

Table 1
Sample characteristics (2013).
Source: Metropolitan Transport Consortium of Granada

Gender	Female	59.20%
	Male	40.80%
Age	18–30 years	44.00%
	31–60 years	46.35%
	More than 61 years	9.65%
Frequency of use	Almost daily	53.15%
	Frequent	24.45%
	Sometimes	12.20%
	Rarely	10.20%
Ticket type	One-way ticket	15.50%
	Consortium pass	76.70%
	Senior pass (+65 years)	7.70%
	Another	0.10%
Vehicle ownership	Yes	44.05%
	No	55.95%

good). To prevent responses due to inertia and reduce the incidence of the common method bias, the scales used to measure the 13 items and overall service quality are different (Babin and Griffin, 1998; Podsakoff et al., 2003).

4. Models and results

The intention of this paper is to model overall satisfaction with the public transport service according to the level of satisfaction with a specific set of factors, taking into account the individual characteristics of users and the differential effect of the use of each of the bus lines.

In our model, the dependent variable is binary and obtained from an ordinal variable that represents the overall quality of service measured on five levels. The Quality variable is set at 1 if the service is considered good or very good and 0 for the other responses. This type of transformation has been used in previous studies on satisfaction with public transport, such as Fiorio et al. (2013), who transmutes the four level individual satisfaction variables into a dichotomous variable. These authors give two reasons for performing this dichotomization and using a binomial model instead of an ordered multinomial model. First, it is easier to interpret and, second, because the results are easier to present.

In this line, some studies have concluded that when the ordinal scale is collapsed into a binary, the logistic regression yields similar results and this involves only a slight reduction in power (Armstrong and Sloan, 1989; Manor et al., 2000). In our case, the main reason for this transformation is that the Metropolitan Transport Consortium of Granada was interested in determining whether or it was necessary not to take corrective measures on certain bus lines according to users' satisfaction or dissatisfaction. This can occur in cases where operating companies and public bodies have to make a decision to adopt measures regarding the service (Ongkittikul and Geerlings, 2006).

The results show that many survey respondents are satisfied with the public transport services. Of those who responded, 67.26% thought that the quality of the service was good or very good. In order to analyse the level of satisfaction with the bus line, a cross-tabulation analysis was carried out. To study the relation between satisfaction and bus line used by customers a Chi-square test was performed (Chi-squared=151.736, df=53, p=0.000). The test showed that there is a significant dependence between Quality and bus line. Fig. 1 displays the level of satisfaction with each bus line in graphic form. Therefore, customers' satisfaction is related to the bus line used.

As can be seen, the cross-tabulation between Quality and bus line for each of the 16 operators (each bus line has been assigned an alphanumeric code using a capital letter to indicate the company and a number for each of its bus routes). In the figure, the width of each rectangle corresponds to the number of surveys administered per bus line. The number of surveys varies depending on the annual number of users; specifically from six for line C1 to 62 for line N11. The bus line which obtained the best results was line J1 with 95% of satisfied users, while the worst was line L3 with 100% of unsatisfied users. The figure also has shaded cells which are individually significant. Residual Pearson's cells below −2.0 and above 2.0 indicate that there are more or fewer observations in that cell than those expected under the null model (independent variable). Therefore, there are more unsatisfied customers on lines I1, I4, L3 and L9 than expected.

These unexpected results are directly related to the number of bus stops for each line and the number of towns they service. Line J1 is a striking case. It provides a circular service between Granada and two other towns and yet it works as a direct service for customers even though it stops 24 times in one direction and 23 in the opposite direction. Line L6 provides direct service between Granada and a nearby town within the metropolitan area. It makes 14 stops in one direction and 15 in the other. On the other hand, line I1 provides service to four towns and makes 24 stops one way and 28 in the other. Line I4 provides service to two neighbourhoods within the same township but among the 26 stops in one direction and 27 in the other, 22 are made in the same township. L3 is the line which provides public transport to more towns (five) and it makes 29 stops one way and 21 on the way back. Finally, L9 line provides service to four townships and makes the largest number of stops: 31 one way and 31 the other.

The perceived quality of the service may depend on the number of stops, which in turn is related to the distance the bus line travels. Therefore, we have considered a variable that is external to the sample: the density of stops. The density of stops reflects the relationship between the number of stops and the distance in kilometres between the origin and the destination. The descriptive statistics of this variable are shown in Table 2. The average bus stop density is 1.57 stops per kilometre. This variable could have both a positive and a negative effect on the quality perceived by the user (Ji and Gao, 2010). On the one hand, a greater number of stops increases the probability that the origin or the destination desired by the user is closer to one of the stops, but, on the contrary, also involves more travel time. In general, it could be expected that a line with a high density of stops could have a negative effect on the users' perceived satisfaction.

4.1. Principal components analysis

A set of 13 ordinal variables were considered to explain users' perception of transport service quality and measure satisfaction with specific services. Moreover, other individual factors such as gender, age, use of own vehicles and frequency of use were also considered.

First of all, a NLPCA analysis was carried out using the categorical principal components analysis (CATPCA) command implemented in SPSS statistical software (Meulman and Heiser, 2001). This methodology was applied to the 13 ordinal variables in order to obtain continuous variables to explain service quality. The principal component structure obtained using varimax rotation yielded 2 factors and extracted 68.912% of the total variance (see Table 3). Factor 1 accounted for 55.98% of the variance and had loading on ten items. Of these 10 items, the most important are temperature, room/space, punctuality and safety. These factors can be characterized as "Comfort". Factor 2 accounted for 12.93% of the variance and had loading on 3 items. This factor can be characterized as "Services Supply" since the significant variables are schedules, frequency and proximity of bus stops. The internal consistency of these factors was tested using Cronbach's alpha and found to be very adequate for Factor 1 (0.935). However, it is low for Factor 2 (0.438). The overall fit was very high for both factors (0.962).

In the previous section, it was shown that the quality perceived by customers is not independent of the bus line. This different behaviour has to do with the different factors related to the unobservable factors affecting each route, the specific operation of each company and the organization of the transport system; namely the route, schedules,

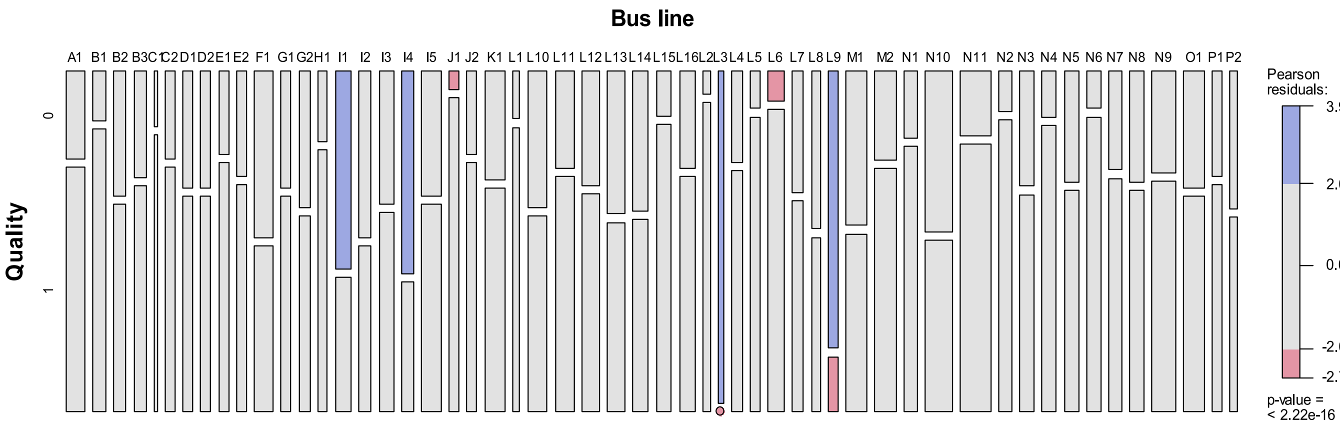

Fig. 1. Mosaic of cross-tabulation between Quality and bus line. Letters refer to operator and numbers refer to bus line.

Table 2
Descriptive statistics of bus stop density.

	Min	1st Qu.	Median	Mean	3rd. Qu.	Max.	SD
Bus stop density	0.281	0.933	1.544	1.570	2.211	3.692	0.815

Table 3
Rotated components matrix.

	Component	
	1	2
Information	0.748	0.196
Punctuality	0.815	0.248
Safety	0.812	0.269
Behaviour	0.731	0.205
Cleanliness	0.800	0.226
Space	0.832	0.183
Temperature	0.837	0.181
Access-Getting on/off	0.772	0.089
Fares	0.669	0.188
Speed	0.737	0.300
Frequency	0.172	0.908
Proximity	0.315	0.792
Schedule	0.196	0.915

Extraction method: Nonlinear principal component analysis.

fares, performance and condition of the vehicle (easy access to get on/off the vehicle, temperature, etc.) and degree of occupation, among others. Special consideration must be given to the comparison between the performance of buses and private vehicles. Often, private vehicle speed (expected performance) differs substantially from commercial bus speed (real service performance) on the same route. In this regard, the ease and the price of parking favour public transport in terms of time spent and the total cost of displacement. The difference between time spent and cost in a private vehicle and on a bus is, for every bus user, a significant benchmark to build a measure of satisfaction with public transport.

Therefore, service quality is explained by variables represented by the factors obtained with NLPCA. Fig. 2 shows the effect of the factors Comfort and Services Supply on the quality of service for the 54 bus lines. In the figure, the graphs of the individual logit models of the 54 bus lines are represented for the two factors. The logit models are shown in each cell of the figure and represent the probability of being satisfied with the service, considering each of the two factors individually. In general, these graphs show that the probability of being satisfied with the quality of service increases with an increase in Comfort and Services Supply. However, some lines show no relationship or even an inverse relationship. It is normally observed that there are more bus lines in which quality is strongly related to Comfort instead of being related to Services Supply. Thus, for example, lines A1, B1 and B2 clearly show a direct relationship between perceived quality and Comfort, while this relationship does not exist or is inverse for Services Supply. The figure shows some specific behaviours that could be due to the differential impact of commercial speed for each bus route, the routes across densely built-up areas with nearer stops (L2 and L3) or routes where there are bus-lanes or non-stop road sections (J1 and L6), which imply staying more time in the vehicle or at the bus stop and hence greater exposure to conditions of pleasure or displeasure with the service received.

4.2. Multilevel analysis

To explain the overall quality considering only the effect of the bus line, a variance components model or "null model" was first estimated (Mod1). This model is an LMLM without explanatory variables (Table 4). The odds that users value the quality of service as good or very good for an "average" line is estimated as exp (0.775)=2.171 with a probability of 2.171/(1+2.171)=0.6846. In addition, we can conclude that there is a significant variation in commuters' satisfaction with different bus lines because the *p*-value of the likelihood ratio test (LR=38.385, df=1, *p*-value=0.000) for testing the null hypothesis that $\sigma_0^2 = 0$ is less than 0.001. The ICC ($\rho = 0.101$) indicates that 10.1% of the variability in commuters is due to the bus line. We can also examine estimates of conditional modes of the random effects caused by satisfaction with the use of different routes.

Fig. 3 (intercept-Mod1) displays the conditional modes of the random effects on bus lines with 95% confidence intervals and the estimated average level of satisfaction with each bus route. As can be observed, for a confidence level of 95%, the effect of some of the lines was below or above the overall average of all the surveys (vertical line zero). In general, it was found that most travellers perceive quality in a similar manner. However, the users' opinions of lines L3, L9, I1, I4, F1 and N10 is significantly below average, thus indicating that their perception of quality is lower than the average of the remaining lines. On the other hand, the quality perceived by line L6 users is above average. L6 is a very direct route with a commercial speed similar to private vehicles because it travels on a highway without stops and in a bus-lane within the urban route. The line L6 terminus is a very central point in the city of Granada where it is difficult to find free parking.

To test whether the observed differences are due to the random effect caused by the performance of individual operators, a model similar to the previous one was estimated but considering the effect of operators (Table 4, Mod2). The *p*-value of the likelihood ratio test (LR=9.930, df=1, *p*-value=0.002) for testing the null hypothesis that $\sigma_0^2 = 0$ is less than 0.01. However, the ICC ($\rho = 0.0215$) is somewhat lower than that obtained for Mod1. In addition, the value of statistical

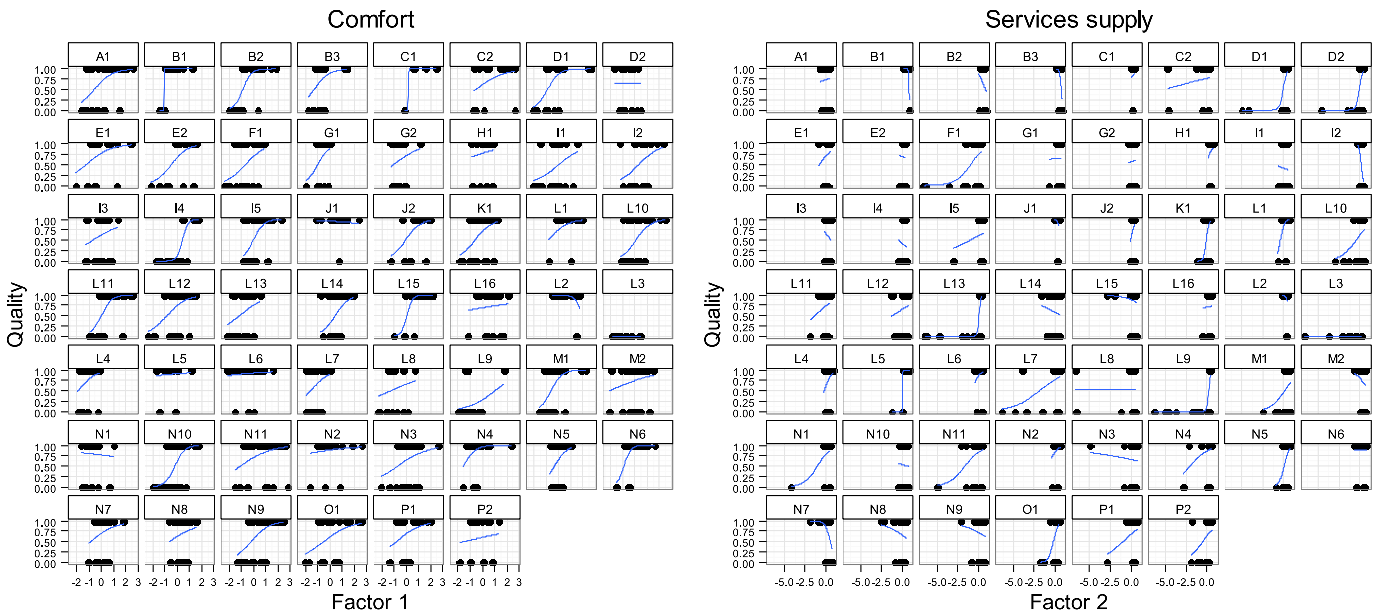

Fig. 2. Relationship between quality, Comfort Factor and Services Supply Factor for each bus line. *Letters (A to P) refers to operator and numbers refer to bus line. The points in the figure represent the quality variable of each respondent (1 or 0) and the line represents the estimated logit model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

AIC is higher, indicating a worse fit.

To consider the possible differential effect not only of the bus lines but also of operators, we have specified a nested model with 3 levels (Table 4, Mod3). In order to determine whether there are significant differences between Mod3 and Mod1, we used a likelihood ratio test with LR=0.158 (df=1) and a p-value=0.691, so by a principle of parsimony we chose Mod1.

The explanatory variables Comfort and Services Supply, which measure specific user satisfactions (Table 4, Mod4), were added to Mod1. The effect that each variable produced on the response probability was given by the odds ratio (OR), which was calculated by exp(β). Table 4 shows the OR values for each variable. If OR is larger than 1, the probability of giving a higher response modality increases with an increase in the value of the explanatory variable. For instance, we would expect the odds that the service quality is rated as good or very good to increase when the Comfort Factor increases. In this way,

Table 4
Multilevel models.

	Mod1	Mod2	Mod3	Mod4	Mod5	Mod6	Mod7
Fixed effects							
Intercept	0.775***	0.723***	0.769***	0.986***	1.017***	0.876***	1.259***
Comfort factor	–	–	–	1.243***	1.267***	1.201***	1.244***
				(3.465)	(3.550)	(3.323)	(3.469)
Services supply factor	–	–	–	0.771***	0.799***	0.873***	0.886***
				(3.323)	(2.223)	(2.394)	(2.423)
Gender	–	–	–	–	0.097	0.116	0.128
					(1.101)	(1.123)	(1.135)
Vehicle ownership	–	–	–	–	0.153	0.171	0.139
					(1.165)	(1.186)	(1.148)
Age	–	–	–	–	−0.441**	−0.446**	−0.397**
					(0.643)	(0.640)	(0.672)
Ticket type	–	–	–	–	0.146	0.143	0.143
					(1.157)	(1.154)	(1.154)
Frequency	–	–	–	–	−0.031	−0.037	−0.043
					(0.969)	(0.964)	(0.958)
Bus stop density	–	–	–	–	–	–	−0.257*
							(0.773)
Random effects (variance)							
Intercept –line	0.369	–	0.345	0.624	0.619	0.389	0.535
Intercept– operator	–	0.072	0.020	–	–	–	–
Slope–comfort factor	–	–	–	–	–	0.148	0.126
Slope– services supply factor	–	–	–	–	–	0.205	0.192
Model fit							
AIC	1866.3	1894.7	1868.1	1552.1	1549.7	1546.9	1527.8
BIC	1876.9	1905.3	1884.0	1573.3	1597.5	1621.3	1607.3
R2m	0.000	0.000	0.000	0.353	0.365	0.354	0.362
R2c	0.101	0.021	0.099	0.457	0.465	0.473	0.492
logLik	−931.1	−945.4	−931.0	−772.0	−765.8	−759.5	−748.9

Significant codes (p-value <): '***' 0.001, '**' 0.01, '*' 0.1. Entries show parameter estimates with odds-ratio in parentheses.

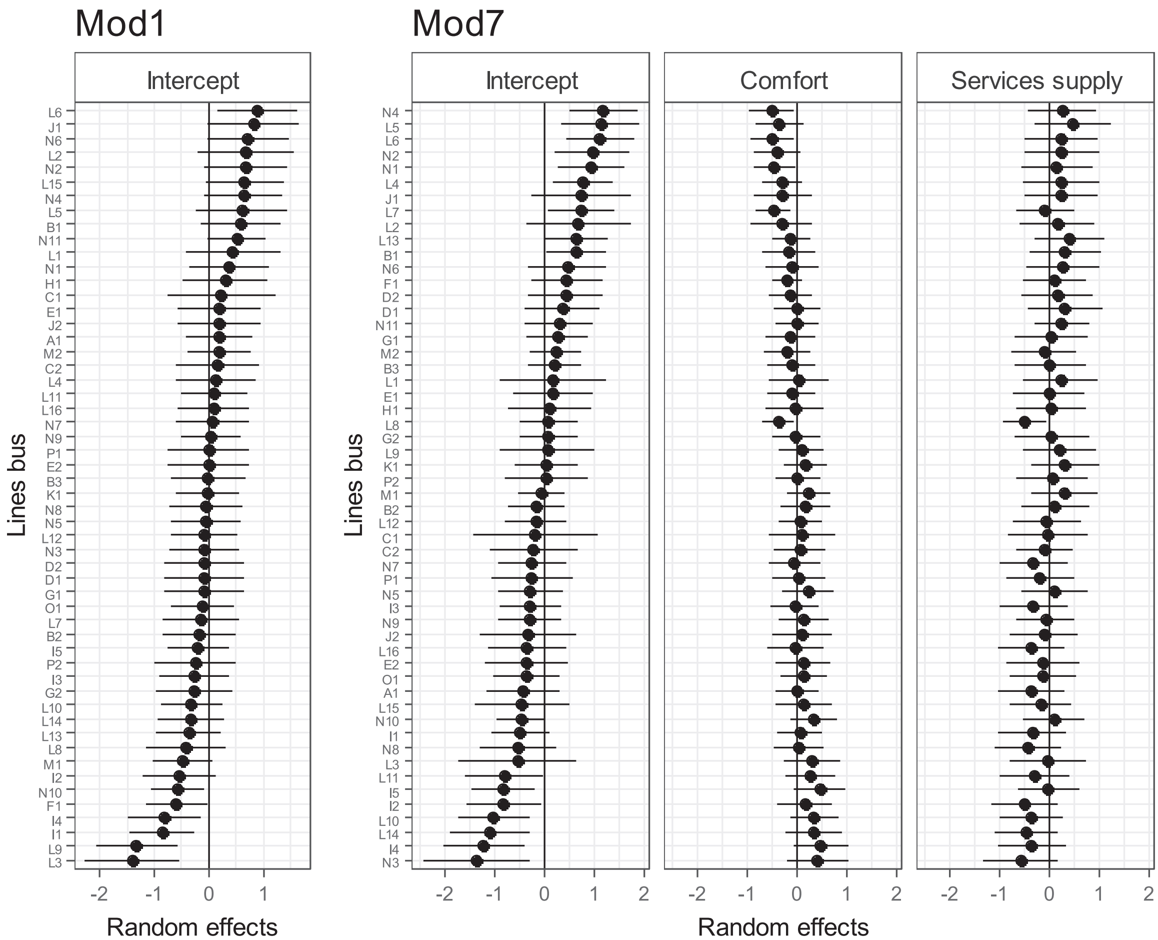

Fig. 3. Conditional modes of the random effects for Mod1 and Mod7 Horizontal lines represent the confidence interval (95%).

for each 1-unit increase in Comfort Factor, the probability of being satisfied with the service increases 246% ([3.465-1]*100).

On the other hand, individual explanatory variables have been added to Mod4 that allow us to determine if there is a change in the perception of quality in terms of gender, having one's own vehicle, age, type of ticket and frequency (Table 4, Mod5). Although the AIC value of this last model is lower than that of Mod4, it does not substantially improve the ICC. However, the R2m and R2c do improve. In addition, we performed an ANOVA to compare Mod5 to Mod4. The ANOVA result (LR=12.403, df=5, p-value=0.029) indicates that Mod5 is preferable to Mod4 since the p-value is less than 0.05.

In Mod5, some other new variables are not significant except age. Therefore, it cannot be said that the quality perceived by users depends on gender, vehicle ownership, type of ticket or frequency of use of public transport. The perception seems to be related more to the real service received than to commuters' personal and economic conditions. Only age, which determines the ease of mobility and requires favourable and differentiated treatment for certain groups, brings meaning to the model.

In Mod6, we estimated a model with random slopes (Table 4, Mod6). This model is relatively better than the Mod5, since the value of AIC is smaller and the R2c is higher, although the R2m is lower. From the ANOVA analysis (LR=12.771, df=5, p-value=0.025) we found that the random effects are indeed significant. In addition to the above, Mod7 (Table 4, Mod7) takes into account the effect of bus stop density. This new model performs better than Mod6 and has been chosen for this reason. We also calculated the coefficients of correlation between the intercept and the two factors. The negative intercept-slope correlation estimate is −0.75 (intercept-Comfort Factor) and 0.55 (intercept-Services Supply Factor). This inverse relationship between the intercept and the Comfort Factor proves that bus lines with above average quality tend to have a below average effect on this factor. That is, the effect of Comfort on service quality in bus lines with high reception is less and vice versa. This is illustrated in Fig. 3. As can be seen (Intercept-Mod7), the users of lines N3, I4, L14, L10, I2, I5, L11 and N10 are significantly below average, indicating that their perception of quality is lower than the average for the rest of the lines. In turn, users of lines N4, L5, L6, N2, N1, L4, L7, L13 and B1 perceive quality above the average. But the effect of the Comfort Factor for lines N4, L6, N1 and L8 is below the average effect of this variable. However, the opposite occurs with the Services Supply Factor, which has a direct relationship.

Moreover, the variance in the random effects of the slopes (Comfort and Services Supply) is lower than that of Mod6, thus indicating that part of the observed variability in Mod6 is due to the effect of bus stop density.

5. Discussion

This research has identified factors which contribute to designing a model for evaluating public transport satisfaction. Previous empirical studies on service satisfaction have proven to be a construct with multiple dimensions. The most common key dimensions regarding transport are reliability, responsibility, receptivity, staff behaviour, attitudes and skills of those involved in provision, security, tangible, information, simplicity in information and capacity for problem solving, frequency, rates, comfort and cleanliness (Parasuraman et al., 1985; Bates et al., 2001; Hensher et al., 2003; Beirão and Cabral, 2007; Morfoulaki et al., 2007; Fellesson and Friman, 2008). These quality components can be summarized into two categories: the technical dimension and the functional dimension (Grönroos, 1984).

In the main component analysis, significant components emerge for the Comfort Factor and the Services Supply Factor. The Comfort Factor, in which the most significant variables are temperature, room/space, safety and punctuality, falls into the category of functional quality. This factor, which is determined by the feeling of comfort, obviously depends on the performance inside the vehicle as a result of both its characteristics of habitability and the behaviour of the driver during the journey. Punctuality may seem out of context but, as mentioned above, we believe it is relevant to this factor since delays imply a longer stay inside the vehicle or at the bus stop and consequently greater exposure to the conditions of pleasure or displeasure with the service received. Moreover, this factor is the most significant explanatory variable in Mod6 and Mod7 as indicated by the high odds-ratio value (see Table 4). The Services Supply Factor is clearly subordinated to the organization and management of metropolitan public transport and corresponds to the dimension of technical quality. Frequencies, schedules and the location of bus stops depend, first, on the concessional obligations of the operating company and, secondly, on the provisions of the transport authority that manages the system.

The results show that there is a direct relationship between the quality perceived by users and the Comfort and Services Supply factors. In addition, both Model 6 and Model 7 indicate that these factors have a differential effect depending on the bus line. As can be seen in Fig. 3, this differential effect is not the same for all bus lines. Thus, for example, the same intervention (i.e. in the same magnitude) either by the operator or the transport authority would have a smaller effect on line N4 than on line L5, although both lines, in general, have an above average rating. This means that there are lines in which improvements would have a larger effect on perceived quality.

Especial features from the specific characteristics of the survey respondents have also been taken into account. Nevertheless, neither gender, own vehicle, type of ticket or frequency of use of public transport are relevant in our model. However, passengers' age has emerged as a variable to be considered.

After realizing that both factors, as well as personal and socio-economic features, do not entirely explain the satisfaction variable, we have assessed some other intangible factors which affect satisfaction and justify the need to model the random effects they cause.

This paper has paid special attention to the effects that other unobserved variables, which affect each transport route, have on users' perception of the quality of the public transport service. These unobserved variables show that the perception of the quality of the service provided depends on the transport route used. Putting the focus on the various bus lines, these unobserved variables may be related to service management, vehicle commercial speed, type of journey, waiting times, etc. In this regard, an inverse relationship between satisfaction and bus stop density has been observed, as well as the fact that the inclusion of this variable improves the goodness-of-fit of the model. Direct lines between towns within the metropolitan area and the city of Granada were evaluated better than those which provide service to more than one township, make more stops or have less direct routes. Lines with routes in high-capacity and high-speed roads obtained a higher satisfaction level than those where sections dominated by densely urbanized areas prevail. In the latter, the presence of more traffic, traffic lights and bus stops, as well as other adverse factors, limit the speed of buses, causing dissonance between the expected service and service received.

We found that perceived satisfaction could also be related to the management of the various private operators. The Spanish concessional system restricts the capacity of transport authorities to intervene directly on how the service is transferred to the public. Operators must comply with the established legal framework, especially with regard to vehicle age and accessibility for people with reduced mobility. However, under their public service obligation contract, operators have full autonomy to organize the services they provide.

6. Conclusions

This research aims to establish the determinants of satisfaction with public transport services in order to develop an interpretative model of perceived quality. To this end, various statistical and econometric techniques have been applied to data from the satisfaction survey conducted in among users in 2013 by the Metropolitan Transport Consortium of Granada, Spain.

The research has shown that quality of service is a multidimensional concept where technical and functional aspects of service provision have considerable importance. The analysis has also revealed the existence of a stable framework for significant variables that explain the perception of quality.

Moreover, the findings of the interpretive model of public transport service quality of are primarily based on unobservable effects that affect each bus line and other observable variables that can be grouped into two categories: the Comfort Factor and the Services Supply Factor. The Comfort Factor constitutes the functional dimension and the Services Supply Factor provides the technical dimension for this model.

The survey from which data was drawn to conduct the analysis shows that even though customers within the metropolitan area of Granada are satisfied with the service received (67.26%), satisfaction is not homogenous across all bus lines. This indicates that the users of different bus lines do not perceive quality the same way; an assumption on which the transport authority intuitively worked. Indeed, the perceived quality of some lines is above or below the average perceived quality. This differential behaviour may be due to different reasons, including the functional and technical performance of the operating companies, the commercial speed of the buses, the type of route and bus stop density, among others. Both operators and the public administration will therefore have to focus their attention on these lines in order to design measures to improve those with below standard compliance (Ongkittikul and Geerlings, 2006).

Acknowledgements

This research was carried out with the collaboration of Metropolitan Transport Consortium of Granada and specially of its Technical Director, Christian Muñoz. The contents of this article reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Metropolitan Transport Consortium of Granada. The authors are very grateful to two anonymous referees for their helpful comments and suggestions.

References

Armstrong, B., Sloan, M., 1989. Ordinal regression models for epidemiologic data. Am. J. Epidemiol. 129 (1), 191–204.
Asensio, J., Matas, A., 2008. Commuters' valuation of travel time variability. Transp. Res. Part E 44, 1074–1085.
Babin, B.J., Griffin, M., 1998. The nature of satisfaction: an updated examination and analysis. J. Bus. Res. 41 (2), 127–136.
Bartholomew, D.J., Steele, F., Galbraith, J., Moustaki, I., 2008. Analysis of Multivariate Social Science Data. CRC Press.
Bates, J.J., Polak, J., Jones, P., Cook, A.J., 2001. The valuation of reliability for personal travel. Transp. Res. Part E 37 (2–3), 191–230.
Beirão, G., Cabral, J.A., 2007. Understanding attitudes towards public transport and private car: a qualitative study. Transp. Policy 14, 478–489.
Borra, S., Chiavarini, S., 2005. Multilevel models for perceived quality analysis: the case of local public transport of Rome. Stat. Appl. 3, 45–52.
del Castillo, J., Benitez, F., 2013. Determining a public transport satisfaction index from user surveys. Transp. A: Transp. Sci. 9 (8), 713–741.
Cordera, R., Canales, C., dell'Olio, L., Ibeas, A., 2015. Public transport demand elasticities during the recessionary phases of economic cycles. Transp. Policy 42, 173–179.
dell'Olio, L., Ibeas, A., Cecín, P., 2010. Modelling user perception of bus transit quality. Transp. Policy 17, 388–397.
dell'Olio, L., Ibeas, A., Cecín, P., 2011. The quality of service desired by public transport users. Transp. Policy 18, 217–227.
Fellesson, M., Friman, M., 2008. Perceived satisfaction with Public transport service in

nine European cities. J. Transp. Res. Forum 47 (3), 93–103.
Ferrari, P.A., Manzi, G., 2014. Citizens evaluate public services: a critical overview of statistical methods for analysing user satisfaction. J. Econ. Policy Reform 17 (3), 236–252.
Ferrari, P.A., Pagani, L., Fiorio, C.V., 2011. A two-step approach to analyze satisfaction data. Social. Indic. Res. 104 (3), 545–554.
Fiorio, C.V., Florio, M., Perucca, G., 2013. User satisfaction and the organization of local public transport: evidence from European cities. Transp. Policy 29, 209–218.
Gifi, A., 1990. Nonlinear Multivariate Analysis. John Wiley and Sons, Chichester, England.
Goldstein, H., Browne, W., Rasbash, J., 2002. Partitioning variation in multilevel models. Underst. Stat.: Stat. Issues Psychol. Educ. Social. Sci. 1 (4), 223–231.
Grönroos, Ch, 1984. A service quality model and its marketing implications. Eur. J. Mark. 18 (4), 36–44.
Hensher, D.A., Stopher, P., Bullock, Ph, 2003. Service quality––developing a service quality index in the provision of commercial bus contracts. Transp. Res. Part A 37, 499–517.
Holmgren, J., 2007. Meta-analysis of public transport demand. Transp. Res. Part A: Policy Pract. 41 (10), 1021–1035.
Jackson, D.A., 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. Ecology, 2204–2214.
Ji, J., Gao, X., 2010. Analysis of people's satisfaction with public transportation in Beijing. Habitat Int. 34 (4), 464–470.
Jilke, S., Van de Walle, S., 2013. Two track public services? Citizens' voice behaviour towards liberalized services in the EU15. Public Manag. Rev. 15 (4), 465–476.
Kroes, E.P., Sheldon, R.J., 1988. Stated preference methods. J. Transp. Econ. Policy, 11–25.
Linting, M., Meulman, J.J., Groenen, P.J., van der Koojj, A.J., 2007. Nonlinear principal components analysis: introduction and application. Psychol. Methods 12 (3), 336.
Litman, T., 2003. Reinventing transportation. In: Exploring the Paradigm Shift Needed to Reconcile Transportation and Sustainability Objectives. Victoria Transport Policy Institute, 27 June 2013. Electronic publication: ⟨http://www.vtpi.org/reinvent.pdf⟩ (5 may 2014).
Lizárraga, C., 2006. Movilidad urbana sostenible: un reto para las ciudades del siglo XXI. Econ. Soc. Territ. VI (22), 283–321.
Manor, O., Matthews, S., Power, C., 2000. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. Int. J. Epidemiol. 29 (1), 149–157.
Meulman, J., Heiser, W.J., 2001. SPSS Categories 13.0: SPSS Incorporated.
Miralles Guasch, C., 2002. Ciudad y transporte: el binomio imperfecto. Ariel, Barcelona.
Morfoulaki, M., Tyrinopoulos, Y., Aifadopoulou, G., 2007. Estimation of satisfied customers in public transport systems: a new methodological approach. J. Transp. Res. Forum 46 (1), 63–72.
Nakagawa, S., Schielzeth, H., 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol. Rev. 85 (4), 935–956.
Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4 (2), 133–142.
de Oña, J., de Oña, R., Eboli, L., Mazzulla, G., 2016. Index numbers for monitoring transit service quality. Transp. Res. Part A 84, 18–30.
Ongkittikul, S., Geerlings, H., 2006. Opportunities for innovation in public transport: effects of regulatory reforms on innovative capabilities. Transp. Policy 13, 283–293.
Parasuraman, A., Zeithaml, V.A., Berry, L.L., 1985. A conceptual model of service quality and its implications for future research. J. Mark. 49, 41–50.
Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J.A., White, P., 2006. The demand for public transport: the effects of fares, quality of service, income and car ownership. Transp. Policy 13, 295–306.
Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J. Appl. Psychol. 88 (5), 879.
Román, C., Martín, J.C., Espino, R., 2014. Using stated preferences to analyze the service quality of public transport. Int. J. Sustain. Transp. 8 (1), 28–46.
Snijders, T.A., 2011. Multilevel Analysis. Springer.