Electronic Journal of Applied Statistical Analysis: Decision Support Systems and Services Evaluation
EJASA:DSS (2011). Vol 2, Issue 1, 65 – 73
e-ISSN 2037-3627, DOI 10.1285/i2037-3627v2n1p65

DISCRETE CHOICE MODELS AS A TOOL FOR TRANSIT SERVICE QUALITY EVALUATION

Laura Eboli*, Gabriella Mazzulla

Dipartimento di Pianificazione Territoriale, Università della Calabria, Italy

Received 22 February 2010; Accepted 27 September 2011
Available online 17 December 2011

Abstract*: In this paper discrete choice Logit models for measuring transit service quality are proposed. Multinomial and Mixed Logit models are used as a tool for evaluating the importance of the different transit service aspects on the overall service quality. Particularly, Mixed Logit models are proposed in order to take into account the heterogeneity of perceptions across individuals. The models were calibrated on the basis of Stated Preferences choice experiments, in which decision makers choose among transit services characterized by different quality levels. The research work is supported by a sample survey addressed to the users of an urban bus service in a medium-sized town.*

Keywords*: Service quality, transit services, discrete choice models, stated preferences*

1. Introduction

Service quality is a focused evaluation that reflects the customer's perception of specific dimensions of service [22]. Among the different aspects characterizing the service, there are characteristics more properly describing the service (e.g. frequency of runs, location of the bus stops, travel time, punctuality and regularity of the runs), and less easily measurable characteristics depending more on customer tastes (e.g. comfort, cleanliness, safety, helpfulness of the personnel). Interested readers may refer to reports published by the Transportation Research Board [20, 21].
Customer opinion on the various service aspects and the overall service allows service quality levels to be measured; so the measure of customer satisfaction provides a measure of perceived

*Corresponding Author. Email: laura.eboli@unical.it

service quality. Service quality measures can be usefully obtained through the customer's point of view by collecting passenger judgements from ad hoc surveys, known in the literature as "customer satisfaction surveys". In the sector of transport services, this kind of information can help transit operators to identify the quality of service factors of greatest importance to their customers. Each passenger expresses an opinion (traditionally in terms of rating) about the main aspects characterizing the services, according to a scale of evaluation specified by the analyst. Customer judgements can be expressed in terms of expectations, which represent what customers expect of the service, and perceptions, which represent what customers receive from the service [15]. Customer satisfaction can be evaluated by collecting only customer perceptions, or through the comparison between expectations and perceptions.

Among the various tools for measuring service quality, statistical models, such as regression models and Structural Equation Models (SEM), can be used in order to relate the attribute representing the main service aspects to the overall service. In these models the dependent variable is represented by the overall customer satisfaction and the independent variables are the service quality attributes, so that the weights of the attributes on the overall customer satisfaction can be estimated. Specifically, SEM permit the introduction of latent variables representing the unobserved and unmeasured factors [2].

An alternative approach for capturing customer judgements in terms of expectations and perceptions is based on the use of conjoint analysis [14], which indirectly captures which service attributes are important and satisfactory to customers. These types of data are usefully analysed through discrete choice Logit models based on the Random Utility Theory (RUT) [1, 4]. The origins of choice modelling can be traced to Thurstone's research into food preferences [18]. According to this approach, all individual decisions involve choice. Individuals choose among different alternatives: shoppers choose between different products, as well as commuters choose between alternative routes and transport modes. Choice modelling considers that human choices stem from a rational decision process, which has a specific functional form. The functional form may be selected as a candidate to model people's behaviour. Human beings try to maximise their total utility. The multinomial logit (MNL) model form is the most commonly used because it is a good approximation to the economic principle of utility maximisation. Although over the last few decades discrete choice models have been widely used for simulating the choice among different transport modes, more recently "within mode" models have been proposed, in which the alternatives relate to a single transport mode, usually public transport mode. Prioni and Hensher [16] first proposed a methodology for measuring transit service quality through the choice-based conjoint analysis. Interviewed passengers were asked to make a choice between two or more transit services, each of which defined by a series of service attributes, varying on predefined levels. Other studies have followed the approach of Prioni and Hensher based on discrete choice models [10, 11, 9, 6, 5, 7]. These models can be calibrated by using the combination of Revealed Preferences (RP) and Stated Preferences (SP) data. SP survey techniques use the statements made by interviewees about their preferences in hypothetical scenarios. SP techniques have several advantages over the traditional RP techniques, which record choices in actual, generally uncontrolled, choice contexts. As an example, SP surveys allow the introduction of choice alternative not available at the time of the survey, the control of the variation of the attribute levels, etc. The major advantage of SP data compared with RP data is that they exploit a more extensive attribute space [12].

This work aims at showing how discrete choice models can be used as a tool for evaluating the importance of the main transit service aspects on the overall service quality. Logit models are calibrated from experimental data collected by a sample survey addressed to the users of an urban bus service, in which choice experiments were proposed to the users.

2. Experimental context

The experimental context is the urban area of Cosenza, a medium-sized town placed in the South of Italy. In this area the University campus is located, which is attended by approximately 32,000 students and 2,000 members of staff (March 2006). The analyzed urban bus service connecting the campus with the town centre is available from 7.30 to 00.30; service frequency is 1 run every 60 minutes. The cost of one-way ticket is 0.77 Euros, while one-day travel card costs 1.55 Euros; the cost of a weekly travel card is about 7 Euros, while a monthly travel card costs about 18 Euros. On a working day, about 8,000 students travel by urban bus.

The survey, realized in the winter of 2006, involved a sample of 470 students; therefore, the sampling rate is approximately equal to 5.8%. An interviewer, located at the bus terminal of the university campus, stopped people while they were waiting for the bus. Each individual was chosen randomly and entirely by chance, according to the simple random sample technique. Respondents were asked to provide information about their trip habits regarding getting to the university and, in addition, about public transport service quality. The interview is divided into three sections: in the first and second section some information about socioeconomic characteristics (gender, age, income and car availability) and travel habits was elicited; the last section of the interview includes an SP experiment proposed to the users, in which they made a choice between the current bus service and two hypothetical bus services. The current service is defined by the user taking into account the bus service used at the time of the interview. The alternatives are defined by nine attributes varying on two levels (0 representing the lowest level of quality, 1 representing the highest level of quality) as reported in Table 1. Each SP alternative is a combination of the attribute levels and represents a bus service. Some levels used in the SP alternatives are not available for the current service. The full factorial design consists of 512 SP alternatives, given that there are nine attributes varying on two levels. In a full factorial design every setting of every factor appears with every setting of every other factor. If there are k factors, each at 2 levels, a full factorial design has 2^k treatments [3]. We established that only three alternatives could be proposed to each user, because several researchers suggested that some difficulties in making a choice between more than three alternatives occur when several attributes define the alternatives (see for example Prioni and Hensher [16]; Hensher and Prioni [10]). Each user was asked to make a choice between his/her current bus service and two SP alternatives. Choosing which alternatives to select among all the 512 of the full factorial design and deciding how to couple them in order to generate the SP experiments was a very difficult process. For this reason, an empirical simulation procedure was proposed by the authors; it is described in Eboli and Mazzulla [5]. By adopting the procedure, we restricted the number of alternatives to 50 and generated 32 couples of SP alternatives. Of the 470 users interviewed, some users made two SP experiments while other users made only one experiment; 633 experiments were completed. When users made two SP experiments, only an SP alternative was replaced in the second experiment in order to reduce the fatigue effect in the respondent.

Table 1. Service quality attributes and levels.

Service quality attributes	*Var.names*	*Levels*
Walking distance to the bus stop	Wtime	same as now (1); 10 minutes more (0)
Frequency	Freq	every 15 minutes (1); same as now (0)
Reliability	Rel	on time (1); late (0)
Bus stop facilities	Stop	bus shelter, seats and lighting (1) no shelter, no seats, no lighting (0)
Bus crowding	Crow	no overcrowded (1); overcrowded (0)
Cleanliness	Clean	clean enough (1); not clean enough (0)
Fare	Fare	same as now (1); 25% more than the current fare (0)
Information	Inf	timetable, map, announcement of delays (1) no timetable, no map, no announcement of delays (0)
Transit personnel attitude	Per	very friendly (1); very unfriendly (0)

By analysing the socio-economic characteristics of the respondents, we can state that the typical bus passenger lives in a family of 4 members, with a medium income level. He/she owns car driving licence, but he/she has not the possibility of using a car to reach the campus, and reaches the bus stops by walking (Table 2).

Table 2. General characteristics of the respondents (n=470).

Characteristics	*Statistics*
1.Gender	Male (46%), female (54%)
2.Age	18-20 (43%), 21-24 (46%), > 24 year-olds (11%)
3.Family members	2 (1%), 3 (11%), 4 (50%), 5 or more members (9%)
4. Monthly family income level	<= 1,000 (17%), 1,000-2,000 (18%), 2,000-4,000 (50%), 4,000-5,000 (11%), > 5,000 Euros (4%)
5. Car driving licence ownership	Did not own car driving licence (16%), own car driving licence (84%)
6. Family members with car driving licence	1 (3%), 2 (17%), 3 (41%), 4 or more members (39%)
7. Family car ownership	1 (26%), 2 (61%), 3 (12%), 3 or more cars (1%)
8. Car availability to reach the campus	Have not car (92%), have car (8%)
9.The way to reach stop	Walking (99%), others (1%)
10. Ticket kind	One-way ticket (25%), one-day travel card (50%), monthly travel card (22%), other (2%)

3. Discrete choice Logit models for transit service quality evaluation

3.1 Model formulation

Following the approach of the discrete choice models based on the RUT, MNL and random coefficients Mixed Logit (ML) models are proposed as a tool for measuring the quality level of an urban bus service, and evaluating the importance level assigned by the regular customers to the various service aspects. More specifically, the choice alternatives are the three bus services (the current service and the two hypothetical SP alternatives) of the different SP experiments proposed to the interviewed users.

The utility functions of the alternatives in the MNL model are linear combinations of the service quality attributes, as shown in the following expression:

$$
\begin{aligned}
U_{ij} = (&\beta_{WTime}WTime_j + \beta_{Freq}Freq_j + \beta_{Rel}Rel_j + \beta_{Stop}Stop_j + \beta_{Crow}Crow_j + \beta_{Clean}Clean_j \\
&+ \beta_{Fare}Fare_j + \beta_{Inf}Inf_j + \beta_{Per}Per_j) + \varepsilon_{ij}
\end{aligned} \tag{1}
$$

with j varying from 1 to m, in which m is the number of the alternative bus services proposed to the i-*th* interviewed, i=1,…,n, ε_{ij} represent the random components of the utilities. Subject i will choose alternative j if U_{ij} is the largest among U_{i1},..., U_{im}.
The link between utility and probability of choice is:

$$
P_{ij} = \frac{\exp(\beta_i x_{ij})}{\sum_j \exp(\beta_i x_{ij})} \tag{2}
$$

in which x_{ij} are the attributes submitted to the user i in the alternative j and β_l are the common parameters.
There are three fundamental hypotheses that underlie the MNL formulation. The first one is that the random components ε_{ij} of the utilities of the bus services are independent and identically distributed. The second one regards the homogeneity across bus passengers in responsiveness to the service aspects. Finally, the third hypothesis is that the error variance-covariance structure of the alternative services is identical across the passengers.
The random coefficients ML model is introduced in order to allow the heterogeneity of bus passengers with respect to the service quality responsiveness to be investigated. Traditionally, the differences in user perceptions and responses were taken into account by introducing some socioeconomic characteristics of the users among the model attributes. According to the random coefficients ML model, some hypotheses of unobserved heterogeneity among passengers are made. This model allows to capture unobserved individual effects by introducing a random term representing peoples' tastes. Random coefficients ML model has the standard form of an MNL model except that one or more parameters are considered as random parameters distributed according to a predefined density function; the standard deviation together with the mean is estimated for each random parameter. In the proposed model, the parameters of the service attributes "Availability of furniture at bus stops", "Bus overcrowding", and "Helpfulness of personnel" are random; the parameters of the remaining attributes are fixed.
The utility function assigned to the *j-th* bus service is specified as follows:

$$
\begin{aligned}
U_{ij} = (&\beta_{WTime}WTime_j + \beta_{Freq}Freq_j + \beta_{Rel}Rel_j + \beta_{Clean}Clean_j + \beta_{Fare}Fare_j + \beta_{Inf}Inf_j) \\
&+ (\alpha_{Stop}Stop_j + \alpha_{Crow}Crow_j + \alpha_{Per}Per_j) + \varepsilon_j
\end{aligned} \tag{3}
$$

$$\alpha \sim G(\alpha), E(\alpha) = b, Var(\alpha) = \Psi, \alpha' = [\alpha_{Stop}, \alpha_{Crow}, \alpha_{Per}]$$

$$\varepsilon_{ij} \sim iid\ extreme\ value$$

in which β represent fixed coefficients, α are the coefficients varying across users according to a $G(\alpha)$ joint distribution and a $g(\alpha)$ density function, ε_{ij} represent the independent and identically distributed random components of the utilities (see Train, [13, 19]). Specifically, the authors assumed that the random coefficients are not correlated (i.e. Ψ is diagonal) and are

distributed according to a normal distribution. The authors supposed that users have more heterogeneous perceptions on the most qualitative service characteristics like bus stop furniture, bus crowding, and personnel helpfulness, for which the relative parameters were considered random in the utility function. The authors also assumed that there is no correlation among the parameters because they retain the relative service aspects as completely different among them.
The proposed model allows the probability to choose each alternative service to be estimated; the higher the advantage received by the passenger using the service, the higher the probability to choose the service is. It is possible to compute estimated choice probabilities from:

$$P_{ij} = \int \left(\frac{\exp(\eta_{ij})}{\sum_{l} \exp(\eta_{ij})} \right) g(\alpha) d\alpha$$
$$\eta_{ij} = \sum_{s} \beta_s x_{sj} + \sum_{k} \alpha_{k,i} x_{kj} \quad (4)$$

in which the fixed parameter attributes are: walking distance to the bus stops, service frequency, service reliability, cleanliness, fare, and information; the *random* attributes are: availability of furniture at bus stops, bus overcrowding, and helpfulness of personnel.
The parameter vector was estimated as the vector value maximizing the log-likelihood function; for the ML models this function involves a multidimensional integral, which was solved by the Monte Carlo numerical simulation method by using N-Logit package [8].

3.2 Model results

In Table 3 the results obtained by using multinomial logit with fixed coefficient (MNL) and random coefficients mixed logit (ML) are presented. In particular, the mixed logit specification allows for random preferences variation, unrestricted substitution patterns, and correlation in unobserved factors over time.
We estimate nine transit level-of-service variables: walking distance to the bus stop, service frequency, service reliability, bus stop facilities, bus crowding, cleanliness, fare, availability of information at the bus stop, and personnel attitude.
All the service quality attributes are defined as dichotomous variables, except “Walking distance to the bus stop” and “Ticket cost” that are continuous, measured in minutes and in Euros respectively.
All the estimated coefficients in the MNL model and in the ML are significant at 95% level of significance (including both means and standard deviations). Although improvements in the final value of the log-likelihood function obtained with mixed logit formulations seem marginal, if one compares the log-likelihoods of the two models by means of LRT, the difference is largely significant (LRT is 18.90 with 3 degree of freedom). Also the Mc Fadden’s adjusted Rho squared [8] shows similar results in both formulation: 0.343 for the MNL model, and 0.357 for the ML model, meaning that the MNL model explain 34% of the information in the sample while the ML model explain 36%.

Table 3. Model results.

	MNL model	*ML model*		
	fixed	*fixed*	*random*	
variable	*coefficient (s.e.)*	*coefficient (s.e.)*	*coefficient (s.e.)*	*parameter*
Wtime	-0.146 (0.017)	-0.191 (0.028)		mean
Freq	2.693 (0.236)	3.963 (0.533)		mean
Rel	1.411 (0.153)	2.128 (0.308)		mean
Stop	0.623 (0.154)		0.833 (0.244)	mean
			1.109 (0.832)	st. deviation
Crow	0.905 (0.186)		1.377 (0.326)	mean
	-		1.559 (0.619)	st. deviation
Clean	0.909 (0.144)	1.357 (0.268)		mean
Fare	-8.549 (1.014)	-11.476 (1.657)		mean
Inf	0.561 (0.154)	0.786 (0.231)	0.907 (0.275)	mean
Per	0.515 (0.143)		2.359 (0.583)	mean
	-			st. deviation
Log-Likelihood	-456.557	-447.107		
Rho squared	0.343	0.357		

The model coefficients indicate that transit users have a positive attitude toward higher frequency, reliable schedules, not overcrowded and clean vehicles, bus stops with adequate furniture and information devices, and helpful personnel. Regarding the service aspects for which the values of the standard deviation must be taken into account being random coefficients, we should specify that helpfulness of personnel is the most heterogeneous service aspect given that it is a positive factor for about 65% of transit users and a negative factor for the other 35%; on the other hand, for the service aspects linked to bus overcrowding and furniture at bus stops about 80% of transit users perceive these factors as positive, and only about 20% as negative.

The findings regarding service frequency can be explained by considering that the service analysed offers 1 run every 60 minutes, which represents a very low level of service for an urban bus line. Also the coefficient linked to walking distance to the bus stop has a negative value. This finding could be expected considering that bus passengers interviewed reach the bus stops by walking; therefore, the utility of the service decreases with the increase of the distance to the bus stop.

The attributes with a random coefficient in the model represent the service aspects characterized by a more qualitative nature compared to the other attributes analysed. The results of the ML model confirmed the hypothesis of the authors about the heterogeneity of some service aspects and show that there is heterogeneity among users in the perception of the most qualitative service characteristics.

In order to interpret the results obtained from the models, subjective willingness to pay were calculated. In both the specifications, users are willing to pay about 0.2 Euro to save 10 minutes walking time, about 0.3 Euro to have a service running every 15 minutes, 0.1 Euro for facilities at the bus stop and 0.12 Euro for cleaner vehicles. The consistency of the results obtained with different specifications indicates that differences in parameters' estimates are due to the different methods of estimation of the parameters; in fact, in the MNL model the parameters are estimated considering the marginal distribution of the underlying utilities, while the ML model estimate the parameters considering a distribution conditional to the random parameter distribution. This implies that the marginal parameters are attenuated with respect to the conditional ones [17]. The low values of WTP are explained by the socioeconomic condition of the sample, which is made up of students from low-middle income household (85% of the sample).

4. Conclusions

The main purpose of this research is to provide a tool for investigating on users' preferences about the different service aspects characterizing a bus service. Traditionally, service quality has been measured by collecting user judgments expressed in terms of rates. On the other hand, when the analyst wants to investigate user behaviour and evaluate the influence of the different service aspects on user choice mechanism, user perceptions about the service in terms of choice should be considered. Although the design of a choice experiment is generally very complex, users have the possibility of expressing more easily their preferences about service aspects by making a choice directly among some services. Therefore, discrete choice models represent a useful tool for investigating on passenger perceptions and understanding user behaviour.
In this work MNL and ML models were calibrated on the basis of the user choices made in SP experimental contexts. Specifically, the ML model allowed investigation on the heterogeneity across individuals about some service quality attributes. The heterogeneity across the individuals of user perceptions about the attribute linked to "Availability of furniture at bus stops", "Bus overcrowding", and "Helpfulness of personnel" was verified. The standard deviation values obtained from the model calibration suggest that there is a notable difference in user perception of these attributes. This model formulation is even more useful when the attributes affecting users' behaviour have a qualitative nature, as in this case, because for this kind of service aspects user perceptions are generally more heterogeneous.
The models calibrated are helpful tools to planners and transit operators for measuring service quality and evaluating the importance of the various service quality aspects from a user's point of view; the identification of the most important aspects is useful to the transit operators for investing on the aspects showing the major weights on overall service quality in order to effectively improve the service. Specifically, the utility of each alternative is an index of service quality of each bus service and the values of the parameters are the attribute weights.
This study is now limited by the sample size and its composition (mainly students); if extended the importance of heterogeneity and tastes across the population can be found to be even more significant for the quality evaluation of transit services.

Acknowledgements
The authors would like to thank the two anonymous reviewers, whose comments helped to improve this article.

References

[1]. Ben-Akiva, M. and Lerman, S. (1985). *Discrete choice analysis: theory and application to travel demand*. Cambridge, Massachusetts: MIT Press.
[2]. Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.

[3]. Box, G.E., Hunter, W.G. and Hunter, J.S. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery, 2nd Edition*. Wiley.
[4]. Domencich, T.A. and McFadden, D. (1975). *Urban travel demand: a behavioural analysis*. New York: American Elsevier.
[5]. Eboli, L. and Mazzulla, G. (2008). An SP Experiment for Measuring Service Quality in Public Transport. *Transportation Planning and Technology*, 31(5), 509-523.
[6]. Eboli, L. and Mazzulla, G. (2008). Willingness-to-pay of public transport users for improvement in service quality. *European Transport*, 38, 107-118.
[7]. Eboli, L. and Mazzulla, G. (2010). How to capture the passengers' point of view on a transit service through rating and choice options. *Transport Reviews*, 30(4).
[8]. Econometric Software Inc. (2009). *NLogit version 4*. Plainview, USA.
[9]. Hensher, D.A. (2001). Service quality as a package: what does it mean to heterogeneous consumers. *In 9th World Conference on Transport Research*. Seoul, Korea, 22-27 July.
[10]. Hensher, D.A. and Prioni, P. (2002). A service quality index for a area-wide contract performance assessment regime. *Journal of Transport Economics and Policy*, 36(1), 93-113.
[11]. Hensher, D.A., Stopper, P. and Bullock, P. (2003). Service quality-developing a service quality index in the provision of commercial bus contracts. *Transportation Research Part A*, 37, 499-517.
[12]. Kroes, E. and Sheldon R.J. (1988). Stated Preference Methods: an introduction. *Journal of Transport Economics and Policy*, 22(1), 11-25.
[13]. McFadden, D. and Train, K. (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, 15, 447-470.
[14]. Orme, B. (2005). *Getting Started with Conjoint Analysis*. WI: Research Publishers LLC.
[15]. Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985). A conceptual model of service quality and its implication for future research. *Journal of Marketing*, 49, 41-50.
[16]. Prioni, P. and Hensher, D.A. (2000). Measuring service quality in scheduled bus services. *Journal of Public Transportation*, 3(2), 51-74.
[17]. Skrondal, A. and Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, 68(2), 267-287.
[18]. Thurstone, L. L. (1927). A Law of Comparative Judgement. *Pychological Review*, 34, 272-86.
[19]. Train, K. (2009). *Discrete Choice Methods with Simulation, Second edition*. Cambridge University Press.
[20]. Transportation Research Board (2003). *A guidebook for developing a transit performance-measurement system. TCRP Report 88*. Washington, D.C.: National Academy Press.
[21]. Transportation Research Board (2003). *Transit capacity and quality of service manual. TCRP Report 100*. Washington, D.C.: National Academy Press.
[22]. Zeithaml, V.A. and Bitner, M.J. (2003). *Services Marketing: Integrating Customer Focus Across The Firm*. New York: McGraw Hill.