SOCIAL INFLUENCES ON TRIP MODE-CHOICES IN NON-DAILY TRIPS

by

Dünya Yavaş

B.S, Civil Engineering, Boğaziçi University, 2010

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Civil Engineering

Boğaziçi University

2013

SOCIAL INFLUENCES ON TRIP MODE-CHOICES IN NON-DAILY TRIPS

APPROVED BY:

Assist. Prof. Ilgın Gökaşar ……
(Thesis Advisor)

Prof. Gökmen Ergün ……

Prof. Cem Ersoy ……

DATE OF APPROVAL: 09.04.2013

ACKNOWLEDGEMENTS

I want to gratefully thank Assist. Prof. Ilgın Gökaşar for giving me the opportunity, time and space to write my thesis. Without her support, patience and tolerance, I would not be able finish this study.

I wish to thank Prof. Gökmen Ergün for his unconditional and perfect guidance at each step of this thesis. His unlimited and impeccable assistance that made my research come into this stage should never be forgotten.

I should also thank Yiğit Çetinel for his great contribution in writing HTML codes for this research. Last but not least, I want to thank my family and my girlfriend, Iklil, who believe in me in the way to achieve this study, for their sincere and continuous support.

ABSTRACT

SOCIAL INFLUENCES ON TRIP MODE-CHOICES IN NON-DAILY TRIPS

Classical choice models in transportation assume that individuals are 100% rational and indifferent to choices of other individuals. Behavioral economics try to relax these assumptions and try to uncover the unobserved factors which affect the decision-making process of individuals. In this study, social influences on trip mode choices in non-daily trips are focused in order to reveal if there is any significant peer influence on decision-makers while choosing a mean of transport. An information cascade experiment is designed to test the hypothesis. 243 people between 18 years old and 30 years old were reached through an Internet survey and they were asked to choose a mode among three alternatives for each three different non-daily trip scenarios. A conditional logit model is calibrated using the data and it is shown that individuals are affected by their peers' choices in the non-daily trip mode choice context. It is also found that females are more affected from their peers than males. Interestingly, a significant difference in the model is found between the subjects who have smartphones and the ones who have not. Subjects without a smartphone only consider their peers' choices and the factors cost and travel time reliability to some extent while choosing a mode of transportation. On the other hand, smartphone users consider every aspect of the trip. Value of time in a vehicle and value of time spend in walking and waiting is calculated as 19.917 TL/hr and 20.314 TL/hr respectively and it is observed that they are in the reasonable limits when they are compared with the previous value of time studies in Turkey.

ÖZET

GÜNLÜK OLMAYAN TÜREL SEÇİM MODELLERİNDE SOSYAL ETKİLER

Klasik türel seçim modelleri bireyleri 100% mantıklı, başkalarının seçimlerinden etkilenmeyen ve faydasını enbüyültmeye çalışan varlıklar olur görür. Davranışsal ekonomi bu varsayımları gevşeterek, seçimlerimizi etkileyen diğer ölçülmesi nispeten zor faktörleri açığa çıkarmayı hedefler. Bu tezde, bireylerin günlük olmayan bir yolculuğa çıkarken seçecekleri ulaşım şeklinin, başkalarının seçimleriyle değişip değişmediği incelenmiştir. Hipotezi test etmek için bir bilgi kaskatı deneyi dizayn edilmiştir. 18-30 yaş aralığındaki 243 kişiye deney internet aracılığıyla ulaştırılıp, deneklerden üç farklı günlük olmayan yolculuk senaryosu için, üç farklı ulaşım şeklinden bir tanesini seçmeleri istenmiştir. Elde edilen data logit modelle kalibre edilmiştir. Buna gore, bireylerin seçim yaparken başkalarının seçimlerinden etkilendiği ortaya çıkmıştır. Ayrıca kadınların başkalarının seçimlerinden daha fazla etkilendiği gözlemlenmiştir. Şaşırtıcı olarak, akıllı telefon kullanmayan bireylerin büyük çoğunlukla başkalarının seçimlerinden etkilendikleri, kısıtlı olmak kaydıyla da yolculuğun ücreti ve yolculuk süresinin güvenirliği faktörlerinden etkilendikleri tespit edilmiştir. Buna karşın akıllı telefon kullananların yolculuğu etkileyen her faktöre dikkat ettikleri gözlenmiştir. Araç içerisinde geçen sürenin değeri 19.997 TL/saat, yürüme ve bekleme sırasında geçen sürenin değeri 20.314 TL/saat olarak hesaplanmıştır. Bu değerler daha önce İstanbul, Mersin ve Kocaeli'nde yapılan çalışmalarla örtüşmektedirler.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT iv
ÖZET v
LIST OF FIGURES viii
LIST OF TABLES ix
LIST OF SYMBOLS x
LIST OF ACRONYMS/ABBREVIATIONS xi
1. INTRODUCTION 1
1.1. Motivation and Problem Statement 1
1.2. Goals and Objectives of the Research 2
1.3. Thesis Contribution 3
1.4. Scope and Limitations 4
1.5. Thesis Outline 4
2. LITERATURE REVIEW 6
2.1. Studies on Social Interaction 6
2.2. Discrete Choice Models 15
2.3. The Conditional Logit Model 20
2.3.1. Some Properties of Conditional Logit Model 20
2.4. Models Using Social Interaction 25
3. RESEARCH METHODOLOGY 27
3.1. Forming Research Hypothesis 27
3.2. Design of the Experiment 28
3.3. Description of the Experiment 30
3.4. Description of the Questionnaire 33
3.5. Data Collection 34
4. ANALYSIS OF DATA AND MODEL CALIBRATION 35
4.1. Preliminary Analysis of the Collected Data 35

4.2. Calibration of the Mode-Choice Model with the Pooled Data 36
4.3. Models with Socio-economic Segments 38
4.3.1. Role of Gender 39
4.3.2. Role of Smartphone Possession 40
4.3.3. Role of Age 42
4.3.4. Role of Driving License 42
4.3.5. Role of AKBIL Possession 43
4.3.6. Role of Auto Ownership 44
4.3.7. Role of Income 45
4.4. Inferred Values of Time from the Models 46
5. CONCLUSIONS AND RECOMMENDATIONS 48
5.1. Conclusions 48
5.2. Recommendations for Further Studies 50
REFERENCES 51

LIST OF FIGURES

Figure 2.1. Equilibrium States of Sugamo.....10

Figure 2.2. The Classical Four-Step Transportation Model. 16

Figure 3.1. Outline of the Research Methodology.....27

Figure 3.2. Information Cascade Experiment Setup.....29

Figure 3.3. Question 1.....31

Figure 3.4. Question 2.....32

Figure 3.5. Question 3.....32

Figure 3.6. Questionnaire Page 1.....33

Figure 3.7. Questionnaire Page 2.....34

LIST OF TABLES

Table 3.1. Correlation Matrix. 30

Table 4.1. Brief Description of Sample Characteristics. 36

Table 4.2. STATA Estimates of the Coefficient of Variables in the Model. 37

Table 4.3. Models for Gender Segments. 39

Table 4.4. Models for Smartphone Possession Segments. 41

Table 4.5. Models for Age Segments. 42

Table 4.6. Models for Driving License Possession Segments. 43

Table 4.7. Models for AKBIL Possession Segments. 44

Table 4.8. Models for Auto Ownership Segments. 44

Table 4.9. Models for Income Segments. 45

Table 4.10. Value of Time. 46

LIST OF SYMBOLS

U	Utility
V	Strict Utility
ε	Error
σ	Scale factor
A	Set of alternatives
C	Choice set
E	Elasticity

LIST OF ACRONYMS/ABBREVIATIONS

BLP	Barry-Levinsohn-Pakes
TAZ	Traffic Analysis Zone
MNL	Multinomial Logit
IIA	Independence of Irrelevant Alternatives
RUM	Random Utility Model
LRT	Likelihood Ratio Test
TUIK	Institution of Statistics, Turkey
HTML	HyperText Markup Language
HBW	Home Based Work
HBS	Home Based School

1. INTRODUCTION

1.1. Motivation and Problem Statement

Understanding travel behavior is an important aspect of the transportation planning: how to predict it and how to influence it over the long term. Choice models in transportation are predominantly based on the neoclassical economic theory of choice which assumes that individuals act 100% rational, are not affected by their peers, are indifferent to welfare of others and try to maximize their utilities [1]. However, for the last forty years, researchers question the assumptions of the theory and try to figure out what really influence individual's decisions in the daily life.

Behavioral economics which is fed by psychology and economics try to answer these questions with simple and cleverly designed experiments. These experiments are used to figure out how much rational we are, to what extend we are or are not affected by our social networks. Choice models in transportation are also affected from this relatively new phenomenon and various researches, which will be covered extensively in the literature review section, have been carried out in this field. However, trip mode choice models do not benefit from behavioral economics adequately yet.

One of the main reasons of such inadequacy is a suspicion that decision-making processes while choosing a mode of transportation are highly dependent on habits [2]. This behavior is also attributed to behavioral phenomena such as "brand loyalty", "modal switching", "brand switching", "resistance to change", and "learning and habit formation" [3]. Another reason is a belief that decision-makers are aware of every alternative and choose the very best one according to their tastes especially for everyday trips such as Home-Based-Work and Home-Based-School trips.

Recent research have shown that decision-makers are more prone to use travel information for non-daily trips [4]. Therefore it is plausible to ask whether if peers' of decision-makers can provide this information or not. Since the nature of non-daily trips may be ambiguous in terms of the final destination and mode alternatives, peers' choices can also surpass the travel information given by other sources.

The main issue with social influences is endogeneity which is the correlation between a variable and the error term. It is also known as Manski's reflection problem in social influences theme and it questions if group behavior affects individual behavior or group behavior is simply the aggression of individual behaviors [5].

In 1992, Banerjee, without dealing with endogeneity, created a one way relation between the group and the individual and showed that decision-makers may rationally ignore their private signals after observing the decisions of previous decision-makers when they are all imperfectly informed [6].

Lately, McFadden discussed why individuals are affiliated with social networks, how sociality affects the choice behavior of decision-makers and how it should enter the choice models [1]. He also suggested Berry-Levinsohn-Pakes (BLP) method to deal with endogeneity.

1.2. Goals and Objectives of the Research

In this thesis, it is hypothesized that decision-makers' are influenced from their peers' choices when they are faced with a trip mode choice for a non-daily trip. So the main goal of this research is to study the hypothesis that individuals' decisions can be altered by revealing the choices of their peers. In addition to this main goal, the study has also aimed to study the following objectives.

(i) To provide a literature review on discrete choice modeling and social interaction in discrete choice modeling,

(ii) To study the relationship between the social interactions and socioeconomic characteristics of the trip makers in order to learn which socioeconomic classes are more socially constrained.

(iii) To investigate the effect of reliability on trip mode choice and the relationship between social interactions and reliability,

(iv) To calculate the value of time in-vehicle and the value of time spend in walking and waiting of the decision-makers.

1.3. Thesis Contribution

The main contribution of this thesis is gaining more insights over the decision-making process of individuals when they are faced with a trip mode choice problem. All contributions of the thesis are listed below:

(i) It has been shown that peers' choices can distort choices of decision-makers in the mode choice setting.

(ii) It has been shown that males and females are not affected equally from their choices of their peers.

(iii) It has been shown that there are significant differences between the model with subjects who have a smartphone and the one with subjects who have not while dealing with the mode choice problem.

(iv) It has been shown that subject without a driving license do not consider their peers' choices in the decision-making process of mode choosing.

(v) Value of time in-vehicle and value of time spend in walking and waiting among the young adults are calculated.

1.4. Scope and Limitations

This research is aimed at showing trip-makers who are between 18 years old and 30 years old, are influenced by choices of their peers when they are asked to choose a mode of transportation for a non-daily trip. Therefore, first of all, the research is constrained with age limits. Secondly, it only considers non-daily trips because it is assumed that decision-makers are more prone to be affected from outside sources such as peers for these kinds of trips.

This research focuses on relationship between given alternative specific variables and decision-makers. Although the setting of the experiments is a mode choice problem, the stated choices of the names of the modes, such as car and transit bus were not mentioned in the survey questionnaire hence the modes are only specified through their characteristics such as travel time and travel cost so as to eliminate the existing biases for such modes.

1.5. Thesis Outline

The remainder of this thesis is organized as follows:

- Chapter 2 discusses the related previous research in the field. In the second part of the chapter, it presents a detailed review of discrete choice models. Then, concentrates on

conditional logit model and gives the main properties of this model This chapter is concluded with a model using social interactions.

- In Chapter 3, the methodology of the research is presented. It starts with forming the research hypothesis and continues with the design of the experiments and data collection procedure.

- Chapter 4 presents first the sample characteristics and the calibration of the model. After that, it gives a detailed analysis of the model for different socio-economic segments.

- Chapter 5 concludes the thesis with a summary and recommendations for further studies.

2. LITERATURE REVIEW

2.1. Studies on Social Interaction

Many economists believe that sociological concepts such as social norms and peer influences are artificial phenomena which can be explained at the level of individuals whereas many sociologists claim that these phenomena are the dominating factors in individual's choices although they have not find any consensus on the nature of social effects, wherein the tendency of an individual's choosing an alternative varies with the choices of other members in his/her reference group. Depending on the context, these effects may be called "social norms", "peer influences", "neighborhood effects", "conformity", "imitation", "contagion", "epidemics", "bandwagons", "herd behavior", "social interactions" or "interdependent preferences" [21].

Manski tried to identify and separate the social effects from other effects and specified a model for the choice behavior. He hypothesized that there exist three effects to explain why individuals belonging to same group behave similarly:

> "a) endogenous effects, wherein the propensity of an individual to behave in some way varies with the behavior of the group;
> b) exogenous (contextual) effects, wherein the propensity of an individual to behave in some way varies with the exogenous characteristics of the group, and
> c) correlated effects, wherein individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments." (pp. 532-533)

He slightly changed the correlated effects later on and added a fourth effect which is ecological effects, wherein individuals in the same group tend to behave similarly because they have similar institutional environments, therefore correlated effects became individual correlated effects and only explains the problem by addressing common individual

characteristics [5]. He defined a linear model using these effects and inferred that observed behavior is always consistent with the hypothesis that individual behavior reflects the mean reference group behavior (tautology). He concluded that there might be realistic opportunities for social effects if the characteristics of reference groups and those directly affecting outcomes are moderately related. However, if these characteristics are either functionally dependent or are statistically independent, the chances are very low. He suggested to develop a tighter theory or to collect richer data in order to explain the nature of the social effects.

The influence of human sociality on choice behavior was also discussed by Daniel Mcfadden in 2010 [1]. He introduced four non-exclusive motivations for affiliation with social networks:

> " a) Mutual support, protection and status
> b) Efficiencies in collection and sharing information
> c) Opportunity-Based Homophily: Joint production and division of effort, and risk-sharing
> d) Preference-Based Homophily: Decision-making economies and approval" (p. 4)

Moreover, he defined reciprocity, exchanges of benefits which are not necessarily synchronous or bilateral, and altruism, conveying of benefits to others without any expectation, as two important features of the social behavior of humans.

He put four mechanisms forward in order to describe how sociality influences the choice process. The first mechanism is defining constraints which are available to the decision-maker. The effect of driving decisions of others on the traffic congestion a commuter faces is an example of this type of mechanism.

The second one is perceptions. Via social networks, individuals collect and spread information on attributes of choices. For example, perceiving a brand of mobile phone is desirable because many network members have acquired one. In addition to that, when individuals face with ambiguities, social networks play an important role in removing ambiguities thus strengthening herd behavior and status-quo effects.

Preferences is the third mechanism for identifying how sociality penetrates individuals' choices. Approval by others, and accountability to network social norms may shape the preferences of decision-makers. And the last mechanism is the process. Individuals behave rationally when stakes are high, but get influenced by their social networks when stakes are low and possibilities for regret is small especially where alternatives are unfamiliar and ambiguous. Moreover, he stated that neighborhood network effects are likely to be related with supply-side constraints and opportunity-based homophily, where social stratum network effects are related with preference-based homophily and economic constraints. Availability and convenience of public transit and bicycle lanes was given as an example for the first situation and for the latter one, he gave the example of similarity of high values of time for high income people. At the last part of his paper, he suggested a BLP method to deal with the endogeneity of field effects and group effects in a discrete choice model.

Herd behavior had been studied by Abhijit V. Banerjee in 1992 [6]. He proposed a simple model for the herd behavior wherein decisions were made sequentially. Noting the fact that all decision makers are imperfectly informed, an information cascade occurs when future decision-makers rationally ignore their private signals after observing the previous decision-makers' decisions [6,22].

Anderson and Holt conducted an experiment in which private signals are draws from an unobserved urn [23]. Individuals were asked to make predictions in a sequence and paid if they correctly predict the urn which is used for draws. They observed that early decision-makers naturally tend to rely on their private signals whereas the rest tend to behave according with the early decision-makers although their private signals inconsistent with them. The same experiment is replicated by Hung and Plott in 2001 [24].

In the same year, Brock and Durlauf [16] added social interactions in binary discrete choice models by letting an individual's choice for a particular alternative to be influenced by the overall share of decision-makers who choose that alternative. If the coefficient on this interaction variable is close to zero and statistically insignificant relative to other variables in the utility, then the decision-makers' choices will not be effectively changed over time with

respect to the choices of other decision-makers. Moreover, in this presumptive situation, choice would not be socially conditioned. However, if the coefficient on this interaction variable is dominant enough relative to other contributions to utility, there may exist an uncontrolled situation over time as all decision-makers flock to one choice alternative. They showed that multiple equilibrium points will exist when social interactions exceed a particular threshold. They concluded their study with proposing a model with social interactions. In 2003, they extended their study on the behavior of binary logit models to multinomial logit models [17].

Fukuda and Morichi studied the illegal parking problem in Japan with Brock and Durlauf's interaction based model [25]. They modeled the effects of conformity in the choice of bicycle parking (on-street and off-street) and proposed an increment in police patrols in order to prevent the illegal bicycle parking behavior. In order to identify the model, they collected data from three towns in Japan using a mail-back survey. Individual's risk attitudes and public moralities were also asked to suggest a policy later on. They found that those who frequently visit stations are likely to park their bicycles off-street. In a similar manner they also found that the duration of parking is a good indicator of off-street parking behavior where off-street parking increases as parking duration increases as well. At the later stages of their research, they observed that the town Sugamo suffers from inferior equilibrium (undesirable equilibrium state of a social group) wherein most people tend to park on-street. Therefore they proposed a policy intervention to take out the town from this deadlocked state by increasing the frequency of police patrols. This corresponds to an increase in private incentive for each individual to park off-street. The solid line in Figure 2.1 represents the present situation of town Sugamo before the intervention and the dashed line represents the situation after intervention. The intersection with the lower average choice level (25.8 %) corresponds to critical mass whereas the one with higher average choice level corresponds to superior equilibrium (desirable equilibrium state of a social group). As one can see before the intervention, the system has only one equilibrium point which is inferior. They conclude their paper by suggesting that minimum six patrols in a month is required for town Sugamo in order to prevent the on-street parking.

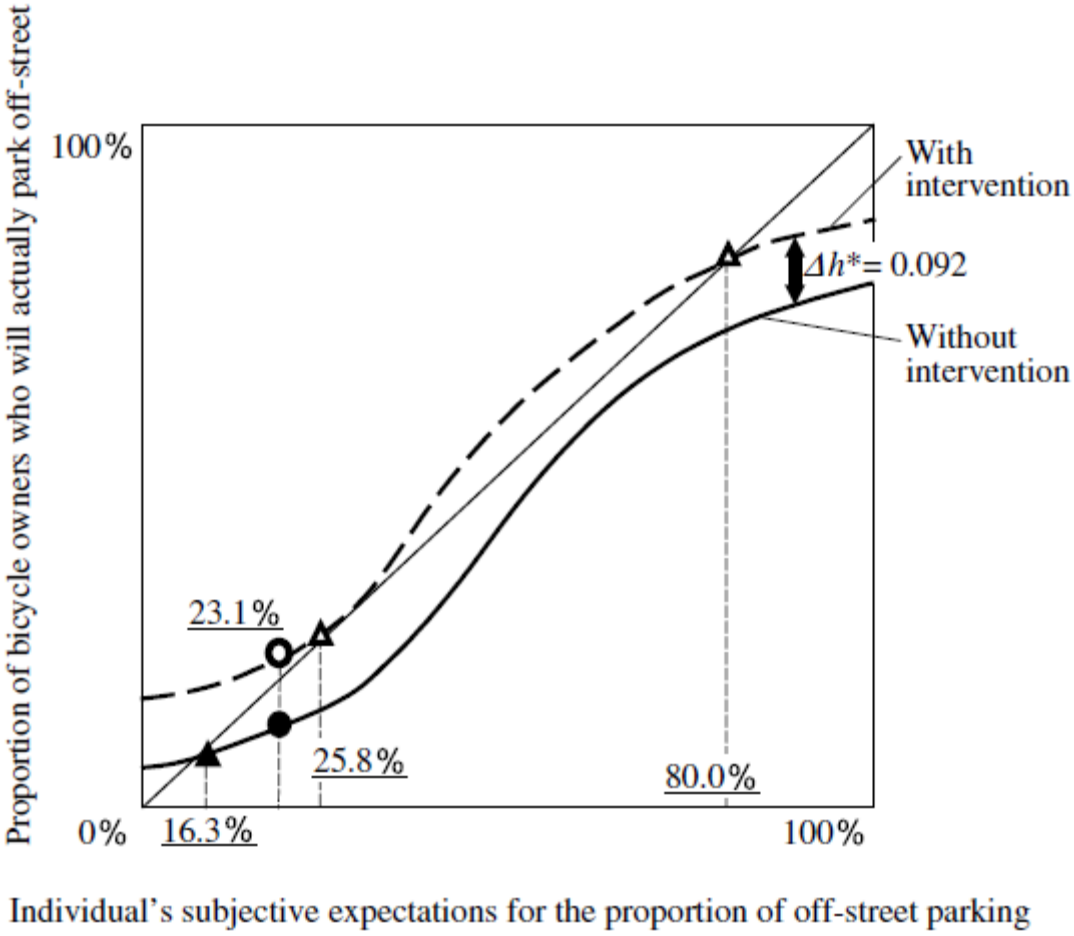


Figure 2.1. Equilibrium States of Sugamo [25].

In 2005, Antonio Paez and Darren M. Scott developed a discrete choice model based on the economic theory of externalities with a social influence variable in addition to attributes of alternatives and characteristics of decision-makers [26]. They applied this model to the case of telecommuting which is a binary choice (telecommute or not) using Monte Carlo simulations. They established a network with 100 to 500 individuals and generated two random orthogonal variables that could represent personal characteristics to each of them. After that they assigned 1 to 10 contacts to individuals and at the end two alternatives were defined, telecommuting and commuting. They concluded that their model can reproduce many behavioral outcomes including reversible decisions. They also mentioned the importance of proximity of individuals in the decision-making process. Moreover their results

showed that networks less than 200 individuals are non-reliable and 500 individuals were needed to have desirable properties of estimators.

In 2006, they extended their research to a multinomial discrete choice approach to analyze individual behavior in social situations where the social position in a network may encourage or discourage different courses of action [27]. By means of a simulation example, they explored the characteristics of the model. According to them, the individualistic or under-socialized approach implied by discrete choice theory is now being brought to task for ignoring the social dimensions of decision-making and thus, for failing the account for externalities and information spillovers that appear in a number of important settings when decisions have social consequences. Therefore they tried to produce a framework within which individuals are not isolated atoms that act in a social void, but structural elements of the social network to which they belong. They designed another Monte Carlo simulation to test their model. The motivation for application is provided by the problem of modeling of residential location choices. They concluded that sample sizes under 200 are unreliable and a sample size about 500 could be necessary to achieve desirable properties of the estimators. In addition to that, they observed significant differences between the model with social influence and the one without social influences.

Walker and Dugundji conducted a research in 2005 to capture the interdependencies in a cross-nested logit model through an empirical application to mode choice problem [18]. In this research, decision-makers were assumed to be influenced by those of similar socioeconomic status and those in spatial proximity. First approach to capture interdependencies was to include variables in the systematic utility that describes choices of others in the decision-maker's social and spatial network. And the second is by allowing for correlation across the disturbances of decision-makers within the same social and spatial network. They tested these two approaches and a combination of two with the mode choice problem. According to them, if there is theoretical or qualitative reason to believe that a feedback effect exists, it can have very important implications for the prediction of (system-wide) results over the course of time. For example, in the introduction of a new transportation mode alternative, if there is a "dominant enough" feedback effect, this can propel the adoption

of the new mode over time. They hypothesized a distinction between social versus spatial interactions and between identifiable versus aggregate interactions, to make an effort to classify different types of demand-side interaction mechanisms. Interaction between identifiable decision-makers means that the links in the network are well-known and explicitly defined on an individual decision-maker by decision-maker basis. Interaction between aggregate decision-makers means that interdependence is assumed to take place only at an aggregate level with links being defined. Spatial network interactions mean that the interdependence represents a confluence of decision-makers in geographic terms. The decision-makers may be connected to each other by a spatial proximity such as shopping, school, work location etc. Social network interaction means that decision-makers are linked based on social circles. They need not to be proximally or tangentially situated in geographic terms and interaction may take place at a distance. Finally, they came up with a framework for conceptualizing the interdependence of decision-makers' choices and showed that it is significantly better than the conventional methods.

In 2008 Dugundji and Gulyas dealt with the interactions between households and generated feedback dynamics in the adoption of various transportation mode alternatives [19]. They developed a nested logit model where a decision-maker's choice is influenced by decision-maker's socio-economic peers and neighbors. They first defined a network based on individual's residential district and socio-economic group. Then, using the travel data from Amsterdam, they tried their model. One of the important findings of the study is that the effect of unobserved heterogeneity through the introduction of scale parameter plays an important role in the empirical application of discrete choice models with social influence variables.

Using Aoki's discrete choice model with social interactions, Dungundji and Gulyas presented a multi-agent-based model with two broad classes of abstract networks: Erdös-Rényi and Watts-Strogatz graphs [28]. Aoki's model assumes uniform, global and perfect information access; therefore they extended it to explicitly model interaction networks. At each time step decision-maker looked at the choices of his/her reference entities made in the previous round and his/her own choice and then calculated the localized values of the

difference in the systematic utility between alternatives. They had carried out their simulations with various densities (ratio of the number of existing links versus the number of all possible links) and with 100 agents. It is observed that both networks behaved in a similar way in the long run with Aoki's model with global interactions.

Walker *et al.* [29] conducted three experiments in 2010 in order to create a bridge between behavioral economics and transportation. The first experiment focused on information. The subjects were asked to choose a route among three alternatives where the attributes of routes were given. In addition to classical attributes such as travel time and cost, green-house gas emissions of each route were also given. In the second experiment, a job scenario with a normal salary was presented to subjects and they were asked to choose whether to buy a conventional car or hybrid car or not to buy a car at all. Moreover, while making decisions the subjects were also informed about number of people that has chosen the alternatives, so that, they were able to collect information about how other's choices affected our choices. The third experiment was a combination of the first two experiments. They had devised different types of information about pedestrian jaywalking behavior and presented only one of them to subjects. The type of information varied from peer behaviors to accident statistics and law. Then, they asked to subjects whether the subject felt that in the coming week he/she would cross against red lights more frequently, less frequently, or the same as the previous week.

It was observed from the first experiment that by providing context and person specific information about the environmental impacts of actions, people can be nudged towards more sustainable behavior. The second experiment showed that subjects were definitely influenced by the decisions of their peers. And the results of third experiment suggested that social norms had significant impact on behavior.

In 2011, following the Mcfadden's aforementioned suggestion, Walker et al. [20] dealt with endogeneity in behavioral choice models with social influences. In order to capture social influences, they used a field effect variable which is the percent of population that has chosen the specific alternative. Using such a variable causes biased and inconsistent

parameter estimates in the model because it is most likely correlated with the error variable or unobserved factors. In order to deal with that, they used Berry, Levinsohn and Pakes (BLP) method since it can be implemented when endogeneity occurs at the market segment level. Therefore they interpreted aggregate peer groups as market segments which is very suitable for social interactions since individual networks are not explicit. In their study, they hypothesized that people are most influenced by both those in spatial proximity and those who are in similar socio-economic groups thus they defined their peer groups as spatial and social reference groups. They assigned spatial reference groups and social reference groups to individuals according to their residential postal codes and income groups respectively. Because of the nature of BLP method they had to use instrumental variable which is correlated with endogenous variable and uncorrelated with the error term. For spatial reference groups, they used mode shares of spatially adjacent zones since mode share of a zone is likely correlated with adjacent zones and assumed that it is not correlated with the error term. Although the first assumption is straightforward, it is difficult to defend the second one. However, they relied on postal code definitions of the Netherlands, which are generally defined as homogeneous within a zone and heterogeneous across zones. For social reference groups, they used the value of the endogenous variable in the socially adjacent zones. They assumed that individuals' choices are most likely to be influenced by those in higher income group after the ones in their own income group due to status seeking behavior. Moreover, they assumed that social and spatial reference groups are independent and even if there is some overlap, it is negligible relative to overall effect. At the end of their study, they applied this method to the mode choice to work problem in Netherlands. They compared the models with uncorrected field variable and corrected field variable. They pointed out that the uncorrected parameter on the field effect variable was 60% higher than the corrected parameter.

There are also different approaches in modeling social interactions in discrete choice models. In 2009 N. A. Ronald *et al.* [30] developed an agent based framework in order to model social influence on travel behavior. They focus on the dynamics of the social network and how it affects individual choice sets of locations, activities, and participants. They were interested in the activities which are generated due to physiological, psychological and economical needs. Activities which emerge from the social and spatial networks can be

categorized as subsistence (work-related), maintenance (keeping the household running), and leisure. Based on these dynamics and background, they developed a model to understand the nature of travel behavior in terms of social influence. They conclude their study by outlining a framework and design for modeling the effects of social networks on travel behavior.

Goetzke [31] tested the network effects in transit using spatial autoregressive logit mode choice model with New York City 1997/98 work trip data. He observed that as the transit mode shares approaching towards one, the indirect utility between transit use and drive-alone trips decreases until transit use utility eventually surpasses the indirect utility of driving alone, since the network effects terms is larger than the coefficient for the constant. He interpreted this as the willingness to pay for transit, which is related to the indirect utility, increases when more people take transit, and thus lowers the relative cost of using transit.

Goetzke and Rave [32] have studied the factors influencing bicycle use in Germany. In addition to individual, trip and municipal characteristics of travels, they considered the a bicycle culture which defined as a social interaction where probability of using bicycle is increasing with the aggregate bicycle share. They regard bicycle culture as a social interaction since bikers send positive signals (biking as a way of traveling) to other people, thus affecting them. In the model, first aggregate bicycle mode share was computed using classical tools. Then they estimated the full binary logit choice model using previously calculate bicycle mode share. They observed that bicycle culture was significant for recreational trips and to some extent, for shopping trips. It was also found that men were more likely to commute with bicycle and bicycle commuting was independent of age and household income.

2.2. Discrete Choice Models

The classical four-step transportation model was introduced in 1950s in order to determine the demand on urban transportation networks [7]. As the name suggests, this process includes four distinctive steps, as shown in Figure 2.2, which are:

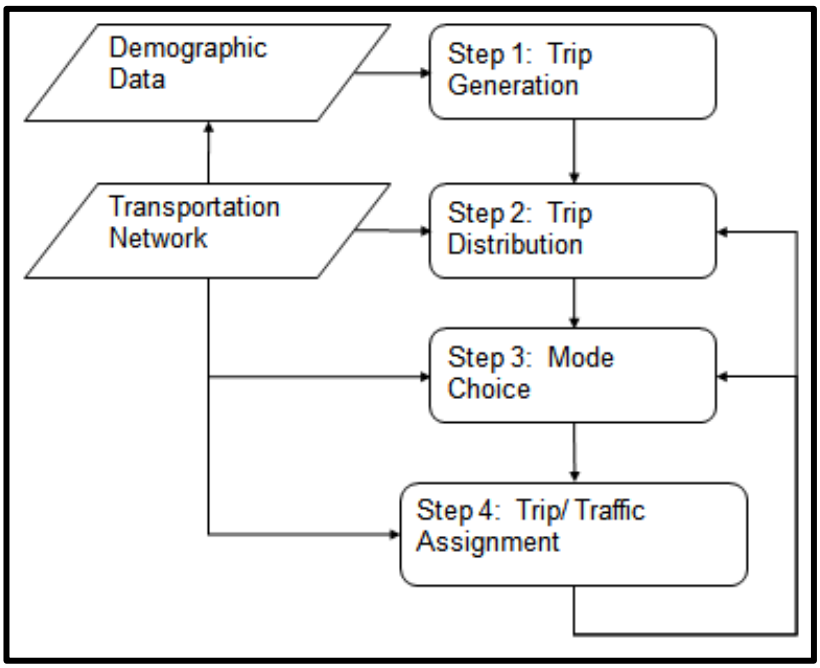


Figure 2.2. The Classical Four-Step Transportation Model.

(i) Trip Generation: This step determines the number of trips that will be produced from and attracted to a traffic analysis zone (TAZ), which are geographical areas primarily defined for transportation planning purposes. Trips going out from one TAZ are referred as *productions*, while trips coming into the TAZ from other zones are referred to as *attractions*. Thus, each trip constitutes of two trip ends, namely *production end* and an *attraction end*. The socio-economic composition and characteristics of the households are the main determinants of the number of trips that will be generated from a TAZ. Apart from the trip numbers, this step also defines each trip by a *trip purpose*, - work trips, shopping trips, etc. The trip purposes can be further subdivided as either home based, non-home based, or other trips types. In addition to the discussed trip purposes, this step also considers locations in a region which generate trips that cannot be captured by the above defined trip purposes.

(ii) Trip Distribution: In this step, the origin and destination of the trips are determined. For example, the shopping trips produced in a residential neighborhood are distributed to other TAZs providing shopping opportunities, and work trips are distributed to TAZs providing employment opportunities. The trip tables are a matrix of origin to destination (*OD matrix)*, showing the number of trips between the TAZ's.

(iii) Mode Choice: This step of the travel demand model takes the origin-destination trip tables from the trip distribution process and determines the mode of transportation that will be used by travelers for each trip. In other words, how many trips will use public transportation, how many trips will be made by either walking or biking, and how many people will drive alone or share a ride. Mode choice determination is based on various factors such as travel time, automobile ownership, accessibility to public transportation, transit fare, and auto operation cost. These factors are the variables in the mode choice model used to determine the most likely mode choice for trips. The outcome of the mode choice model is the vehicular trip tables.

(iv) Trip Assignment: In this final step, the vehicular trips are assigned to the road network based on a number of factors, including travel times, congestion, trip length, and travel mode. Upon completion of this step, the result is a roadway network with traffic volumes assigned to each individual roadway segment. The traffic volumes are then used to identify the required number of travel lanes and to help determine the location of congested roadways.

The gravity model showed good performance in describing flows in the highway network, however it failed in forecasting demand along some of the dimensions needed to evaluate major projects being planned through the 1960s in the U.S.A. The impact on behavior of small changes in the transportation system such as bus route density and or bus headways could not be handled with the models that were used that time.

In the years just before 1970s, a number of research studies are funded to develop new travel demand analysis methods that would be more manageable and sensitive to factors

controlled by a transportation policy. As a result, Domencich and McFadden [8] came up with a new disaggregate urban travel demand model that included work and shopping trips, mode choice, destination choice and trip generation using multinomial logit (MNL). MNL has its roots back to 1927, the study of Thurstone [9] who proposed that an alternative *i* with true stimulus level V_i being perceived as $V_i + \boldsymbol{\varepsilon}_i$, where the $\boldsymbol{\varepsilon}_i$ are independent normally distributed perception errors in the paper of *"Law of Comparative Judgment"*. Moreover, he observed that *i* would be preferred over *j* would satisfy

$$P_{(i,j)}(i) = \Pr(V_i + \boldsymbol{\varepsilon}_i \geq V_j + \boldsymbol{\varepsilon}_i), \tag{2.1}$$

where $P_{(i,j)}(i)$ is the probability of alternative *i* chosen between *i* and *j*,
V_i is strict utility of alternative *i*,
$\boldsymbol{\varepsilon}_i$ is error term related with alternative *i*.

In the paper of *"Individual Choice Behavior"* in 1959, Luce [10] postulated that the ratio of choice probabilities for *i* and *j* is the same for every choice set C that includes *i* and *j*; i.e.,

$$P_C(i) / P_C(j) = P_{\{i,j\}}(i) / P_{\{i,j\}}(j), \tag{2.2}$$

where $P_C(i)$ is the probability of alternative *i* chosen among every choice set C,
$P_{\{i,j\}}(i)$ is the probability of alternative *i* chosen between *i* and *j*.

He called this *Independence from Irrelevant Alternatives* (IIA). He showed that if this axiom holds, then one can associate with each alternative a positive "strict utility" w_i such that:

$$P_C(i)=w_i/\sum_{k\in C} w_k, \tag{2.3}$$

where $P_C(i)$ is the probability of alternative *i* chosen among choice set C,
w_i is the utility of alternative *i*.

In 1960, Marschak and Block, and Marschak [11, 12] generalized the theory of Thurstone to stochastic utility maximization over multiple alternatives. He called this RUM (Random Utility Maximization) and proved that in a finite set of alternatives that choice probabilities satisfying Luce's IIA axiom were satisfying RUM as well. As a result of this study, it is found that in order to satisfy IIA, necessary and sufficient condition for RUM with independent errors $\boldsymbol{\varepsilon}_{i,}$ should be identically distributed with a Type I Extreme Value Distribution which is

$$\Pr(c \geq \varepsilon_i) = \exp(e^{-c/\sigma}), \tag{2.4}$$

where σ is a scale factor.

The random utility theory is based on the following assumptions [8, 13]:

- An individual belongs to a certain population Q act rationally and has the perfect information such that he/she is able to select the option which maximizes his/her personal utility subject to legal, social, physical, time and money constraints. That kind of an individual is also known as *"homo economicus"*.

- An individual makes his/her choice among a certain set of available alternatives A = $\{A_1, A_2, \ldots, A_j, \ldots, A_N\}$.

- There is a set X of vectors of measured attributes of the individuals and their alternatives.

- Any alternative *j* in the choice set has an associated net utility U_{jq} for individual *q*. V_{jq} is a function of measurable attributes *x* and is the systematic, representative and measurable part of utility. $\boldsymbol{\varepsilon}_{jq}$ is the random part of utility and represents the idiosyncrasies and particular tastes of each individual plus any measurement or observational errors made by modeler.

It is important to emphasize two points: firstly, the individual rationally weighs all the elements of interest without indiscrimination and selects the most convenient alternative. And secondly, residuals $\boldsymbol{\varepsilon}$ are needed to explain so called irrationalities which is why two identical individuals facing with same choice set may select two different alternatives.

2.3. The Conditional Logit Model

Conditional Logit Model is the most simple and popular discrete choice model. It assumes that random residuals $\boldsymbol{\varepsilon}$ are distributed IID Gumbel, such that [7]:

$$P_{iq} = \frac{\exp(\beta V_{iq})}{\sum_{A_j \in A(q)} \exp(\beta V_{jq})}, \tag{2.5}$$

where P_{iq} is the probability of alternative *i* chosen by individual *q*,

V_{iq} is the strict utility of alternative *i* for individual *q*,

β is a scale parameter which is assumed to be constant for all individuals but may vary across alternatives.

2.3.1. Some Properties of Conditional Logit Model

2.3.1.1 Independence of Irrelevant Alternatives (IIA). The most important property of Conditional Logit Model is the *independence of irrelevant alternatives (IIA)* which is stated by Luce and Suppes in 1965 [14] as: where any two alternatives have a non-zero probability of being chosen, the ratio of one probability over the other is unaffected by the presence or absence of any additional alternative in the choice set.

It can be easily shown as follows:

$$\frac{P_i}{P_j}=\frac{\exp(\beta V_i)/\sum_{k\in C}\exp(\beta V_k)}{\exp(\beta V_j)/\sum_{k\in C}\exp(\beta V_k)}=\frac{\exp(\beta V_i)}{\exp(\beta V_j)}=\exp\{\beta(V_i-V_j)\}, \quad (2.6)$$

where P_i is the probability of alternative *i* chosen,

V_i is the utility of alternative *i*,

β is a scale factor.

While IIA property is realistic in some choice situations, it is clearly not appropriate in others. Consider the notorious red bus - blue bus problem: A traveler has a choice of going to work by car or taking a blue bus. For simplicity assume that the representative utility of the two modes are the same, such that the choice probabilities are equal: $P_c = P_{bb} = 1/2$, where *c* is car and *bb* is blue bus. In this case, the ratio of probabilities is one: $P_c/P_{bb} = 1$. Now suppose that a red bus is introduced and that the traveler considers the red bus to be exactly like the blue bus. The probability that the traveler will take the red bus is therefore the same as for the blue bus, so that the ratio of their probabilities is one: $P_{rb}/P_{bb} = 1$. However, in the MNL model the ratio P_c/P_{bb} is the same whether or not another alternative, in this case the red bus, exists. This ratio therefore remains at one. The only probabilities for which $P_c/P_{bb} = 1$ and $P_{rb}/P_{bb} = 1$ are $P_c = P_{bb} = P_{rb} = 1/3$, which are the probabilities that the MNL model predicts.

In real life, however, we would expect the probability of taking a car to remain the same when a new bus is introduced that is exactly the same as the old bus. We would also expect the original probability of taking bus to be split between the two buses after the second one is introduced. That is, we would expect $P_c = 1/2$ and $P_{bb} = P_{rb} = 1/4$. In this case, the MNL model, because of its IIA property, overestimates the probability of taking either of the buses and underestimates the probability of taking a car. The ratio of probabilities of car and blue bus, P_c/P_{bb}, actually changes with the introduction of the red bus, rather than remaining constant as required by the MNL model.

The core of the problem is the assumption that the disturbances are mutually independent. In the case of red buses and blue buses, this is wholly implausible since both

these alternatives share all the unobserved characteristics of buses. In fact, rather than being independent, the disturbance of the red bus and blue bus can be assumed to be perfectly correlated.

However, thanks to the IIA property, it is possible to estimate model parameters consistently on a subset of alternatives for each sampled decision maker. Consider a situation with 100 alternatives; the researcher may use 10 of those alternatives so as to reduce computing time instead of dealing with 100. Since relative probabilities are not affected from the existence or absence of the alternatives not in the subset, the consistency of the estimator will not be affected.

2.3.1.2. Elasticity. Another useful property of MNL models is the elasticity concept. Elasticity is the change in demand when a specific variable is changed 1%. The direct elasticity of demand measures the responsiveness of the quantity demanded of an alternative to a change in attribute of the same alternative. The cross elasticity of demand measures the responsiveness of the quantity demanded of an alternative to a change in attribute of another alternative [15]. It can be shown as:

$$E_{X_{jkq}}^{P_{iq}} = \beta_{jk} X_{jkq} (\delta_{ij} - P_{jq}), \tag{2.7}$$

where q is the individual,
k is the parameter,
i, j are modes,
δ_{ij} is 1 for i=j (direct elasticity) and 0 for i≠j (cross elasticity).

Elasticity is the concept that makes MNL models sensitive to the policy changes. With the help of elasticity, the results of price changes or any enhancement in the alternative can be observed. Yet, elasticity is not the only way to predict the changes. There is also incremental logit for the linear-in-parameters logit model [15]. For choosing *i*, among a choice set C:

$$P(i:C)=\frac{\exp(U_i)}{\sum_{m\in C}\exp(U_m)}, \quad (2.8)$$

If utilities of an alternative changes by U_m, the probability of selecting *i* then;

$$P(i:C)=\frac{\exp(U_i+\Delta U_i)}{\sum_{m\in C}\exp(U_m+\Delta U_m)}, \quad (2.9)$$

Dividing both the numerator and the denominator by $\sum_{m\in C}\exp(U_m)$ to obtain the incremental logit model:

$$P(i:C)=\frac{P(i:C).\exp(\Delta U_m)}{\sum_{m\in C}P(m:C).\exp(\Delta U_m)}, \quad (2.10)$$

Thus, without recalculating utilities changes in the shares can be predicted using base choice probabilities and changes in utilities due only to the affected variable.

2.3.1.3. Market Segmentation Method. In some circumstances, it can be thought that all subjects are not affected equally from the model variables. Subjects with high-income could be less sensitive to the cost parameter than the low-income subjects. In these situations, the sample is divided into two or more segments and the model is calibrated separately for each segment and it is tested if the market segmentation enhances the model using the likelihood ratio test below [15].

$$LRT=-2\left[L_N(\beta)-\sum_{g=1}^{G}L_{N_g}(\beta^g)\right], \quad (2.11)$$

where $L_N(\beta)$ is the log-likelihood of the pooled model,

$L_{N_g}(\beta^g)$ is the log-likelihood of the segment g,

LRT is likelihood ratio test,

β is a constant.

If the result is larger than the corresponding X^2 value, it is said that the market segmentation method increases the overall significance of the model.

In order to see which coefficients of variables are significantly different from each other in the market segments, t-test below is used.

$$t=\frac{\beta_{11}-\beta_{12}}{\sqrt{var(\beta_{11})+ var(\beta_{12})}}, \quad (2.12)$$

where β_{11} is the coefficient of variable 1 for segment 1,

β_{12} is the coefficient of variable 1 for segment 2,

$var(\beta_{11})$ is the variance of the coefficient of variable 1 for segment 1,

$var(\beta_{12})$ is thevariance of the coefficient of variable 1 for segment 2.

If the *t* value found is larger than the corresponding t-value in the t-tables, it is said that coefficients of a variable in the segments are significantly different from each other.

2.3.1.4. Value of Time or Willingness to Pay. It is common to compare the cost variable with other variables in the model in order to calculate decision-makers' value of in-vehicle time, value of waiting time or value of walking etc. Thus, it can be found that how much decision-makers are willing to pay for a reduction in travel time [15].

$$V = \frac{\beta_i}{\beta_{cost}}, \quad (2.13)$$

where β_i is the coefficient of the travel time variable,

β_{cost} is the coefficient of cost variable,

V is the value of time.

2.4. Models Using Social Interaction

In the last decade, several articles have developed different approaches to model these social network effects econometrically. Here, a model with social interaction is proposed based on the early theoretical works of Brock and Durlauf [16,17] and Dugundji and Walker [18].

A conditional logit model was proposed with the following utility function:

$$U_j = \alpha + \beta * \text{trip characteristics} + \omega * \text{social network effects} + \boldsymbol{\varepsilon}, \quad (2.14)$$

where U_j is the utility of mode j,

α, β, ω are constants.

Individual characteristics are gender and income where trip characteristics are travel time, reliability of travel time, price and access time. Social network effects are defined as the percent of reference group that has chosen the specific generic mode alternative and $\boldsymbol{\varepsilon}$ is the error term.

Brock and Durlauf's model considered the social interactions as global which means that a decision-maker is influenced by all other decision-makers. Later on a more general behavior is derived by Dugundji and Gulyas for the case that a decision-maker is only influenced by only a subset of the whole population (non-global or local interactions). Here, in this study, reference group refers to the local interactions [19].

Another important issue about the model is endogeneity which is defined as the correlation between a variable and the error term. Here, the model may suffer from correlation between "social interaction variable" and "error term". In MNL models, error term stands for the factors which are unobserved to the modeler. These factors can be the influence of land-

use, spatial factors, individual tastes, etc. In order to diminish the effects of endogeneity, stated preference approach was used in the experiment. Walker *et al.* [20] give a great example for this case:

> "… people who live proximally to the decision maker will face the same attributes of alternative modes as the decision maker. The attributes that make transit or driving attractive or unattractive to a particular commuter may similarly affect his or her neighbors. Or, if the decision maker lives along an attractive bike path, then her neighbor also lives along or near the bike path, which makes bicycling arguably easier or more fun for both neighbors. Inevitably in this scenario, both the observed factors (i.e. travel time) and unobserved factors (i.e. aesthetics of the bike route or perceived safety of bus ride) will be similar for the decision maker and those who live nearby. …some (social) classes may have a stigma surrounding transit use—a la "only poor people take transit." Alternatively, if bicycling is considered hip in certain social circles, this would represent a shared unobserved preference." (p.365)

Using stated preference method enables modeler to place decision-makers into a hypothetical space so that they will not be aware of any unobserved factors. All of them will face with same alternatives with fixed attributes. Therefore in the model, the error term will only stand for decision-makers' idiosyncrasies.

3. RESEARCH METHODOLOGY

The outline of the methodology chapter is structured as shown in Figure 3.1: Section 3.1 describes the forming of the research hypothesis. Section 3.2 covers the design of the experiment in order to test the research hypothesis. This section is followed by section 3.3 which describes the experiments in a detailed form. Section 3.4 describes the questionnaire and the last section 3.5 explains the data collection procedure.

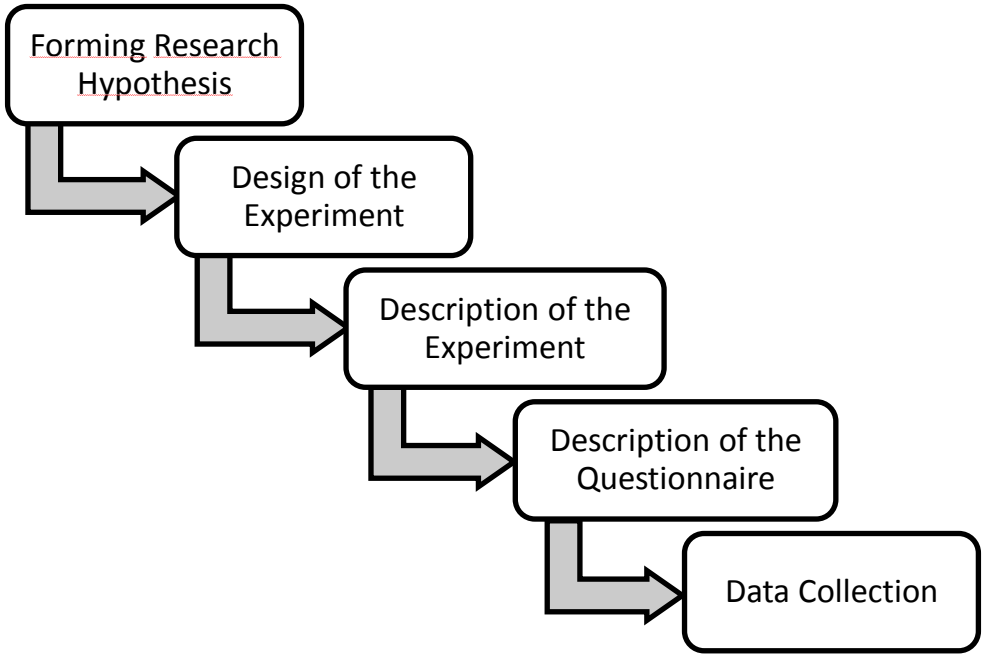


Figure 3.1. Outline of the Research Methodology.

3.1. Forming Research Hypothesis

The question wanted to be studied in this thesis is whether people are affected by their peers' choices in the mode choice setting when they are going to a destination not familiar to

subjects. These types of trips are also called non-daily trips. The reason behind working with non-daily trips is McFadden's influential paper [1] in which he postulated that individuals behave rationally when stakes are high, but get influenced by their social networks when stakes are low and possibilities for regret is small especially where alternatives are unfamiliar and ambiguous. It is perceivable that non-daily trips and means of these kinds of trips are unfamiliar to decision-makers. Furthermore, the role of gender and income of the decision-makers' will be examined and their effect on dependent variables will be investigated. An information cascade experiment is developed in order to test these hypotheses.

3.2. Design of the Experiment

In an information cascade experiment, subjects make decisions in a pre-determined order and their choices can be seen by subsequent subjects. Subjects were given three scenarios and three mode options for each and ask to choose one of the options for each scenario. Moreover, except the first group, they were also given the distribution of the choices of their peers who decided prior to them. Being able to see the choices of the previous peers is an important element of information cascade experiments but besides that it allows experimenters to design experiments which are not endogenous by creating a one way relation between subjects as seen in Figure 3.2.

A HTML code was written to create an information cascade experiment. Each group consisted of eight decision-makers and each session consisted of four groups which made 32 decision-makers in total for each session. The first group was a control group who decided only using attributes of alternatives. The second group see the choices of the first group third group see the choices of the first two groups and the fourth group see the choices of all previous groups besides attributes of alternatives. When all 32 people made their decisions, another session starts and continues in the same way. Note that peer information does not roll over into succeeding sessions. Every decision-maker had to complete the experiments and questionnaire in 12 minutes otherwise they were dropped out of the system.

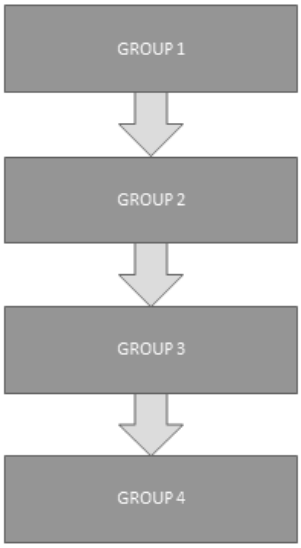


Figure 3.2. Information Cascade Experiment Setup.

The provided attributes for scenarios are travel time, reliability of travel time, access and egress time (walking and waiting time) and cost. Travel time and access time are in minutes, cost is in Turkish Liras and reliability is unitless and presented as percentages. Although the units and presentation of conventional variables are straightforward, the presentation of reliability is not very easy. Here, the percentage is used since Traffic Control Center of Istanbul used it in their website for the route planning application; therefore it is assumed that subjects are more familiar with this type of representation. Decision-makers were also informed with the definitions of attributes such that when the cursor points an attribute on the screen, an information balloon pops up to help decision-makers in what this attribute means.

In order to ensure the competitiveness of choices and avoid decision-makers flock to an alternative, a pilot study has been carried out with 32 people and under the light of this study the attributes of choices were updated to increase the competitiveness.

One last aspect in design of the experiment is the orthogonality of attributes. Although in real life, attributes of choices are not supposed to be orthogonal, in order to avoid unwanted correlations in the calibration of the model, a correlation test between variables has been

carried out in SPSS and the results are given in Table 4.1. The correlations between variables are not significant at the 5% confidence level and the highest absolute correlation is in between “accegtime” and “ttime” with -0.589.

Table 3.1. Correlation Matrix

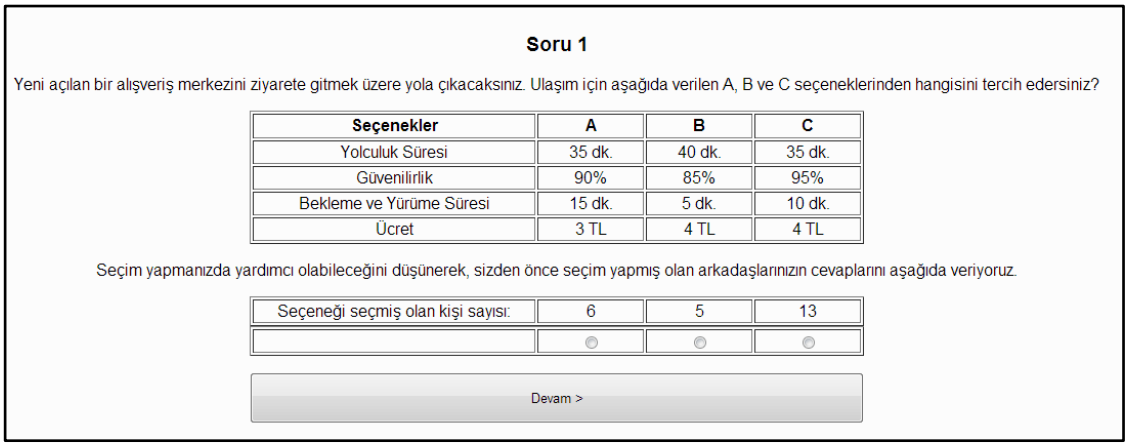
		ttime	**cost**	**accegtime**	**reli**
ttime	Pearson Correlation	1	0.333	-0.589	0.009
	Sig. (2-tailed)		0.382	0.095	0.982
	Sum of Squares and Cross-products	4100.000	103.333	-416.667	25.000
	Covariance	512.500	12.917	-52.083	3.125
cost	Pearson Correlation	0.333	1	-0.487	0.276
	Sig. (2-tailed)	0.382		0.184	0.471
	Sum of Squares and Cross-products	103.333	23.556	-26.111	60.000
	Covariance	12.917	2.944	-3.264	7.500
accegtime	Pearson Correlation	-0.589	-0.487	1	-0.354
	Sig. (2-tailed)	0.095	0.184		0.350
	Sum of Squares and Cross-products	-416.667	-26.111	122.222	-175.000
	Covariance	-52.083	-3.264	15.278	-21.875
reli	Pearson Correlation	0.009	0.276	-0.354	1
	Sig. (2-tailed)	0.982	0.471	0.350	
	Sum of Squares and Cross-products	25.000	60.000	-175.000	2000.000
	Covariance	3.125	7.500	-21.875	250.000

3.3. Description of the Experiment

In the experiments, the alternatives represent different travel modes but their names are not specified. Therefore the modes are “generic” modes and the choices will only be affected

by the stated characteristics given here. If the names of the modes were specified the modes would no longer be "generic" and hence the choices might be affected by the unspecified characteristics of them through the experience of the respondents. For instance, some might favor car because of its "privacy" which is not studied here. Furthermore, since the names of the modes are not specified using "alternative-specific" constants in the choice models are not needed. By using generic alternatives the choices were, in a way, simplified and the decisions would only have to be made by the consideration of the stated characteristics under each scenario.

The first question in the experiment, as shown in Figure 3.3, is about a hypothetical trip of going to a recently opened shopping mall. The travel time of alternatives are relatively short if the average duration of a trip in Istanbul that is 50 minutes, is considered.



Soru 1

Yeni açılan bir alışveriş merkezini ziyarete gitmek üzere yola çıkacaksınız. Ulaşım için aşağıda verilen A, B ve C seçeneklerinden hangisini tercih edersiniz?

Seçenekler	A	B	C
Yolculuk Süresi	35 dk.	40 dk.	35 dk.
Güvenilirlik	90%	85%	95%
Bekleme ve Yürüme Süresi	15 dk.	5 dk.	10 dk.
Ücret	3 TL	4 TL	4 TL

Seçim yapmanızda yardımcı olabileceğini düşünerek, sizden önce seçim yapmış olan arkadaşlarınızın cevaplarını aşağıda veriyoruz.

Seçeneği seçmiş olan kişi sayısı:	6	5	13
	○	○	○

Devam >

Figure 3.3. Question 1.

.

The second question, as given in Figure 3.4, is about a hypothetical meeting with friends for coffee. Durations of the journeys varies around average travel time of a single trip in Istanbul.

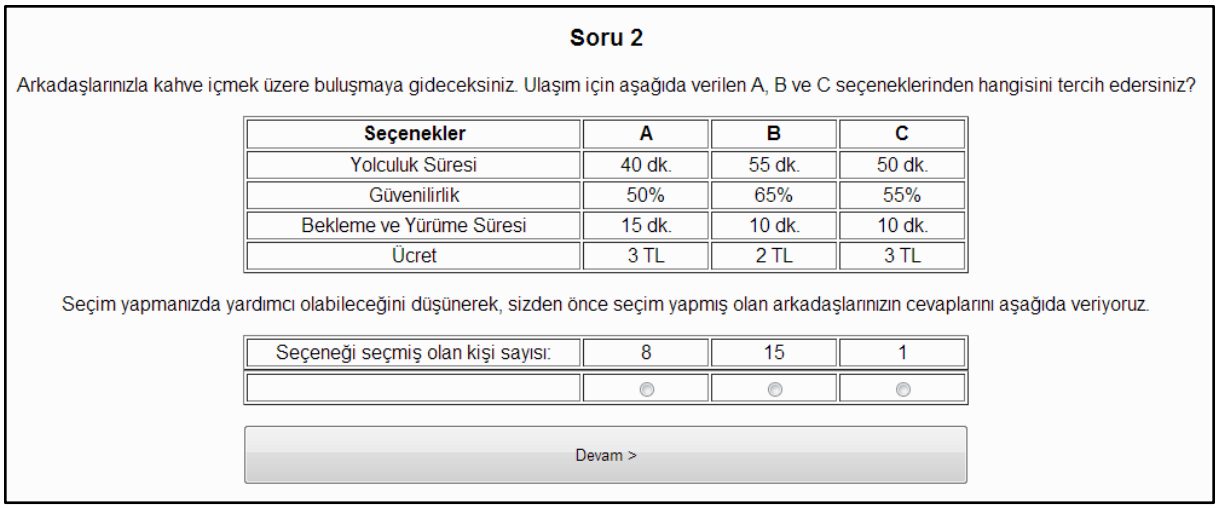


Figure 3.4. Question 2.

The third question, as shown in Figure 3.5, is about a hypothetical trip of going to a governmental institution to get some documents. In this question, travel times are over one hour which makes this trip, the longest trip in the experiment.

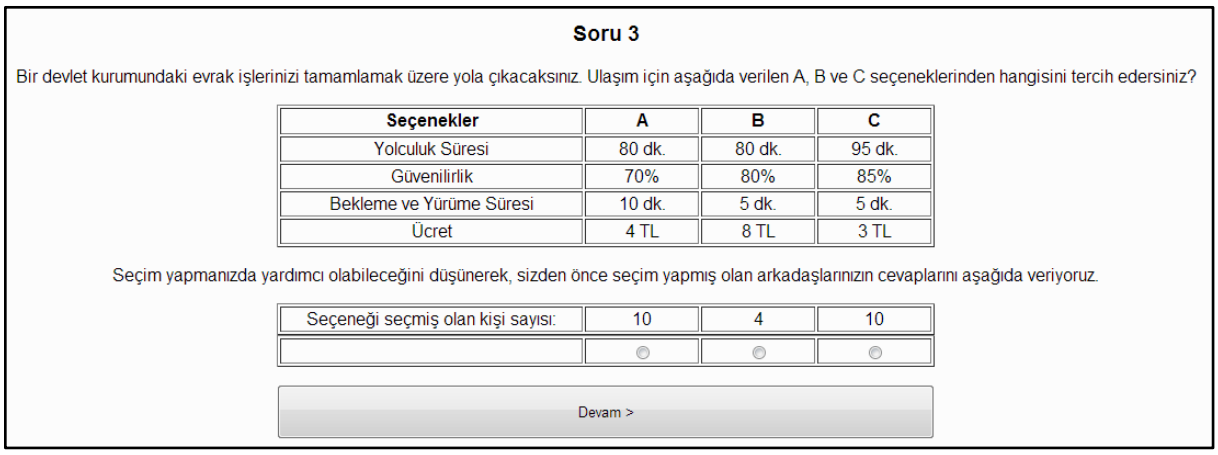


Figure 3.5. Question 3.

3.4. Description of the Questionnaire

After all subjects finish the questions, a questionnaire will pop-up on the screen as shown in Figure 3.6. The first page that the subjects will see, consists of basic information, which are name, surname and e-mail address, about the subjects. Nevertheless, these questions are not necessary to be filled.

The second page consists of ten questions related to the socio-economic characteristics as shown in Figure 3.7. In this part, subjects were asked to report their gender, age, driving license possession, Akbil (Istanbul pass-card) possession, personal car possession, smart-phone possession, frequency of internet use, frequency of pre-trip information use, monthly income and monthly expenditure on transportation.

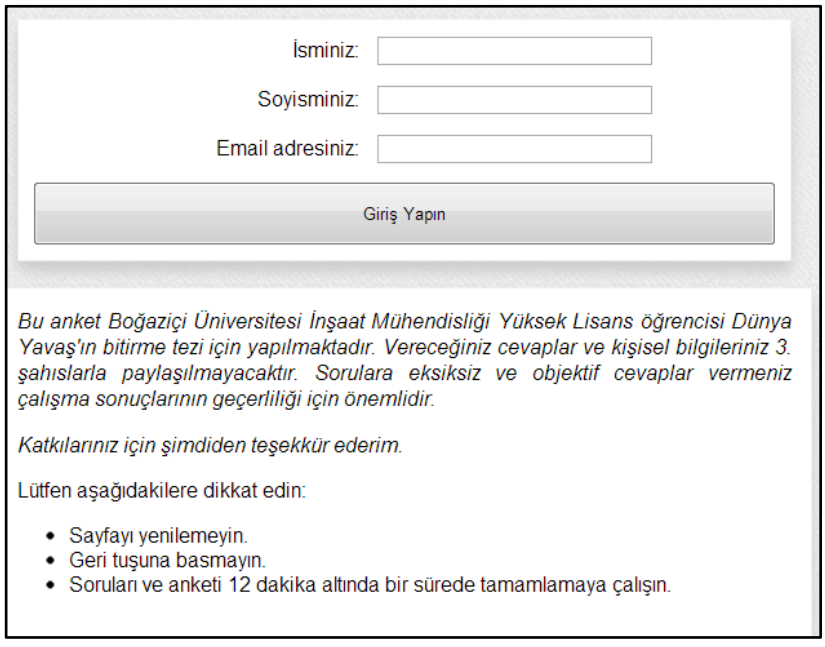


Figure 3.6. Questionnaire Page 1.

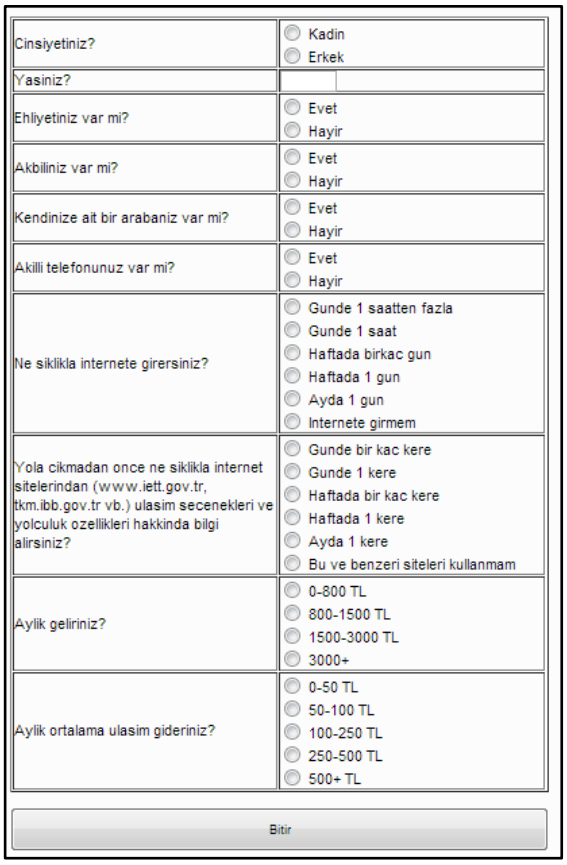


Figure 3.7. Questionnaire Page 2.

3.5. Data Collection

Data required to test the research hypothesis was collected via an Internet survey. The survey was online for two weeks and it has been spread via social media to subjects. 254 people answered the questions and filled the survey and 243 of the respondents were considered for the analysis because the scope of the research is limited with people who are older than 18 years old and younger than 31 years old. Since each respondent was asked to respond to three different scenarios and select one alternative under each of these scenarios a sample size of 729 cases was obtained.

4. ANALYSIS OF DATA AND MODEL CALIBRATION

In this chapter, first the socio-economic characteristics of the subjects were analyzed. In the second part, the pooled model was calibrated with the collected data and the results were presented. The third part of the chapter gave the results of the market segmentation over various socio-economic variables and in the last part of the chapter, inferred values of in-vehicle time and values of time spent in walking and waiting were presented.

4.1. Preliminary Analysis of the Collected Data

Mean age of the 243 respondents is 25.53. As presented in Table 4.1, mode age is 24 with 46 people. Sample represents the population characteristics in terms of gender. 49% of the subjects were female where 49.8% of the population of Turkey is female according to TUIK. 69% of the sample has a valid driving license however only 21% of them have their own personal cars. 159 of the total subjects have a smart-phone. 12% of the respondents use pre-trip information more than once in a day, 10% of them use it only once per day where 19% of them use it more than once in a week and 16% of them use it only once in a week. 26% of the subjects use pre-trip information once in a month and 17% of them never use this kind of information. 14% of the subjects spend less than 50 TL in a month, 41% of them spend between 51 TL and 100 TL. 25% of the subjects spend between 101 TL and 250 TL whereas 13% of them spend between 251 TL and 500 TL. Only 7% of the subjects spend more than 500 TL on transportation in a month. 85% of the subjects possess AKBIL. 91% of the subjects are online more than 1 hour per day. 8% of them use internet around 1 hour and only 1% of the subjects reported that they use internet less than 1 hour per day. 18% of the subjects earn less 800 TL in a month, 25% of them earn between 801 TL and 3000 TL, 39% of them earn between 1501 TL and 3000 TL and 18% of them earn more than 3001 TL in a month.

Table 4.1. Brief Description of Sample Characteristics

Number of Respondents:	243
Age:	100 % between 18-30 (mean 25.53)
Gender:	49 % Female
Have Driving License:	69 %
Have AKBIL:	85 %
Have an Auto:	21 %
Have Smart-phone:	65 %
Use Pre-trip Info more than once in a Week	57 %
Spends more than 100 TL for transportation	45 %
Spends more than 1 hour/day in Internet	91 %
Earns less than 1500 TL per month	43 %

4.2. Calibration of the Mode-Choice Model with the Pooled Data

The conditional logit model was calibrated using STATA [33] software in order to test the hypothesis that decision-makers are influenced by their peers when they are asked to choose a mode for non-daily trips.

The utility function of the model that was hypothesized was given in Equation 2.14. This formula has the following open form:

$$U_j = \gamma_1 * ttime + \gamma_2 * reli + \gamma_3 * accegtime + \gamma_4 * cost + \omega * inter + \boldsymbol{\varepsilon}, \quad (4.1)$$

where $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ and ω are constants,

U_j is the utility of alternative j,

ttime is the travel time in-vehicle (minutes),

reli is the reliability of the travel time (%),

accegtime is the total access and egress time (minutes),

cost is the cost of the trip (TL),

inter is the percentage of time the previous peers have chosen the specific alternative,

ε is the error term.

Here, in this model, alternative specific coefficients were not used because given mode alternatives are hypothetical ones namely; A, B and C. Therefore only the coefficients of variables and relationship among each other will be taken into account for following sections. The general model was calibrated using *asclogit* command of the STATA software. The output of the STATA including the calibrated model and the associated test statistics are presented in Table 4.2. The null hypothesis that all the model coefficients are equal to zero can be rejected with a very high confidence level (Prob > chi2 = 0.000).

Table 4.2. STATA Estimates of the Coefficient of Variables in the Model.

```
Alternative-specific conditional logit          Number of obs      =        2187
Case variable: id                               Number of cases    =         729

Alternative variable: alternative               Alts per case: min =           3
                                                               avg =         3.0
                                                               max =           3

                                                   Wald chi2(5)    =      109.27
Log likelihood = -743.39207                        Prob > chi2     =      0.0000
```

choice	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
alternative						
ttime	-.1641512	.0320167	-5.13	0.000	-.2269028	-.1013995
reli	9.162525	1.41618	6.47	0.000	6.386863	11.93819
accegtime	-.1667506	.0419381	-3.98	0.000	-.2489477	-.0845535
cost	-.4925226	.0959518	-5.13	0.000	-.6805846	-.3044606
inter	1.009691	.3001729	3.36	0.001	.4213626	1.598019

The signs of all variables are correct. Probability of choosing a specific alternative should decrease with increase in travel time (therefore minus), decrease in reliability, increase

in access and regress time, increase in cost and decrease in percentage of previous peers choosing specific alternative.

All variables are significant with 99.9% probability (P>|z|), therefore the null hypothesis that coefficients of variables are equal to zero can be rejected with a very high probability for all of the variables. In particular, the social interaction variable "inter" is proved to be significant in the mode choice problem and it can be said that choices of the peers do affect the choices of the travelers. It can also be interpreted as peers choosing a specific alternative give decision-makers positive signals that this specific alternative is feasible. Therefore the probability of latter decision-makers choosing this alternative increases. This model, with a high degree of certainty, proves our a priori hypothesis that "social network effect" do play a role in the travelers' behavior when they select a transportation mode.

The utility function has now the following form:

$$U_j = -0.164151*ttime + 9.162525*reli - 0.166751*accegtime - 0.492252*cost + 1.009691*inter + \varepsilon, \quad (4.2)$$

where U_j is the utility of alternative j,

ε is the error term.

4.3. Models with Socio-economic Segments

Market segmentation method was used to test if variables in the model equally affect the decision-makers with different socio-demographic backgrounds. It can be checked whether market segmentation increases the overall efficiency of model or not using the LRT formula in Section 2.3.1.3. T-test is used in order to check if a coefficient of variable in a segment is significantly different from the ones in the other segments.

4.3.1. Role of Gender

First, role of gender was examined and Table 4.3 presents the calibrated models for the gender segment and the associated test statistics.

Applying Equation 2.11:

$$LRT = -2\left[-743.39207-\left(-383.36371-355.98093\right)\right] = 8.09486 \quad (4.3)$$

Table 4.3. Models for Gender Segments.

	MALE		FEMALE		POOLED MODEL		T-TEST
	Coef. (sign.)	**Std. Error**	**Coef. (sign.)**	**Std. Error**	**Coef. (sign.)**	**Std. Error**	
TTIME	-0.2098465 (0.000)	0.04386	-0.0967682 (0.043)	0.04789	-0.164152 (0.000)	0.03202	-1.7413**
RELI	10.92251 (0.000)	1.8938	6.426256 (0.003)	2.1621	9.162525 (0.000)	1.416	1.5643**
ACCEGTIME	-0.233466 (0.000)	0.05872	-0.0772534 (0.206)	,06106	-0.166750 (0.000)	0.04194	-1.8440**
COST	-0.6519756 (0.000)	0.13233	-0.269818 (0.058)	0.14235	-0.492522 (0.000)	0.09595	-1.9663**
INTER	0.3387642 (0.397)	0.40036	1.891347 (0.000)	0.46183	1.009691 (0.001)	0.30017	-2.5402*
NO. of OBS.	125		118		243		
NO. of CASES	375		354		729		
LIKELIHOOD	-383.36371		-355.98093		-743.39207		
LRT FOR SEGMENTS	8.09486						
* significant in the 5% confidence level ** significant in the 10% confidence level							

The value found through LRT, 8.09486, is smaller than the appropriate X^2 value which is 11.07 for 5% significance level and degree of freedom 5; therefore market segmentation with gender does not increase the overall effectiveness of the model. However, there are significant differences in the coefficients of the two models, as explained below, although this sounds contradiction, this is possible as stated by Ortuzar and Willumsen [7].

It is seen that, in the only-female model, access and egress time variable is not significant anymore and significance of cost is only 0.058 where other variables still play an important role in the decision-making process.

In the only-male model, however, all the variables except social interaction variable keep their significance. The significance of the coefficient of social interaction variable became 0.603 and hence the null hypothesis that it is equal to zero cannot be rejected. All of the coefficients of the variables in the two models are significantly different from each other in the 10% confidence level and the coefficient of the social interaction variable "inter" is significantly different from the one in the other model in the 5% confidence level.

To conclude the role of gender in the model, separating the model into two segments, male and female, does not improve the model. Females do not consider cost and access and egress times while choosing their mode of transport. It seems that females are only affected by the travel time, reliability of travel time and their peers' choices as only the coefficients of these variables are significantly different from zero. Males, on the other hand, do not consider their peers' choices at all (the variable "inter" is not significant at the 5% level) but they are influenced by all the other travel related variables i.e. travel time, reliability, cost and access and egress times. This finding is important because it reveals very significant differences in the choice behavior of genders.

4.3.2. Role of Smartphone Possession

The market segmentation method is applied on different socio-economic segments. Although there is no evidence that these market segmentations increase the overall effectiveness of the model except, interestingly, smart-phone possession. Table 4.4 below shows the result of this market segmentation.

Table 4.4. Models for Smartphone Possession Segments.

	HAVE SMARTPHONE		HAVE NO SMARTPHONE		POOLED MODEL		T-TEST
	Coef. (sign.)	**Std. Error**	**Coef. (sign.)**	**Std. Error**	**Coef. (sign.)**	**Std. Error**	
TTIME	-0.22587 (0.000)	0.04220	-0.06593 (0.203)	0.05183	-0.16415 (0.000)	0.03202	-2.39290*
RELI	1.20263 (0.000)	1.79927	4.15687 (0.082)	2.38929	9.16250 (0.000)	1.41600	-0.98771
ACCEGTIME	-0.25521 (0.000)	0.05693	-0.03189 (0.625)	0.06521	-0.16675 (0.000)	0.04194	-2.57985*
COST	-0.63699 (0.000)	0.12503	-0.27576 (0.079)	0.15723	-0.49252 (0.000)	0.09595	-1.79820**
INTER	0.81543 (0.029)	0.37351	1.40835 (0.006)	0.51213	1.0097 (0.001)	0.30017	-0.93539
NO. of OBS.	159		84		243		
NO. of CASES	477		252		729		
LIKELIHOOD	-474.83715		-262.36899		-743.39207		
LRT FOR SEGMENTS	12.37186*						
* significant in the 5% confidence level ** significant in the 10% confidence level							

The model with smartphone users has similar characteristics with the general model in terms of significance of variables. In the segment with smartphone users, all the model coefficients are significant. However, in the model of subjects without smartphone, travel time and access and egress times become insignificant. The coefficients of these variables are significantly different from the ones in the model with smartphone users as shown in the significance (sign.) of the coefficients. Moreover, reliability and cost variables are no longer significant in the 5% confidence level but 10% confidence level and the cost variable is significantly different between the models in the 10% confidence level. Therefore, the only remaining significant variable is the interaction variable in this model of subjects without a smartphone.

Difference in the two models can be interpreted as smart-phone users are more conscious decision-makers who think over every aspect of the trip whereas non-smart-phone users are relatively less conscious decision-makers and depend very heavily on their peers'

choices although the reasons of why there is such a significant difference between two models is quite ambiguous.

4.3.3. Role of Age

Table 4.5 presents the results of the market segmentation over age. Subjects under 25 years old are not affected by their peers' choices although subjects over 25 years old are affected by choices of their peers. However, statistically there is no significant difference between the variables of the two models and the market segmentation does not improve the significance of the general model.

Table 4.5. Models for Age Segments.

	ABOVE 25 YEARS OLD		**BELOW 25 YEARS OLD**		**POOLED MODEL**		**T-TEST**
	Coef. (sign.)	**Std. Error**	**Coef. (sign.)**	**Std. Error**	**Coef. (sign.)**	**Std. Error**	
TTIME	-0.20296 (0.000)	0.04728	-0.13604 (0.002)	0.04414	-0.16415 (0.000)	0.03202	-1.03465
RELI	9.01749 (0.000)	2.02164	9.74535 (0.000)	2.01164	9.16253 (0.000)	1.41600	-0.25521
ACCEGTIME	-0.21427 (0.001)	0.06355	-0.13310 (0.018)	0.05644	-0.16675 (0.000)	0.04194	-0.95496
COST	-0.57121 (0.000)	0.14004	-0.43984 (0.001)	0.13371	-0.49252 (0.000)	0.09595	-0.67847
INTER	1.25205 (0.003)	0.42230	0.65676 (0.129)	0.43308	1.00969 (0.001)	0.30017	0.98413
NO. of OBS.	116		127		243		
NO. of CASES	348		381		729		
LIKELIHOOD	-352.24333		-386.46665		-743.39207		
LRT FOR SEGMENTS	9.36418						

4.3.4. Role of Driving License

There is a significant difference at the 10% confidence level in the social effect variable, "inter", between two segments namely, subjects with a driving license and subjects without a

driving license. Subjects without a driving license do not consider their peers' choices when they face with a mode-choice problem whereas subjects with a driving license although there is not any significant difference in the variables of the segments, except "inter" which is significantly different in the 10% confidence level. Table 4.6 presents the result of the market segmentation.

Table 4.6. Models for Driving License Possession Segments.

	HAVE LICENSE		**HAVE NO LICENSE**		**POOLED MODEL**		**T-TEST**
	Coef. (sign.)	**Std. Error**	**Coef. (sign.)**	**Std. Error**	**Coef. (sign.)**	**Std. Error**	
TTIME	-0.14321 (0.000)	0.03915	-0.20689 (0.000)	0.05610	-0.16415 (0.000)	0.03202	0.93096
RELI	8.02698 (0.000)	1.72312	11.53356 (0.000)	2.50762	9.16253 (0.000)	1.41600	-1.15250
ACCEGTIME	-0.14098 (0.006)	0.05129	-0.21906 (0.003)	0.07338	-0.16675 (0.000)	0.04194	0.87214
COST	-0.39977 (0.001)	0.11614	-0.69797 (0.000)	0.17354	-0.49252 (0.000)	0.09595	1.42803
INTER	1.40253 (0.000)	0.36590	0.13110 (0.805)	0.53237	1.00969 (0.001)	0.30017	1.96819**
NO. of OBS.	167		76		243		
NO. of CASES	501		228		729		
LIKELIHOOD	-509.86395		-230.09246		-743.39207		
LRT FOR SEGMENTS	6.87132						
** significant in the 10% confidence level							

4.3.5. Role of AKBIL Possession

Interaction variable is insignificant in the segment with subjects who do not possess AKBIL. Moreover, access and egress time variable and cost variable is barely significant in the 5% confidence level. There is no significant difference between the variables of the two models and there is no evidence that segmentation over AKBIL possession increase the overall significance of the model as seen Table 4.7.

Table 4.7. Models for AKBIL Possession Segments.

	HAVE AKBIL		HAVE NO AKBIL		POOLED MODEL		T-TEST
	Coef. (sign.)	Std. Error	Coef. (sign.)	Std. Error	Coef. (sign.)	Std. Error	
TTIME	-0.15732 (0.000)	0.03422	-0.21283 (0.023)	0.09349	-0.16415 (0.000)	0.03202	0.55757
RELI	8.69709 (0.000)	1.52064	11.94201 (0.003)	3.96873	9.16253 (0.000)	1.41600	-0.76350
ACCEGTIME	-0.15596 (0.000)	0.04465	-0.24669 (0.050)	0.12597	-0.16675 (0.000)	0.04194	0.67880
COST	-0.49799 (0.000)	0.10325	-0.50404 (0.066)	0.27375	-0.49252 (0.000)	0.09595	0.02066
INTER	1.03019 (0.002)	0.32478	0.98068 (0.223)	0.80554	1.00969 (0.001)	0.30017	0.05701
NO. of OBS.	207		36		243		
NO. of CASES	621		108		729		
LIKELIHOOD	-634.35448		-106.70853		-743.39207		
LRT FOR SEGMENTS	4.65812						

4.3.6. Role of Auto Ownership

Table 4.8. Models for Auto Ownership Segments.

	HAVE AN AUTO		HAVE NO AUTO		POOLED MODEL		T-TEST
	Coef. (sign.)	Std. Error	Coef. (sign.)	Std. Error	Coef. (sign.)	Std. Error	
TTIME	-0.23657 (0.002)	0.07527	-0.15077 (0.000)	0.03591	-0.16415 (0.000)	0.03202	-1.02885
RELI	9.65164 (0.002)	3.05553	9.18139 (0.000)	1.61940	9.16253 (0.000)	1.41600	0.13598
ACCEGTIME	-0.28171 (0.007)	0.10498	-0.14398 (0.002)	0.04629	-0.16675 (0.000)	0.04194	-1.20049
COST	-0.57546 (0.008)	0.21855	-0.48787 (0.000)	0.10869	-0.49252 (0.000)	0.09595	-0.35884
INTER	0.86267 (0.146)	0.59271	1.03862 (0.003)	0.34978	1.00969 (0.001)	0.30017	-0.25565
NO. of OBS.	50		193		243		
NO. of CASES	150		579		729		
LIKELIHOOD	-153.17748		-585.76055		-743.39207		
LRT FOR SEGMENTS	8.90808						

As seen in Table 4.8, travel time, reliability of travel time, access and egress time and cost are significant in both models; however interaction variable is only significant in the model with the subjects who have a personal car.

4.3.7. Role of Income

All variables except social interaction variable are significant in both models. The social interaction variable is only significant in the high income segment which constitutes subjects earning more than 1500 TL. There is no statistical evidence that segmentation over income increases the overall efficiency of the model as presented in Table 4.9.

Market segmentation method is also used over monthly transportation expenditure and pre-trip information usage but both of them fail at increasing the overall efficiency of the model with pooled data. Moreover, no evidence is found such that coefficient of a variable is different from the one in the general model.

Table 4.9. Models for Income Segments.

	HIGH INCOME		**LOW INCOME**		**POOLED MODEL**		**T-TEST**
	Coef. (sign.)	**Std. Error**	**Coef. (sign.)**	**Std. Error**	**Coef. (sign.)**	**Std. Error**	
TTIME	-0.17943 (0.000)	0.04366	-0.15021 (0.002)	0.04748	-0.16415 (0.000)	0.03202	-0.45300
RELI	9.74906 (0.000)	1.87957	8.80982 (0.000)	2.17836	9.16253 (0.000)	1.41600	0.32645
ACCEGTIME	-0.18379 (0.002)	0.05819	-0.15130 (0.013)	0.06088	-0.16675 (0.000)	0.04194	-0.38575
COST	-0.51417 (0.000)	0.12935	-0.48572 (0.001)	0.14461	-0.49252 (0.000)	0.09595	-0.14664
INTER	1.30940 (0.001)	0.38613	0.48115 (0.321)	0.48445	1.00969 (0.001)	0.30017	1.33694
NO. of OBS.	138		105		243		
NO. of CASES	414		315		729		
LIKELIHOOD	-414.55934		-326.95563		-743.39207		
LRT FOR SEGMENTS	3.7542						

4.4. Inferred Values of Time from the Models

A value of time analyses has been carried out in order to gain more insight into variables which influence mode choice behavior. Because the alternatives are hypothetical, calculating sensitivity of alternatives to the variables is not meaningful, but calculating value of time and value of access and regress time can give us further information about the decision-makers.

Table 4.10 shows the comparison of values of time for in vehicle time and access egress time spend in walking and waiting calculated and those obtained from previous mode choice models from different cities of Turkey. Note that home-based-other trip values are used for Istanbul, Mersin and Kocaeli because our scenarios are non-daily trips which are mostly covered in home-based-other trips.

Table 4.10. Value of Time.

	MALE	FEMALE	POOLED	ISTANBUL	MERSIN	KOCAELI	UNIT
TTIME	19.312	21.519	19.997	29.408	11.555	12.095	TL/hour
ACCEGTIME	21.485		20.314				TL/hour

Source: 34, 35, 36

Our sample has slightly lower values of time for in vehicle time (TTIME) than Istanbul and higher than Mersin and Kocaeli data [34-36]. Because the scope of this research is limited with people between 18 years old and 30 years old, it is plausible to have lower values than Istanbul data. Mersin and Kocaeli are industrial cities where life is much slower than Istanbul. Therefore it is not a surprise that they have lower values of time. Hence these values are quite reasonable. Value of time spend in access and egress to the mode which includes walking and waiting resulted with higher values of time than in-vehicle travel time which is usually the case in mode choice models [37]. This is mainly because walking and waiting time are more bothersome than the time spend inside the vehicle.

Males are willing to pay 19.312 TL for 1 hour decrease in travel time where females are willing to pay 21.519 TL for the same decrease. For a 10 minute decrease in travel time, decision-makers are willing to pay 3.333 TL which sounds over-priced. However it should be noted that these are values of in-vehicle travel time.

Decision-makers are willing to pay for 1 hour decrease in access and regress times more than 1 hour decrease in travel time. It is very reasonable to say that instead of waiting and walking to the stops they would prefer to spend same time in vehicle because value of time for access and egress times are higher than the value of in-vehicle travel time. Decision-makers are willing to pay 22.576 TL for one hour of waiting and walking time.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

For the last forty years, decision-makers are assumed to be *homo-economicus* who are individualistic and egocentric in their tastes and beliefs. They are immune to the influence of their peers and indifferent to the prosperity of others. Choice models in transportation make the same assumptions and treat decision-makers as isolated individuals in space where there is no inter-relationship between decision-makers.

Behavioral economics question these assumptions with simple and cleverly designed experiments and try to show that decision-makers are not 100% rational, they often fail at maximizing their utilities and these failures follow a pattern such that they cannot be treated as individual mistakes. In the case of social interactions, decision-makers can be directly influenced from their peers or they can use their peers to acquire facts and update their expectations.

In this thesis, the role of social interactions on mode choice context was selected and whether decision-makers were affected from choices of their peers was studied. Only the non-daily trips were considered firstly because daily HBW and HBS trips were suspected to be less affected by social interactions and secondly it would have been difficult to analyze all the trip purposes within a MS thesis. Moreover, it was assumed that unfamiliarity of final destination and travel modes of non-daily trips will increase the need for extra information which is, in this case, observing choices of peers. This research also concentrated on young adults who are between 18 years old and 30 years old with an assumption that they are more into social media where social interactions hold an importance and can often be observed.

An information cascade experiment was designed to test the hypotheses. In an information cascade experiment, subjects make decisions in a pre-determined order and their choices can be seen by subsequent subjects. Decision-makers were asked to choose between three modes of transportation for each given non-daily trip scenario, then they were asked to fill a questionnaire with socio-demographic questions. The survey was online for two weeks and 243 people validly filled it and each had worked out three different choice scenarios.

Conditional logit models were calibrated with the collected data for the mode choice. All travel related variables were significantly different from zero with significance level of at least 0.01%. The social interaction variable "inter" was significant at 0.1% level. Hence it was shown that observing choices of peers influences trip-makers' mode choices, which proves the research hypothesis of the thesis.

Moreover, the segmentation over the gender revealed that males and females were not equally affected from the variables of the model. It was observed that females were only affected by travel time, reliability of travel time and the social interaction variable significantly. Cost variable was significant at 5.8% level. Social interaction variable was insignificant in the only-male model where all the other variables remained significant.

It was also observed that separating the subject pool into smart-phone users and non-smart-phone users significantly increased the overall significance of the model. In the model of non-smart-phone users, only the social interaction variable was significant at 5% confidence level where travel time, reliability of travel time, access and egress time and cost were all insignificant. In the model with smart-phone users, all of the variables remained significant.

Value of time for in-vehicle travel time and access and egress times were calculated in this research. Males were found to be willing to pay 19.312 TL for 1 hour decrease in travel time where females were willing to pay 21.519 TL for the same decrease. The value of time for both genders was found to be 19.997 TL and these values were quite comparable to the values of time found in other cities of Turkey. Trip-makers were found to be willing to pay

20.314 TL for one hour of waiting and walking, which is slightly more from the in-vehicle travel time as expected.

5.2. Recommendations for Further Studies

In this research, it was found that social interactions play a role in the decision-making process of individuals while choosing a mode of transportation for a non-daily trip. Further research directions are recommended below:

- Removing the age limit of the sample such that it represents the whole population.
- Extending the trip purpose from non-daily trips to all other trip purpose categories in order to observe whether social interactions still play a role in these choice contexts.
- The role of social interactions was only studied within the mode choice context. Studying the role of social interactions in other trip choice contexts such as trip generation (the decision to make a trip or not), destination choice, and the route choice.
- Checking the validity of these findings by applying the using social interaction analysis in the real-world experiments.
- Investigating the possibility of correlation between social interaction and reliability.

REFERENCES

1. McFadden, D., "Sociality, Rationality, and the Ecology of Choice", *Choice Modelling: The State-of-the-art and The State-of-practice : Proceedings From The Inaugural International Choice Modelling Conference,* Bingley, UK: Emerald, 2010.

2. Verplanken, B., H. Aarts and A. van Knippenberg, "Habit, Information Acquisition, and the Process of Making Travel Mode Choices", *European Journal of Social Psychology*, Vol. 27, pp. 539-560, 1997.

3. Ergün, G., P. R.. Stopher, H. M. Al-Ahmadi, "Captivity Revisited", *ASCE Journal of Transportation Engineering*, Vol. 125, No.1, pp. 1-7, 1999.

4. Khattak A., A. Kanafani, and E. Le Colletter, "Stated and reported route diversion behavior: Implications on the benefits of ATIS", *Transportation Research Record,* 1464, TRB, National Academies, Washington, D.C., pp. 193-212, 1994.

5. Manski, C. F., "Identification of Endogenous Social Effects: The Reflection Problem", *The Review of Economic Studies*, Vol. 60, No. 3, pp. 531-542, 1993.

6. Banerjee, A. V., "A Simple Model of Herd Behavior", *The Quarterly Journal of Economics*, Vol. 107, No. 3, pp. 797-817, 1992.

7. Ortúzar, J. d. D. and L. G. Willumsen, *Modelling Transport*, Fourth Edition, John Wiley & Sons, Ltd, Chichester, UK, 2011.

8. Domencich, T. and McFadden, D., *Urban Travel Demand: A Behavioral Analysis*, North-Holland Publishing Co., 1975.

9. Thurnstone, L L., “The Method of Paired Comparisons for Social Values”, *Journal of Abnormal and Social Psychology*, 21, pp. 384-400, 1927.

10. Luce, R. D., *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley, 1959.

11. Marschak, J., “Remarks on the Economics of Information”, *In Contributions to Scientific Research in Management*, University of California Printing Department, Berkeley, pp. 79-98, 1960.

12. Marschak, J. and Block, H. D., “Random Orderings and Stochasitc Theories of Responses”, In Olkin, Ghurye, Hoeffding, Madow and Mann, eds., *Contributions to Probability and Statistics*, Stanford University Press, pp. 97-132, 1960.

13. Williams, H. C. W. L., “On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit. *Environment and Planning A* 9, pp. 285–344, 1977.

14. Luce, R. D., and Suppes, P., “Preference, utility, and subjective probability”. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.) *Handbook of Mathematical Psychology*, Vol. III. New York: Wiley. pp. 252-410, 1965.

15. Ergün G., *Ulaşım Türel Seçim Modelleri El Kitabı,* Istanbul Metropoliten Planlama ve Kentsel Tasarım Merkezi (İMP), 2008.

16. Brock, W. A. and Durlauf, S. N., “Discrete Choice with Social Interactions”, *The Review of Economic Studies*, Vol. 68, No. 2 (Apr., 2001), pp. 235-260, 2001.

17. Brock, W. A. and Durlauf, S. N., "Multinomial Choice with Social Interactions", *NBER Technical Working Papers 0288*, National Bureau of Economic Research, Inc., 2002.

18. Dugundji, E. R. and Walker, J. L., “Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed GEV Models with Field and “Panel” Effects”, *Transportation Research Record: Journal of the Transportation Research Board*, Issue Number: 1921, Transportation Research Board, pp. 70-78, 2005.

19. Dugundji, E. R. and Gulyás, L., "Socio-dynamic Discrete Choice on Networks in Space: Impacts of Agent Heterogeneity on Emergent Outcomes", *Environment and Planning B: Planning and Design,* 35(6), pp. 1028 – 1054, 2008.

20. Walker, J. L., E. Ehlers, I. Banerjee and, E. Dugundji, “Correcting for Endogeneity in Behavioral Choice Models with Social Influence Variables”, *Transportation Research Part A: Policy and Practice*, Volume 45, Issue 4, pp. 362–374, 2011.

21. Manski, C. F., “Identification Problems in the Social Sciences”, *Sociological Methodology,* Vol. 23, pp. 1-56, American Sociological Association, 1993.

22. Bikhchandani, S., D. Hirshleifer and I. Welch, “A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades”, *Journal of Political Economy*, Vol. 100, No. 5, pp. 992-1026, 1992.

23. Anderson, L. R. and Holt, C. A., "Information Cascades in the Laboratory," *American Economic Review*, Vol. 87, pp. 847–862, 1997.

24. Hung, A. A. and Plott, C. R., "Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions," *American Economic Review*, American Economic Association, Vol. 91, pp. 1508-1520, 2001.

25. Fukuda, D. and Morichi, S., "Incorporating aggregate behavior in an individual's discrete choice: An application to analyzing illegal bicycle parking behavior," *Transportation Research Part A: Policy and Practice*, Elsevier, Vol. 41, pp. 313-325, 2007.

26. Páez, A. and Scott, D. M., "Social Influence on Travel Behavior: A Simulation Example of the Decision to Telecommute" *84th Transportation Research Board Annual Meeting*, Washington D.C., pp. 9 – 13, 2005.

27. Páez, A. and Scott, D. M., "A Discrete-choice Approach to Modeling Social Influence on Individual Decision Making", *Environment and Planning B: Planning and Design* 35(6), pp. 1055 – 1069, 2008.

28. Dugundji, E. R. and Gulyás, L., "Discrete Choice on Networks: An Agent-Based Approach", *North American Association for Computational Social and Organizational Science Conference* (NAACSOS 2003), pp. 22-25, Omni William Penn, Pittsburgh, PA, 2003,

29. Walker, J. L., D. Welch and Y. Zheng, "Experimental Economics in Transportation: A Focus on Social Influences and the Provision of Information", *Transportation Research Record: Journal of the Transportation Research Board*, Issue 2156, Transportation Research Board, pp. 47-55, 2010.

30. Ronald, N.A., T. A. Arentze and H. J. P. Timmermans, "An Agent-based Framework for Modelling Social Influence on Travel Behaviour.", *Proceedings of the 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, pp. 2955-2961, 2009.

31. Goetzke, F., "Network Effects in Public Transit Use: Evidence from a Spatially Autoregressive Mode Choice Model for New York", *Urban Studies*, Vol. 45, pp. 407-417, 2008.

32. Goetzke, F. and Rave, T., "Bicycle Use in Germany: Explaining Differences between Municipalities with Social Network Effects", *Urban Studies*, Vol. 48, pp. 427–437, 2011.

33. StataCorp., *Stata 11 Base Reference Manual*, College Station, TX, Stata Press, 2009.

34. Istanbul Metropolitan Municipality and Japan International Cooperation Agency (JICA), *The Study on the Integrated Transportation Master Plan for Istanbul Metropolitan Area in the Republic of Turkey, Progress Report*, October 2008.

35. Boğaziçi Proje Mühendislik, Planlama ve İnşaat San. Tic. Ltd. Şti., *Mersin Büyükşehir Belediyesi Ulaşım Master Planı Revizyonu Çalışması*, 2010.

36. Boğaziçi Proje Mühendislik, Planlama ve İnşaat San. Tic. Ltd. Şti., *2025 Kocaeli Ulaşım Ana Planı ve Öncelikli Toplu Taşıma Sistemleri Ön Proje ve Fizibilite Etütlerinin Hazırlanması*, 2012.

37. Ben-Akiva, M. and Lerman, S., *Discrete choice analysis*, The MIT Press, Cambridge Massachusetts, 1985.