Analysis of Preferences for the Use of a Bicycling Sharing System in Athens

G. Yannis, P. Papantoniou, E. Papadimitriou, A. Tsolaki

Department of Transportation Planning and Engineering
National Technical University of Athens
5, Heroon Polytechniou Str, GR-15772, Athens
geyannis@central.ntua.gr,
ppapant@central.ntua.gr
nopapadi@central.ntua.gr
athinatsolaki24@gmail.com

ABSTRACT

As nowadays sustainable urban mobility is becoming more and more critical for a balanced combination of economic development and living standards, the new policy of Athens includes the promotion of cycling and the implementation of a bike-share program, as increasing cycling might lead to several advantages such as reducing congestion, improving air quality, providing complementary services to public transport, improving city’s image and branding as well and offering residents an active mobility option. The objective of the present research is the analysis of the parameters influencing the use of a bicycle sharing system in Athens, a city without a strong culture of cycling. Among other things, the influence of the existence of bicycle lines was studied, in order to assume whether safety plays an important role in user’s decision or not. For this purpose, a specially designed on-line survey to a sufficient sample of commuters and residents of the Municipality of Athens was carried out in order to collect their characteristics, attitudes and behaviour as well as their preferences through the stated preference methodology. A logistic regression model was developed describing travellers’ choices based on the characteristics of the travellers and their journeys. Both multinomial and mixed logit models were developed. Results indicate that the probability for a traveller using bicycle instead of passenger car or public transport depends primarily on travel time, cost and comfort level, along with the age and the gender of the traveller. More specifically, travel time was found to be the most important parameter for the participants in order to choose using the bicycle instead of the passenger car, while the absence of specific bicycle lanes constitutes the basic deterrence reason for the use of bicycles.

Keywords: Bicycle sharing systems, stated preference survey, multinomial logit model, mixed logit model.
1 INTRODUCTION

An alternative way of everyday travels, gaining more followers worldwide each day, bicycle is becoming part of our lives. Environmental friendly, needing no fuel and the minimum of space, no noise pollutant, cycling is the safest vehicle regarding both pedestrians and its rider. In an era when the environment is in the limelight and efforts on health, safety, standards of living and economic prosperity are being made, it is commonly acknowledged that cycling is a very effective and contemporary way of commuting [1].

During the recent years, Bicycle Sharing Systems (BSSs) have gained a lot of attention, being a transport alternative for urban city centres that improves the quality of public transport and co-exists harmonically with the rest of the transport systems. The ways that a BSS can benefit a modern city are various [2]:

- Reduction in the level of congestion, and air quality improvement, as it provides an alternative to travelling by car.
- Increased accessibility, to areas where otherwise would be difficult to approach.
- Increased convenience in covering the distance from the stop/station to the final destination.
- Improvement of bicycle’s public image, as BSSs shape the cycling culture of a city and bring a modern attribute to it.
- Provides an alternative way of travelling in the city by car or by public transport.
- Healthier individuals, as it brings both physical and spiritual benefits for the users.

However, BSSs have limitations as well. Bicycles at the Stations being sensitive to thievery and vandalism, the need of frequent station maintenance, and the difficulty in achieving equilibrium between demand and supply are the main ones [2].

There has been a number of studies regarding the history and evolution of BSSs [3, 4] marketing strategies and safety issues [5, 6] the distribution of bicycles at the stations [7,8], or even time and space pattern analysis for the use of bicycles [9]. Moreover, a lot of studies examine network and facility issues [10, 11], whereas others focused on the influence BSSs have on the rest of the transport alternatives [12, 13].

Romero et.al. [12], in the context of designing an effective BSS, investigate, among others, the interaction between cars and BSSs, while conducting a behaviour analysis of the user, emphasizing on the characteristics of the potential bicycle user. Interestingly, a study from Jäppinen et.al. [13] presented a model for the potential influence of a BSS in Helsinki to the travel times of public transport, concluding that the implementation of a BSS can reduce that time over 10%.

An interesting study comes from Liu et.al. [14], which examined the causes of usage reduction of the first generation of the BSS of Beijing and tried to find ways to re-establish it. In addition to BSS studies, there is quite a big number of studies that have been concerned with stated preference (SP) method, with this of Axhausen et.al. [15], being one of the earliest. This study presents the primal effort to formulate a logistic model for the choice of the kind of parking space.

The importance of BSS geographical coverage and price in shaping the profile of users is also acknowledged based on data from the use of Barclays Cycle Hire (BCH) in London, U.K. [16,17]. Overall, this evidence indicates that residents in less affluent areas can and do use bike sharing systems if these are made available in their local areas. In fact, Ogilvie and Goodman [17]
found that trip rates amongst registered users were higher amongst residents in poorer areas after adjusting for the fact that these poorer areas were less likely to be near a BCH docking station.

One more stated preference study, that examines the threshold of private parking space cost in the area of the business centre of Athens, comes from Tsampoulas [18]. By studying the driver’s behaviour over the change of an already chosen parking space, this study seeks the threshold of increase on the current price of parking spaces that would make drivers choose something other than their cars. The main conclusion of the study is that drivers parking at more peripheral spaces are more sensitive to price changes, while a tendency of changing location when faced with increased prices is observed. Moreover, an overall reduction of 72% was recorded in the number of vehicles parked the whole day at municipal parking spaces for increases up to 123% in the prices of those spaces. That is without indication of moving over to a competitive private parking space.

The present study aims to evaluate the characteristics influencing the choice of a BSS for travelling in the Municipality of Athens over the other alternatives (car, public transport, other) with the use of a mathematically logistic model. The structure is the following: after the literature review of relevant studies comes the data collection and processing which involves the design of the questionnaire, the completion of the research and the characteristics of the sample. Consequently, the statistical models previously developed are recorded. Finally, based on the analysis of the statistical models outcomes, the main conclusions are presented, along with points for further study.

2 DATA COLLECTION AND PROCESSING

The present study is trying to document and analyse the response of the users of Municipality of Athens’ transportation network after the addition of a BSS. The aim is a sensitivity analysis of travellers’ stated preference towards the new travelling conditions that come after the implementation of a BSS and whether they would choose it over their current transport mode or not. Stated preference method was chosen as a suitable method of analysis. Moreover, a very common way of data collection for such studies, which was used here as well, is by a specially designed questionnaire, which was filled out in the form of an e-survey.

In stated behaviour surveys, a reference questionnaire is built, based on a list of selected topics and a representative sample of population is interviewed. The survey approach can employ a range of methods to answer the research questions such as postal questionnaires, face-to-face interviews, and telephone interviews.

They produce data based on real-world observations allowing investigating new situations, outside the current set of experiences. Furthermore, the breadth of coverage of many people or events means that it is more likely than some other approaches to obtain data based on a representative sample, and can therefore be generalizable to a population. Moreover, surveys can produce a large amount of data in a short time for a fairly low cost, making it easier to planning and delivering end results. On the other hand, the nature of questions is often hypothetical and the actual behaviour is not observed, while the data that are produced are likely to lack details or depth on the topic being investigated [19].
2.1 Questionnaire design

During the design of the questionnaire, special attention was given at its structure, in order to will satisfy the needs of the study and at the same time follow certain basic principles in order to ensure the validity of the results deriving from the study. The questionnaire consists of four parts.

The first part of the questionnaire comprises eight questions regarding the driving behaviour and habits of the participants. Its purpose is to introduce the respondents to the study’s concept and contents. The recording of those characteristics can help derive various outcomes when combined with the responses of the third and fourth part of the questionnaire. The data collected from this first part can also be used for assessing sample representativeness.

The second part includes three questions targeted to make the interviewee familiar with the concept of using bicycle and the reasons of choosing it for a daily transport mode or not. In the first question, the respondent has to state the criteria with which he chooses the transport mode he uses. The second question gives considerations regarding the disadvantages of using bicycle for commuting, whereas the third question presents some of the advantages.

The third part of the questionnaire aims to investigate users’ switching behaviour through a specific stated preferences (SP) survey. It consists of three questions which are the core of the statistical analysis that follows. In the first question, the interviewee is presented for the first time with the suggested BSS, along with a map showing some indicative stations. He has to state how many times he would use it, and justify. The reasoning aims to make him consider of the advantages and disadvantages presented to him in the second part, instead of answering biased or recklessly. In the second question, some suggested prices for a rental hour are given and the respondent has to consider how much he would be willing to pay. Finally, the third question of this part includes 8 possible scenarios with the choice of A, B, C or D, asking him to denote his preference between those four alternatives (stated preference method).

Before the introduction of the scenarios comes the description of the concepts that are going to follow. First and foremost, it is made clear that the scenarios are referring to net travel time duration from 10 to 30 minutes. Additionally, the concepts of travel time, cost and comfort are explained, as met at the scenarios of the third part. And finally, public transport is defined, in order to be clear for the respondent when asked to choose. Each of the eight scenarios describes some transport conditions with fixed time, cost and level of comfort, distinct for every mode. The interviewee is asked to choose between the modes for each of the scenarios.

Last but not least, the fourth part contains seven questions with regard to the demographic characteristics of the sample, i.e. sex, age, marital condition, educational level, annual family income, as well as the nature of the working schedule (fixed or flexible).

The most important part of the questionnaire, on which the study was based and was used for the implementation of the stated preference method, is the hypothetical scenarios included in the third part of it. Orthogonal design was applied to the present study, according to which there should not be correlations between the attributes of the alternatives. That stands for this study as well, if the metro and overground railway are considered the most representative modes of public transport for the Municipality of Athens. Thus, the values of time, cost and comfort of one alternative can be considered with no correlation to another. The SP experiment was conducted through a selection of all of the possible scenarios starting from the Full Factorial Design scenario. This means that there is a scenario for every possible combination of
the variables. To put it into perspective, for this study there were four alternatives (Car, Public Transport, BSS, None of the above), each of then having three variables (time, cost, comfort). Two of those variables (cost, comfort) can receive two different values, whereas the third can receive three, resulting in a total of 1728 scenarios. Then a subset of scenarios was generated introducing the partialisation techniques of the experiment known as Fractional Factorial Design [20]. This eliminates completely some scenarios while retaining orthogonal comparisons which allow for the estimation of the main effects. Each of the 252 surveyed users responded to 8 SP scenarios, thus 2016 observed behaviours were obtained.

2.2 Survey Implementation

The survey data were collected from a sample of 252 individuals aged 18 and over. Respondents were randomly selected users of the Municipality of Athens’ transport system, travelling at least once a week inside the Municipality. In that final form of the questionnaire, a map was included with the suggested BSS stations. The stations proposed follow the main routes of the Municipality, where the number and distances between followed the example of Brussels’ BSS stations, as they serve approximately the same population per square metre as the one in the present study. The purpose of the map was to make the network format of the BSS familiar to the interviewees, and help realise that bicycle gives access to places where car can be faced with major congestion, or parking difficulties, and public transport would require two or even three transfers.

The website was spread via pages of social network, personal e-mail, and even personal contact at central areas of the Municipality of Athens, aiming to obtain a sample with various socioeconomic characteristics, as objective and representative as possible. The survey lasted almost two months resulting in the collection of the data base necessary for the statistical analysis and formation of the final mathematical model that was to follow.

2.3 Sample Characteristics

In order for the results to be as representative and valid as possible, an effort was made for the sample of the present study to depict the following basic principles:

- **Goal-oriented**: For a survey regarding the use of bicycle as an alternative transport mode, the sample should be from the travelling population that has the ability to use a bicycle, when given as a choice. More specifically, when regarding travels inside the Municipality of Athens, the sample must derive from the population of users of the transport system of the Municipality, at least in a weekly basis.
- **Law of Inertia of Large Numbers**: The larger the size of the sample, the more accurate the results are likely to be. A total of 252 responses is a sufficient number regarding the nature of the present study.
- **Accurate representative of the universe**: For the present study, special attention was given so that the sample should consist of people with various socioeconomic background. For that reason, high importance was given to the first part of the questionnaire (socioeconomic questions) in order for the sample to contain individuals from every category in sensible proportions.
- **Proportional**: A sample should be large enough to represent the universe properly. The sample size should be sufficiently large to provide statistical stability or reliability.
- **Random selection**: Any item in the group has a full and equal chance of being selected and included in the sample.
Some indicative characteristics of the sample used for the purpose of the study are shown in the charts below in the form of a percentage distribution.

**Chart 1.** Classification of scenario responses regarding occupation

![Chart 1](chart1.png)

**Chart 2.** Classification of responses in “How often would you use the BSS” regarding the type of work schedule

![Chart 2](chart2.png)

From the descriptive analysis and the charts given it appears that the participants of the study seem to be younger at age (the majority being under 35 years old), resulting in a bigger percentage of singles, and a relatively big percentage of students, with a small proportion of them possessing a car. As a consequence, there is a higher than anticipated interest in bicycles, depicted in the results of both simple and complex statistical processing.

At the same time, a good distribution regarding the sexes is observed, along with a good distribution concerning the weekly travels inside the Municipality of Athens, while the majority of the sample (51%) takes an average amount of time (approximately 30 minutes) for a typical daily travel.
2.4 Theoretical background

For the statistical analysis, logistic regression was applied. The model parameters were estimated with the application of logistic regression analysis, which is commonly employed in transport mode choice situations, to identify those parameters that are significant in affecting these choices. Before deciding for the choice of logistic regression, other methods were also considered, like linear regression, discriminant, probit and logit analysis. Logistic regression analysis was chosen in this work not only because it allows for the development of models on alternative choice probabilities for discreet dependent parameters [21], but also because it makes easier the identification of the sensitivity of the impact of the parameters examined. The logistic regression constitutes a suitable method for the elaboration of data resulted from independent observations or statements of the public and is considered appropriate for analysis of answers to stated preference surveys.

The estimators of the logistic regression model parameters are calculated with the maximum likelihood method. The mathematic model that results initially from the analysis provides the utility function, which is based on the random utility theory. This logistic regression model associates in a linear way the parameters influencing the decision. The probability of this choice is directly calculated through appropriate transformation of the utility function. The relation between probability and utility function is not linear. The logistic regression can be used for the development not only of binary models, but also of models with more alternative choices.

3 RESULTS

Logistic regression models were developed as part of the statistical analysis. To be more precise, four models were created, one for each alternative option. In the beginning, the analysis was conducted under an increasing number of independent variables in order to clarify the relationships between them and to identify those that influence the four dependent variables, i.e. the bicycle, the car, the public transportation and the ‘none of the above’ option. The results of these models were evaluated by applying the nested logit model, which showed that the hierarchical choice between motorized and non-motorized means of transportation is not relevant for this research. Afterwards the existence of dynamic panel data, i.e. the influence of the heterogeneity of the people who gave answers, was examined. The influence was found to be important for the analysis of the answers on the stated preference scenarios of this research. Hence it was used on the final models.

The following table depicts the results of the final models.
### Table 1. Results of the model

<table>
<thead>
<tr>
<th>Independent variables</th>
<th></th>
<th>BSS</th>
<th></th>
<th></th>
<th>Car</th>
<th></th>
<th></th>
<th>Public Transport</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\beta_i$</td>
<td>Wald</td>
<td></td>
<td>$\beta_i$</td>
<td>Wald</td>
<td></td>
<td>$\beta_i$</td>
<td>Wald</td>
<td></td>
</tr>
<tr>
<td><strong>DISCRETE VARIABLES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenience</td>
<td></td>
<td>0.953</td>
<td>9.44</td>
<td></td>
<td>0.537</td>
<td>6.25</td>
<td></td>
<td>0.537</td>
<td>6.25</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>0.905</td>
<td>5.61</td>
<td></td>
<td>2.78</td>
<td>2.30</td>
<td></td>
<td>3.25</td>
<td>2.69</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td>2.48</td>
<td>2.06</td>
<td></td>
<td>0.600</td>
<td>3.33</td>
<td></td>
<td>3.25</td>
<td>2.69</td>
<td></td>
</tr>
<tr>
<td><strong>CONTINUOUS VARIABLES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td>-0.083</td>
<td>-15.99</td>
<td></td>
<td>-0.075</td>
<td>-6.97</td>
<td></td>
<td>-0.083</td>
<td>-15.99</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
<td>-0.274</td>
<td>-1.36</td>
<td></td>
<td>-0.025</td>
<td>-0.184</td>
<td></td>
<td>0.870</td>
<td>5.26</td>
<td></td>
</tr>
</tbody>
</table>

It should be noted regarding the summary statistics of the whole statistical model that the $R^2=0.368$ and the Likelihood ratio test is $L_{10}=2.059,13$ confirming the suitability of the model. Furthermore, from the above table the utility function of the bicycle can be examined. It is observed that the function correlates with time, cost, comfort, gender and age.

The time coefficient of the bicycle’s utility function is -0.083, which means that an increase of time by one unit causes the decrease of $U$ by 0.083. In other words, the longer the time required to travel by bicycle, the less likely it will be chosen for transportation. This observation is not surprising, as it is a known fact that minimizing the travel time is one of the more basic criteria in choosing a means of transportation.

The cost coefficient of the bicycle’s utility function is -0.274, meaning that for an increase of cost by one unit causes the decrease of $U$ by 0.274. Specifically, increasing the cost of the bicycle will decrease the chances of it being selected. This is also expected, since by increasing the cost of a product or service leads to lower demand. It is imperative to note that the cost parameter in the bicycle function was found to be statistically not important. Since the Wald index for the cost was -1.41, with absolute value smaller than 1.7. This does not mean that the cost parameter is not important for the majority of sample, it means it is important to a degree (barely) less than 95%.

Subsequently, the comfort coefficient is examined. It is observed to be positive (0.953), which means that the increase of comfort by one unit entails the increase of the probability function by 0.953. That is, the more the comfort for the bicycle the better the chances of it being selected. This may be due to the fact that the definition of high levels of comfort implies the existence of bicycle lanes. From the answers to question B10 (how much do the following factors prevent you from using a bicycle), the lack of bicycle lanes was answered ‘A lot’ by 60.8%, making it first in importance, followed by lack of respect from other drives with 53.2%, followed by less security compared to other means of transportation with 44.7%.

At this point, it is worth noting that, in absolute value, the comfort coefficient is significantly higher than the time and cost coefficients. Meaning that the change in comfort affects the probability of choosing a bicycle, more than that of time and cost. This can be explained by the way comfort is defined, as it contains the notion of security, the existence (or nonexistence) of bicycle lanes and the good (or bad) traffic environment. From this perspective, it is logical that security affects more than cost and time.
The gender on the utility function of the bicycle also has a positive coefficient (2.48). Thus, it becomes apparent that for women the utility function increases by 2.48, which comes into contradiction with the change of probability. The probability of a woman choosing a bicycle is less than that of a man, as shown on the diagrams on the next subsection. This happens because the bicycle’s probability equation is a function of all three utility equations. So even though the bicycle’s utility function increases, the utility functions of the car and public transportations increase even more, as a result the probability of a bicycle being chosen decreases. That men are more likely to choose bicycles is consistent with the percentages of the scenarios in which the bicycle was chosen, which don’t take into consideration the heterogeneity of the individuals and therefore may differ from the results of the statistical analysis. This result is rational, considering that the athletic capabilities and stamina of men are higher compared to that of women, and also the use of a bicycle requires a certain dressing type, which might not necessarily comply with a woman’s attire.

The positive sign of age (0.864) indicates that for a value equal to 1 of the Age variable, the utility function of the bicycle increases by 0.864. The value 1 is assigned to Age if the respondent belongs to the 18-24 age group, otherwise 0 is assigned. Meaning that if the respondent is between 18 and 24 years old the probability of choosing a bicycle increases. This may be due to the fact that the fitness required by the bicycle occurs more frequently in younger ages. It could also be because younger people are more familiar with the bicycle, either for transportation or for fun, or because they are more open minded to alternative means of transportation.

4 CONCLUSIONS

The present study aims to evaluate the characteristics influencing the choice of a BSS for travelling in the Municipality of Athens. Based on the results of the mathematical models, it is concluded that:

The probability of choosing a BSS for travels in the Municipality is highly affected by time, cost and travel comfort of the travel while the traveller’s characteristics that affect the choice is gender and age. These results agree with the findings of international references.

The influence of time over the choice, and more precisely the reduction of the probability of the choice due to increase of the travel time, can be explained given the fact that bicycle is considered as mean of transport and not as mean of entertainment or fitness. Therefore, as for every mean of transport, the minimum possible travel time constitutes a strong incentive for choosing it, while the opposite stands for increasing travel time.

Cost was found as a factor as well, even marginally, for choosing a BSS. The influence of the cost can be easily explained since renting a bicycle is a paid service and the demand of every service usually increases with the reduction of the cost. The correlation between choosing the bicycle and travel comfort (where safety is also included) can be explained when regarding the bicycle as a mean of transport, as well when considering that the existence of bicycle roads is in the context of safety. It is expected that as travel comfort increases and the provided services are optimized, the demand also increases.

At the same time, the statistic results of the B part of the questionnaire indicate that the majority of the users consider dissuasive the absence of bicycle lanes, which shows that the increased risks and reduced safety measures regarding the use of bicycle are valued high in decision making. As a result, travel comfort proved to be a really important factor for using this system.
Furthermore, the increased probability of men choosing a bicycle is possibly related to their greater physical strength and stamina compare to women, as well to the restriction in women’s clothing.

People between 18 and 24 years old seem to prefer bicycle, since the probability to choose a bicycle increases when answering inside this age group. This phenomenon can be due to good physical condition, which is most common to young ages. Another reasoning can be that young people are more familiar with riding a bicycle, for exercise or for fun, or even that they are more open to alternative ways of transport.

The nested logit model proved to be not applicable for this research. That is, the nested logit model between motorized and non-motorized means of transport was not a criterion for the answers received. On the contrary, the study of panel data existence had positive results and helped to the optimization and accuracy of the final models, as well as to their interpretation. The panel data take into consideration the error due to heterogeneity of people and it can be estimated when having multiple answers from a respondent.

Finally, it should be noted that the highest than anticipated interest for bicycle, due to the online nature of the survey, has direct influence to the results both of the simple and complex statistic process.

Further investigation should be made on other factors such as weather condition, the reason and frequency of traveling, as well as the possession of a car. At the same time, it would be really interesting to examine the change of travel time and mainly cost in the stated preference scenarios, as it would be very likely to give different outcomes.

5 REFERENCES


