Examination of Crowded Public Transport Service Perception of Users by using Binary Logit Model

Mahmut Esad Ergin

Istanbul Commerce University, Logistics Management Department, Istanbul, TURKIYE

Abstract: *The livability of cities may be described by the inexpensive infrastructure and social possibilities they provide, as well as the user's perception of these options. The major goal of this study is to employ a binary logit model to capture road users' perceptions of the crowdedness of public transportation. Aside from transportation mode characteristics and user socioeconomic factors, the study looked at how the "being crowded" condition of public transit influences consumers' perceptions of service quality. A survey containing revealed preferences (RP) and stated preferences (SP) sections was done in Istanbul in 2015 with a total of 175 randomly selected respondents as part of this study. The acquired data is modeled using a binary logit model that compares the existing mode (old) with the new mode (new) that is determined using fictional variables. Furthermore, the users' willingness to pay is investigated. As a consequence, people do not want to pay much more for transportation, and one of the reasons for this is their familiarity with the present system.*

Keywords: Binomial logit model, Crowdedness, Public transport, Willingness to pay

1. Introduction

The livability of cities can be explained by the affordable infrastructure and social opportunities they offer, and especially by how these opportunities are perceived by the user. Today, there is a rapidly increasing urbanization situation. Today, 55 percent of the world's population lives in cities, and by 2050, that number is predicted to rise to 68 percent. According to projections, urbanization, or the steady movement in human population from rural to urban regions, combined with global population increase may add another 2.5 billion people to urban areas by 2050, according to a new UN data set (UN, 2019; Demographia, 2019). These data and forecasts show that 7 out of 10 people will live in cities in the near future.

In urban plans, development decisions are made and strategies are developed with a certain foresight and some other assumptions. However, sometimes, unexpected occurrences such as mass migration, politics, etc. the prepared urban plan cannot meet the real demand that emerges above the expected. As a result of this, especially infrastructure services may not be sufficient for that unforeseen demand. While transportation is one of the most important infrastructure services of the city, which is also one of the important variables of the city's economy, cannot fulfill the demand. Transportation is one of the services that can be most affected by such unplanned growth.

It can be said that Istanbul is one of the best examples of such a city. According to TUIK (2020), it has a population of 15.6 million. The daily population increases with domestic and foreign tourists. Especially in recent years, the population has increased unexpectedly due to the waves of mass migration from different countries to Turkey. Due to the fact that it is very difficult to intervene in the city with its historical structures, topography and protected areas, there is an inadequacy in urban infrastructure services for such a great population. Whether the provided service is sufficient, or in other words, its quality is also related to how users perceive these infrastructure services.

The main purpose of this study is to catch the perception of the road users of the crowdedness of the public transport system by using a binary logit model. In the study, besides the features of transportation modes and socio-economic characteristics of users, it was investigated how the "being crowded" criteria of the public transport affects users' perception of public transportation service quality. In addition to this investigation, willingness to pay of the users are also examined in return for the improvement that can be made within the framework of the crowd of the public transport. Consequently, the perceived quality of the public transport under various income effect is also analyzed.

The remainder of this paper is structured as follows. Following section describes the survey analysis and study area. The next part, the model formulation and the methodology is discussed, and then results of the modelling are presented, followed by conclusion as a lastpart of the studywith the recommendations for the furhter studies.

2. Lıterature Review

Seco and Gonçalves (2007) refers a study called as Project QUATTO that focuses on the quality of the public transport. In this study, the quality is divided into four parts in accordance with the various point of view. Expected quality is defined as the degree of quality that customers expect, the targeted quality is the level of quality that the operator aims to offer, the delivered quality is the quality level experiencing today, and the percieved quality is the level that passengers percieve on their daily routine trips. As a result, if quality is defined as being able to respond to the wishes of the customers of users, it can be said that the quality of the public transport system is "as long as it meets the wishes and expectations of the passengers using this system".

Rather than the perception of the entire transportation system, it should be more important to determine the public transportation perception of the users in terms of more effective traffic management. Del Castillo and Benitez

(2021) offers that the quality of the public transport system can be obtained by users scoring variables such as punctuality, size of the network, integration of the lines. It was observed that public transport users care more about waiting time, cleanliness and comfort variables and give less value to variables such as driver's polite behavior, crowding and travel time (Dell'Olio, vd., 2011). Li (2003) looks at how people felt about how much time they spent commuting in the city and establishes a model to explain the travel time perception of the users. In another study, the opinions of urban commuters on the interplay between this public transportation system's accessibility, mobility, and seamless connectedness (Cheng and Chen, 2015). The mode choice decision is not only dependent by the measurable cirteria. It also affected by other criteria that cannot be measured easily and precisely such as stress, crowdedness, behavior of the driver, passenger profile, line route effect on the user, etc. In this context, Li and Hensher (2011) revealed that users do not tend to wait for the next public transport vehicle despite being crowded, due to the increased value of their time. Overcrowding might reduce the perceived attractiveness of public transport in comparison to other transportation modes. Furthermore, if congestion causes long travel times, it might restrict public transport accessibility in an urban region. Understanding how (over)crowding in public transport networks affects passengers' nominal and actual trip times is critical for public transportation planning (Yap and Cats, 2021).

3. Methodology

Mode selection models in transportation are one of the important pillars of transportation planning studies. These models can be used in studies such as forecasting and scenario analysis for the future, as well as providing clues about the behavior of individuals using the current system. In this context, issues such as how the user will react to a price increase in public transportation vehicles, how the users of the system will behave in case of an increase or decrease in the frequency of flights, whether the behavior of private vehicle owners will change as a result of an improvement in the public transportation system are the issues that mode selection models can respond to. The flexibility method cannot present the exact results of changes in the transportation system such as price increases, public transportation frequency and route (Horowitz et al., 1986). Collective derivative discrete models, on the other hand, are not sensitive enough because they collectively examine the observed average travel relationship of a group of travelers or residents of an area. On the other hand, discrete demand models deal with the individual preferences of passengers. Collective demand models were used until the 1980s, and after these dates, discrete demand models began to come to the fore (Williams, 1981).

The logic in discrete choice models is that users choose the mode that will benefit them the most which is the one of the basic assumptions of the Random Utility Theory. The users' benefits are defined by a function and this function is called the utility function. The utility function is usually a linearly expressed function as it is shown in Equation **[1]**.

$$U_i = V_i + \varepsilon_1 \tag{1}$$

The U in the formula represents the person's utility. The variable V in the equation, on the other hand, expresses the quantities that can be measured by observation (age, travel time, travel cost, income, etc.). This part constitutes the determining components of the model and is the part that the model can explain. However, the ε is an option-specific error term. This section describes features that are not easy or possible to measure or observe, but have an impact on mode or alternative selection. This part is called as random or stochastic part of the model. It includes variables such as the person's perception, comfort and convenience. However, due to the error term, it is accepted that these elections take place with a certain probability of being selected, not under certainty. Therefore, if the issue is a comparison of being selectedamong alternatives, the mathematical process can take place as in Equation **[2]**:

$$P_r(1) = P_r(U_1 > U_2)$$
$$P_r(1) = P_r(V_1 + \varepsilon_1 > V_2 + \varepsilon_2) \tag{2}$$
$$P_r(1) = P_r(V_1 - V_2 > \varepsilon_2 + \varepsilon_1)$$

Here, the P value takes a value in the range of 0-1.

The logit and probit models are the two most popular forms of binomial choice models used in practice. The probit model assumes a normal distribution of errors, whereas the logit model assumes a logistic distribution (Koppelman and Bhat, 2006). In this study, the logit model is used and the mathematical expression of the possibility of to be chosen of an alternative is given in equation **[3]**.

$$P_r(1) = \frac{e^{V_1}}{e^{V_1} + e^{V_2}}$$
$$P_r(1) = \frac{e^{V_1}}{e^{V_1} + e^{V_2}} \times \frac{e^{V_1}}{e^{V_1}} \tag{3}$$
$$P_r(1) = \frac{1}{1 + e^{(V_2 - V_1)}}$$

$P_r(1)$ representsprobability of being choosen of alternative 1.

4. Survey Analysis and Data

Within the scope of this study a survey conducted in Istanbul in 2015 with a total of 175 randomly selected people, 90 on the European side and 85 on the Asian side. The survey was carried out by face-to-face interview method and it was tried to ensure that the participants understood the questions completely and correctly. For the survey, the four most preferred districts to participate in the activity in Istanbul were selected and presented in Figure 1(Şişli, Beşiktaş, Üsküdar and Kadıköy) and were completed regardless of day and time, usually on the way out of work. Survey data consisting of Revealed Preference (RP) and Stated Preference (SP) parts were used in the model estimation used for statistical evaluations and interpretation of perception status.
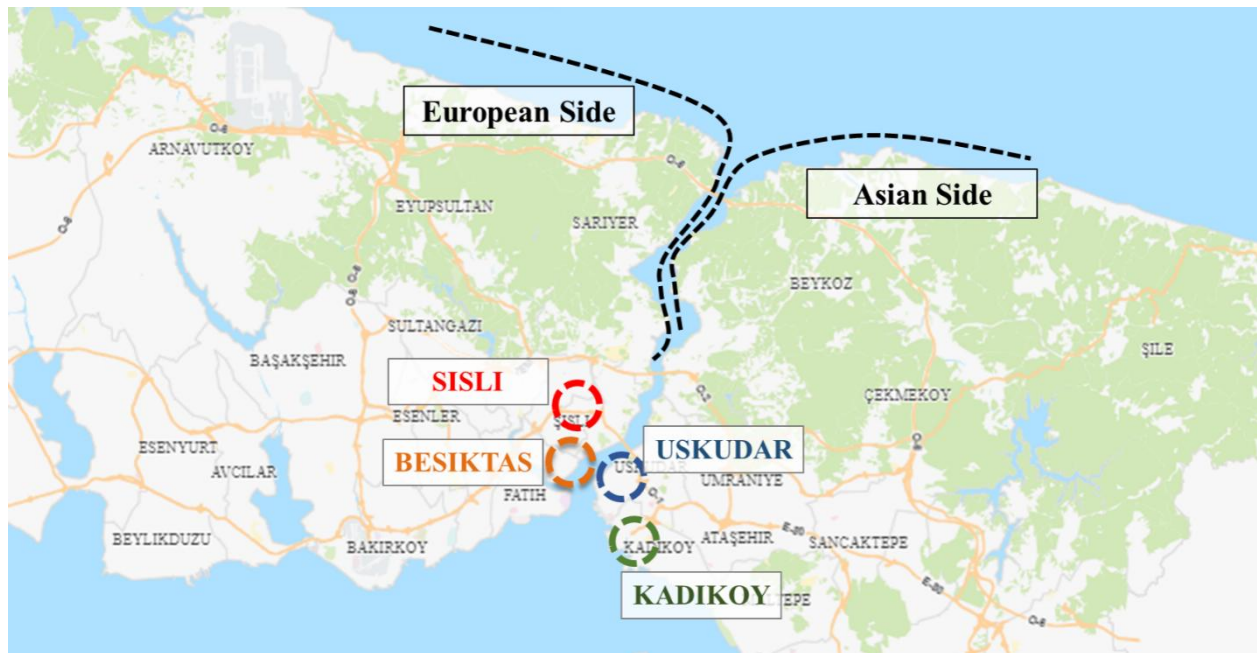
Figure 1: The selected districts in Istanbul

In this context, the perception of private vehicle and public transportation users about the public transportation system was measured. In this measure, the respondents were asked to score between 1 and 5, and they were asked how much they could afford to pay to improve the system performance in terms of crowdedness. 1-5 likert scale means 1-very bad, 2-bad, 3-moderate, 4-good, 5-very good. As a result of the examination method, the respondents were grouped into two groups such asprivate vehicle users and public transport system users. Accordingly, the binary logit model is considered as the best choice to establish a model to assest the perception of the users of the public transport quality.

In this study, the data collected with 175 questionnaires were modeled with binary logit modeling approach. A total of 8 variables were used in the models. These variables were collected in 4 groups. Among the variables used in the models, the travel time, the ratio of the journey cost to the income (Travel Cost/Income) and the use of private vehicles are taken in terms of showing how the users travel and constitute the first group variables of the model. The variable only "income" was not taken as a variable in the model. Because what matters is the ratio of the travel cost in the user's income. Age, Gender and Marriage Status variables describe the passenger profile and constitute the second group variables of the model. The Home-Work variable describes the purpose of the trip, and the "AVG" variable is taken as a variable that takes a value of 1 if the user gives a score higher than the middle value (value 3) to the mode they choose, and 0 otherwise. "Additional Payment" is classified as the fourth group variable. The additional payment was created to reveal how much money users are willing to pay for better conditions for the crowding factor, in other words, for a more comfortable public transportation system.

Table 1: Groups of Variables

1. Group Variables: Travel profile	1)Travel time,
	2) Travel Cost/Income,
	3) Private car usage
2. Group Variables: Passenger profile	1)Age,
	2)Gender,
	3)Marriage status
3. Group Variables: Travel purpose and quality	1)Home – Work,
	2)Average score (AVG)
4. Group Variables: Fictitious variable	1)Additional payment

According to the survey data, survey participants stated that they do not want to pay extra for a more comfortable transportation service (26.8%). Considering the payments that can be made, it is seen from the data that stated payment amount of money densify around 1 Turkish Lira (TL), 2 TL, and 5 TL. The proportions of the payment amounts are, on average, 16.6%, 18.9%, and 8%, respectively. The remaining approximately 30% is also distributed into various values changing between 0 – 10 TL. As a result of the reasons above, the fictitious variables have been determined as 1 TL, 2 TL and 5 TL as the monetary amounts that users can give in addition to the current transportation cost for better public transportation system service conditions.

While obtaining the model data, it has been examined whether each individual will stay in the currentmode (it is called as OLD in the model) or switch to the new mode (it is called as NEW in the model) according to the price they specify against the accepted additional payment amounts. In this case, the same 175 individuals were taken separately for Additional Payment = 1 TL, Additional Payment = 2 TL and Additional Payment = 5 TL and the total number of data was obtained as 175 x 3 = 525, finally. In the modeling study, a binary logit model established.To see whether people remained in the OLD mode or they chose the NEW mode, based on the values they were willing to pay. In addition, the

following model reveals which independent variables have a more active role in the mode selection of individuals.

5. Results

Descriptive Statistics
The average age of the people surveyed is 29.7. 67% of the people surveyed are male and 67% are single. In addition, the average household size was 3.09. This value is close to the average household size of 3.53 given in the Istanbul Transportation Master Plan (IUAP) 2006 Report. It has been determined that 69% of the survey participants are working, 50% are homeowners and 45% have a private vehicle.

According to the survey results, the monthly average household income is 5,300 TL. According to the data, 45% of the trips are originatedfrom the European side, 35% of the

trips are made between the two continents. As a result of the analysis of the purpose of the trips, it is seen that 41% of the trips are home based work trips. Although 45% of the participants have a private vehicle, the rate of those who travel only with their private vehicle is 20%, and the rate of those use park and ride system is only 6%. Moreover, the average travel time is 48.8 minutes in one direction with the average travel cost is 6.25 TL. Likewise, the average travel time is 48.9 minutes according to the İUAP report.

Modelling Results
The utility function was estimated only for the NEW mode. Thus, the estimated coefficients express the comparative effect with respect to the OLD mode. The model was estimated 3 times for the entire sample and for two tiers, private vehicle and public transport users.

Table 3: The binary logit model results

Variables	Whole Sample		Private Car Users		Public Transport Users	
	Coefficient	t-stat.	Coefficient	t-stat.	Coefficient	t-stat.
Travel time	0,020	4,957(a)	0,013	1,698(b)	0,028	5,083(a)
Travel cost/Income	-7,717	-1,245	2,989	0,261	-14,394	-1,866(b)
Additional payment	-0,475	-7,080(a)	-0,204	-1,809(b)	-0,630	-7,025(a)
Age	0,072	4,811(a)	0,082	3,077(a)	0,070	3,416(a)
Gender	-0,424	-1,886(b)	-0,899	-1,737(b)	-0,235	-0,889
Marriage status	-0,587	-1,936(b)	-0,600	-1,096	-0,645	-1,631
Private car usage	0,198	0,746	-	-	-	-
Home – Work trips	-0,188	-0,811	-0,093	-0,202	-0,317	-1,091
Average score	-0,557	-1,540	-1,659	-2,712(a)	0,357	0,753
Constant	-1,816	-3,895(a)	-2,029	-1,910(b)	-1,768	-3,098(a)
No. of observations	525		138		387	
LL()	289,318		81,674		197,659	
LL(M)	348,387		94,724		252,103	
- 2LL	118,138		26,100		108,888	
ρ^2	0,170		0,138		0,216	

*(a) significant at 95% (1,960), (b) significant at 90% (1,645), significants are written in **bold**.*

If the journey time increases, users tend to switch to the new type. This is normal as users will want to take shorter trips. As the Travel Cost/Income ratio increases, the probability of public transport users staying in the old style increases. While this variable is significant only in the 90% confidence interval for public transport users, it is not statistically significant for the sample and private vehicle users. As the Additional Payment value increases, the probability of users staying with the old mode increases. In other words, as the amount of money users sacrifice increases, users generally tend to stay with the old mode. Users do not want to pay much more money for transportation, the habituation to the current system is one of the reasons for this situation. When the age variable is analyzed within the framework of crowding, it is an important variable for the general sample, private vehicle users and public transportation users, and it is statistically significant at the 95% confidence interval. As age increases, users are more likely to switch to the new mode. It can be said that this situation arises as the age increases, as the crowding makes the journey more difficult. The gender variable is not statistically significant only for public transport users. A negative coefficient of the variable means that if the gender is male, the probability of staying in the old type increases. While Marriage Status is statistically significant in the 90% confidence interval, it has no statistical significance for private vehicle and public

transportation users. Across the sample, users are more likely to stay in the old style if they are married. The Average Score variable, which is determined according to the points given by the users to the existing system, was only significant for private vehicle users. This situation can be considered as a reason why private vehicle owners do not use the public transportation system or as a reason why they find the public transportation system unattractive.

6. Conclusion

Participants generally assess the urban transportation system as "moderate", "poor" or "very bad". Only in the light of this information, it can be said that it would be appropriate to make investments in the urban transportation system within this framework in order to increase the customer / user satisfaction. In particular, the fact that the average score variable was obtained with a negative coefficient in the model results means that the Average Score value will decrease as the density to be experienced in the public transport system increases, which means that there are conditions that deter users from using the public transport system. As a result of the potential enhancement, it can be said that private vehicle users can be encouraged to use public transportation. For example, public transportation modes can serve without being stuck in the existing traffic,

therefore, due to the possibility of shorter trips, private car users may tend to choose public transport modes. The travel cost over income is a more comparable variable. This variable is not statistically significant for private vehicle users, however, as the ratio of the travel cost over income increases for public transportation users, it is predicted that they tend to remain in the old mode in the context of crowdedness of the mode. As the ages of the users increase, their desire to have a more comfortable and safe trip emerges. Female users want to get improved public transportation service more than male users.

First of all, public transportation system quality should be offered for individuals who have no choice but to use public transportation and then in order to attract users from using private vehicles. Increasing the frequency of public transport according to the results of this analysis, increasing the number of seats in public transportation vehicles, and not allowing more than a limited number of passengers to get into the vehicles are some examples of improvements that can be done.As can be seen from the examples, in order to erase some bad impressions that have been ingrained in individuals' memories, and to encourage individuals to use the public transportation system, first of all, an improved and comfort public transport service should be provided with an affordable fee.

References

[1] Cheng, Y.,and Chen, Y. 2015. Perceived accessibility, mobility, and connectivity of public transportation systems, Transportation Research Part A: Policy and Practice, 77, (386-403. https://doi.org/10.1016/j.tra.2015.05.003

[2] Demographia. (2020). Demographia World Urban Areas: 16th Annual Edition June 2020. (05.11.2020, http://www.demographia.com/db-worldua.pdf)

[3] Del Castillo, J. M. ve Benitez, F. G. (2012). "A methodology for modeling and identifying users satisfaction issues in public transport systems based on users surveys", Procedia-Social and Behavioral Sciences, 54, 1104-1114.

[4] Dell'Olio, L., Ibeas, A. veCecin, P.(2011). "The quality of service desired by public transport user", Transport Policy, 18, Sf. 217-227.

[5] Horowitz, J. L., Koppelman, F. S. veLerman, S. R. (1986). "A self – instructing course in disaggregate mode choice modelling", Washington, Federal Transit Administration.

[6] Koppelman, F.S. and Bhat, C. (2006). A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models, U.S. Department of Transportation Federal Transit Administration.

[7] Li, Y., 2003. Evaluating the Urban Commute Experience: A Time Perception Approach. Journal of Public Transportation, 6 (4).

[8] Li, Z. veHensher, D. A. (2011). "Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal", Transport Policy, 18, 880-887.

[9] Ortuzar, J. D. veWillumsen, L. G. (2011). "Modelling transport: 4th Edition", ISBN: 9781119993315. BirleşikKrallık. John Wiley & Sons, Ltd.

[10] Seco, A., and Gonçalves, J.H. (2007). The quality of public transport: Relative importance of different performance indicators and their potential to explain modal choice, Urban Transport and the Environment in the 21st Century – Urban Transport, Coimbra.

[11] Türkiyeİstatistik Kurumu (TUIK). (2020). (05.11.2020, https://data.tuik.gov.tr/tr/main-category-sub-categories-sub-components2/)

[12] United Nations, (2019). World Urbanization Prospects 2018: Highlights, Department of Economic and Social Affairs. https://population.un.org/wup/Publications/Files/WUP 2018-Highlights.pdf Accessed: 27.12.2021.

[13] Yap, M., and Cats, O. 2021. Taking the path less travelled: Valuation of denied boarding in crowded public transport systems,Transportation Research Part A 147, 1–13.

[14] Williams, H. C. W. L. (1981). "Travel demand forecasting: an overview of theoretical development. In D.J. Banister and P.G. Hall (eds.)", Transport and Public Policy Planning,Londra.