

An Ordered Regression Model to Predict Transit Passengers' Behavioural Intentions

Juan de Oña, Rocio de Oña

TRYSE Research Group, University of Granada, Spain

Laura Eboli, Carmen Forciniti, Gabriella Mazzulla

University of Calabria, Italy

SUMMARY

Passengers' behavioural intentions after experiencing transit services can be viewed as signals that show if a customer continues to utilise a company's service. Users' behavioural intentions can depend on a series of aspects that are difficult to measure directly. More recently, transit passengers' behavioural intentions have been just considered together with the concepts of service quality and customer satisfaction. Due to the characteristics of the ways for evaluating passengers' behavioural intentions, service quality and customer satisfaction, we retain that this kind of issue could be analysed also by applying ordered regression models. This work aims to propose just an ordered probit model for analysing service quality factors that can influence passengers' behavioural intentions towards the use of transit services. The case study is the LRT of Seville (Spain), where a survey was conducted in order to collect the opinions of the passengers about the existing transit service, and to have a measure of the aspects that can influence the intentions of the users to continue using the transit service in the future.

1. INTRODUCTION

Knowing transit passengers' behavioural intentions to use transit services can represent a useful support for transit managers and marketers for defining the most convenient strategies to satisfy existing passengers and attract new ones. Behavioural intentions represent a concept investigated for a long time. As an example, Zeithaml et al. (1996) defined behavioural intentions as signals that show whether a customer continues to utilize a company's service or switch to a different provider. More recently, there has been an interest in analysing the concept of behavioural intention together with the concepts of service quality and customer satisfaction. All these concepts are very complex and need to be analysed by sophisticated techniques such as structural equation models. There are studies proposing models investigating relationship between passengers' perceptions about service quality factors and satisfaction with the overall service (e.g. Eboli and Mazzulla, 2007, 2012, 2014; de Oña et al., 2013). Other studies instead mix the concept of service quality with behavioural intentions, such as Chen (2008), Lai and Chen (2011), Chen and Chao (2011), Chowdhury and Ceder (2013).

We retain, as it was proved by the above mentioned authors, that the concept of behavioural intentions to use transit service is strongly linked to the concept of service

quality, that is passengers' satisfaction with the several factors characterizing the quality of a transit service have surely an influence in the decision of passengers to continue to use a transit service. For this reason, in this paper we propose a methodology to analyse the influence of service quality factors on passengers' behavioural intentions towards the use of transit services. The light rail transit (LRT) of Seville (Spain) offers the service supporting our work (de Oña et al., 2016). We collected through an ad-hoc survey the opinions of the passengers about the used LRT system, and we propose an ordered probit (OP) model to explain how passengers' opinions influence their intentions to use the LRT again. The authors of this paper have already adopted in the past this kind of models for analysing service quality of airport transit services (Eboli and Mazzulla, 2009), and also in the field of road safety (Cardamone et al., 2014, 2015, de Oña et al., 2014).

In the following, we propose a section about the case study, where we describe the transit service and the survey supporting the research, the sample characteristics, and the opinions expressed about the LRT system. Then, we propose the section about the OP model, by providing a theoretical framework, and successively the description of the proposed model and a discussion of the results. The paper ends with brief conclusions about the work.

2. STUDY CASE

2.1 The transit service and the survey

The analysed transit system is the LRT transit service of Seville, a city of about 700,000 inhabitants located in the south of Spain. The analysed new LRT system, which came into operation in 2009, consists of a sole line characterized by a length of 18 km (10.08 km underground) and 21 stations. In 2013, the LRT carried more than 13.7 million passengers. A questionnaire was addressed online to the users of the LRT system, via a web-based platform. The main part of the questionnaire aimed to collect users' perceptions of different service attributes and the overall service. This part contains 37 questions related to various aspects of the LRT service, such as availability of the service, accessibility, information, timeliness, attention to client, comfort, safety and environmental pollution. Users had to express their opinions about 37 service quality attributes and the overall service on an 11-numeric scale from 0 to 10 (0 being of poor quality and 10 being of the highest quality). Behavioural intentions were collected by presenting to the users this sentence "Surely, I will use the LRT service again". Respondents had to express a rate on a 11-numeric scale from 0 (totally disagree) to 10 (totally agree). Other two parts of the questionnaire are about users' travel habits and socioeconomic information.

2.2 The sample

2.2.1 General information and socioeconomic characteristics

The survey was conducted in a period of two weeks (May-June 2014) at LRT stations. 3,365 responses were registered, of which 3,211 were valid for subsequent analysis. The sample is almost equally spread between females (53.3%) and males (46.7%). Most of

the passengers are aged between 18 and 25 years (41.6%), followed by people between 26 and 40 years (28.8%) and 41 and 65 years (25.5%). The major part of respondents has a bachelor's degree at university (48.5%) or a degree of high school or professional education (41.9%). Respondents are mainly employees (43.7%) and students (41.5%). Most of the sample has a net monthly income equal or lower than 1,200 Euros (28.7%). More than one half of the sample travels by LRT every day (52.1%), and only 16.4% travels occasionally. Passengers mainly travel by LRT to go to school (38.8%) or to work (35.5%). Travellers reach or move from the LRT station mainly by walking (62.6% and 86.3%, respectively), making a trip 12 minutes long on a total trip of about 34 minutes.

2.2.2 Analysis of the service quality factors

As above mentioned, 37 service quality factors belonging to 8 service aspects were judged by the users. Factors concerning availability of the service, which was judged by considering some factors such as service regularity and number of trains per day, registered a discrete average satisfaction rate, equal to 6.7. On the other hand, users very positively judged service accessibility, which includes service factors such as access to stations or operations of elevators; this aspect is appreciated with an average rate of 7.9. Also information services were appreciated enough, registering an average satisfaction rate of 7.6; this aspect includes information on the vehicles and in stations as well as information through internet, phone, etc. Respondents are very satisfied with timeliness of LRT, described in terms of punctuality, speed of the trip, and waiting time on the platform, registering an average rate of 8. Also the conduct of the employees was appreciated enough (average rate of 7.6). The aspect concerning comfort was represented by several service factors, including cleanliness and lighting both in vehicle and at stations, comfort linked to the seats, temperature on vehicle and in stations, and so on. An average satisfaction rate equal to 7 was registered for this variegated aspect; this result is mainly due to the fact that comfort included also a factor concerning coverage cell-phone and 3G for which users expressed a very low average rate of 2.7. "Safety" aspect registered an average of 7.3. Finally, "Environmental Pollution" aspect reached a sufficient judgment (average rate of 6.4). In addition to the judgements on the various service quality characteristics, users expressed also their opinions about the overall service; an average rate equal to 7.6 was registered for the overall service quality.

Just the judgment about the overall service was adopted in order to better analyse the service quality factors and define which are the most important for the users. Specifically, we conducted a correlation analysis between the satisfaction rates expressed about the service quality factors and the satisfaction with the overall service. In this way, we had the possibility to select the attributes to be included in the OP model. Specifically, we identified 9 attributes (table 1) out of the 37 investigated ones. The selected attributes belong to all the service aspects, with the exception of environmental pollution, which is not included in the OP model. We retain that correlation analysis gave satisfactory results because the nine attributes represent characteristics which can be considered as fundamental factors of a transit system like LRT. As an example, an aspect like operation

of elevators is peculiar of a metro system, as well as it is right that metro users give more importance to information at stations than on vehicle or internet, because they need clear and effective indications when they are at station to take the metro. It is also justifiable that users give more importance to the possibility to have a seat during the run than having clean seats, due to the relative shortness of the trip, as well as safety aspect is perceived more in terms of security against aggression than safety of the journey, because metro is maybe considered safer in terms of vehicle as regards a system travelling on the road. Table 1 shows the average rates of satisfaction expressed by the users about the nine service quality factors and the overall service, together with the correlation coefficient.

Service quality factors	Mean	Correlation coefficient
A. Number of trains per day (frequency of the service)	7.3	0.377
B. Regularity of the service (absence of interruptions caused by breakdown or incidents)	7.7	0.316
C. Easy access to stations and platforms from the street	8.2	0.303
D. Operation of elevators, escalators, etc.	8.2	0.311
E. Updated, precise and reliable information in stations (price, operating hours, stops, etc.)	7.8	0.348
F. Speed of the trip	8.1	0.387
G. Performance of the Customer Service (offices, web site, phone, deal with complaints, etc.)	7.0	0.359
H. Level of comfort on vehicle (seat availability or enough room while standing up)	6.4	0.384
I. Sense of security against theft and aggression in stations and on vehicles	7.3	0.301
Overall service	7.6	

Table 1. Average satisfaction rates

From table 1 we can observe that the most appreciated service factors concerns service accessibility, speed of the trip, information and regularity. On the other hand, the most correlated with the overall service, which we can define as the most important for the users, are speed of the trip, level of comfort, and number of trains per day. If correlation analysis give an indication of the most relevant service factors, in order to know how service quality factors can influence users' intentions to continue to use the LRT system, we need another kind of analysis provided by the OP model described in the following.

3. ORDERED PROBIT MODEL

3.1 Theoretical framework and proposed model

The OP model was originally developed by McKelvey and Zavoina (1975). In the OP model there is an observed ordinal variable Y , which is, in turn, a function of another variable Y^* that is not measured. Specifically, in the ordered model there is a continuous unmeasured latent variable Y^* , whose values determine what the observed ordinal variable Y matches. The continuous latent variable Y^* has various threshold points. The value Y_i of the observed variable depends on whether or not the value of Y^* crossed a particular threshold, as showed by the following formulas (1).

$$\begin{aligned}
 Y_i &= 1 \text{ if } Y_i^* \leq \mu_1 \\
 Y_i &= 2 \text{ if } \mu_1 < Y_i^* \leq \mu_2 \\
 &(\dots) \\
 Y_i &= j \text{ if } \mu_{j-1} < Y_i^* \leq \mu_j
 \end{aligned} \tag{1}$$

(...)

$$Y_i = m \text{ if } Y_i^* > \mu_{m-1}$$

In the population, the continuous latent variable Y^* is equal to (formula 2):

$$Y_i^* = \sum_{k=1}^K \beta_k X_{ki} + \epsilon_i = Z_i + \epsilon_i \quad (2)$$

where there is a random disturbance term ϵ_i normally distributed. The error term reflects the fact that the variables may not be perfectly measured, and some relevant variables may be not introduced in the equation.

By means of the OP we can estimate the expected average value of the Y_i^* (formula 3):

$$E(Y_i^*) = Z_i = \sum_{k=1}^K \beta_k X_{ki} \quad (3)$$

Once we have estimated β coefficients and the $(m-1)$ k cutoff terms, we can estimate the probability that Y will have a particular value. The formulas are the following (4):

$$P(Y_i = j) = \Phi(\mu_j - \mathbf{x}_i \boldsymbol{\beta}) - \Phi(\mu_{j-1} - \mathbf{x}_i \boldsymbol{\beta}) \quad (4)$$

$$P(Y_i = m) = \Phi(\mu_m - \mathbf{x}_i \boldsymbol{\beta}) - \Phi(\mu_{m-1} - \mathbf{x}_i \boldsymbol{\beta}) = 1 - \Phi(\mu_{m-1} - \mathbf{x}_i \boldsymbol{\beta})$$

Finally, the OP model can be used to estimate the probability that the unobserved variable Y^* falls within the various threshold limits.

Before deciding the structure of the models, we analysed the nature of the variables which we defined as response. Our intent is to focus on the behavioural intentions of the users to continue to use the LRT service again; we decided to assume just this variable as dependent variable of the proposed model. We introduced 27 independent variables representing “high”, “medium” and “low” satisfaction level with the nine service quality factors selected through the correlation analysis. All the variables representing responses of our model showed a natural ordering and then they are ordinal variables. We decided to transform the responses of the users given on 11-numeric scales into categorical variables having three levels: 0, 1, 2, where the level 0 groups the rates from 0 to 3, the level 1 from 4 to 6, and the level 2 from 7 to 10.

3.2 Model results

In order to calibrate the coefficients, the model was based on a particular reference case, which corresponds to level “2” that means “high” level of quality (satisfaction rate from 7 to 10). The statistics on the goodness of fit are adequate (table 2). Based on the p-values of the Wald tests, 14 variables are found to be significant with $p < 0.1$, and one variable with $p < 0.2$; the other three variables cannot be considered as significant. By observing the estimated probabilities we can state that the probabilities of the reference case are about 95% for “high” level and about 0.36% for “low” level. This was to be expected because the reference case corresponds to the case where satisfaction is “high” for any service factor. The result means that when all the independent variables assume a “high” level, the intentions to continue to use LRT has a probability of 0.95 to assume a “high” level.

The signs of the variables are negative, meaning that low and medium levels of satisfaction with the various characteristics have a negative effect on the intentions to use LRT as regards high levels of satisfaction with the same attributes. In other words, the analysis of probability values suggests that when the level of satisfaction with a certain attribute is

medium or low, the probability to have higher levels of intentions to use LRT system decreases than the reference case and the probability to have lower level tends to increase.

Service quality factor	Variable	Estimated Coefficient (β)	Wald	p-value	Estimated probability		
					0	1	2
	Reference case				0.0036	0.0452	0.9512
Number of trains per day	[V1=0]	-0.243	3.680	0.055	0.0403	0.1836	0.7761
	[V2=1]	-0.260	8.500	0.004	0.0281	0.1494	0.8225
	[V3=2]	0	.	.	0.0029	0.0412	0.9559
Regularity of the service	[V4=0]	-0.316	5.080	0.024	0.0490	0.1939	0.7571
	[V5=1]	-0.339	14.082	0.000	0.0339	0.1758	0.7903
	[V6=2]	0	.	.	0.0033	0.0441	0.9526
Easy access to stations and platforms from the street	[V7=0]	-0.246	1.225	0.268	0.0645	0.2296	0.7060
	[V8=1]	-0.250	6.361	0.012	0.0454	0.2035	0.7511
	[V9=2]	0	.	.	0.0044	0.0506	0.9450
Operation of elevators. escalators. etc.	[V10=0]	0.058	0.083	0.773	0.0451	0.1841	0.7709
	[V11=1]	-0.177	3.092	0.079	0.0436	0.1923	0.7641
	[V12=2]	0	.	.	0.0051	0.0537	0.9412
Updated, precise and reliable information in stations	[V13=0]	-0.302	3.653	0.056	0.0610	0.2246	0.7145
	[V14=1]	-0.236	6.871	0.009	0.0326	0.1679	0.7995
	[V15=2]	0	.	.	0.0037	0.0461	0.9503
Speed of the trip	[V16=0]	-0.721	20.913	0.000	0.0893	0.2810	0.6297
	[V17=1]	-0.392	17.753	0.000	0.0408	0.2014	0.7578
	[V18=2]	0	.	.	0.0038	0.0483	0.9479
Performance of the Customer Service	[V19=0]	-0.208	2.486	0.115	0.0428	0.1779	0.7793
	[V20=1]	-0.177	4.536	0.033	0.0206	0.1227	0.8567
	[V21=2]	0	.	.	0.0030	0.0410	0.9560
Level of comfort on vehicle	[V22=0]	-0.282	7.588	0.006	0.0331	0.1556	0.8113
	[V23=1]	-0.042	0.225	0.635	0.0141	0.0962	0.8896
	[V24=2]	0	.	.	0.0029	0.0394	0.9577
Sense of security against theft and aggression in stations and on vehicles	[V25=0]	-0.271	4.768	0.029	0.0354	0.1648	0.7998
	[V26=1]	-0.296	12.838	0.000	0.0249	0.1377	0.8374
	[V27=2]	0	.	.	0.0034	0.0427	0.9539
Number of observations		3211					
k_1 (threshold)		-3.322					
k_2 (threshold)		-2.129					
ρ^2 (Cox and Snell)		0.126					
ρ^2 (Nagelkerke)		0.268					
ρ^2 (McFadden)		0.212					
log likelihood		-547.609					

Table 2. Model results

By observing the estimated probabilities, the variables mainly influencing the intentions to use LRT are those corresponding to a low level of satisfaction about the factor. Specifically, the variable V16=0 has the biggest influence; if users reveal a low satisfaction level about “Speed of the trip”, the probability to have low level of intentions to use LRT system is equal to 8.93% and it is higher than the reference case. Instead, for the same factor, the probability to have high level decreases to the value of about 63%, lower than the reference case (95.12%). Other important variables regard “Regularity of service”

(V4=0) and “Updated, precise and reliable information in stations” (V13=0). Even in these cases, the probabilities to have low and medium levels of intentions to use LRT are higher than the probabilities for the reference case, and, at the same time, the probability to have high levels of intentions to use LRT decreases compared to the reference case.

Concerning the medium level of satisfaction, the variable V8=1 relating to “Easy access to stations and platforms from the street” presents the greatest impact on intention to use LRT. In fact, the probability to have low level of intentions to use LRT is equal to 4.54%, whereas the probability to have medium level is 20.35%. Both the probabilities are higher than the values representing the reference case. Instead, for the same factor, the probability to have high level decreases (75.12%) and is lower than in the reference case.

The variables V11=1 and V17=1, related respectively to the factors “Operation of elevators, escalators, etc.” and “Speed of the trip”, have a relevant impact on the users’ intentions and produce the decrease of high level of the probability and the rise of the low and medium level compared to the reference case.

Ultimately, the factor “Speed of trip” mainly influences users’ intentions to use LRT system, because both the two levels of satisfaction (low and medium) significantly impact on the probability to have higher or lower levels of intentions to use LRT. “Speed of trip” is also the factor showing the highest correlation with the overall service; so, it mostly influences users’ overall satisfaction as well as the intention to continue to travel by metro.

4. CONCLUSION

The main aim of this paper was to analyse the effects of users’ perceptions about some service quality factors on their intentions to continue to use transit service, which is represented in this case by the service offered by a new LRT system in the city of Seville. The model provided the evaluation of passengers’ behavioural intentions to continue to use LRT among three levels (low, medium, and high) by varying the level of satisfaction with any service quality factor considered in the model. In this way, the variables that cause the increase of the probabilities to have higher levels of intention to use LRT system can be identified. These findings can be useful for the operators to identify the service characteristics that play a role in the decision of the users to use the transit system.

ACKNOWLEDGEMENTS

The authors would like to thank the FEDER of European Union for financial support via project “Mejora de la calidad del TP para fomentar la movilidad sostenible: Metro de Sevilla (G-GI3002/IDIK)” of the “Programa Operativo FEDER de Andalucía 2007-2013”. We also thank all Agency of Public Works and the Ministry of Furtherance and Housing of Andalusia Regional Government staff and researchers for their dedication and professionalism. We thank the experts who participated in the questionnaire development and the team of trained interviewers who collected the data.

REFERENCES

CARDAMONE, A.S., EBOLI, L., FORCINITI, C., and MAZZULLA, G. (2014),

Willingness to use mobile application for smartphone for improving road safety, *International Journal of Injury Control and Safety Promotion*. DOI:10.1080/17457300.2014.945463.

CARDAMONE, A.S., EBOLI, L., FORCINITI, C., and MAZZULLA, G. (2015), How usual behaviour can affect perceived drivers' psychological state while driving, *Transport*. DOI: 10.3846/16484142.2015.1059885.

CHEN, C.F. (2008). Investigating structural relationships between service quality, perceived value, satisfaction, and behavioral intentions for air passengers: evidence from Taiwan. *Transportation Research Part A* 42(4), pp. 709-717.

CHEN, C.F., and CHAO, W.H. (2011). Habitual or reasoned? Using the theory of planned behavior, technology acceptance model, and habit to examine switching intentions toward public transit. *Transportation Research Part F* 14, pp. 128-137.

CHOWDHURY, S., and CEDER, A. (2013). A psychological investigation on public-transport users' intention to use routes with transfers. *International Journal of Transportation* 1(1), pp. 3-20.

DE OÑA, J., DE OÑA, R., EBOLI, L., and MAZZULLA G. (2013). Perceived service quality in bus transit service: A structural equation approach. *Transport Policy* 29, pp. 219-226.

DE OÑA, J., DE OÑA, R., EBOLI, L., FORCINITI, C., and MAZZULLA G. (2014), Key factors affecting drivers' perception of accident risk. *Accident Analysis and Prevention* 73 (1), pp. 225-235.

DE OÑA, J., DE OÑA, R., EBOLI, L., FORCINITI, C., and MAZZULLA G. (2016), Key Transit passengers' behavioural intentions: the influence of service quality and customer satisfaction. *Transportmetrica A*, DOI: 10.1080/23249935.2016.1146365.

EBOLI, L., and MAZZULLA, G. (2007). Service Quality Attributes Affecting Customer Satisfaction for Bus Transit. *Journal of Public Transportation* 10(3), pp. 21-34.

EBOLI, L., and MAZZULLA, G. (2009). An ordinal logistic regression model for analysing airport passenger satisfaction. *Euromed Journal of Business* 4(1), pp. 40-57.

EBOLI, L., and MAZZULLA, G. (2012). Structural Equation Modelling for Analysing Passengers' Perceptions about Railway Services. *Procedia-Social and Behavioural Science* 54, pp. 96-106.

EBOLI, L., and MAZZULLA, G. (2014). Relationships between rail passengers' satisfaction and service quality: a framework for identifying key service factors. *Public Transport* 7(2), pp. 185-201.

LAI, W.T., and CHEN, C.F. (2011). Behavioral intentions of public transit passenger. The roles of service quality, perceived value, satisfaction and involvement. *Transport Policy* 18, pp. 318-325.

MCKELVEY, R.D., and ZAVOINA W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology* 4, pp. 103-120.

ZEITHAML, V.A., BERRY, L.L., and PARASURAMAN, A. (1996). The behavioral consequences of service quality. *Journal of Marketing* 60(2), pp. 31-46.