ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

Development of Modal Choice Model For Home-Based Work Trips-The Case of Jemo Area Residents In Addis Ababa

A Thesis in Road and Transport Engineering

By Firdews Ali

August 15, 2018

Addis Ababa

A Thesis

Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science

The undersigned have examined the thesis entitled '**Development of modal choice model for home-based work trips-the case of Jemo area residents in Addis Ababa**' presented by **FIRDEWS ALI**, a candidate for the degree of **Master of Science** and hereby certify that it is worthy of acceptance.

Advisor	Signature	Date
Internal Examiner	Signature	Date
External Examiner	Signature	Date
Chair person	Signature	Date

UNDERTAKING

I certify that research work titled "development of modal choice model for home-based work trips-the case of Jemo area residents in Addis Ababa" is my own work. The work has not been presented elsewhere for assessment. Where material has been used from other sources it has been properly acknowledged / referred.

Signature of Student

Firdews Ali

ABSTRACT

Addis Ababa is one of the cities that are plagued by a continuous increase in traffic congestion which resulted in longer travel times, increased accidents, more fuel consumption, and deterioration in air quality. Providing and improving urban public transport service is highly important in mitigating these challenges as well as creating sustainable transport systems. The corridor from Jemo to Mexico is one of the most congested corridors in the city and the trips in the morning peak hour are mostly home-based work trips. The study of modal choice behavior of residents traveling to work will guide policy makers to improve the service quality of public transportation systems in a way the majority of the commuters will use them.

Discrete choice models can be used to analyze and predict a decision maker's choice of one alternative from a set of alternatives. This study aims to contribute for the improvement of the service quality of public transport systems in the city by determining the basic factors that affect the preference of modes and how each factor affect the mode choice of trip makers to travel to work place. 475 questionnaires were collected in which $4/5^{th}$ of the data were used for calibration and $1/5^{th}$ for validation of the revealed and stated preference models.

The collected data of variables related to the trip maker and characteristics of the mode are fitted in special type of multinomial logit model known as conditional logit model using SPSS. Revealed preference model was developed for preferences among large bus, minibus (taxi) and private cars. Whereas the stated preference model was developed for preferences among the existing modes (large bus, minibus (taxi) and private car) and the BRT system.

The developed models for the revealed preference survey proved that the total travel time and total travel cost of the transport modes are the primary factors that statistically significantly affect in addition to car ownership, time of leaving home and education level interaction variable with taxi. The stated preference models indicated that the preference towards the BRT increases with the decrease in total travel cost, in-vehicle travel time and out of vehicle travel time. The models are validated to be capable of predicting 80% and 92% of the preference of trip makers among the available revealed and stated modes, respectively.

Key words: Mode choice modeling, conditional logit model, stated preference, BRT, Addis Ababa

ACKNOWLEDGEMENT

I would first like to acknowledge the Ethiopian Roads Authority for funding the scholarship.

Then my sincere gratitude goes to my thesis advisor Dr. Bikila Teklu for the helpful, constructive and valuable advises and comments throughout my thesis work. I am thankful for AAiT, school of civil engineering staff members Kidus Ayalneh, Mahlet Demere and Raeed Ali for their advises, encouragement and fruitful comments.

I would also like to thank my friends for their valuable support during data collection of the study.

This thesis work is a result of a loving, courageous, motivating family that provide unfailing support and continuous encouragement throughout my study and thesis work. So I am deeply thankful to them.

I would also like to acknowledge the people of Jemo area for their patience during the interview, and all the people, whom I would have listed all by names if it was possible, that encouraged, supported, and advised me throughout this thesis work.

Finally, I would like to thank the Almighty for all the blessings to accomplish this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	Artificial neural network
BRT	Bus rapid transit
CMNL	Conditional multinomial logit
COXREG	Cox regression command in SPSS
ELM	Easy logit model
ETB	Ethiopian birr
FIS	Fuzzy inference system
GDP	Gross domestic product
GEV	Generalized extreme value
IIA	Independence of irrelevant alternatives
IID	Identically and independently distributed
MNL	Multinomial logit
MNP	Multinomial probit
NL	Nested logit
NOMREG	Nominal regression command in SPSS
SPSS	Statistical package for the social sciences
STATA	Statistics and data
XLSTAT	Statistical software for excel

CHAPTER 1. INTRODUCTION

1.1 Background

The contribution of transportation sector to a country's development is high. Its share of contribution to the gross domestic product (GDP) of a country is incontrovertible and it has been confirmed that its share in this respect in many developed countries is as high as 26%; whereas, in developing countries like Ethiopia it is 5% (Temesgen, 2006). Transportation services is a precondition for economic activities as well as for leisure activities. At the same time, transportation, especially in cities, causes a number of environmental problems and negative impacts on human health due to emissions, noise or vibrations. Accidents and congestion are other visible negative manifestations of transport. Both represent a social as well as economic problem.

Congestion is becoming a common experience in all parts of Addis Ababa due to lack of consistent concerted efforts from various stakeholders and policy makers (Eshete, 2015) resulting in longer travel times, lost productivity, increased accidents and automobile insurance rates, more fuel consumption, increased freight transportation costs, and deterioration in air quality. Moreover, congestion and auto pollution mitigation strategies are strongly associated with mode choice and composition of the traffic stream moving on the road network.

The city uses mostly mass transport modes such as Light Train, buses, taxis and midi buses. These providers are hardly able to cope with the passengers' demand for transportation. As a result, residents of Addis Ababa had to face great inconveniences, as well as additional costs during their daily trips to their destinations. The city's growth in economy, geographical area and population, brings urgent attention and needs additional mass transport service provision supported by capacity and technology (Ethiopian Ministry of Transport, 2011).

Public transportation systems are useful means of transportation which will improve the environmental pollution and congestion of the city. According to (Ortuzar & Willumsen, 2011), traveling in almost all public transport modes uses road space more efficiently and produce fewer accidents and emissions than using private cars. Consequently, to balance the demand and supply gap, the government focuses on mass transport service. Most importantly the government introduces government employees transport service buses, encourages private taxi operators to shift to busses by incentivize custom duty free and 70% bank loans facility. With these incentives, they have established two new companies; Sheger Express bus and Alliance bus having 500 buses. The public Anbessa bus company has introduced intelligent transport system and increased its fleet by new and clean buses. In addition, bus rapid transit (BRT) is under construction along one corridor, i.e. from Jemo residential area to Mexico, and other five routes are identified (H.Mariam, 2017).

These needs, introductions and provisions of new modes of mass transit systems in the city make it necessary to investigate the attributes of the characteristics of the trip makers, the modes, and the trips, whose values may be changed through these decisions and implementations. Developing discrete choice models for modal choice also guides the most effective improvements that should be employed to the existing public transport systems and to those in plan, so that the modal shares of these systems is maximized in different corridors of the city.

Dealing with the present bottlenecks as well as creating long lasting and sustainable transport systems has been the greatest challenge of urban transportation planning. Calibrating the present need and forecasting the future demand is the underlying agenda of travel demand forecasting. Mode choice forms an integral part of this process as it gives a complete insight to the mode choice preferences of the commuters validating the introduction of new transport systems to existing ones. Mode choice analysis is the procedure to study the factors and decision making process of the trip maker and to be able to model it (Ch.Ravi Sekhar, et al., 2016).

The fundamental goal of similar researches is to build a mode choice model, which would deliver beneficial information to the policy-makers and transport planners (Manssour A. , et al., 2013). The proposed model will be beneficial to the planners in the sense that, it is responsive to a number of parameters, and consequently, can be used to estimate the impact of changes in the several features or the particular policy changes of the demand for each of modes. Similarly proposed models to Addis Ababa should investigate the travelling modes used in the city and the model can be used in order to test the modal share of public transport alternatives along with observing the need for new public transport options.

A discrete choice model predicts a decision made by an individual (such as mode or route choice) as a function of any number of variables. Dimitry and Viara investigated the influence of a wide range of factors that could affect passenger choice, and estimated their marginal effects. The ways these factors affect passenger choice can be used to improve the services of public transport systems carriers and stations in the city (Dmitry Pavlyuk & Vaira Gromule, 2010).

According to the study conducted by Mintesinot and Shin-ei in 2007, going to work was the main purpose of the trips in Addis Ababa followed by involvement in educational activities (Mintesnot & Shin-ei, 2008). In addition, work trips are the major types of trips that cause peak hour congestion typically in the morning and evening. This research will provide an overview of the impact of variables on the mode preference of the people commuting in Addis Ababa.

1.2 Statement of the problem

Addis Ababa is one of the cities in the world with congested routes that result in longer travel times, lost productivity of travelers, increased accidents and deterioration in air (UNHABITAT, 2017). The travel time of the route from Jemo to Mexico area was initially observed to be 20-30 minutes during off-peak periods and 90-120 minutes during peak periods for large buses. This makes the route one of the congested routes in the city. Moreover there is a huge demand for public transportation services in the city (Ethiopian Ministry of Transport, 2011) in which, in the study area there are longer queues of people waiting for buses during morning and evening peak periods in the weekdays.

The improvement of the existing public transportation systems and provision of new systems in the study area are essential measures to mitigate congestion, longer travel time and other consequences in the study area; and to satisfy the need of the people in the study area for public transportation service. The characteristics of the public transit systems that area to be provided for Jemo-Mexico route should be designed so that the intended purposes in solving congestion problems and serving travelers of the study area are satisfied.

Modal choice models provide the notion to the designers and policy makers on how a particular policy (such as fare reduction or increment) will affect modal share of the available modes in the corridor. However, practically, most of the design of characteristics of public transport systems in

Addis Ababa are not based the modal choice analysis of the target population. The developed models in this study will guide the designers in the determination of the right values for the characteristics of the new BRT system or the existing large bus systems; like fare of service, arrival frequency, station spacing, and travel time.

Moreover the developed modal choice models will be advantageous for public transit enterprises and companies; Anbessa bus, Alliance bus and Sheger bus, to quantify the impact of various factors like modal characteristics and the socio economic characteristics of the trip makers on the individuals' decision making processes in choosing among the available modes.

1.3 Research questions

The following are questions that are envisaged to be answered by this study.

- What are the main factors that affect the travelers' modal preferences?
- How do the above determined factors affect the modal choice of the commuters?
- How will the trip makers behave towards the stated modes which include the BRT system?
- What are the significant factors that are responsive and can be used to increase the modal shares of the public transport systems in the city?

1.4 Objectives of the Research

1.4.1 General Objectives

The general objective of the study is to determine the most efficient options policy makers and designers can consider in the development of new public transportation systems and improvement of the existing systems so that the mode choice towards public transportation system is increased. Generally, the researcher aims to contribute for an improved and sustainable transportation system of the city with the increased number and quality of public transportation systems that people use with better satisfaction.

1.4.2 Specific objectives

The following are specific objectives of the study.

- To determine the factors that affect the preference of means of transportations by people who make work trips from Jemo to Mexico area;

- To develop models that describe the modal choice of trip makers over the current available modes of transport and to investigate which of the factors have statistically significant effect to the modal choice of transportation modes; and
- To develop models that describe the modal choice of trip makers over a stated modes that primarily includes the BRT system, large bus, minibus (taxi) and private car.

1.5 Scope and Limitations of the Study

This study covers the modeling of work trips from Jemo residential area to Mexico downtown area. The trips included in this study are only home-based work trips whose origin is Jemo and destination is Mexico, and starting between 6:00 AM and 9:00 AM.

The modeling of home-based non-work trips and non-home-based trips has received less attention in the urban travel mode choice literature. However, the increasing number of these trips and their contribution to traffic congestion has recently led to development of models for these trip purposes (Koppleman & Chandra, 2006). In this study, the trips in Addis Ababa other than the home-based work trips are not included; therefore, the researcher recommends the development of travel mode choice model for home based non-work trips so that the study will be complete.

The factors like comfort, convenience and reliability of transport modes are necessary in the modal preference of modes, however in this study these factors are not considered. Other limitation of this study is that in the travel cost determination for private cars, only fuel cost and maintenance cost were considered. Other costs like parking cost were not included.

1.6 Organization of the study

This thesis is organized in five chapters. The first chapter is the introduction of the thesis stating the back ground of the study, statement of the problem, the questions aimed at the research responses, general and specific objectives, scope and limitations of the study.

The second chapter deals with review of literatures. In this chapter, starting from transportation planning and modal choice, through modal choice modeling techniques and factors considered in modal choice modeling to the multinomial logit modeling concept, description, and estimation method of parameters are reviewed step by step.

Chapter three is devoted to present the research materials and methods utilized in this study. The research material part reports the type of data required, how it is was sampled and how many samples were required along with the description of the study area and the target group. The research method part emphasizes the parameter estimation methods, evaluation and comparison techniques of the calibrated models, and the validation of the true models.

Chapter four presents the results and discussion of the research. First descriptive analysis of the collected data is summarized in tabular, and graphical forms. Then, the calibration, validation and interpretation of revealed and stated preference models are presented in detail.

Conclusions and recommendations of the study make up chapter five of the report. This chapter finalizes the organization of the study with the summary of the drawn inferences and forwarded recommendations.

CHAPTER 2. LITERATURE REVIEW

2.1 Transportation Planning and Modal Choice

Transportation planning is a process of identifying opportunities to improve transportation systems so that the access and mobility of people and businesses are accelerated. It involves problem definition, search for solutions and performance analysis as well as evaluation and choice of project. It has a huge role in describing the effects of a proposed transportation alternatives and explaining the benefits to the traveler of a new transportation system and its effect to the community (Nicholas & Lester, 1997).

Transport problems have become more widespread and severe than ever in both industrialized and developing countries alike. The general increase in road traffic and transport demand has resulted in congestion, delays, accidents and environmental problems well beyond what has been considered acceptable so far (Ortuzar & Willumsen, 2011). These problems are not likely to disappear in the near future unless efforts are not made in improving most forms of transport, in urban and inter-urban contexts.

One part of transportation planning aspect is the transportation congestion management process which is a way of systematically considering issues related to congestion using technical tools and evaluations based on a discrete set of performance measures (FHWA, 2016). A congestion management process, according to (FHWA, 2016), provides for the systematic review of performance of multimodal transportation systems and identification of strategies to address congestion focused on both the use and operation of facilities and services**.** Addis Ababa is plagued by a continuing increase in traffic congestion resulting in serious consequences (UNHABITAT, 2017), which imply that examining and implementing transportation congestion management policies is essential.

According to (Eluru, et al., 2012), the increase in traffic congestion not only impacts travel delays and stress of drivers, but also adversely affects the environment as a result of greenhouse gas emissions and rising air pollution. An effective means of reducing the over reliance on the auto mode is to encourage public transportation ridership. The importance of quantifying the impact of various exogenous factors such as individual and household socio-demographics, transit level of service measures and accessibility to public transportation on the individual decision making process for public transit agencies was also reported.

Mode choice is that aspect of the demand analysis process that determines the number (or percentage) of trips between zones that are made by the different modes of transportation in the area (Nicholas & Lester, 1997). The choice of transport mode is probably one of the most important classic model stages in transport planning because of the key role played by public transport in policy making (Ortuzar & Willumsen, 2011). Almost without exception, travelling in public transport modes uses road space more efficiently and produce fewer accidents and emissions than using a private car (Ortuzar & Willumsen, 2011). Modelling modal choice is an important part of urban transportation planning in evaluating the effectiveness of Transport Congestion Management policies in shifting single-occupancy vehicle users to high-occupancy vehicle modes (Koppleman & Chandra, 2006). Generally, the issue of mode choice, therefore, is the single most important element in transport planning and policy making.

2.2 Modal Choice Modeling

2.2.1 General overview

Transport models are tools that assist improved decision making and planning in the field of transportation so are the modal choice models (Ortuzar & Willumsen, 2011). Mode choice models generally form a critical part in analyzing the travel demand of a study area. One of the different ways to model modal choice in transportation is discrete choice modelling, which can be used to analyze and predict a decision maker’s choice of one alternative from a finite set of mutually exclusive and collectively exhaustive alternatives (Kohlová & Markéta, 2008). The ultimate interest in discrete choice modeling lies in being able to predict the decision making behavior of a group of individuals (Moshe & Michel, 1999).

It is possible to model the modal choice for more than two alternatives by considering both the travel maker’s and the travel mode’s characteristics. Moreover, considering both socio-economic and travel mode variables will lead to a better understanding and simulation of the people’s behavior in choosing among the competing modes. According to (Mekidim, 2007), multinomial logit modelling technique was employed for work purpose trips made by Bole Sub-City residents, which resulted in satisfactory models describing the modal split of the area.

To understand the behavioral process that leads to the commuter's choice, we take a causal perspective and come up with factors that collectively cause the commuter's choice. Some of these factors are observed by the analyst and some are not. Different discrete choice models are derived under different specifications of the density of unobserved factors, therefore the issues are what distribution is assumed for each model, and what the motivation for these different assumptions is (Kenneth, 2009).

Logit model is by far the most widely used discrete choice model and is derived under the assumption that the factors that affect the utility of alternative i but not included or considered are independently and identically distributed extreme value for all alternatives i (Koppleman & Chandra, 2006). The critical part of the assumption is that the unobserved factors are uncorrelated over alternatives, as well as having the same variance for all alternatives (Kenneth, 2009). This assumption, while restrictive, provides a very convenient form for the choice probability.

2.2.2 Factors affecting the modal choice of trip makers

The factors that are assumed to impact the mode preference of trip makers are used to estimate parameters of a discrete modal choice model. Traveler modal choice is generally a function of three basic factors: characteristics of the journey, e.g., length, time of day, and purpose, the socio-economic characteristics of the traveler, and the modal attributes (Ashu S. Kedia & B.K. Katti, 2014).

The commonly used explanatory variables in mode choice models according to (Koppleman & Chandra, 2006) include:

i. Traveler (Decision Maker) Related Variables
    - Income of traveler or traveler's household,
    - Number of automobiles in traveler's household,
    - Number of workers in traveler's household,
    - Gender of the traveler,
    - Age group of the traveler,
    - Functions of these variables such as number of autos divided by number of workers, and

- Trip Context Variables
  - Trip purpose,
  - Employment density at the traveler's workplace,
  - Population density at the home location, and
  - Dummy variable indicating whether the traveler's workplace is in the Central Business District (CBD).

ii. Mode (Alternative) Related Variables

- Total travel time,
- In-vehicle travel time,
- Out-of-vehicle travel time,
- Walk time,
- Wait time,
- Number of transfers,
- Transit headway, and
- Travel cost

iii. Interaction of Mode and Traveler or Trip Related Variables

- Travel cost divided by household income,
- Travel time or cost interacted with sex or age group of traveler, and
- Out-of-vehicle time divided by total trip distance.

Development of mode choice model for Gaza city by Sadi I. S. Alraee (Essam & Sadi , 2013) considers gender, marital status, occupation, age, family monthly income, family size, means of transport ownership, trip length, means of transport usually used by respondents collected from employees in Gaza city as a socio-economic variables of decision makers and considered the following attributes for private car and taxi modes.

- Private car mode; travel time, average monthly fuel consumption, number of kilometers driven using 1 liter, maintenance cost, license fees, insurance fees, and
- Taxi mode; waiting a taxi, in-vehicle travel time, fare.

Comfort, safety and reliability are the also the types of quality attributes under mode related data requirements and if omitted might result in endogeneity error in the model (Stephane & Andrew, 2010). In addition, if these quality attributes were independent across individuals the endogeneity problems could have been solved by including alternative modal specific constants, however comfort, safety and reliability might depend on the distance traveled by each individual (Stephane & Andrew, 2010).

2.2.3 Disaggregate and aggregate modeling

There are two basic ways of modeling group behaviors as in modal choice; aggregate and disaggregate modeling. The first approach directly models the aggregate share of more than one individual choosing each alternative as a function of the characteristics of the alternatives and socio-demographic attributes of the individuals, whereas the second approach assumes that aggregate behavior is the result of numerous individual decisions and models individual choice responses as a function of socio-demographic attributes of each individual and the characteristics of the alternatives available (Minal & Ch. Ravi , 2014).

The disaggregate approach has several important advantages over the aggregate approach to modeling the decision making behavior of a group of individuals. According to (Koppleman & Chandra, 2006), disaggregate data are better able to reflect changes in choice behavior, due provide substantial variation enabling the efficient estimation of model parameters, if properly specified, disaggregate modeling will obtain un-biased parameter estimates and it is likely to be more transferable to a different point in time and to a different geographic context.

Disaggregate models offer substantial advantages over the traditional aggregate models, however, they demand quite a high level of econometric and statistical skills for their use, which is much higher than the case of aggregate models (Ortuzar & Willumsen, 2011).

2.2.4 Stated preference and revealed preference

There are two types of modelling approaches used to model travel demand and travelers' behavior; revealed preference and stated preference survey. These two modelling techniques are able to provide data for the development of disaggregate travel demand models, including modal choice models.

Revealed Choice data describes current observed travel conditions and costs, which has practical limitations in expressing policies in terms of changes in attributes, in order to evaluate projects

(Ortuzar & Willumsen, 2011). The stated preference techniques allow the deficiencies of revealed preference data to be overcome by testing hypothetical transport alternatives using interviews. 'If you are faced to these situations, what will be your preferred mode?' kind of questions are involved in stated preference survey (Essam & Sadi , 2013), and it becomes an attractive option in transportation modeling since it presents the decision-makers choice and behavioral pattern under different hypothetical scenarios.

In a research conducted for Yangon city (Akimasa, et al., 2003), people's preferences for the selected mode choice has been stated for several hypothetical situations with different levels of service including travel time, travel cost, waiting time and punctuality are asked after the stated preference questions. Essam Almasri and Sadi Alraee (Essam & Sadi , 2013) stated that a type of questionnaire that can be used in stated preference studies is a choice questionnaire, in which the respondent simply chooses the hypothetical combination of attributes that is most favorable and the researcher can actually predict the respondent's choice in the hypothetical situation.

2.2.5 Different methods used to model modal choice

There are two general methods of modeling transportation modal choices of people in urban areas; statistical mode choices (discrete choice models) and the use of soft computing techniques. The statistical mode choice method include; multinomial logit (MNL), nested logit (NL), multinomial Probit (MNP), and generalized extreme value (GEV) models. Multinomial logit model is the simplest and easiest discrete choice model developed under the following assumptions; the error components are extreme-value (or Gumbel) distributed, the error components are identically and independently distributed across alternatives, and the error components are identically and independently distributed across observations/individuals (Koppleman & Chandra, 2006).

The derivation of the nested logit model is based on the assumption that the random component is identically, non-independently distributed with type-1 extreme value distribution and it allows the estimation of proportions among selected modes prior to the estimation of proportions between modes (Minal & Ch. Ravi , 2014)

Generalized extreme value models relax the assumption that the error component is independently distributed across alternatives allowing the random components of the alternatives to be correlated.

Models that belong to this class result in closed form expression of the choice probabilities which makes estimation possible without estimation (Chandra R.Bhat, 2000).

The normal distribution assumption for error terms leads to the MNP Model which has some properties that make it difficult to use in choice analysis. When applied to sequences of choices over time, the unobserved factors are assumed to be jointly normal over time as well as over alternatives, with any temporal correlation pattern. The flexibility of the Probit model in handling correlations over alternatives and time is its main advantage. Its only functional limitation arises from its reliance on the normal distribution. In some situations, unobserved factors may not be normally distributed (Kenneth, 2009).

The most recent modeling techniques that researchers developed are with the use of soft computing techniques; artificial neural network (ANN), fuzzy logic and hybrid models which are the combinations of ANN and fuzzy logic. ANN is the important component of intelligent algorithm which is famous for its powerful information processing ability. Neural network provides the capability without the complication of having to formalize all the complex causal variables and relationships which other models (Juan Li, et al., 2010). Generally, the major constraints, drawbacks and levels of accuracy of the different modal choice modelling techniques are as summarized in the Table 2-1 below.

Table 2-1: Comparison of various mode choice models.

Model Name	Basic Hypothesis	Major constraints	Major Drawbacks	Accuracy
Logit Model	Extreme value distribution (type I GEV)	Error term should be identically and independently distributed (IID)	Suffers from IIA assumption	High
Probit Model	Normal distribution	Error term is non-identically and non-independently distributed.	Difficult interpretation of results, computational problems	Low
GEV Model	Multivariate extreme value distribution	Independent, but not identically distributed error terms.	Perceived utility of any alternative should not exceed an upper bound	Low
ANN Model	Inspired by human neural system.	Performance depends on architecture, training and activation function.	Difficult to interpret the result in terms of natural language	High
Fuzzy Logic Model	Based on underlying assumption of commensurability	Input values should be fuzzy	No learning ability and difficulty in deciding membership function (Tuning).	High
Neuro Fuzzy Model	ANN and FIS models complement each other.	Pre-processing of input training data is required	Lack of common framework, different NF models can't be Compared	High

Source 2-1: (Minal & Ch. Ravi , 2014)

2.3 Multinomial Logit Modelling

2.3.1 General overview

The multinomial logit model is by far the most widely used discrete choice model due to the following advantages of the Gumbel distribution, as stated in (Koppleman & Chandra, 2006).

i. Gumbel distribution has computational advantages in a context where maximization is important,
ii. It closely approximates the normal distribution as can be seen in Figures 2-1 and 2-2 below; and
iii. It produces closed form probabilistic choice model; for which the probability can be calculated without the use of numerical integration or simulation method.

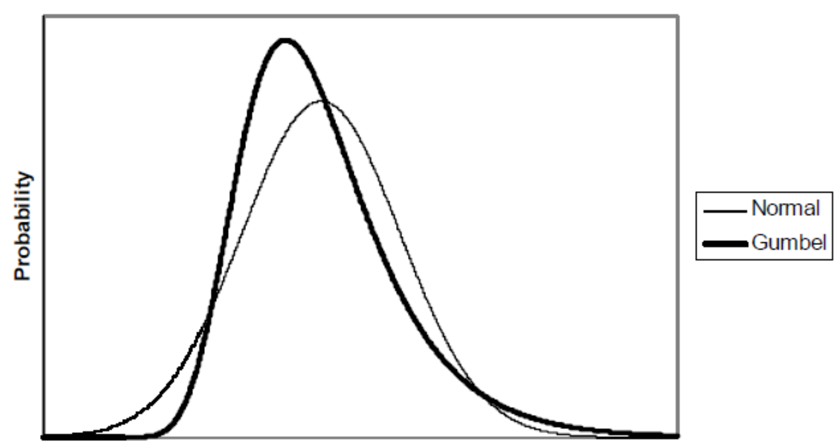

Figure 2-1: Probability density function for Gumbel and normal distributions

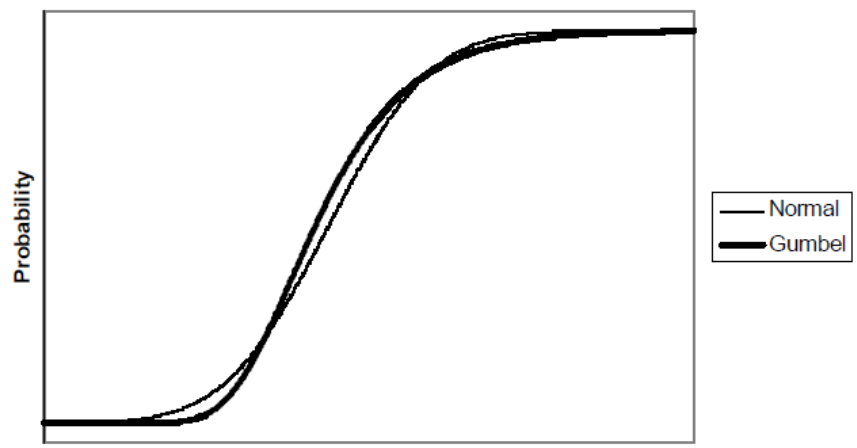

Figure 2-2: Cumulative distribution function for Gumbel and normal distribution

Moreover the assumption that the unobserved factors are uncorrelated over other alternatives as well as having the same variance for all alternatives, is restrictive but provides a convenient form for the choice probability (Kenneth, 2009).

Multinomial logit model assumes independence of irrelevant alternatives (IIA) which can be stated 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 (R.D. Luce, et al., 1965). Initially this was considered an advantage of the model, as it allows us to treat quite neatly the new alternative problem: however, nowadays this property is perceived as a potentially serious disadvantage which makes the model fail in the presence of correlated alternatives (the red bus/ blue bus paradox) (Ortuzar & Willumsen, 2011).

2.3.2 Utility maximization theory

The utility maximization theory states that a trip maker will select the transport mode from his/her set of available modes which maximizes his or her utility. The utility function, U, has the property that an alternative is chosen if its utility is greater than the utility of all other alternatives in the individual's choice set (Essam & Sadi , 2013).

There are two different choice concepts; deterministic and probabilistic choice concepts. In deterministic utility models the utility maximization rule implies no uncertainty in the individual's decision process; that is, the individual is certain to choose the highest ranked alternative under the observed choice conditions. Probabilistic choice theory arise because analysts do not have the knowledge; they do not fully understand the decision process and they have no realistic possibility of obtaining this information (Koppleman & Chandra, 2006).

As with deterministic choice theory, the individual is assumed to choose an alternative if its utility is greater than that of any other alternative. The probability prediction of the analyst results from differences between the estimated utility values and the utility values used by the traveler. This difference is represented by two decomposed terms of the utility of the alternative. One component represents the utility observed by the analyst, called the deterministic utility and the other component is known as random error (Minal & Ch. Ravi , 2014). Formally, this is represented as:

$$U_{it} = V_{it} + \varepsilon_{it} \tag{2-1}$$

Where, U_{it} is the true utility of the alternative I to the decision maker t,

V_{it} is the deterministic or observable portion of the utility estimated by the analyst, and

ε_{it} is the error or the portion of the utility unknown to the analyst

The total error which is the sum of errors from many sources is represented by a random variable. Different assumptions about the distribution of the random variables associated with the utility of each alternative result in different classes of statistical modeling methods described in Section 2.2.5.

Components of the deterministic portion of the utility function

The deterministic or observable portion of the utility of an alternative is a mathematical function of the attributes of the alternative and the characteristics of the decision maker. The systematic portion of utility can have any mathematical form but the function is most generally formulated as additive to simplify the estimation process (Koppleman & Chandra, 2006). This function includes unknown parameters which are estimated in the modeling process.

$$V_{t,i} = V(S_t) + V(X_i) + V(S_t, X_i) \qquad (2\text{-}2)$$

Where, V_{ti} is the systematic portion of utility of mode i for trip maker t,

$V(S_t)$ is the portion of utility associated with characteristics of trip maker t,

$V(X_i)$ is the portion of utility of mode i associated with the attributes of mode i, and

$V(S_t, X_i)$ is the portion of the utility which results from interactions between the attributes of mode i and the characteristics of trip maker t.

2.3.3 Description and functional form of MNL modeling

The three assumptions that lead to MNL stated in section 2.2.5, gives the choice probabilities of each available modes as a function of the systematic portion of the utility of all the modes. The general expression for the probability of choosing a mode ‘i’ (i = 1, 2… j) from a set of J available modes, based on (Moshe & Michel, 1999) is:

$$Pr(i) = \frac{exp(V_i))}{\sum_{j=1}^{J} exp(V_j)} \quad (2\text{-}3)$$

Where $Pr(i)$ is the probability of the decision-maker choosing alternative I, and

V_j is the systematic component of the utility of alternative j.

This formulation implies that the probability of choosing an alternative increases monotonically with an increase in the systematic utility of that alternative and decreases with increases in the systematic utility of each of the other alternatives.

The property of independence of irrelevant alternative/ IIA property of logit model may not properly reflect the behavioral relationships among groups of alternatives, like in the classic “red bus/blue bus paradox.” However, this property of MNL allows the addition or removal of an alternative from the choice set without affecting the structure or parameters of the model. Moreover this property allows modeling in cases where members of the population face different sets of alternatives, simplifies the estimation of the parameters, and is advantageous when applying a model to the prediction of choice probabilities for anew alternative (Ortuzar & Willumsen, 2011).

The multinomial logit model is chosen for this study, also because its error distribution is Gumbel or extreme value which has computational advantages where maximization is important, closely approximates the normal distribution and produces a closed form probabilistic choice model which can be calculated without integration or simulation it produces.

Gumbel or extreme value distribution for each unobserved component of the utility has the following density and cumulative distribution functions, for a traveler n having j alternatives. (Kenneth, 2009)

$$f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}} \quad (2\text{-}4) \qquad F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}} \quad (2\text{-}5)$$

The graphic representation of the gumbel distribution and the cumulative gumbel distribution of the error term is as shown in the Figures 2-1 and 2-2. The difference between two extreme value variables is distributed logistic, i.e , if ε_{nj} and ε_{ni} are iid extreme value, then $\boldsymbol{\varepsilon^*_{nji}} = \varepsilon_{nj} - \varepsilon_{ni}$, follows the logistic distribution,

$$F\left(\varepsilon_{nji}^{*}\right) = \frac{e^{\varepsilon_{nji}^{*}}}{1+e^{\varepsilon_{nji}^{*}}} \tag{2-6}$$

From McFadden's law, the probability that decision maker n chooses alternative i is,

$$P_{ni} = Prob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \,\forall j \neq i$$
$$= Prob(\varepsilon_{nj} < \varepsilon_{ni} + V_{ni} - V_{nj} \forall\, j \neq i \tag{2-7}$$

If ε_{ni} is considered given, this expression is the cumulative distribution for each ε_{nj} evaluated at ε_{ni} + V_{ni} − V_{nj} , which is exp(−exp(−(at ε_{ni} + V_{ni} − V_{nj}))). Since the ε's are independent, this cumulative distribution over all j different from i is the product of the individual cumulative distributions:

$$P_{ni} \mid \varepsilon_{ni} = \prod_{j \neq i} e^{-e^{-\left(\varepsilon_{ni}+V_{ni}-V_{nj}\right)}} \tag{2-8}$$

Since ε_{ni} is not given, and so the choice probability is the integral of $P_{ni} / \varepsilon_{ni}$ over all values of ε_{ni} weighted by its density,

$$P_{ni} = \int \left(\prod_{j\neq i} e^{-e^{-\left(\varepsilon_{ni}+V_{ni}-V_{nj}\right)}}\right) e^{-\varepsilon_{ni}}\, e^{-e^{-\varepsilon_{ni}}}\, d\varepsilon_{ni} \tag{2-9}$$

Algebraic manipulation of the integral finally resulted in the following closed form expression of the probability statement result of multinomial logit model.

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \tag{2-10}$$

2.3.4 Parameter estimation method

MNL model development consists of, first determining model specifications and then estimating the parameters of the different variables specified in each utility function by fitting the models to the collected data using different methods. Least squares method and maximum likelihood methods are the two widely known parameter estimation techniques, maximum likelihood and least squares estimation methods.

Maximum likelihood estimation methods

The procedure for maximum likelihood estimation involves two important steps: developing a joint probability density function of the observed sample, called the likelihood function, and estimating parameter values which maximize the likelihood function (Omer , et al., 2004). The likelihood function for a sample of 'T' individuals, each with 'J' alternatives is defined as follows:

$$L(\beta) = \prod_{\forall t \in T} \prod_{\forall j \in J} \left(P_{jt}(\beta)\right)^{\delta_{jt}} \quad (2\text{-}11)$$

Where, $\delta_{jt} = 1$ is chosen indicator (=1 if j is chosen by individual t and 0, otherwise) and

P_{jt} is the probability that individual t chooses alternative j.

The numerical values of the parameters which maximize the likelihood function are determined by finding the first derivative of the likelihood function and equating it to zero. Since the log of a function yields the same maximum as the function and is more convenient to differentiate, we maximize the log-likelihood function instead of the likelihood function itself. In most practical problems, this involves significant mathematical computations and specialized statistical software to find the desired solution (Koppleman & Chandra, 2006).

Least squares estimation methods

This method, usually known as ordinary least square method, is based on determining values which minimize the sum of squared differences between the observed and predicted values of the observations. However it is reported that the least square method works consistently and efficiently for linear models only, and may result in erroneous coefficient values if used for the estimation of complex models (Essam & Sadi , 2013). Thus, the maximum likelihood method is found to be practical for the estimation of modal choice models.

2.3.5 General overview Conditional and multinomial logit modelling

Conditional logit models are a special type of multinomial logit model which are appropriate when the choice among alternatives is modeled as a function of the characteristics of the alternatives, rather than (or in addition to) the characteristics of the individual making the choice. The conditional logit model is widely used in transportation demand studies. However, its main disadvantage is that the effects of all individual and the socio-demographic characteristics of the trip makers are expressed in terms of interactions (D. Hoffman & J. Duncan, 1988).

The conventional MNL regression only considers factors that vary with the trip maker known as case-specific or alternative-invariant variables whereas the conditional logit regression considers both case-specific and “alternative-specific” variables. Most of the statistical packages consider the multinomial logistic regression as a regression of dependent variable with case-specific variables or decision-maker specific variables. This is the situation in which case-specific variables are typically observed for the chosen alternative, not for other alternatives (Cameron & Trivedi , 2009).

In modal choice modeling, case-specific variables are the socio-economic characteristics of the trip maker such as Gender, Age and income. The choice process of the modal choice is through the utility maximization of the modes in which all related variables has to be observed for all available alternatives of each trip maker, which includes the alternative specific variables (variables which have different values for different alternatives like travel time and travel cost of travel modes). Therefore conditional logit modeling is the type of MNL that suits modal choice modeling best.

The dataset for conditional logit modeling of modal choice which can be shown in Table 3-2, makes it easy to use it in software packages that can perform conditional logit parameter estimation. SPSS performs conditional logit regression with a command called COXREG.

In Stata the Clogit command is used to regress a conditional logit model which is designed for grouped data used in matched case-control group studies and is similar to the Xtlogit command used for panel data grouped over time for an individual. Similarly in SPSS, fitting models with multiple controls is possible using COXREG procedure. Since these commands does not have an option for case-specific variables, those variables are entered interacted with dummies for m-1 alternatives, and the m-1 variables are entered as regressors (David & Stanley, 2000).

In most literatures the term multinomial modeling is used instead of conditional logit modeling. In this study also, all the terms “multinomial logit modeling” doesn’t express the conventional logit modeling, instead it expresses the conditional logit modeling.

2.4 Methods of testing, comparison and validation of models

There are different ways to review the results for the estimated models using informal judgment-based tests, goodness-of-fit measures, and statistical tests. These elements, taken together, provide a basis to evaluate each model and to compare models with different specifications (Koppleman & Chandra, 2006).

2.4.1 Method of testing and comparison of models

Checking signs of parameters

Variety of informal tests are designed to examine the reasonableness of the results of estimated parameters. The most common test includes test of the estimation results to assess the signs of the estimated parameters with theory, intuition and judgment based on the expected impact of the corresponding variables (Essam & Sadi , 2013). For example, it is logical for signs of travel time and travel cost parameters to be negative because the utility of modes will decrease with an increase of travel time and travel cost.

Over all goodness of fit measures

Rho-squared value (ρ^2): this is similar to the R^2 measure for linear regression models. The value of rho-squared measurements lie between 0 and 1. A value of zero implies that the model is no better than the reference model; whereas, a value of one implies a perfect model; that is, every choice is predicted correctly (Koppleman & Chandra, 2006).

i. $\rho_0^2 = 1 - \left(\frac{LL(\beta)}{LL(0)}\right)$; measures how explanatory the model is relative to a model with all the parameters zero where LL(β) is the log likelihood of the estimated model or the model at convergence and LL(0) is the log likelihood when all the parameters are zero.

ii. $\rho_c^2 = 1 - \left(\frac{LL(\beta)}{LL(c)}\right)$; measures how the estimated model explains the preference of modes relative to the constants only or intercept model. LL(c) is the log likelihood of the constants only model.

Basically, this rho-squared value measures the explanatory power of the discrete choice model which includes behavioral assumptions. However, all the relevant predictors cannot be included; as a result most of the data generated will have lower rho-squared values. Rho-squared even with small values can be significantly different from zero, implying that the model has statistically significant explanatory value.

The basic problem with these measures is that they improve no matter what variable is added to the model independent of its importance. Replacing the rho-squared measure with an adjusted measure which takes into account the degree of freedom of the model is one approach to such problem (Koppleman & Chandra, 2006).

iii. $\rho_0'^2 = 1 - \left(\frac{LL(\beta)-K}{LL(0)}\right)$; is the adjusted rho-squared measure with respect to the zero model where K is the degree of freedom or number of parameters in the estimated model.

iv. $\rho_c'^2 = 1 - \left(\frac{LL(\beta)-K}{LL(c)-K'}\right)$; is the adjusted rho-squared measure with respect to the constants only model where K' is the degree of freedom or number of parameters for the constants only model (Koppleman & Chandra, 2006).

Tests for the individual parameters

The null hypothesis in most software packages are that the parameter is equal to zero. In cox regression, too, Wald statistics is used to test the null hypothesis that a parameter is zero by measuring the statistical significance of the parameter, with the following formula.

$$Wald\ statistic = \frac{\beta}{S.E.} \tag{2-12}$$

Where β is the covariate regression coefficient (parameter) estimated value and S.E is the standard error associated with the parameter. SPSS results in a P-value for the null hypothesis that the parameter is zero. For 95% confidence level, if the P-value is less than 0.05, the null hypothesis is rejected which means that the estimated covariate significantly improves the fit model. P-values greater than 0.05, provided 95% of level of confidence, imply that the variable does not contribute significantly to the explanatory power of the model and can be considered for exclusion (Essam & Sadi , 2013).

Tests for the entire model

a. Log likelihood ratio test: nested hypothesis test

The likelihood ratio test is used to compare two models provided that one is the restricted version of the other by formulating a test statistics. If all the restrictions that distinguish between the restricted and unrestricted models are valid, it is expected for the difference in log-likelihood values at convergence of the two models to be small. If some or all of the restrictions are invalid, the difference in log-likelihood values of the two models will be sufficiently large to reject the hypothesis (Koppleman & Chandra, 2006). The test statistics is;

$$\chi^2 = -2 * (LL_R - LL_U) \tag{2-13}$$

Where LL_R is the log likelihood of the restricted model at convergence and LL_u is the log likelihood of the unrestricted model at convergence. The test statistics is chi-squared distributed with critical chi-squared (χ^2) values for different level of confidence by number of restrictions, which can be referred in Appendix-B.

b. Non-nested hypothesis test

This model test is used to compare two rival models that do not have those types of restricted and unrestricted type of relationships as in log likelihood ratio tests, and with equal numbers of parameters. This non-tested hypothesis test uses the adjusted likelihood ratio index, $\rho_0'^2$, to set the hypothesis that the model with the lower adjusted rho squared value is the true model.

The null hypothesis that the model with the lower $\rho_0'^2$ value is the true model is rejected at the significance level determined by the following equation (Koppleman & Chandra, 2006). The null hypothesis will be rejected if the calculated significance level is less than 0.05 for 95% confidence level.

$$Significance\ level\ = \Phi\left[-\left(-2 * \ (\rho_H'^2 - \rho_L'^2) * \ LL(0) + \ (K_H - \ K_L)\right)^{0.5}\right] \tag{2-14}$$

Where; $\rho_H'^2$ and $\rho_L'^2$ are the adjusted likelihood ratio index at zero with the higher and lower value,

LL(0) is the log likelihood of the models at zero

K_H and K_L are the numbers of parameters in the models with higher and lower $\rho_0'^2$ value, respectively.

Φ is the standard normal cumulative distribution function.

2.4.2 General guidelines

In addition to the tests and different techniques used to compare developed models, the following three basic guidelines are followed in determining the true model.

- Any specifications which result in coefficients that are inconsistent with our understanding of travel behavior should be rejected (Essam & Sadi , 2013). For example, any specification which results in a positive sign for travel time or travel cost should be rejected.
- Variables which are statistically insignificant are dropped out of the model except for the alternative specific constants. The lack of significance of the alternative specific constants is immaterial since the constants represent the average effect of all the variables not included in the model and should always be retained despite the fact that they do not have a well understood behavioral interpretation (Koppleman & Chandra, 2006).
- Moreover, if it is found reasonable to retain variables such as the level of service (travel time variables) or policy variables (travel cost) in the model, those variables will be kept even if it is statistically insignificant (Ortuzar & Willumsen, 2011).

2.4.3 Methods of model validation

Evaluation of the performance of the calibrated model and its ability to predict modal split for data other than used for calibrating the model is one of the important stages of modeling which can be executed with different model validation methods. There will be samples prepared for both fitting and validation of models; the two samples shall be samples from the same population but at the same time, are distinct and independent from one another. With the fitting sample we are evaluating how well the model fits the data on which it was developed and for the validation sample we evaluate how valid the model is as it attempts to predict an independent sample (Rosa & Luigi , 2003).

One of the methods of validation is the test of reasonableness of the signs of parameters in the model, as described in section 2.4.1. At this phase, models with opposite signs of parameters from the expected sign are to be dropped. Usually this phase is completed during comparison of models in the calibration process.

The most important method of validation is the assessment of the prediction capability of the model (Essam & Sadi , 2013). The data that was reserved for validation (one third or one fourth of the data used for calibration of the models) will be used to predict the mode choice of the trip makers using the calibrated model. First, the utility of all the available modes for each trip maker will be computed. Then it is assumed that a trip maker will choose a mode with highest utility and the predicted chosen models are determined. Finally, the prediction ratio is computed by dividing the number of modes that are correctly predicted by the total number of validation data.

2.5 Review of similar case studies

Different researchers around the world have contributed for the state of the art of the modeling techniques of modal preference. In this section development of models of modal choice for different types of trips in different cities are reviewed starting from those that are done for trips in Ethiopia.

- (Mekidim, 2007) developed logit model of work trips for Addis Ababa Bole-sub city residents using STATA. The independent variables in the final optimal model include total travel time, travel cost, gender, age, child in the house or not and peak hour dummy variable. The validation of the model estimated on the data at which the model was calibrated was found to be 100%.
- An ordered logit model was developed to examine travelers' perceptions of the bus condition as determining factor for the choice of transportation of trips in Addis Ababa. In addition, to analyze traveler choice behavior (Mintesnot & Shin-ei, 2008) developed a binary logit model. According to this work trip makers' perception of fare, convenience, and frequency of a bus have significant influence on public transport mode choice in Addis Ababa.
- The MNL model developed in Addis Ababa for the work and school work trips along the railway route corridor, based on (Etsub , 2015), considered 183 and 100 trip makers were considered for the work and school trips respectively. Job category, travel time, income, out of vehicle travel time and gender are the factors that are found to be statistically significant to regress the models. In this study both revealed and stated preference models which considers the light rail service were included.
- A study in Kolkata, India, (Ram, et al., 2016), aims in determining weight vector of different attributes, developing utility function and modal share for a corridor in Kolkata city of India using logit modeling techniques. The factors included in this study are travel cost, travel time, convenience, comfort, reliability, and dust and noise. The weighted vectors determined based on perceptions of trip makers determined the importance of each

attribute. The most relevant factors used in the utility functions total travel time, total travel cost and waiting time of the modes. Finally based on the utility functions they were able to determine the modal sharing of the modes for work trips in that corridor.

- Most of the developed modal choice models in different cities of the world use MNL modeling technique. One of the reason for this is that it is simple in terms of formulation and estimation in addition to the accuracy of the outcomes (Essam & Sadi , 2013). Similar models has been developed for Gaza city (Essam & Sadi , 2013), for a revealed and stated preference surveys using the ELM software package. The prediction capacities for the revealed and stated preference models is 69% and 80%, respectively.
- With 395 observations (Amit, 2017) used MNL regression methods through XLSTAT software to model modal preference for school trips of college aged students in Panipat city, India. Age, department, residential area, gender and family class are the factors considered while developing the modeling.

CHAPTER 3. RESEARCH METHODS AND MATERIALS

3.1 Research materials

3.1.1 Data requirements and Data Collection Methods

There are many factors that affect the choice of the people among the available modes of transportation in the city. The following are some of the relevant variables chosen for these study according to related study results. The data are classified into three categories and listed as shown below.

I. **Traveler related data**;
   - Gender,
   - Age,
   - Marital status,
   - Household monthly income,
   - House hold size,
   - Car ownership,
   - Education level, and
   - Occupation category.

II. **Trip related data**;
   - Total trip length or distance/ trip origin and destination,
   - Mode usually used by traveler,
   - Out of vehicle time; access time, waiting time, egress time,
   - In vehicle travel time, and
   - Travel cost (travel cost/ km) including paid for other access mode.
   - Time of leaving home for work

The traveler and trip related data are to be collected by the use of questionnaire and interview at the home of the respondent.

III. **Mode related data**;
   - Travel cost,
   - Travel time for each trip length/ speed, and

- Waiting time/ Transit frequency or headways.

These data were collected for the considered modes in this model; private car, large buses and minibus (taxi), from the collected questionnaire and secondary data from transportation sector offices. The secondary mode related data of Bus Rapid Transit system are collected from Addis Ababa Transport Programs Management Office which is in control of the design and construction of the system.

Travel cost of a private car alternative is determined as a summation of fuel and total maintenance expense. Fuel expense of a trip is determined as a product of fuel efficiency of the vehicle and the length of the trip in kilometers. Total maintenance cost is the summation of maintenance cost and service cost which are paid approximately every month and every 5000 KMs, respectively. The computation is shown with an example in section 3.1.3.

3.1.2 Study area and target group

The first issue to consider in sample design is the population of interest and the study area in the study. The longer travel times during the peak hours (90 to 120 minutes during the peak periods and 20-40 minutes during non-off peak periods) and the queue of vehicles during peak periods shows that the route from Jemo to Mexico is one of the congested routes in the city. Moreover bus rapid transit service is under design and construction in that route for the first time in Addis Ababa and Ethiopia.

The area is under Nifas-Silk-Lafto sub city and Woreda 1 Jemo administration. The Woreda 1 Jemo administration encompasses 337 blocks under Jemo 1 site, 186 blocks under Jemo 2 site, 187 blocks under Jemo 3 site, each block containing average of 25 households, and 16,000 villa households. Moreover, one trip per household is assumed for the households in the condominium houses. However during the data collection it was found that the type of trips in this study were almost 1 out of 5 households for the case of villa houses. Therefore the total villa households that have target trip makers are assumed to be 5000. This will make up a total of 22,750 households. The google map of the target population house hold area is shown in Figure 3-1, below.

Figure 3-1: Map of the household area of the target population of Woreda 1 Jemo.

3.1.3 Alternative transport modes in the study area

The available alternatives on the route from Jemo to Mexico include private vehicles, Taxi or minibus, midi-buses (Higer or Lonchinas), large buses, such as Anbessa bus, Sheger bus and Alliance bus, governmental and non-governmental company service vehicles, and BRT system in the stated preferences. People who use company services as their mode don't comply with utility maximization theory because they use service vehicles just because they are available for free or the transport allowance is deducted from their salary in advance. In addition, during the pilot survey, midi-buses (Higer and Lonchinas) were found to be unavailable between 6:00 AM and 9:00AM. Hence midi-buses were removed from the choice set. Therefore, the choice set will only include; large bus, minibus (taxi), private car and BRT system.

Large Bus

There are three different public bus transport enterprises that provide transport services in the study corridor; Anbessa bus, Sheger Bus and Alliance bus. It was observed on site that Anbessa buses along that corridor are of three types; double decker, mid floor and articulated or harmonica buses. Whereas, Sheger buses are of double decker and flat floor bus types, and Alliance buses are of mid floor bus type only. The double decker buses shown in Figure 3-2, below, have a capacity of 80 seated and 60 standing passengers, the articulated or harmonica buses have capacity of transporting

300 passengers (Anon., 2016). The mid and low floor buses shown in Figure 3-4, and Figure 3-3 respectively, have capacity of 100 and 130 people, respectively.

Figure 3-2: Anbessa Double-Decker bus

The cost of traveling with these three different buses is similar from the origin to destination of the study trip. Hence, these three different buses were merged as one in this study, and they are named as large buses.

Figure 3-3: Sheger low floor buses

Figure 3-4: Alliance mid floor buses

Minibus (taxi)

In this study, the name taxi is used to refer the minibuses shown in Figure 3-5, below, with carrying capacity of 12 passengers seated for normal minibuses and 15 passengers seated for Dolphin cars. The traveling cost of this mode is a higher than that of large buses which makes it compulsory to consider these two modes independently.

Figure 3-5: Minibuses or taxis

Private car

Private cars are the third class of transport modes in this study in which the trip maker drives himself/herself to work with his own car. The alternative ‘private car’ expresses both situations in which the trip maker is a driver or a passenger in a private vehicle. Even though, the data collection included private car users as passenger as an independent mode, it was found that trip makers who travel with private car as passenger are small in numbers (6 out of 475 samples) compared to other modes and were excluded from the study. Therefore, in this study the alternative ‘private car’ refers private car user driving him/herself.

The travel cost of private car includes only fuel cost and maintenance cost as stated in section 3.1.1. Travel cost calculation of a trip maker who makes an approximated total trip of 60 work trips per month and whose home is 9.5 KMs away from Mexico, whose car with fuel efficiency of

10 KM/L uses gasoline fuel and is maintained 400 ETB/month and serviced 1000 ETB every 5000 KMs; is given below.

Travel Cost = Fuel Cost + Total Maintenance Cost

$$= \frac{18.77\ \mathrm{ETB/L} * 9.5\ \mathrm{KM}}{10\ \mathrm{KM/L}} + 9.5\ \mathrm{KM} * \frac{1000\ \mathrm{ETB}}{5000\ \mathrm{KM}} + \frac{400\ \mathrm{ETB/Month}}{60\ \mathrm{trips/month}}$$

$$= 26.40\ \mathrm{ETB/trip}$$

BRT

The bus rapid transit system in this study corridor, which is the first of its kind in Addis Ababa and also in Ethiopia, that operates with high speed buses on back and forth lanes restricted only to those buses, as shown in Figure 3-6. Based on the data collected, the in vehicle travel time of BRT is expected to vary from 25 minutes to 35 minutes and the travel cost will comply with large buses which are currently giving service.

The operation manual of the BRT line, B2 corridor pilot project states that the normal passengers carried per vehicle per day for buses with a capacity of 80-100 passengers is between 1000 and 2000. Whereas, for articulated buses with a capacity of 150-175 passengers a normal range would be between 1800 and 3000 passengers carried per day. This is based on the assumption that 85% of the fleet is operational.

The arrival frequency of the BRT stated in the feasibility and operational planning study is 3 minutes (peak), 6 minutes (off-peak) and 12 minutes (late evening) for the inner routes. For the outer end sections, and all day frequency of 6 minutes is stated. However, in this research a general frequency of 5 minutes was assumed. Moreover, the stations are typically to be located approximately one kilometer apart and the stations are located in the center of the road for the inner routes.

The basic bus design options in BRT that the feasibility and operational planning study considers are;

i. Standard/Rigid bus, 12m length, 60-80 passengers capacity,
ii. Articulated bus, 18m length, 120-170 passengers capacity, and
iii. Bi-articulated bus, 21m length, 240-270 passengers capacity.

During the data collection to show how the bus rapid transit system works for the respondents Figure 3-6 was inserted in the questionnaire paper. The modal characteristics of the BRT in the study corridor was collected from the Addis Ababa city transport programs management office. The office is responsible of design, construction and operation of the BRT.

Figure 3-6: Graphical plan of one of the station in the BRT corridor

3.1.4 The home-based work trip in this study

This research basically studies the mode preferences of work trips during the rush hours; target population being people traveling to work during the rush hour from Jemo to Mexico which is considered between 6:00 AM and 9:00 AM. Large buses, taxis and private cars use the route through Sarbet and Mekenisa, as shown in Figure 3-7, while the BRT, when it starts giving service, follows the route from Jemo to Mexico through Gofa and Kera areas, as shown in Figure 3-8.

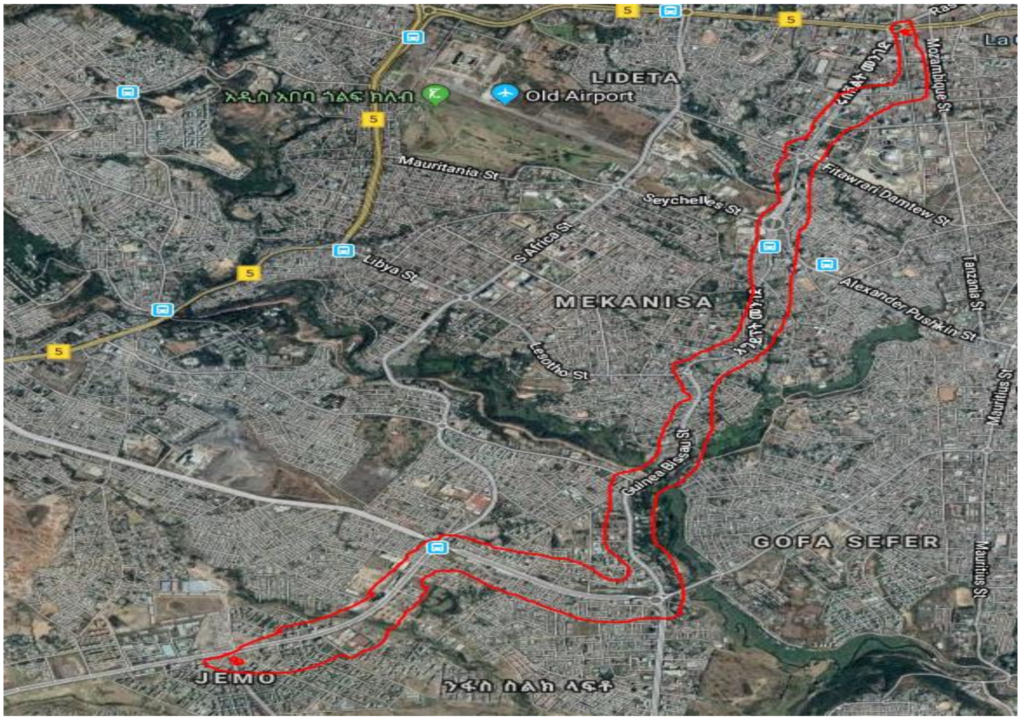

Figure 3-7: Map of route between Jemo and Mexico Square through Sarbet and Mekenisa.

Figure 3-8: Map of route for the BRT system between Jemo and Mexico Square

3.1.5 Design of Questionnaire

The questionnaire was developed based on the most relevant traveler, mode and trip attributes listed in the data requirements in section 3.1.1. Before the collection of the whole data, pilot survey was done and some questions of the questionnaire were modified and some were excluded. The questionnaire that was finally distributed to respondents can be referred in the appendix part. The questionnaire is generally presented in three parts as detailed below.

Part 1; The questions under this part include the social and economic information of the respondent such as gender, education level, type of occupation, household monthly income, family size, and car ownership.

Part 2; This part of the questionnaire includes trip related variables related to mode choice such as; trip length, mode of transportation usually used and mode related variables such as in vehicle travel time, out of vehicle travel time, walk time, waiting time or arrival frequency and travel cost.

Part 3; The last part of the questions include the preference of the respondents for a new mode of transport which is going to start giving service on the study route with different travel time, travel cost, and arrival frequency.

3.1.6 Sample size

The next issue in survey sampling was the sample size to be used. The size of the sample required to adequately represent the population is a function of the level of statistical accuracy and confidence desired from the survey. The precision of parameter estimates and the statistical validity of estimation results improve with sample size. However, the cost and time of the survey also increases with sample size and in many cases it is necessary to restrict the sample size to ensure that the cost and time of the survey remains within budgetary and time constraints (Koppleman & Chandra, 2006).

The following formula is usually used to determine the sample size in different research areas with population sizes greater than 100,000 (Glenn , 1992).

$$n = \frac{Z^2 * p(1-p)}{e^2} \tag{3-1}$$

Where; n is the minimum sample size, Z is the z-score for desired level of confidence, p is the estimated proportion of an attribute that is present in the population, and e is the desired level of precision. For 95% level of confidence Z = 1.96, maximum variability p = 0.5 and e = 5% level of precision or sampling error, the minimum sample size is calculated to be 385.

On the other hand, based on (Green, 1991), the minimum sample size recommended for multinomial logistic regression is,

$$N = 50 + 8 * m \tag{3-2}$$

Where m is the number of factors considered in the model. Therefore for maximum number of 30 factors to be considered in this study, the minimum sample size is 290.

The total sample size required will include the data for validation of the model which will be one fourth of the minimum sample size. A response rate of 95% can be acceptable since the questionnaires were going to be collected through an interview at the households of the respondents. For a minimum sample size of 350, therefore, a total of 475 samples were collected for this study ((350 *1.25)/0.95) = 461). Sample of respondents that were selected proportionally from condominium and villa households are as shown in Table 3-1 below.

Table 3-1: Sample of respondents proportional to the number of households

	Jemo 1	Jemo 2	Jemo 3	Villa	Total
Number of households	8425	4650	4675	5000	22750
Percent of households from total	37.03	20.44	20.55	21.97	100
Proportional sample size as a summation for calibration and validation	176	98	97	104	475

Source: Nifas-silk Lafto Sub city administration

3.1.7 Sampling method

Sampling design was used to select samples. The most common class of sampling procedures is probability sampling, where each sampling unit has a pre-determined probability of being selected into the survey sample (Koppleman & Chandra, 2006). However it may be difficult to obtain a random sample as it may not be possible to get an up to date and complete list of potential respondents from which to select a sample. This is the primary reason for using "deterministic rules" of sampling.

As long as there is no inherent bias in setting up the deterministic rule, the systematic sampling plan is essentially equivalent to random sampling, and the choice of one over the other may be a matter of operational convenience. From the total 22,750 households of the study area discussed above a sample of 461 respondents were to be selected with deterministic rule of choosing every $(22{,}750/475)^{th}$ or every 48^{th} household number from the study area of every division, described in Table 3-1.

3.2 Research Methods

3.2.1 Utility function

The basic assumption of logit modelling is that, a traveler will choose a certain mode of transport if its utility is maximum of all other available modes of transport for that trip (Kenneth, 2009). This can be described with utility function as traveler t will choose mode i if the utility of mode i is larger than other available modes j. This underlying concept will help us rank the available alternatives and choose the alternative with highest utility. This implies that there is no absolute reference or zero value for utility values.

The total utility function is described as the sum of the deterministic and the probabilistic portion of the utility as shown below.

$$U_{it} = V_{it} + \varepsilon_{it} \qquad (3\text{-}3)$$

Where; U_{it} is the total utility of traveler t while using mode i

V_{it} is the deterministic utility portion where the analyst can determine in terms of attribute values.

ε_{it} is the probabilistic portion or the error term analyst doesn't fully understand the decision process which the analyst cannot understand or include

The deterministic portion of the utility function for mode choice can be described with multinomial regression involving the attribute variables related to the modes, trips and traveler characteristics.

$$V_{it} = \beta_0 + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \cdots + \beta_n * X_{ni} \qquad (3\text{-}4)$$

Where X_{ni} is the value of attribute n for mode i; these are the variables, to be collected using questionnaire from the sampled travelers in the study area; and

β_n is the parameter for attribute n defining direction and importance of effect of the attribute on the utility of the alternative; these are the values to be determined using maximum likelihood methods.

3.2.2 Methods of analysis

Analysis of logit models is the estimation of the parameters in the utility function with the method of maximum likelihood. Method of maximum likelihood involves developing joint density probability function or likelihood function for the random term in the utility function and estimating values for parameters which maximize the likelihood function (Koppleman & Chandra, 2006). The likelihood function for sample of t individuals with *j* alternatives is as shown below.

$$L(\beta) = \prod_{\forall t \in T} \prod_{\forall j \in J} \left(P_{jt}(\beta)\right)^{\delta_{jt}} \tag{3-5}$$

Where $\delta_{jt} = 1$ is chosen indicator (=1 if j is chosen by individual t and 0, otherwise) and P_{jt} is the probability that individual t chooses alternative *j*; in addition, the probability that individual t chooses alternative *j* is;

$$P_{jt} = \frac{\exp\left(X'_{jt}\beta\right)}{\sum_{j'} \exp\left(X'_{j't}\beta\right)} \tag{3-6}$$

The parameter values which maximize the likelihood function are obtained by finding the first derivative of the likelihood function and equating it to zero (Koppleman & Chandra, 2006). The log of a function yields the same maximum as the function and it is more convenient to differentiate. Therefore, the log-likelihood function is maximized instead of the likelihood function itself. The log likelihood function and its derivative is as shown below, after substituting the probability function and its derivative.

$$\frac{\partial P_{jt}}{\partial \beta_k} = P_{jt}\left(X'_{jkt} - \sum_{j'} P_{j't} X_{j'kt}\right), \forall \text{k} \tag{3-7}$$

$$LL(\beta) = Log\left(L(\beta)\right) = \sum_{\forall t \in T} \sum_{\forall j \in J} \delta_{jt} \times \ln\left(P_{jt}(\beta)\right) \tag{3-8}$$

The maximum likelihood is obtained by setting the first derivative of the log likelihood function, shown below, equal to zero and solving for the best values of the parameter vector, β, which is computed with the use of SPSS.

$$\frac{\partial(LL)}{\partial\beta_k} = \sum_{\forall t \in T} \sum_{\forall j \in J} \delta_{jt} \left(X'_{jt} - \sum_{j't} P_{j't} X_{j't} \right)$$

$$= \sum_{\forall t \in T} \sum_{\forall j \in J} \left((\delta_{jt} - P_{j't}) X'_{jt} \right) \text{ , } \forall \text{k} \quad (3\text{-}9)$$

3.2.3 Conditional multinomial logistic regression in SPSS

The special type of logit modeling which works best for modeling of discrete choices whose one or more independent variables are alternative specific; meaning whose values are different for each alternative, is known as conditional multinomial logit modeling. (Cameron & Trivedi , 2009)

In SPSS, the conditional logit modeling which primarily includes the variables that vary among the transport modes, for example; travel time and travel cost in modal choice modeling, cannot be handled by the NOMREG (conventional multinomial logistic regression) command which can only handle the variables that vary with the decision makers. Instead of the NOMREG command, the conditional MNL is computed with a COXREG command in SPSS in which the data set is as illustrated below in Table 3-2.

Table 3-2: Example of the data view in SPSS user interface

Trip #	Alternative number	Number of alternatives	HH Income in ETB	Travel time in minutes	Travel cost in ETB/trip	Alternative chosen
1	1/LBUS/	3	5000-15000	60	3	0
1	2/TAXI/	3	5000-15000	45	5	0
1	3/PC/	3	5000-15000	30	20	1
2	1	2	Below 5000	45	3.5	0
2	2	2	Below 5000	35	6	1
3	1	3	Above 15000	65	3.5	0
3	2	3	Above 15000	30	6	0
3	3	3	Above 15000	30	17	1
4	1	3	Above 15000	50	3	1

The data view example in Table 3-2 above shows that a trip number shows the trip made by a trip maker. For example, the first trip maker has three alternatives, hence there are three different values of travel time and travel cost related variables which have different values for the available modes. However the socio-economic variables have similar values repeated 3 times since these variables have different values for different trip maker. In the same manner, all the variables related to the trip maker and to the available modes are inserted in the data set of SPSS for modeling.

The outputs from SPSS and other software packages typically include, at least, the following estimation of results:

- Parameter names, parameter estimates, standard errors of these estimates, the corresponding test statistics and statistical significance values for each variables and constant parameters,
- Log-likelihood values at zero, LL(0); a value when all the parameters are set equal to zero which is usually equivalent to having no model at all,
- Log likelihood value at constants, LL(c); that value of the model when only the alternative specific constants are estimated and
- Log likelihood value at convergence, LL(β), which is the value of the likelihood function at the estimated parameters.

3.2.4 Model specification

The decision makers in this study are the people living in Nifas-Silk Lafto- sub city Woreda1-Jemo area making work trips from Jemo to Mexico between 6:00 AM and 9:00 AM local time. The alternatives included are large bus, Taxi and private car. In addition to these the bus rapid transit system corridor number 6 which is at construction stage at present time, in Addis Ababa is included under the stated preference survey. The trip makers are assumed to choose the mode which maximizes their utility. Therefore, all the data about the trip maker and the modal attributes of each mode for each trip maker is recorded in the data set of SPSS.

The attributes are as follows with their syntax used on SPSS.

i. Age [AGE], in years has 2 categories; AGE=1 for younger than or equal to 45 and Age=2 for older than 45 years.

ii. Gender [Male], 1 for male and 0 for female

iii. Education level [EDU]; 1 for 12th grade and below, 2 for diploma or certificate, 3 for degree and 4 for Masters or above level of education.

iv. Occupation category [OCC]; 1 for government employee, 2 for private or international organization employee and 3 for self-employed or other occupation category.

v. Household monthly income [INC], 1 for less than 5000 ETB, 2 for 5,000-15,000 ETB, and 3 for above 15,000 and above ETB.

vi. Household size [HH_SIZE], 1 for 1 member only, 2 for 2-4 members, and 3 for more than 4 members

vii. Car ownership [CAR_OWN], 0 for no car ownership and 1 for car ownership.

viii. Time of leaving home [TIME], 1 for those who leave their homes between 6:00-6:30, 2 between 6:31-7:00, 3 between 7:00-8:00, and 4 between 8:00-9:00 AM.

ix. Chosen mode for revealed preference [MODE_RCHOICE], 1 for large buses, 2 for taxi, and 3 for private car.

x. Chosen mode for stated preference [MODE_SCHOICE], 1 for large buses, 2 for taxi, 3 for private car, and 4 for BRT.

xi. Out of vehicle travel time including, walk time or access time with Bajaj and waiting time [OVTT] in minutes.

xii. In-vehicle travel time [IVTT], in minutes.

xiii. Total travel time [TTT] is the summation of out of vehicle time and in-vehicle travel time, in minutes.

xiv. Access cost [ACC_COST], cost of Bajaj as access for the major mode in ETB.

xv. Travel cost [TC], cost of travel with the major mode, in ETB.

xvi. Total travel cost [TTT] is the summation of the access cost and the travel cost, in ETB.

3.2.5 Methods of comparison and test of estimated models

There are different methods to test and compare estimated models as stated in section 2.4 in detail. In this study informal tests and tests for the individual parameters were used to test parameters of each variables. The overall all goodness of fit measures using rho-squared and adjusted rho-squared values, log likelihood ratio tests and non-nested hypothesis tests are used to test each developed model and compare one model from another.

3.2.6 Methods of validation

The validation of the finally chosen most satisfactory model is assessed using samples that were collected for validation purpose which is about 21% of the total collected data. These data were used to evaluate how well the model fits the data by determining the prediction capability of the model. The prediction capability of the model is determined by dividing the number of correctly predicted modes to the total number of validation data.

CHAPTER 4. RESULTS AND DISCUSSION

4.1 Descriptive data analysis

4.1.1 General socio-demographic data presentation

There were 475 valid responded questionnaires that were collected from randomly selected work trip makers between Jemo and Mexico areas with interviews at their households. The socio-economic collected data include gender, age, education level, type of occupation, monthly household income, number of people in the household or household size, car ownership of the trip maker and the time at which the trip maker leaves his/her home. These data are summarized as shown in Table 4-1, below.

According to the data summarized in Table 4-1, above, 58.11% of the sample were men and 41.89% were female. Around 91% of the trip makers in the sample are aged 45 years and below, and only 9% of them were above 45 years of age. The larger proportion of the trip makers had education level of degree (36%), following by 12th grade and below (31.37%), diploma or certificate (18.32%) and masters and above (14.32%). Occupation category of private company or international organization employees had the largest percentage of the respondents (45.68), while trip makers who are self-employed or in other employment category have 32% proportion and governmental employees account for the 22.32% of the sample.

Table 4-1 also illustrated that the majority (around 45%) of the respondents have monthly house hold income between 5,000 ETB and 15,000 ETB and the remaining 25.89 % and 28.84% have below 5,000 ETB and above 15,000 ETB, respectively. Almost half of the respondents have 2 to 4 members of house hold size; whereas, 12.63% and 36% of them have only 1 member and above 4 members, respectively. The data summary tells that most of the trip makers in the study area leave their homes early between 6:00 AM and 6:30 AM which account for 44.42% of the respondents. On the other hand 8.42% of them leave their home for work between 8:01 AM and 9:00 AM. The car ownership attribute shown in Table 4-1 illustrated that 71.58% of the respondents do not own private car and 28.42% of them own private car.

Table 4-1: Summary of the socio-economic profile of the trip makers in the study area

Socio Demographic Attributes	Label	Category	Frequency	Percentage
Gender	1	Male	276	58.11
	0	Female	199	41.89
Age group	1	Less than or equal to 45	431	90.74
	2	Older than 45	44	9.26
Education level	1	12^{th} grade or below	149	31.37
	2	Diploma or certificate	87	18.32
	3	Degree	171	36.00
	4	Masters or above	68	14.32
Occupation category	1	Governmental employee	106	22.32
	2	Private or international org. employee	217	45.68
	3	Self-employed or other	152	32.00
HH Income	1	Below 5000	123	25.89
	2	5000-15000	215	45.26
	3	Above 15000	137	28.84
HH size	1	1 member	60	12.63
	2	2-4 members	244	51.37
	3	Above 4 members	171	36.00
Local time of leaving home	1	12:00-12:30	211	44.42
	2	12:31-1:00	120	25.26
	3	1:01-2:00	104	21.89
	4	2:01-3:00	40	8.42
Car Ownership	0	No	340	71.58
	1	yes	135	28.42

Moreover, the modal share of the corridor is presented in Table 4-2, below, and the stated preference survey resulted in 36.42% taxi, 41.89% large bus and 21.68% private car users. Out of 173 taxi users, 86.71% and out of 199 large bus users 96.48% were willing to divert to the stated mode of BRT. About 60% of private car users were found to be unwilling to divert to BRT. The table below also illustrates that, 80.84% of the total respondents were willing to divert to BRT.

Table 4-2: Modal split of work trips between Jemo and Mexico areas.

Available modes of transportation	Number of users	Percentage	Willing to divert to BRT	Willing to divert to BRT in %
Taxi/minibus	173	36.42	150	86.71
Large bus	199	41.90	192	96.48
Private car	103	21.68	42	40.78
Total	475	100	384	80.84

4.1.2 Revealed preference survey data presentation

The relation between gender of the trip maker and modal preference is illustrated in Figure 4-1. Comparing the ratio of male users to female users of the three modes; as can be seen in Figure 4-1, the ratio of male to female users of private car is larger than those values for large bus and taxis. In contrast, large bus and taxi users have ratios of male to female users close to 1. Consequently private cars are mostly used by male trip makers whereas large bus and taxis have closer numbers of male and female users.

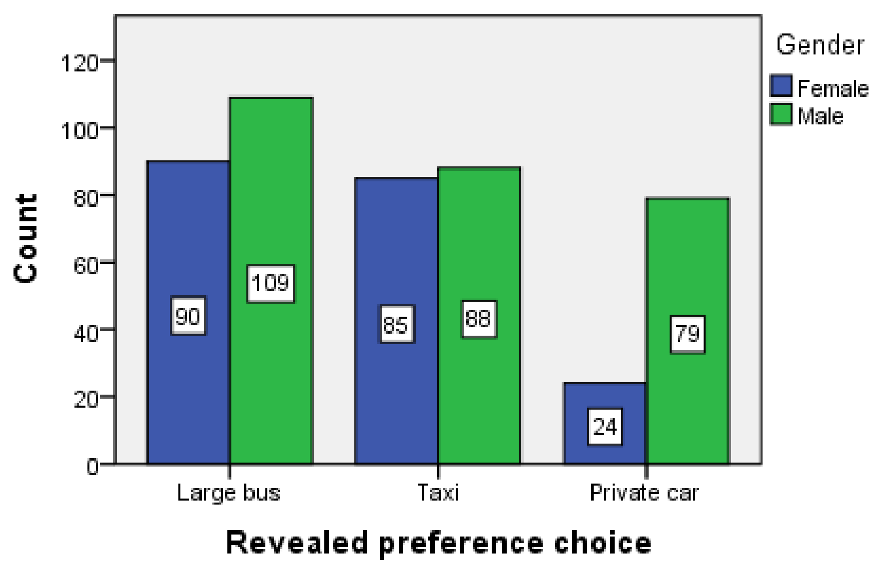

Figure 4-1: Revealed mode preference of trip makers with respect to the gender of the trip maker

According to figure 4-2, most of the work trip makers with education level of 12^{th} grade or below and with diploma or certificate prefer large bus, most of the trip makers with education level of degree prefer taxi, and most of the trip makers with masters or above education level prefer to use private car. It is also apparent from the figure that large bus is mostly preferred by trip makers with education level of 12^{th} grade and below, and taxis and private cars are mostly preferred by trip makers with degree.

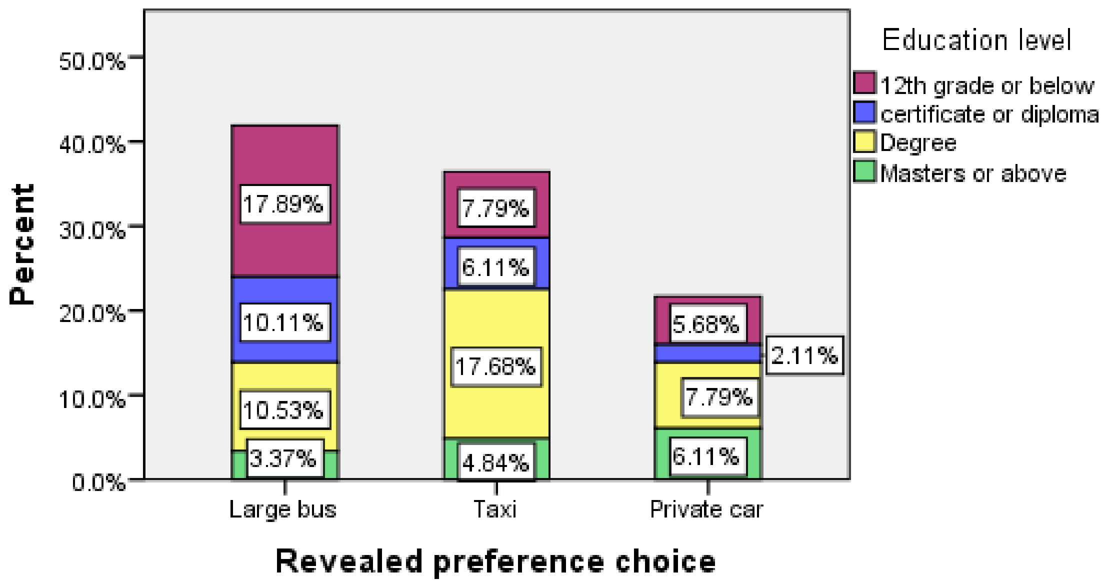

Figure 4-2: Revealed mode preference with respect to the education level of the trip maker

As Figure 4-3 shows, most of the trip makers who preferred to use their private car or to make work trips are those whose monthly household income is above ETB 15,000, and those trip makers whose monthly household income is less than ETB 5,000 do not prefer to use private car. On the other hand, for large bus and taxi (minibus), most of their customers are those with ETB 5,000 to 15,000 monthly family income. In addition to this, most of the people who own personal car prefer to use their personal car, and those who do not prefer to use their personal cars prefer more to use taxi than large buses, as presented in Figure 4-4.

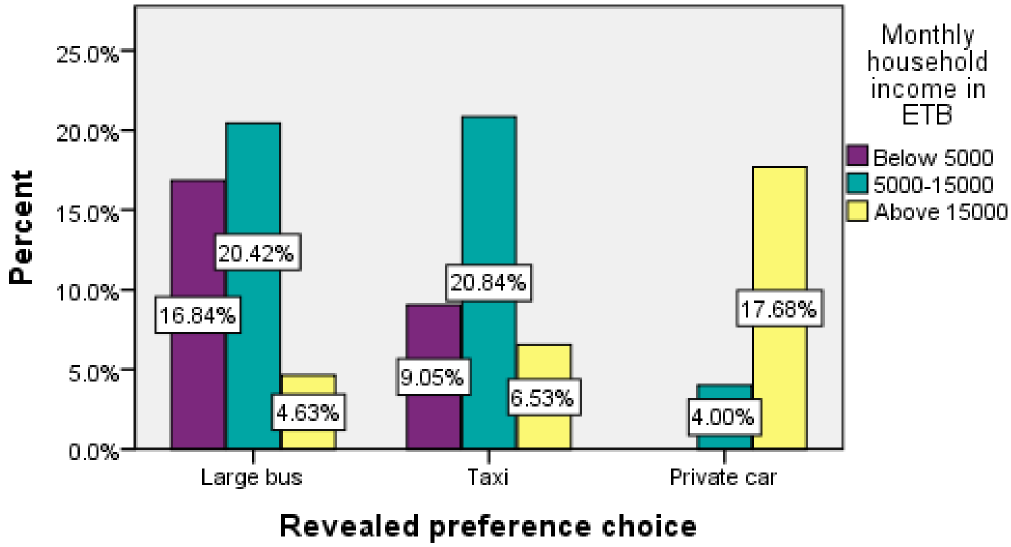

Figure 4-3: Revealed mode preference with respect to the monthly household income of the trip maker

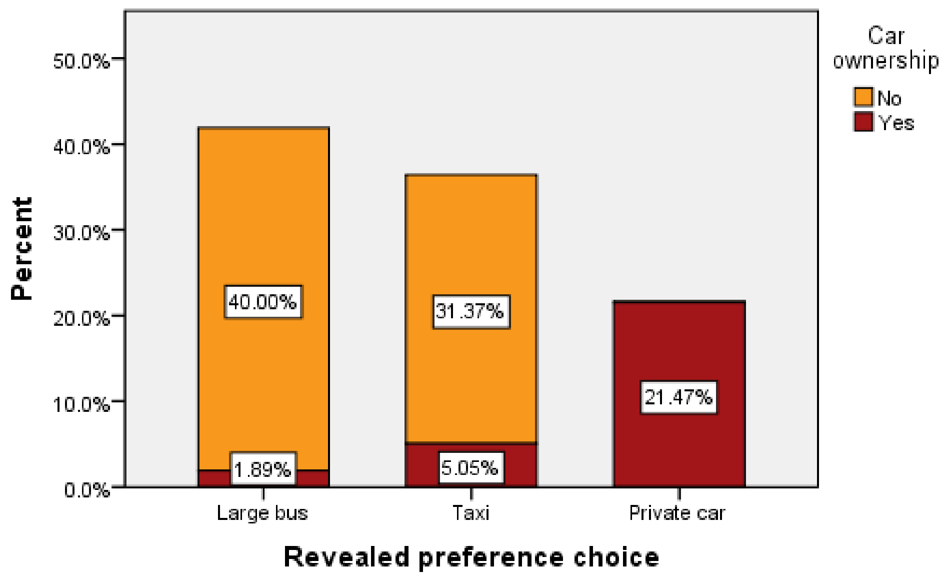

Figure 4-4: Revealed mode preference with respect to the car ownership of the trip maker

Most of the trip makers of Jemo area who preferred to use large bus and taxi (minibus) leave their home from 6:00 and 6:30 local time, whereas more number of the people whose preferred mode is private car, leave their home between 7:01 and 8:00 AM local time, as shown in the graph below. Trip makers who leave their home between 8:00 AM and 9:00 AM are those whose preferred mode is private car, as shown in Figure 4-5, below.

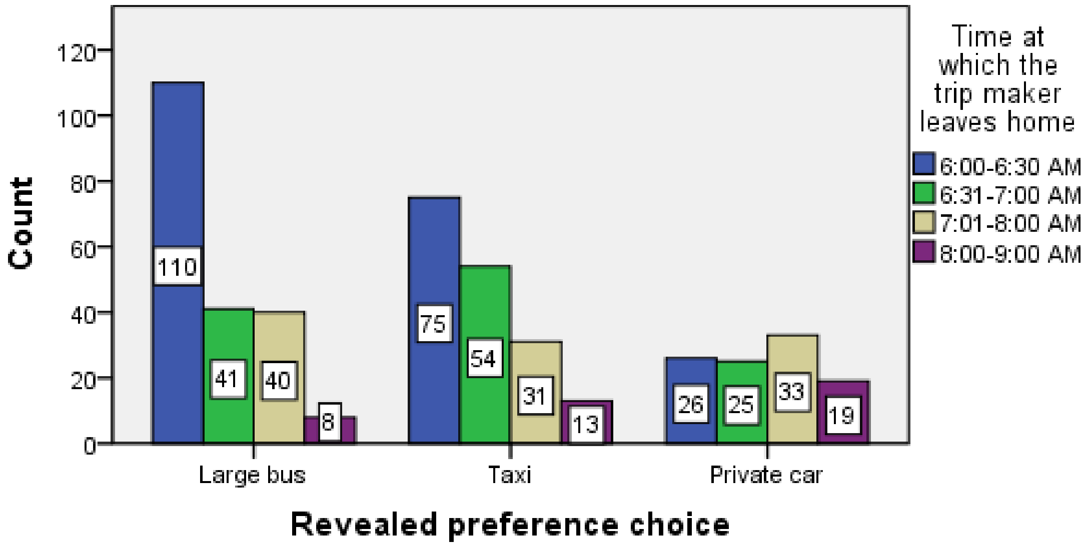

Figure 4-5: Mode preference of trip makers with respect to the time at which the trip maker leaves home.

4.1.3 Stated preference survey data presentation

The stated preference survey basically tries to estimate the diverted trip makers to a stated mode in the questionnaire with its probable modal characteristics. The stated mode in this study is bus rapid transit system and based on the pie chart shown in Figure 4-6, below, around 81% of the trip makers are willing to divert to the stated bus rapid transit system, while the remaining 19 % carried on with the preferred mode illustrated in the revealed preference survey.

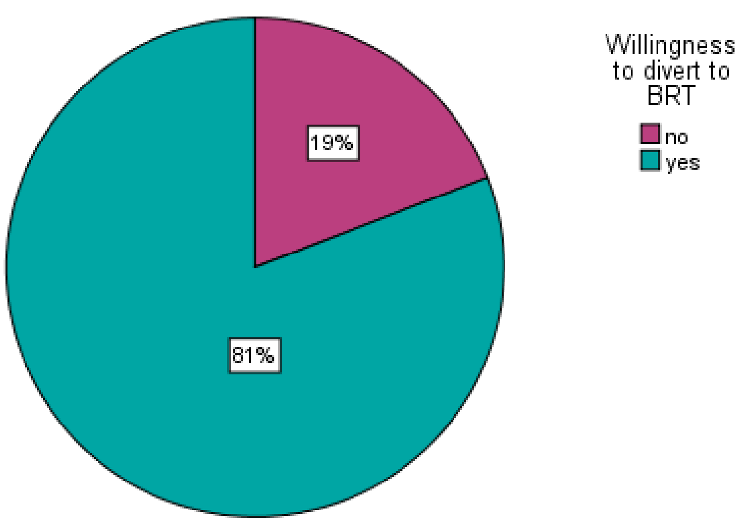

Figure 4-6: Percentage description of trip makers who are willing and not willing to divert to BRT.

Furthermore, Figure 4-7 below, presents most of trip makers whose revealed mode choices are large bus and taxi (minibus) are willing to divert to the stated bus rapid transit system mode. On the other hand, most of private car users are not willing to divert to bus rapid transit system. Similarly, Figure 4-6 above, show that the trip makers who are not willing to divert to BRT are those who own private cars.

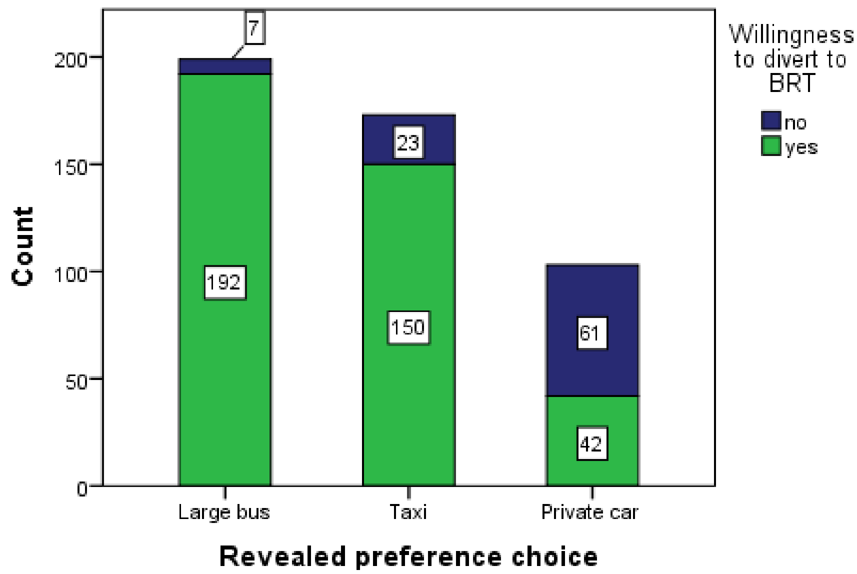

Figure 4-7: Willingness to divert to the BRT system with respect to the revealed mode preference.

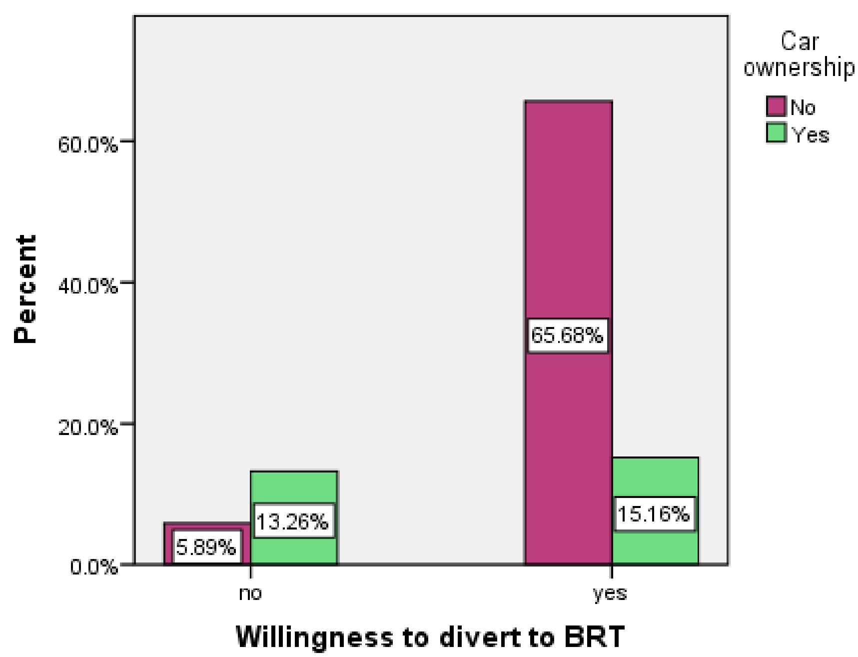

Figure 4-8: Willingness to divert to BRT with car ownership of the trip makers

4.2 Calibration, interpretation and validation of revealed preference model

4.2.1 Parameter estimation, test and comparison of revealed preference models

Based on the descriptive analysis of the collected data; large bus, minibus (taxi) and private car are the modes that people in the study area use to make trips from home to work. Travel time, travel cost and other trip related independent variables are included as generic variables meaning that an increase of one unit of these variables has the same impact on all the three modes. The socio economic independent variables are considered alternative specific; with different parameter for all the three modes or an increase of one unit of these variables has different impact for all the three modes. Variables and corresponding descriptions are shown in Table 4-3 below.

The large bus mode is considered as the base or reference alternative in SPSS for all the parameter estimation of the independent variables except for the generic variables (like TTT and TTC) and for variables with categories more than one, always the last category is the reference category. There are 840 observations and 375 cases out of which 90 of the cases have the three alternatives, private car, large bus and taxi (minibus), and the remaining 285 of the cases have two alternatives, large bus and minibus (taxi). Different models were developed and analyzed step by step to determine which one of the developed models is the fit or true model that replicates the data better.

Table 4-3: Description of explanatory variables

Variable		Description
LBUS		Constant parameter for large bus utility
BRT		Constant parameter for bus rapid transit system utility
TAXI		Constant parameter for taxi utility
PC		Constant parameter for private car utility
TTT		Total travel time in minutes (generic variable)
TTC		Total travel cost in ETB (generic variable)
IVTT		In-vehicle travel time in minutes (generic variable)
OVTT		Out of vehicle travel time in minutes (generic variable)
INC	INC_TAXI	Household monthly income and taxi interaction (alternative specific)
	INC_PC	HH monthly income and private-car interaction (alternative specific)
AGE	AGE_TAXI	Age and taxi interaction (alternative specific variable)
	AGE_PC	Age and private car interaction (alternative specific variables)
MALE	MALE_TAXI	Gender male and taxi interaction (alternative specific variable)
	MALE_PC	Gender male and private car interaction (alternative specific variables)
HHS	HHS _TAXI	Household size and taxi interaction (alternative specific variable)
	HHS _PC	HH size and private car interaction (alternative specific variables)
CAR	CAR _TAXI	Car ownership and taxi interaction (alternative specific variable)
	CAR _PC	Car ownership and private car interaction (alternative specific)
JOB	JOB _TAXI	Job category and taxi interaction (alternative specific variable)
	JOB _PC	Job category and private car interaction (alternative specific variables)
EDU	EDU _TAXI	Education level and taxi interaction (alternative specific variable)
	EDU _PC	Education level and private car interaction (alternative specific)
TIME	TIME _TAXI	Time of leaving home and taxi interaction (alternative specific)
	TIME _PC	Time of leaving home and private car interaction (alternative specific)

The different levels of models estimated with their deterministic portion of the utility of each mode are presented in detail below.

I. Model 1 and 2

The first model (Model 1) includes travel time (TTT) and travel cost (TTC) in addition to the constant parameters of taxi and private car. From Table 4-5, parameters of TTT, TTC and the alternative specific constant of private car have a P-value less than 0.05, which can reject the null hypotheses that those parameters are zero. The P-value for constant parameter of taxi is greater than 0.05. In spite of this, alternative specific constants were retained in the model since the constants represent the average effect of all the variables not included in the model.

Table 4-4: Parameter estimation results for Model 1 and Model 2

		MODEL 1			**MODEL 2**		
Parameters		**B**	**Sig.**	**Exp(β)**	**B**	**Sig.**	**Exp(β)**
TTT		-0.043	0.000	0.958	-0.043	0.000	0.958
TTC		-0.371	0.001	0.690	-0.376	0.001	0.686
LBUS		Base mode					
TAXI		-0.246	0.516	0.782	-0.069	0.888	0.934
PC		9.884	0.001	2.0E+4	10.081	0.001	2.4E+4
	INC_TAXI[1]	--------	-------	--------	-0.608	0.116	0.544
	INC_TAXI[2]	--------	-------	--------	0.023	0.949	1.023
	INC_TAXI[3]	Reference category					
-2LL(0)		592.844			592.844		
-2LL(C)		445.982			445.982		
-2LL(β)		404.752			399.240		
Rho squared w.r.t. zero (ρ_0^2)		0.317			0.327		
Rho squared w.r.t. constants (ρ_c^2)		0.092			0.105		
Degree of freedom (k)		4			6		
Adjusted rho squared w.r.t. zero ($\rho_0'^2$)		0.304			0.305		
Adjusted rho squared w.r.t. constants ($\rho_c'^2$)		0.075			0.076		

In addition, as expected the travel time and cost parameters have negative coefficients which imply that with the increase in travel time and travel cost of modes their utilities decrease. This is consistent with our understanding of travel behavior, hence the model is not rejected. Table 4-4 above shows parameter results for Model 1 and Model 2.

The utilities for each mode model 1 can be written as follows, according to the output summarized in Table 4-4.

$$\boldsymbol{U_{LBUS}} = -0.043 * \boldsymbol{TTT} - 0.371 * \boldsymbol{TTC} \quad (4\text{-}1)$$

$$\boldsymbol{U_{TAXI}} = -0.246 - 0.043 * \boldsymbol{TTT} - 0.371 * \boldsymbol{TTC}$$

$$\boldsymbol{U_{PC}} = 9.884 - 0.043 * \boldsymbol{TTT} - 0.371 * \boldsymbol{TTC}$$

The monthly household income variable interaction with taxi was added to improve the statistics of Model 1 and is labeled as Model 2**.** The utilities of the modes can be formulated as follows based on the result of estimations in Table 4-4.

$$\boldsymbol{U_{LBUS}} = -0.043 * \boldsymbol{TTT} - 0.376 * \boldsymbol{TTC} \quad (4\text{-}2)$$

$$\boldsymbol{U_{TAXI}} = -0.069 - 0.043 * \boldsymbol{TTT} - \boldsymbol{0.376} * \boldsymbol{TTC} - 0.608 * \boldsymbol{INC}\,[\boldsymbol{1}] + 0.023 * \boldsymbol{INC}\,[\boldsymbol{2}]$$

$$\boldsymbol{U_{PC}} = 10.081 - 0.043 * \boldsymbol{TTT} - 0.376 * \boldsymbol{TTC}$$

As shown in Table 4-4, the coefficients for income-taxi interactions have p-values of 0.116 and 0.949 which are both greater than $\alpha = 0.05$, which cannot justify to reject the null hypothesis that the coefficients for income-taxi interactions are zero. This means the effect of income-taxi interaction on the taxi utility may not make it different from the value in Model 1. So this value is considered for exclusion. To confirm the exclusion, comparison is made between Models 1 and 2 by computing the chi-squared value between the two models.

Taking Model 1 as restricted model and Model 2 as unrestricted model, the chi-squared value is,

$$\chi^2 = -2 * (LL_R - LL_U) = -2LL_R - (-2LL_U) \quad (4\text{-}1)$$

$$= 404.752 - 399.240 = 5.512 < 5.99.$$

This shows that the null hypothesis that the income-taxi interaction coefficients are zero cannot be rejected since the chi-squared value is not sufficiently large. Therefore, the exclusion of income-taxi interaction coefficients is confirmed. Therefore, Model 1 is apt for comparison with other models in the next step.

II. Models 3 and 4 compared to Model 1

In Model 3, the coefficient for monthly household income variable interaction with private car was included and the income-private car interaction results in a p-value of 0.045 < 0.05, as shown in Table 4-5. This makes the null hypothesis rejected; the income-private car interaction effect to the utility of the private car makes the model different from Model 1.

Checking for its chi-squared value (taking Model 1 as the restricted and Model 3 as the unrestricted model) results in the following value, which makes it possible to reject the null hypothesis that the parameter of income variable interaction with private car is zero.

$$\chi^2 = 404.752 - 397.765 = 6.987 > 3.84$$

Where 3.84 is the critical chi-squared value from Table 4-3 at 95% confidence level and the restricted parameter being 1. Table 4-5 below presents parameter results of Model 3 and Model 4.

Model 4 was developed by adding male variable interactions with taxi and private car to Model 3. However, the chi-square value for this model (by taking Model 3 as the restricted and Model 4 as the unrestricted model) results with a value that cannot reject the null hypothesis that the parameters of male variable interactions with taxi and private car should be excluded.

$$\chi^2 = 397.765 - 396.208 = 1.557 < 5.99$$

Where 5.99 is the critical chi-squared value for 95% confidence level and the restricted parameter being 2. According to Table 4-5, the utilities of the three modes with the statistically significant values of parameters in model 3 are as follows.

$$U_{LBUS} = -0.041 * \boldsymbol{TTT} - 0.410 * \boldsymbol{TTC} \tag{4-3}$$

$$U_{TAXI} = -0.131 - 0.041 * \boldsymbol{TTT} - 0.410 * \boldsymbol{TTC}$$

$$U_{PC} = 13.281 - 0.041 * \boldsymbol{TTT} - 0.410 * \boldsymbol{TTC} - 3.604 * \boldsymbol{INC\_PC}[\boldsymbol{2}]$$

Table 4-5: Parameter estimation results for Model 3 and Model 4

		MODEL 3			MODEL 4		
Parameters		B	Sig.	Exp(β)	β	Sig.	Exp(β)
TTT		-0.041	0.000	0.959	-0.041	0.000	0.960
TTC		-0.410	0.001	0.664	-0.443	0.000	0.642
LBUS		Base mode					
TAXI		-0.131	0.736	0.878	0.054	0.900	1.056
PC		13.281	0.001	5.9E5	13.085	0.001	4.8E+5
INC_PC	INC_PC[1]	There is no data for INC[1] and PC interaction					
	INC_PC[2]	-3.604	0.045	0.027	-3.464	0.060	0.032
	INC_PC[3]	Reference category					
MALE_TAXI		----------	---------	---------	-0.170	0.489	0.844
MALE_PC		----------	---------	---------	1.356	0.330	3.879
-2LL(0)		592.844			592.844		
-2LL(C)		445.982			445.982		
-2LL(β)		397.765			396.208		
$\boldsymbol{\rho_0^2}$		0.329			0.332		
$\boldsymbol{\rho_c^2}$		0.108			0.112		
k		5			7		
$\boldsymbol{\rho_0'^2}$		0.312			0.308		
$\boldsymbol{\rho_c'^2}$		0.086			0.080		

III. Models 5 and 6 compared to Model 3

The addition of Age variable interaction with taxi to Model 3 resulted in a model (Model 5) with statistically significant parameter for the interaction. Therefore, this interaction variable retains in the model. The addition of the interaction variable with private car, which is Model 6, resulted in the estimation result shown in Table 4-6 below. However the computed chi-squared value, taking Model 5 as restricted and Model 6 as unrestricted model, confirmed its exclusion from Model 6.

It was excluded, because the resulted chi-square value computed below, was smaller than the critical value.

$$\chi^2 = 392.171 - 392.156 = 0.015 < 3.84$$

Where 3.84 is the critical chi-squared value from Table 4-1 below for 95% confidence level and the restricted parameter being 1.

Table 4-6: Parameter estimation results for Model 5 and Model 6

		MODEL 5			MODEL 6		
Parameters		β	Sig.	Exp(β)	β	Sig.	Exp(β)
TTT		-0.044	0.000	0.957	-0.044	0.000	0.957
TTC		-0.420	0.000	0.657	-0.422	0.000	0.655
LBUS		Base mode					
TAXI		-1.045	0.065	0.352	-1.036	0.069	0.355
PC		13.547	0.001	7.7E+5	13.881	0.004	1.1E+6
INC_PC	INC_PC[1]	No data					
	INC_PC[2]	-3.682	0.045	0.025	-3.696	0.046	0.025
	INC_PC[3]	Reference category					
AGE_TAXI	AGE_TAXI[1]	0.985	0.023	2.679	0.982	0.024	2.668
	AGE_TAXI[2]	Reference category					
AGE_PC	AGE_PC[1]	---------	--------	--------	-0.293	0.901	0.746
	AGE_PC[2]	Reference category					
-2LL(0)		592.844			592.844		
-2LL(C)		445.982			445.982		
-2LL(β)		392.171			392.156		
$\boldsymbol{\rho_0^2}$		0.338			0.339		
$\boldsymbol{\rho_c^2}$		0.121			0.121		
k		6			6		
$\boldsymbol{\rho_0'^2}$		0.318			0.318		
$\boldsymbol{\rho_c'^2}$		0.094			0.094		

According to Table 4-6, the utilities of the three modes with the statistically significant values of parameters in Model 5 are as follows.

$$U_{LBUS} = -0.044 * \boldsymbol{TTT} - 0.420 * \boldsymbol{TTC} \quad (4\text{-}4)$$

$$U_{TAXI} = -1.045 - 0.044 * \boldsymbol{TTT} - 0.420 * \boldsymbol{TTC} + 0.985 * \boldsymbol{AGE\_TAXI}[\mathbf{1}]$$

$$U_{PC} = 13.547 - 0.042 * \boldsymbol{TTT} - 0.420 * \boldsymbol{TTC} - 3.682 * \boldsymbol{INC\_PC}[\mathbf{2}]$$

IV. Models 7 and 8 compared to Model 5

Car ownership variable interactions with taxi and private car are added to Model 6 to check whether it can improve the statistics of the model as the previous variables did. Model 7 is the model with car ownership variable interaction to taxi and the coefficient for this interaction has a p-value of 0.006 < 0.05 (Table 4-7), and its chi- square value, taking Model 6 as restricted and Model 7 as unrestricted model, is shown below.

$$\chi^2 = 392.171 - 383.550 = 8.621 > 3.84$$

This value rejects the exclusion of the added interaction variable from the model.

The effect of adding car ownership variable interaction with private car to Model 7, Model 8, results with a p-value of 0.926 > 0.05, Table 4-7, for the coefficient of the interaction. The computed chi-squared value, taking Model 7 restricted and Model 8 unrestricted model, confirmed the exclusion of this interaction variable from the model.

$$\chi^2 = 383.566 - 383.550 = 0.016 < 3.84$$

According to Table 4-7, the utilities of the three modes with the statistically significant values of parameters in Model 7 are as follows.

$$U_{LBUS} = -0.043 * \boldsymbol{TTT} - 0.449 * \boldsymbol{TTC} \quad (4\text{-}5)$$

$$U_{TAXI} = -1.101 - 0.043 * \boldsymbol{TTT} - 0.449 * \boldsymbol{TTC} + 0.982 * \boldsymbol{AGE\_TAXI}[\mathbf{1}] + 1.388 * \boldsymbol{CAR\_TAXI}[\mathbf{1}]$$

$$U_{PC} = 15.114 - 0.043 * \boldsymbol{TTT} - 0.449 * \boldsymbol{TTC} - 3.682 * \boldsymbol{INC\_PC}[\mathbf{2}]$$

Table 4-7: Parameter estimation results for Model 7 and Model 8

		MODEL 7			MODEL 8		
Parameters		**β**	**Sig.**	**Exp(β)**	**β**	**Sig.**	**Exp(β)**
TTT		-0.043	0.000	0.958	-0.043	0.000	0.958
TTC		-0.451	0.000	0.637	-0.438	0.000	0.645
LBUS		Base mode					
TAXI		-1.101	0.058	0.332	-1.129	0.051	0.323
PC		15.191	0.000	3.9E+6	16.816	0.503	2.0E+7
INC_PC	INC_PC[1]	No data					
	INC_PC[2]	-3.735	0.057	0.024	-3.361	0.063	0.035
	INC_PC[3]	Reference category					
AGE_TAXI	AGE_TAXI[1]	1.038	0.021	2.823	1.035	0.020	2.816
	AGE_TAXI[2]	Reference category					
CAR_TAXI	CAR_TAXI[0]	Reference category					
	CAR_TAXI[1]	1.382	0.006	3.984	1.374	0.006	3.950
CAR_PC	CAR_PC[0]	Reference category					
	CAR_PC[1]	--------	--------	------	-2.324	0.926	0.098
-2LL(0)		592.844			592.844		
-2LL(C)		445.982			445.982		
-2LL(β)		383.550			383.566		
$\boldsymbol{\rho_0^2}$		0.353			0.353		
$\boldsymbol{\rho_c^2}$		0.140			0.140		
k		7			8		
$\boldsymbol{\rho_0'^2}$		0.329			0.326		
$\boldsymbol{\rho_c'^2}$		0.109			0.104		

V. Model 9 *and* 10 compared to Model 7

The interaction variable for time of leaving home, with taxi and private car added to Model 7 resulted in Models 9 and 10, respectively. As shown in Table 4-8, the variables for the interaction of time of leaving home with taxi have two statistically significant coefficients and one not statistically significant coefficient, shown in Table 4-8 below of Model 9. To check the exclusion of these coefficients, chi-squared value for this model was computed as shown below, taking Model 7 as restricted and Model 9 as unrestricted model.

$$\chi^2 = 383.550 - 373.796 = 9.754 > 7.81$$

where 7.81 is the critical chi-squared value at 95% confidence level and three restricted parameters Based on this result the null hypothesis is rejected and the time of leaving and taxi interaction variable retains in the model. Table 4-8 presents the parameter results of Models 9 and 10.

However, as the computed chi-squared value below, those results for the inclusion of time of leaving interaction variable with private car does not satisfy the criteria and the interaction variable was dropped.

$$\chi^2 = 373.796 - 372.085 = 1.711 < 7.81$$

According to Table 4-8, the utilities of the three modes with the statistically significant values of parameters in Model 9 are as follows.

$$U_{LBUS} = -0.046 * TTT - 0.460 * TTC \quad (4\text{-}6)$$

$$U_{TAXI} = 0.056 - 0.046 * TTT - 0.460 * TTC + 0.904 * AGE_{TAXI[1]} + 1.277 * CAR_{TAXI[1]} - 1.377 * TIME_{TAXI[1]}$$

$$U_{PC} = 15.370 - 0.046 * TTT - 0.460 * TTC$$

Table 4-8: Parameter estimation results for Models 9 and 10

		MODEL 9			**MODEL 10**		
Parameters		**β**	**Sig.**	**Exp(β)**	**β**	**Sig.**	**Exp(β)**
TTT		-0.046	0.000	0.955	-0.046	0.000	0.955
TTC		-0.460	0.000	0.631	-0.480	0.000	0.619
TAXI		0.056	0.945	1.058	.074	0.930	1.076
LBUS		Base mode					
PC		15.370	0.000	4.7E+6	15.030	0.001	3.4E+6
INC_PC	INC_PC[1]	No data					
	INC_PC[2]	-3.820	0.066	0.022	-3.898	0.039	0.020
	INC_PC[3]	Reference category					
AGE_TAXI	AGE_TAXI[2]	0.904	0.042	2.470	0.902	0.043	2.464
	AGE_TAXI[1]	Reference category					
CAR_TAXI	CAR_TAXI[0]	Reference category					
	CAR_TAXI[1]	1.277	0.012	3.586	1.297	0.011	3.659
TIME_TAXI	TIME_TAXI[4]	Reference category					
	TIME_TAXI[1]	-1.377	0.020	0.252	-1.333	0.026	0.264
	TIME_TAXI[2]	-0.655	0.286	0.519	-0.581	0.352	0.559
	TIME_TAXI[3]	-1.211	0.056	0.298	-1.214	0.059	0.297
TIME_PC	TIME_PC[4]	--------	---------	-------	1.208	0.658	3.347
	TIME_PC[1]	--------	---------	-------	2.013	0.462	7.483
	TIME_PC[2]	--------	---------	-------	-0.203	0.939	0.816
	TIME_PC[3]	Reference category					
-2LL(0)		592.844			592.844		
-2LL(C)		445.982			445.982		
-2LL(β)		373.796			372.085		
$\boldsymbol{\rho_0^2}$		0.369			0.372		
$\boldsymbol{\rho_c^2}$		0.162			0.166		
k		10			13		

	MODEL 9	MODEL 10
$\boldsymbol{\rho_0'^2}$	0.336	0.329
$\boldsymbol{\rho_c'^2}$	0.117	0.107

VI. Models 11 and 12 compared to Model 9

Model 11 is developed by adding interaction variable between education level and taxi, the results of parameter estimation for Models 11 and 12, presented in Table 4-9. The parameters of the interaction has p-value greater than 0.05, however, the chi-squared value rejected the null hypothesis and the interaction variable retain to the model, as computed below.

$$\chi^2 = 373.796 - 355.659 = 18.137 > 7.81$$

Model 12 was developed but the interaction between education level and private car have a chi-squared value (computed below) that cannot reject the null hypothesis and the parameters are excluded from the model.

$$\chi^2 = 355.659 - 355.494 = 0.165 < 7.81$$

Table 4-9: Parameter estimation results for Models 11 and 12

		MODEL 11			MODEL 12		
Parameters		**β**	**Sig.**	**Exp(β)**	**β**	**Sig.**	**Exp(β)**
TTT		-0.050	0.000	0.951	-0.050	0.000	0.951
TTC		-0.462	0.000	0.630	-0.462	0.000	0.630
LBUS		Base mode					
TAXI		0.228	0.812	1.256	0.217	0.822	1.242
PC		15.494	0.001	5.4E+6	19.034	0.551	1.9E+8
INC_PC	INC_PC[1]	No data					
	INC_PC[2]	-3.766	0.092	0.023	-3.698	0.108	0.025
	INC_PC[3]	Reference category					
AGE_TAXI	AGE_TAXI[1]	0.711	0.115	2.037	0.706	0.117	2.027
	AGE_TAXI[2]	Reference category					
CAR_TAXI	CAR_TAXI[0]	Reference category					
	CAR_TAXI[1]	1.215	0.020	3.371	1.208	0.021	3.348
TIME_TAXI	TIME_TAXI[1]	-1.638	0.009	0.194	-1.616	0.011	0.199
	TIME_TAXI[2]	-0.998	0.124	0.369	-0.978	0.134	0.376
	TIME_TAXI[3]	-1.511	0.024	0.221	-1.488	0.027	0.226
	TIME_TAXI[4]	Reference category					
EDU_TAXI	EDU_TAXI[1]	-0.361	0.444	0.194	-0.360	0.445	0.697
	EDU_TAXI[2]	-0.072	0.883	0.369	-0.084	0.865	0.919
	EDU_TAXI[3]	0.920	0.046	0.221	0.913	0.049	2.492
	EDU_TAXI[4]	Reference category					
EDU_PC	EDU_PC[1]	--------	--------	--------	-3.204	0.920	0.041
	EDU_PC[2]	--------	--------	--------	-3.947	0.901	0.019
	EDU_PC[3]	--------	--------	--------	-3.641	0.909	0.026
	EDU_PC[4]	Reference category					
-2LL(0)		592.844			592.844		
-2LL(C)		445.982			445.982		

	MODEL 11	MODEL 12
-2LL(β)	355.659	355.494
$\boldsymbol{\rho_0^2}$	0.400	0.400
$\boldsymbol{\rho_c^2}$	0.203	0.203
k	13	16
$\boldsymbol{\rho_0'^2}$	0.356	0.346
$\boldsymbol{\rho_c'^2}$	0.144	0.131

VII. Model 13 compared to Model 11

The above models, using the total travel time one generic variable, assume that the in vehicle travel time and out of vehicle travel time have the same utility, by adding them in to one variable TTT. Model 13 was developed by substituting the total travel time with in-vehicle travel time and out of vehicle time, to check the sensitivity of trip makers towards the two variables, independently.

The developed model was evident that trip makers are more sensitive to out of vehicle travel time than in-vehicle travel time since the parameter for IVTT is -0.036 and the parameter for OVTT is -0.078. This implies that a minute of out of vehicle travel time will generate a higher disutility than a minute of in-vehicle travel time. The comparison between Models 11 and 13 is done using the non-nested hypothesis test using the outputs of the parameter estimation in Tables 4-9 above and Table 4-10 below.

The null hypothesis that the model with the lower $\boldsymbol{\rho_0'^2}$**,** Model 11, is the true model, will be rejected if the significance level computed below is less than 0.05, for 95% confidence level.

$$Significance\ level\ = \Phi\left[-\left(-2 * (\rho_H'^2 - \rho_L'^2) * LL(0) + (K_H - K_L)\right)^{0.5}\right] \quad (4\text{-}2)$$

$$= \Phi\left[-\left((0.358 - 0.356) * 592.844 + (14 - 13)\right)^{0.5}\right]$$

$$= \Phi\,[-1.486] = 0.0681 > 0.05$$

Since the significance level calculated above is greater than 0.05 then the null hypothesis that Model 11 is the true model cannot be rejected. Therefore, at this stage, the true model is Model 11.

Table 4-10: Parameter estimation results for Model 13

		MODEL 13		
Parameters		**β**	**Sig.**	**Exp(β)**
IVTT		-0.036	0.009	0.965
OVTT		-0.078	0.000	0.925
TTC		-0.454	0.000	0.635
LBUS		Base mode		
TAXI		0.342	0.724	1.408
PC		14.836	0.001	2.8E+6
INC_PC	INC_PC[1]	No data		
	INC_PC[2]	-3.785	0.084	0.023
	INC_PC[3]	Reference category		
AGE_TAXI	AGE_TAXI[1]	0.721	0.112	2.056
	AGE_TAXI[2]	Reference category		
CAR_TAXI	CAR_TAXI[0]	Reference category		
	CAR_TAXI[1]	-1.213	0.020	3.364
TIME_TAXI	TIME_TAXI[1]	-1.798	0.006	0.166
	TIME_TAXI[2]	-1.063	0.108	0.345
	TIME_TAXI[3]	-1.621	0.018	0.198
	TIME_TAXI[4]	Reference category		
EDU_TAXI	EDU_TAXI[1]	-0.455	0.343	0.634
	EDU_TAXI[2]	-0.097	0.845	0.907
	EDU_TAXI[3]	0.918	0.048	2.504
	EDU_TAXI[4]	Reference category		
-2LL(0)		592.844		
-2LL(C)		445.982		
-2LL(β)		352.584		
ρ_0^2		0.405		
ρ_c^2		0.209		

	MODEL 13
K	14
$\boldsymbol{\rho_0'^2}$	0.358
$\boldsymbol{\rho_c'^2}$	0.147

Model 11 has better values 1for $\boldsymbol{\rho_0^2}$, $\boldsymbol{\rho_c^2}$, $\boldsymbol{\rho'^2_0}$, and $\boldsymbol{\rho'^2_c}$ than all other models and it has more statistically significant parameters at 95% confidence level. Therefore, Model 11 is chosen as the most satisfactory model to represent the modal preference of trip makers between Jemo and Mexico areas of Addis Ababa. The three utility functions of Model 11 can be written as follows, based on the estimation results in Table 4-9.

$$U_{LBUS} = -0.050 * \boldsymbol{TTT} - 0.462 * \boldsymbol{TTC} \tag{4-7}$$

$$U_{TAXI} = 0.228 - 0.050 * \boldsymbol{TTT} - 0.462 * \boldsymbol{TTC} + 1.215 * \boldsymbol{CAR\_TAXI}[\boldsymbol{1}] - 1.638 * \boldsymbol{TIME\_TAXI}[\boldsymbol{1}] - 1.511 * \boldsymbol{TIME\_TAXI}[\boldsymbol{3}] + 0.920 * \boldsymbol{EDU\_TAXI}[\boldsymbol{3}]$$

$$U_{PC} = 15.494 - 0.050 * \boldsymbol{TTT} - 0.462 * \boldsymbol{TTC}$$

According to the developed models, total travel time and total travel cost of the modes, car ownership and education level of the trip maker and the time at which the trip maker leaves home interacted with taxi are the factors that statistically significantly affect the reveled preference of trip makers among large bus, taxi (minibus) and private car. The parameters of total travel time and total travel cost both have negative parameters. This indicates that the decrease in total travel time and total travel cost of a mode will improve the utility of the mode for a trip maker.

The socio-economic variables are expressed as interaction with the chosen modes. Based on the resulted revealed preference model, trip makers who own car are highly likely to choose taxi over bus than those who do not own a car. This means; other factors kept similar; the utility of using taxi for a trip maker with private car is higher than that of a trip maker with no private car.

The trip makers who leave their home between 6:30 and 8:00 AM are less likely to choose taxi over bus than those who leave their home between 8:00 and 9:00 AM local time. The parameters also tells us that the utility of using a taxi decreases for a person who leave their home between 6:30 and 8:00 AM.

4.2.2 Validation of revealed preference model

The validation of the most satisfactory model is done in two stages; informal test for reasonableness of basic parameters and computation of the ratio of prediction. The parameters of the generic variables, total travel time and travel cost in Model 11 are negative which is the same as the expected signs since the utility of modes towards the trip makers decrease as the travel time and travel cost increase. Table 11 which is shown below, presents validation data of the revealed preference model partially.

Table 4-11: Example of the validation of revealed preference model

Trip #	Available modes	Utility	Probability of choosing a mode	predicted choice	Observed choice	Observed versus predicted choices
1	large bus	-3.636	0.965	1	1	Similar
1	Taxi	-6.948	0.035	0	0	
2	Taxi	-5.721	0.282	0	1	Not similar
2	large bus	-4.786	0.718	1	0	
3	Taxi	-5.16	0.581	1	1	Similar
3	large bus	-5.486	0.419	0	0	
4	Taxi	-4.56	0.581	1	1	Similar
4	large bus	-4.886	0.419	0	0	
5	large bus	-3.886	0.916	1	1	Similar
5	Taxi	-6.278	0.084	0	0	
21	private car	4.128	1.000	1	1	Similar
21	large bus	-4.886	0.000	0	0	
21	Taxi	-6.031	0.000	0	0	

For the purpose of validation, 100 samples of data were reserved and the prediction ratio of model 11 are computed as described in detail in Section 3.2.7. The resulting utility and probability of choosing the modes for each trips, the observed preference and the predicted preferences is as shown in the table above, partially. The number of preferences that are correctly predicted by the model are 80 out of 100 validation data. This implies that the calibrated model is capable of predicting 80% of the preference of trip makers, as computed below. The observed and predicted modes are presented briefly in Table 4-11 above.

As shown in Table 4-12 below, the shaded cells show the samples that were correctly predicted using the developed models and the remaining cells show the samples that the developed models predicted different chosen modes from the observed chosen modes.

Table 4-12: Observed and predicted modal split of validation data

		Predicted modal split			
Observed modal split		Large bus	Minibus / taxi	Private car	Total
	Large bus	39	4	0	43
	Minibus/ taxi	15	22	1	38
	Private car	0	0	19	19
	Total	54	26	20	100

$$Prediction\ accuracy\ = \frac{39 + 22 + 19}{100} = 80\%$$

4.2.3 Application of the developed models to predict the modal shifts at different scenarios of the revealed preference condition

The developed modal choice models can be applied to analyze and predict the effects of the possible remedy actions to be applied in the increment of the modal share of large buses. The reduction of the total travel time and total travel cost of large buses are some of the possible remedial actions and their possible effects on the modal share of large buses are presented below in Table 4-13.

The effects of decreasing total travel time and total travel cost of large buses were determined using the developed models and the data collected for validation purposes. A total of 16 scenarios were analyzed including no TTT and TTC reductions.

Table 4-13: Percent of large bus users at scenarios with different combinations of TTT and TTC

TTC reduction	TTT reduction			
	0%	20%	40%	60%
0%	54%	54+17=71%	54+23=77%	54+24=78%
20%	54+7=61%	54+18=72%	54+24=78%	54+24=78%
40%	54+14=68%	54+21=75%	54+24=78%	54+27=81%
60%	54+17=71%	54+24=78%	54+26=80%	54+27=81%

The average TTT and the average TTC of large buses of the collected samples are 75 minutes and 3.40 ETB. At these values, using the developed models, the resulted modal share of large buses were 54%. According to the results shown in Table 4-13 above, for example, a scenario resulted in reducing 20% of TTT (decreasing TTT to 60 minutes) and keeping the TTC the same, and another scenario resulted in reducing 60% of TTC (decreasing TTC to 3.40 ETB) and keeping the TTT the same result in similar additional 17% modal share to large buses.

Reducing 20% of both TTT and TTC will eventually increase the modal share of large buses by 18%. Similarly decreasing these two values of large buses by 40 % will result in 78% modal share for large buses. Increment of the modal share of large buses by 27% can be achieved by decreasing 40% of the TTC and 60% of the TTT of large buses. Generally, the scenario analyses results will guide policy makers and transport planners which actions are more efficient and which actions to consider as policies in increasing the modal share of large buses.

4.3 Calibration, interpretation and validation of stated preference model

4.3.1 Parameter estimation, test and comparison of stated preference model

Bus rapid transit mode is a new system in Addis Ababa and in Ethiopia. The probable modal characteristics of this new bus system was stated in the questionnaire and respondents were asked if they would prefer the stated mode keeping the modal characteristics of other alternatives the same as they are. Similar parameter estimation method is used as for the revealed preference modeling.

The new bus rapid transit system is considered as the reference alternative of all the parameter estimation of the independent variables except for the generic variables and for variables with categories, always the last category is the reference category. There are 840 observations and 375 cases out of which 90 of the cases have the four alternatives, BRT, large bus, minibus(taxi), and private car, and the remaining 285 of the cases have three alternatives, BRT, large bus and minibus (taxi).

Similar technique of model estimation and comparison between estimated models is performed as in the revealed modeling. The different levels of models estimated with their deterministic portion of the utility of each mode are presented in detail below.

I. Model 1 and Model 2

Model 1 results of total travel time, total travel cost and the alternative specific constant parameters, and Model 2 results of alternative specific constants and other alternative specific variable parameters are shown in Table 4-14 below. The chi-squared value for Model 2, considering Model 2 as the unrestricted model and Model 1 as the restricted one, is computed as;

$$\chi 2 = 292.603 - 291.706 = 0.897 < 12.565,$$

Where 12.565 is the critical chi-squared value at 95% level of confidence and 6 restricted parameters, above which the null hypothesis that the restricted parameters should be excluded is rejected. Therefore the income variables interacted with the three modes are excluded since the null hypothesis cannot be rejected. Similar to the outputs in Model 2 parameter estimation, other socio economic characteristics of the trip makers resulted in statistically insignificant parameters.

Table 4-14: Parameter estimation results for stated preference, Models 1and 2

		MODEL 1			MODEL 2		
Parameters		**β**	**Sig.**	**Exp(β)**	**B**	**Sig.**	**Exp(β)**
TTT		-0.084	0.000	0.920	-0.084	0.000	0.919
TTC		-0.100	0.032	0.905	-0.106	0.028	0.899
BRT		Base mode					
LBUS		-1.771	0.001	0.170	-1.842	0.087	0.158
TAXI		-1.946	0.000	0.143	-2.217	0.000	0.109
PC		1.376	0.121	3.959	1.395	0.118	4.035
INC_LBUS	INC_LBUS[1]	---------	---------	---------	0.248	0.843	1.281
	INC_LBUS[2]	---------	---------	---------	0.018	0.988	1.019
	INC_LBUS[3]	Reference category					
INC_TAXI	INC_TAXI[1]	---------	---------	---------	0.172	0.822	1.188
	INC_TAXI[2]	---------	---------	---------	0.433	0.522	1.542
	INC_TAXI[3]	Reference category					
INC_PC	INC_PC[1]	No data					
	INC_PC[2]	---------	---------	---------	0.327	0.568	1.386
	INC_PC[3]	Reference category					
-2LL(0)		875.742			875.742		
-2LL(C)		333.511			333.511		
-2LL(β)		292.603			291.706		
$\boldsymbol{\rho_0^2}$		0.666			0.667		
$\boldsymbol{\rho_c^2}$		0.123			0.125		
K		5			11		
$\boldsymbol{\rho_0'^2}$		0.654			0.642		
$\boldsymbol{\rho_c'^2}$		0.093			0.059		

II. Model 1 and Model 3

Total travel time is the sum of in-vehicle travel time and out of vehicle time and trip makers have different sensitivity to these two variables. In order to show this, another model were estimated, Model 3, and compared to Model 1. The results of estimation is presented in Table 4-15, as compared to Model 1. The output resulted in the following significance level at which, if it resulted in a value than α = 0.05, the null hypothesis that the model with lower adjusted rho-squared value is the true model is rejected.

Table 4-15: Parameter estimation results for stated preference, Models 1and 3

	MODEL 1			**MODEL 3**		
Parameters	**β**	**Sig.**	**Exp(β)**	**β**	**Sig.**	**Exp(β)**
IVTT	---------	---------	---------	-0.100	0.000	0.905
OVTT	---------	---------	---------	-0.045	0.182	0.956
TTT	-0.084	0.000	0.920	---------	---------	---------
TTC	-0.100	0.032	0.905	-0.104	0.028	0.901
BRT	Base mode					
LBUS	-1.771	0.001	0.170	-1.919	0.000	0.147
TAXI	-1.946	0.000	0.143	-2.040	0.000	0.130
PC	1.376	0.121	3.959	2.073	0.048	7.951
-2LL(0)	875.742			875.742		
-2LL(C)	333.511			333.511		
-2LL(β)	292.603			289.409		
ρ_0^2	0.666			0.670		
ρ_c^2	0.123			0.132		
K	5			6		
$\rho_0'^2$	0.654			0.656		
$\rho_c'^2$	0.093			0.096		

$$Significance\ level\ = \Phi\left[-\left(-2 * (\rho_H'^2 - \rho_L'^2) * LL(0) + (K_H - K_L)\right)^{0.5}\right] \quad (4\text{-}3)$$

$$= \Phi\left[-\left(\left((0.656 - 0.654) * 875.742\right) + (6 - 5)\right)^{0.5}\right]$$

$$= \Phi\,[-1.659] = 0.0485 < 0.05$$

The result showed that the null hypothesis that the model with lower adjusted rho-squared value is the true model, is rejected. Therefore the model with higher adjusted rho-squared is the true model, i.e., Model 3 is the improved model.

In Model 3, the out-vehicle travel time is not statistically significant since the significance value of the parameter is larger than 0.05. However, this variable is one of the variables that determine the level of service of transport systems, hence the variable retain in the model. As can be seen from Table 4-15 above, the p-value of the parameters in Model 3 has been improved and the all the rho-squared value w.r.t zero model, w.r.t constants only model, and the adjusted rho-squared values has increased. This makes Model 3 the true model for the stated preference model of the study and the utilities of the four modes can be formulated as below.

$$U_{BRT} = -0.100 * \boldsymbol{IVTT} - 0.045 * \boldsymbol{OVTT} - 0.104 * \boldsymbol{TTC} \quad (4\text{-}4)$$

$$U_{LBUS} = -1.919 - 0.100 * \boldsymbol{IVTT} - 0.045 * \boldsymbol{OVTT} - 0.104 * \boldsymbol{TTC}$$

$$U_{TAXI} = -2.040 - 0.100 * \boldsymbol{IVTT} - 0.045 * \boldsymbol{OVTT} - 0.104 * \boldsymbol{TTC}$$

$$U_{PC} = 2.073 - 0.100 * \boldsymbol{IVTT} - 0.045 * \boldsymbol{OVTT} - 0.104 * \boldsymbol{TTC}$$

According to the above formulated stated preference of modes, the decrease in in-vehicle travel time and travel cost of a particular mode apparently increases the preference of the mode. Since their magnitudes are more or less similar, a unit decrease of in-vehicle travel time has similar impact on the increase of the utility of the mode as the impact in decrease of a unit of total travel cost of the mode. This can be applied in the fare structure and arrival frequency design of the BRT system which is under design and installation now.

4.3.2 Validation of stated preference model

The estimation of the stated preference model output showed that the sign of the parameters of the generic variables; travel time and the travel cost is negative implying that the utility of modes towards the trip makers decrease as the travel time and travel cost increase. Table 4-16 below, shows the validation data of the stated preference model partially.

Table 4-16: Example of validation of the stated preference model

Trip #	Stated modes	Utility	Probability of choosing a mode	predicted choice	Observed choice	Observed versus predicted choices
1	large bus	-6.299	0.05	0	0	
1	taxi	-6.354	0.02	0	0	
1	BRT	-2.915	0.93	1	1	Similar
2	taxi	-5.834	0.06	0	0	
2	large bus	-7.755	0.01	0	0	
2	BRT	-2.915	0.93	1	1	Similar
3	taxi	-6.354	0.02	0	0	
3	large bus	-8.379	0.00	0	0	
3	BRT	-2.915	0.97	1	1	Similar
4	taxi	-5.834	0.06	0	0	
4	large bus	-8.119	0.01	0	0	
4	BRT	-2.915	0.93	1	1	Similar

According to the validation of the stated preference model shown in Table 4-16, and the prediction ratio of the model computed below, the model is capable of predicting 92% of the stated preference of the trip makers. The observed and predicted stated modes are presented in Table 4-17 below. The cells that correctly predicted the modal choice of the validation samples are shaded.

$$Prediction\ accuracy\ = \frac{82 + 0 + 0 + 10}{100} = 92\%$$

Table 4-17: Observed and predicted modal split of validation of stated preference data

		Predicted modal split				
		BRT	Large bus	Minibus / taxi	Private car	Total
Observed modal split	BRT	82	0	0	2	84
	Large bus	1	0	0	0	1
	Minibus/ taxi	3	0	0	0	3
	Private car	2	0	0	10	12
	Total	88	0	0	12	100

4.3.3 Application of the developed models to predict the modal shifts at different scenarios of the stated preference condition

The stated preference in this study includes the bus rapid transit system in addition to the present available transport modes. The bus rapid transit system is expected to have an average in vehicle travel time, out of vehicle travel time and total travel cost of 28.35 minutes, 13.99 minutes and 3.37 ETB, respectively. The BRT was found to have 88% modal share at these values. In this section a total of 6 scenarios were analyzed with cases that increase and decrease the modal share of BRT system, using the collected data for validation purposes and the developed model to predict the preferred mode of the trip makers.

Scenario #1: decrease the IVTT of the BRT by 70%

The decrease in the IVTT of the BRT by 70% results in an IVTT near to 20 minutes. Keeping other characteristics of the mode similar, the modal share of the BRT will increase by 1% only.

Scenario #2: decrease the IVTT by 70%, the OVTT by 40%and the TTC by 70%

The resulting IVTT, OVTT and TTC will be 20 minutes, 5.6 minutes and 2.4 ETB, respectively. This will eventually increase the modal share of the BRT system by 4%, making the total modal share of the BRT 92%. The further decrease in these values can be easily analyzed with the use of the developed model and will result in further increment of the modal share of the BRT system.

Scenario #3: increase the IVTT by 70%

The resulting IVTT is 48.2 minutes, and keeping all other characteristics the same, the modal share of the BRT system will decrease by 18% and result in 70% modal share.

Scenario #4: increase the OVTT by 60%

The resulting OVTT is 22.4 minutes which includes the access time and waiting time at stations. Keeping all other characteristics the same, the modal share of the BRT system will be 86% with a decrease in 2% only.

Scenario #5: increase the TTC by 50%

The resulting TTC is near to 5 ETB. Keeping all other characteristics the same, at this scenario the modal share of the BRT system will not vary and it will stay 88%.

Scenario #6: increase the IVTT by 50%, OVTT by 60%, TTC by 30%

The combination of the above three conditions will eventually decrease the modal share of the BRT system by 24%. This means the BRT with average IVTT of 48.2 minutes, OVTT of 22.4 minutes and TTC of 5 ETB, its modal share is expected to be 61%.

CHAPTER 5. CONCLUSIONS, RECOMMENDATION AND FUTURE RESEARCH AREAS

5.1 Conclusions

The research aims generally to model the modal preference of Jemo area residents and determine how the service quality of public transportation system in that area can be improved so that commuters' choice will shift to the public transportation systems. This generally contributes for the improvement of the service quality of public transportation in the study area and in Addis Ababa as a whole.

Based on the data analysis, 42% of the trip makers use large buses, 36 % of the trip makers use taxis or minibuses and the remaining 22% of the trip makers use private cars under revealed preference survey. For the stated preference models, the data summary showed that the percentage modal share of BRT, private cars, taxis and large buses are 80.84%, 12.84%, 4.48% and 1.48%, respectively.

The developed models for revealed preference survey using the conditional multinomial logit modeling techniques in SPSS are tested and compared with a series of goodness of fitness tests at 95% level of confidence level.

The resulted revealed preference models showed that the total travel time and total travel cost of the mode, interaction variable of car ownership, local time of leaving home for work and education level interacted with taxi are the factors that are statistically significant at 95% level of confidence. Moreover, the developed models showed that decreasing a unit travel cost has a better effect than decreasing a unit travel time of a transport mode to increase its utility.

The developed models for the revealed preference data are validated to be capable of predicting 80% of the preference of modes of trip makers of the study area.

In the stated preference models, the in-vehicle travel time, out of vehicle travel time and total travel cost are the only factors that are found to be statistically significant at 95% level of confidence for the stated preference of modes.

The models indicated that for the situation in which BRT is included in the choice set of trip makers, to increase the utility of the modes, decreasing a unit in-vehicle travel time has more effect than decreasing a unit out of vehicle travel time. It is also apparent that the in-vehicle travel time and total travel cost has similar effects on the utility of the modes.

For the stated preference of modes, the model is validated to be capable of describing 92% of the preference of trip makers among the stated modes.

5.2 Recommendations

- Discrete choice models as developed in this research are tools that are beneficial to transport planners and policy makers in estimation of the impact of changes of the modal attributes and the socio economic characteristics of the trip makers in the area.
- The developed models also guide how the modal attributes; travel time and travel cost of public transport should be modified so that the preference of trip makers towards them will be increased.
- It is also recommended for the collected data to be fitted with other statistical mode choice methods and with the use of soft computing techniques such as artificial neural network (ANN), fuzzy logic and hybrid models, and compare them with the developed models in this study.
- The variables related to comfort, safety, availability and reliability are also factors of transport mode choice. Therefore, the researcher highly recommends the inclusion of these factors in further studies related to this research.
- For developing countries like Ethiopia in which there are extensive projects planned and going on in improving urban transport system, modal choice behavior of the users should be studied.
- For practical applications of modal choice models, the researcher recommends development of the models for the whole city, considering all types of trips and the whole combinations of modes using this study as a starting point.

5.3 Future research areas

The following are research areas that are recommended to be conducted in the future.

- Determination of the level of importance of variables related to modal choice of different types of trips for trip makers in Addis Ababa;
- Development of modal choice models for home and non-home based, non-work trips in Addis Ababa;
- Development of modal choice models for home based work trips including the variables related to comfort, safety, availability and reliability of transport modes; and
- Development of modal choice models using statistical modeling techniques and soft computing techniques like artificial neural network techniques and compare the two results.

REFERENCE

Akimasa, F., Janyi, Z., Toshiyuki, O. & Soe, T., 2003. Analysis of mode choice preferences under changes in travel and socio-economic environments in Yangon city based on stated preference survey. *JSCE*, pp. 1-4.

Amit, V., 2017. A study on travel choices of college aged students. *International Journal for Scientific Research & Development*, 5(07-2017), pp. 66-70.

Anon.,2016.*AddisAbabaOnline.com.*[Online] Available at: http://addisababaonline.com/anbessa-plans-to-shell-out-950m-br-for-new-buses/ [Accessed 10 6 2018].

Ashu S. Kedia & B.K. Katti, 2014. DISCRETE MODE CHOICE MODELLING THROUGH FRBS: A CASE STUDY OF METROPOLITAN CITY. *Research Gate*, p. 11.

Baltes , M. R., 2002. *Customer Surveying for Public Transit: a design manual for on board surveys,* Florida: s.n.

Cameron, C. A. & Trivedi , P. K., 2009. Microeconometric using stata. In: *Multinomial models.* Texas: Stata Press, pp. 477-519.

Ch.Ravi Sekhar, Minal & Errampalli Madhu, 2016. MULTIMODAL CHOICE MODELING USING RANDOM FOREST. *Research Gate,* p. 13.

Chandra R.Bhat, 2000. *Random Utility-Based Discrete Choice Models for Travel Demand Analysis.* 1 ed. Texas: s.n.

D. Hoffman , S. & J. Duncan, G., 1988. *Multinomial and Conditional Logit Discrete-Choice Models in Demography.* vol. 25 ed. Michigan: Population Association of America.

David, H. W. & Stanley, L., 2000. *Applied Logistic Regression.* 2nd ed. Canada: A Wiley-Interscience.

Dmitry Pavlyuk & Vaira Gromule, 2010. *DISCRETE CHOICE MODEL FOR A PREFERRED,* Latvia: Transport and Telecommunication Institute.

Eluru, N., Chakour, V. & El-Geneidy, A. M., 2012. Travel Mode Choice and Transit Route Choice Behavior in Montreal: Insights from McGill University Members Commute Patterns. *Public Transport: Planning and Operations,* 4(2), pp. 129-149.

Eshete, M., 2015. PUBLIC TRANSPORTATION SYSTEM: THE CASE OF ADDIS ABABA. p. 100.

Essam , A. & Sadi , A., 2013. Factors Affecting Mode Choice of Work Trips in Developing Cities—Gaza as a Case Study. *Journal of Transportation Technologies,* June .pp. 247-259.

Ethiopian Ministry of Transport, 2011. *Transport Policy of Addis Ababa,* Addis Ababa: s.n.

Etsub , M. H., 2015. *Logit Model for School and Work Trip Mode Choice On Railway Route Corridor.* Addis Ababa: s.n.

FHWA, 2016. *The Transportation Planning Process Key Issues.* Washington: s.n.

Glenn , I. D., 1992. *Fact Sheet PEOD-6,* Florida: Institute of Food and Agricultural Sciences, University of Florida.

Green, S. B., 1991. How many subjects does it take to do a regression analysis?. In: *Multivariate behavioural research.* s.l.:s.n., pp. 499-510.

H.Mariam, K., 2017. *Ethiopian Initiatives to Cleaner Public Transport.* Accra, Ghana, s.n., p. 28.

Juan Li, et al., 2010. *Travel Mode Choice Modeling Based on Improved Probabilistic Neural Network.* Beijing , Traffic and Transportation studies.

Kenneth, E. T., 2009. *Discrete Choice Methods with Simulation.* 2nd ed. New York: Cambridge University Press.

Keuchela, S. & Cornelia Richter, 2011. Applying Integrated Hierarchical Information Integration to. *Elsevier,* p. 11.

Kohlová, B. & Markéta, 2008. Everyday travel mode choice and its determinants: trip attributes versus lifestyle. p. 11.

Koppleman, F. S. & Chandra, B., 2006. *A Self Instructing Course in Mode Choice Modeling:*. Georgia: s.n.

Manssour A. , A. B. M., Ahmed , . A. M. & Riza Atiq , A. B. O. R. ., 2013. Modeling a multinomial logit model of intercity travel mode choice behavior for all trips in Libya. *International Scholarly and Scientific Research & Innovation*, p. 10.

Mekidim, T., 2007. *LOGIT MODEL OF WORK TRIP MODE choice for bole-subcity residents*, Addis Ababa: s.n.

Memon, I. A., Madzlan Napiah, Mir Aftab Hussain Talpur & M. Rehan Hakro, 2016. MODE CHOICE MODELLING METHOD TO SHIFT CAR TRAVELERS. *Research Gate*, p. 8.

Minal & Ch. Ravi , S., 2014. Mode choice analysis: the data, the models and future ahead. *International Journal for Traffic and Transport Engineering*, pp. 269-285.

Mintesnot, G. & Shin-ei, T., 2008. Diagnostic evaluation of public transportation mode choice in Addis Ababa. *Transportation Research Forum*, 10(4), pp. 27-50.

Moshe, B.-A. & Michel, B., 1999. *Discrete Choice Methods and Their Applications to Short Term Travel Decisions.* Massachusetts: s.n.

Nicholas, J. G. & Lester, A. H., 1997. *Traffic and Highway Engineering.* Massachusetts: PWS publidhing company.

Omer , A. K., Luis , F., Jonathan , B. & Partha, P., 2004. Modelling Passenger Mode Choice Behaviour Using Computer Aided. *Research Gate*, p. 96.

Ortuzar, J. d. D. & Willumsen, L. G., 2011. *Modelling Transport.* 3rd ed. United Kingdom: John Wiley & Sons, Ltd.

R.D. Luce, R.R. Bush & E. Galanter, 1965. *Preference, utility and subjective probability.*. New York: John Wiley & Sons, Inc.

Ram, C., Septarshi, s. & Sudip, K. R., 2016. *Mode choice modeling of work trips: a case study of Kolkata.* Kolkata, s.n.

Rosa, A. G. & Luigi , S., 2003. Model performance analysis and model validation. *STATISTICA*, pp. 375-396.

Stephane , H. & Andrew, D., 2010. *Choice modelling the state of the art and the state of practice.* Bingley, Emerald Group Publishing Limited.

Temesgen, A., 2006. The role of the ransport sector in Ethiopia's economic development. *Economic Focus of Ethiopia Economic Association*, p. 30.

UNHABITAT, 2017. *The State of Addis Ababa,* Nairobi: s.n.

APPENDIX- A

I. Questionnaire in English

Addis Ababa University
Addis Ababa Institute of Technology
School of Civil and Environmental Engineering

Questionnaire for a research titled with '**Development of Mode Choice Model for Home-Based Work Trips-The Case of Jemo Area Residents in Addis Ababa, Ethiopia'**

This questionnaire aims to study the situation of transport system and the factors that affect the employed people's choice for transportation modes in Addis Ababa by developing a mathematical mode choice model for work trips from Jemo area.
Please cooperate by filling the following questionnaire with facts, having in mind that this information will be used for scientific study only and will be treated confidentially.

The respondent chosen to fill this questionnaire should satisfy the following criteria.

1) The person should have a job and should not have a transport service from the company he/she works.
2) The person mostly leaves home between 12:00-3:00 AM local time.
3) The trip of the respondent should be from Jemo to Mexico or beyond.

State special name of the respondents HH location ………………………………………………….

PREPARED BY: FIRDEWS ALI
SUPERVIZED BY: DR. BIKILA TEKLU

Part I. Socio-economic data

1. Age ……………… years

2. Gender Female [ ] Male [ ]

3. What is your marital status?

Married [ ] Single [ ] Divorced [ ] widowed [ ] widower [ ]

4. Education level?

8th grade and below [ ] high school to preparatory / 9th -12th grade [ ] Diploma [ ] Degree [ ] Masters and above [ ]

5. What is your occupation;

Government Employee [ ] Private organization Employee [ ] NGO and International Organization Employee [ ] Employer [ ] Self-Employed [ ] Other Employee [ ]

6. How many people live in your family? ………………………………

7. What is your family monthly income?

Less than 1,000 Birr [ ] 1,001-3,000 Birr [ ] 3,001-5,000 Birr [ ] 5,001-10,000 Birr [ ] 10,001-15,000Birr [ ] 15,001-20,000 Birr [ ] 20,001-30,000 Birr [ ] Greater than 30,000Birr [ ]

8. How many cars are there in your household? **…………………………**

9. Do you have driving license? Yes [ ] No [ ]

Part II. Trip characteristics

1. At what time of the day do you usually leave your home for work? **…………………………..**

2. What are the available alternatives for you to travel from Jemo to Mexico?

Private car as driver [ ] Private car as passenger [ ] Taxi [ ]

Anbessa Bus [ ] Alliance Bus [ ] Sheger Bus [ ]

Which mode of transportation do you take to go to work from your home **most of the time**? **………**

Please specify why you chose that mode **……………………………………………………………………**

3. From the alternatives listed below choose the alternative you use most of the time and answer the respective questions about your trip from Jemo to Mexico using that alternatives.

a) Anbessa Bus [ ] Alliance Bus [ ] Sheger Bus [ ]

- Walk time and access time (if you use other modes like Bajaj before your major mode) from home to nearby station in minutes ……………………………………………………………
- Access cost (if you use other modes like Bajaj) in Birr and Cents**………………………………………**
- Time waiting a bus in minutes **…………………………………………………………………**
- Travel time in the bus in minutes **…………………………………………………………**
- Fare in Birr **…………………………………………………………………………………………….**

b) Taxi/ Minibus

- Walk time and access time (if you use other modes like Bajaj before your major mode) from home to nearby station in minutes ………………………………………………………….
- Access cost (if you use other modes like Bajaj) in Birr and Cents……………………………………….
- Time waiting for a taxi in minutes ………………………………………………….
- Travel time in the taxi in minutes ………………………………………………….
- Fare in Birr ……………………………………………………………………….

c) Private car as driver

- Travel time in minutes …………………………………………………………………….
- Type of fuel the car uses ………………………….
- How many kilometers you travel with 1 liter fuel? ……………………….
- How often do you give your car for maintenance and how much do you pay in Birr? ……………………………………………………………………………………………….
- Which route do you usually take to go to work from Jemo to Mexico?

d) Private car as passenger

- Travel time in minutes ……………………………………………………………

Part III. Stated preference survey

I. Are you willing to shift to Bus Rapid Transit with the following attributes for trips to work from Jemo to Mexico?

Attributes	Bus Rapid Transit
Travel time	Less than 25-35minutes
Travel cost	Similar to current Bus fare
Arrival frequency	Every 5 minutes
Route	From Jemo through German Square, Gofa-Gebriel Roundabout, Kera Junction to Mexico Junction.

Yes [ ] **No** [ ]

II.Questionnaire in Amharic

አዲስ አበባ ዩኒቨርሲቲ

አዲስ አበባ የቴክኖሎጂ ተቋም

የሲቪል እና የኢንቫይሮመንታል ምህንድስና ትምህርት ቤት

'አዲስ አበባ ዉስጥ ከቤት ወደ ስራ ለሚደረጉ ጉዞዎች መጓጓዣ አመራረጥ ሞዴል ዝግጅት' በሚል ርእስ ለሚደረገው የመመረቂያ ፅሁፍ ጥናት የተዘጋጀ መጠይቅ/ መጋቢት 2010 ዓ.ም

- ይህ መጠይቅ በአዲስ አበባ ዉስጥ ያለውን የትራንስፖርት ሁኔታ እና ወደ ስራ የሚሄዱ ሰዎች መጓጓዣ ሲመርጡ ተፅእኖ ሊያሳድሩ ይሚችሉ ነገሮችን ሂሳባዊ ሞዴል በማዘጋጀት ጥናት ለማድረግ ያለመ ነው.
- ይህ መረጃ ለሳይንሳዊ ጥናት ብቻ እንደሚያገለግል በማሰብ እባክዎ የሚከተለውን መጠይቅ በእውነታዎች በመሙላት ይተባበሩን.

ይህን መጠይቅ ለመሙላት የተመረጡት ምላሽ ሰጪዎች የሚከተሉትን መስፈርቶች ሊያሟሉ ይገባል.

1) ግለሰቡ ስራ ያለው እና ከሚሰራበት ድርጅት የመጓጓዣ አገልግሎት የሌለው መሆን አለበት.
2) ግለሰቡ በአብዛኛው ከጠዋቱ 12 ሰአት እስከ 3 ሰአት ጀምሮ ከቤቱ ወደ ስራ የሚወጣ መሆን አለበት
3) ጉዞው ከጀሞ እስከ ሜክሲኮ ወይም ሜክሲኮን አልፎ የሚሄድ መሆን አለበት. ነገር ግን መጠይቁ ከጀሞ እስከ ሜክሲኮ ያለውን ጉዞ ብቻ ይመለከታል.

ይህን መጠይቅ ለመሙላት የተመረጡትን ምላሽ ሰጪ የመኖሪያ አድራሻ ………………

አዘጋጅ: ፊርደውስ አሊ

ሱፐርቫይዘር: ዶ/ር ቢቂላ ተክሉ

ክፍል 1. ማህበራዊ እና ኢኮኖሚያዊ መረጃ

1. ዕድሜ .... ዓመት

2. ጾታ

   ሴት [ ]

   ወንድ [ ]

3. የጋብቻ ሁኔታ

   ያገባ [ ]

   ያላገባ [ ]

   የተፋታ [ ]

   ባል የሞተባት/ ሚስቱ የሞተችበት [ ]

4. የትምህርት ደረጃ

   8 ኛ ክፍል እና ከዚያ በታች [ ]

   ከሁለተኛ ደረጃ እስከ መሰናዶ /9 ኛ -12 ኛ ክፍል/ [ ]

   ሰርቲፊኬት እና ዲፕሎማ [ ]

   ዲግሪ [ ]

   ማስተርስ እና ከዛ በላይ [ ]

5. የስራ አይነት

   የመንግስት ሰራተኛ [ ]

   የግል ድርጅት ሠራተኛ [ ]

   NGO እና ዓለምአቀፋዊ ድርጅት ሰራተኛ [ ]

   ቀጣሪ [ ]

   የግል ስራ [ ]

   ሌላ ስራ [ ]

6. በቤተሰብዎ ውስጥ ስንት ሰዎች ይኖራሉ? ....

7. ጠቅላላ የቤተሰብዎ ወርሃዊ ገቢ

   ከ 1,000 ብር በታች [ ]

   ከ 1,000 እስከ 3,000 ብር [ ]

   ከ 3,001 እስከ 5,000 ብር [ ]

   ከ 5,001 እስከ 10,000 ብር [ ]

   ከ 10,001 እስከ 15,000 ብር [ ]

   ከ 15,001 እስከ 20,000 ብር [ ]

   ከ 20,001 እስከ 30,000 ብር [ ]

   ከ 30,000 ብር በላይ [ ]

8. በቤትዎ ውስጥ ስንት መኪናዎች አሉ? ………………………………………………..

9. የመንጃ ፈቃድ አለዎት? አዎ [ ] አይ [ ]

ክፍል 2. የጉዞ ባህሪያት

1. አብዛኛውን ጊዜ ወደ ስራ ለመሄድ ከቤቶ ስንት ሰአት ይወጣሉ? ....

2. ከጀሞ እስከ ሜክሲኮ ድረስ ለመጓዝ ለርስዎ ያሉ አማራጮች የትኞቹ ናቸው?

በግል መኪና እየነዱ [ ] በግል መኪና እንደ ተሳፋሪ [ ]

ታክሲ/ ሚኒባስ [ ] አንበሳ አውቶቡስ [ ]

አሊያንስ አውቶቡስ [ ] ሸገር አውቶቡስ [ ]

አብዛኛውን ጊዜ የትኛውን የትራንስፖርት አማራጭ ከቤትዎ ወደ ስራ ለመሄድ ይጠቀማሉ? …………………...

እባክዎን ያንን አማራጭ ለምን እንደመረጡ ያስፍሩ ………………………….……………...

3. ከታች ከተጠቀሱት አማራጮች ውስጥ **አብዛኛውን ጊዜ ከጀሞ ወደ ሜክሲኮ** ለመጓዝ የሚጠቀሙበትን አማራጭ **ብቻ** መርጠው ጥያቄዎቹን ይመልሱ

ሀ) በአንበሳ አውቶቡስ [ ] በአሊያንስ አውቶቡስ [ ] በሸገር አውቶቡስ [ ]

- በእግር እና/ ወይም በባጃጅ ከበቤትዎ እስከ አውቶቡስ መሳፈሪያ ድረስ የሚጉዋዙበት ጊዜ በደቂቃ …………
- ከቤትዎ እስከ አውቶቡስ መሳፈሪያ ድረስ ባጃጅ ከተጠቀሙ ምን ያህል ይከፍላሉ በብር እና በሳንቲም…………..
- አውቶቡስ የሚጠብቁበት ጊዜ በደቂቃ………………………………..................
- በአውቶቡስ ውስጥ ሆነው ጠቅላላ የጉዞ ጊዜ ከጀሞ እስከ ሜክሲኮ (በደቂቃ)………..
- የአውቶቡስ ክፍያ በብር………………………………………………………

ለ) በታክሲ / ሚኒባስ

- በእግር እና/ ወይም በባጃጅ ከበቤትዎ እስከ ታክሲ መሳፈሪያ ድረስ የሚጉዋዙበት ጊዜ በደቂቃ ……………….
- ከቤትዎ እስከ ታክሲ መሳፈሪያ ድረስ ባጃጅ ከተጠቀሙ ምን ያህል ይከፍላሉ በብር እና በሳንቲም …………...

- ታክሲ የሚጠብቁበት ጊዜ በደቂቃ ………………………………………………
- በታክሲ ውስጥ ሆነው ጠቅላላ የጉዞ ቆይታ ከጀሞ እስከ ሜክሲኮ (በደቂቃ)……
- የታክሲ ክፍያ በብር …………………………………………………………

ሐ) በግል መኪና እየነዱ

- ከጀሞ እስከ ሜክሲኮ ስንት ደቂቃ ይፈጃል? ……………………………………
- መኪናዎት ምን አይነት ነዳጅ ይጠቀማል? ……………………………………
- በ 1 ሊትር ነዳጅ የሚጉዋዙት ርቀት መጠን በኪሎሜትር…………………….
- መኪናዎትን በየ ስንት ጊዜው ያስጠግናሉ እና ስንት ይከፍላሉ?......................
- አብዛኛውን ጊዜ የትኛውን መንገድ ከቤት ወደ ሥራ ለመሄድ ይጠቀማሉ? ....….

መ) በግል መኪና እንደ ተሳፋሪ

- ከጀሞ እስከ ሜክሲኮ ስንት ደቂቃ ይፈጃል?…………………………………..

ክፍል 3. አዲስ የአውቶቡስ አማራጭ በተለየ አቀራረብ በቅርቡ አገልግሎት መስጠትይጀመራል. ከጀሞ እስከ ሜክሲኮ ለመጓዝ የሚከተሉትን ባህርያት ይይዛል.

ባህርያት	አዲስ የአውቶቡስ አማራጭ
የጉዞ ርዝመት በደቂቃ	ከ30 ደቂቃዎች ያነሰ
የጉዞ ክፍያ	ከሌሎች አውቶቡሶች ጋር ተመሳሳይ
በየስንት ደቂቃ አውቶቡስ ይመጣል	በየ 3 ደቂቃ
ከጀሞ እስከ ሜክሲኮ ለመጓዝ አውቶብሱ የሚከተለው መንገድ	ከጀሞ ተነስቶ በጀርመን አደባባይ፣ በጎፋ-ገብርኤል ማዞሪያ፣ በቄራ መስቀለኛ አድርጎ ሜክሲኮ

ከጀሞ እስከ ሜክሲኮ ወደ ስራ ለመጓዝ አሁን የሚጠቀሙበትን አማራጭ ወደ አዲሱ አማራጭ ለመቀየር ፈቃደኛ ይሆናሉ?

አዎ [ ] አይ [ ]

APPENDIX-B

Critical chi-squared (χ^2) values for selected level of confidence by number of restrictions

Level of Confidence	Number of Restrictions								
	1	2	3	4	5	7	10	12	15
90%	2.71	4.61	6.25	7.78	9.24	12.01	15.99	18.54	22.31
95%	3.84	5.99	7.81	9.49	11.07	14.06	18.31	21.02	25.00
99%	6.63	9.21	11.34	13.28	15.09	18.48	23.21	26.21	30.58
99.5%	7.88	10.60	12.84	14.86	16.75	20.28	25.19	28.30	32.80
99.9%	10.83	13.82	16.27	18.47	20.51	24.32	29.59	32.91	37.70

Source: (Koppleman & Chandra, 2006)

APPENDIX-C

I. Revealed preference modal choice SPSS parameter estimation results

Case Processing Summary

		N	Percent
Cases available in analysis	Event	375	44.6%
	Censored	465	55.4%
	Total	840	100.0%
Cases dropped	Cases with missing values	0	0.0%
	Cases with negative time	0	0.0%
	Censored cases before the earliest event in a stratum	0	0.0%
	Total	0	0.0%
Total		840	100.0%

Base model

Omnibus Tests of Model Coefficients

-2 Log Likelihood (-2 LL_0)
592.844

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
445.982	147.078	2	.000	146.862	2	.000	146.862	2	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
TAXI	-.145	.118	1.513	1	.219	.865	.687	1.090
PC	3.263	.426	58.630	1	.000	26.115	11.330	60.197

I. Model 1

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
404.752	169.288	4	.000	188.092	4	.000	188.092	4	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
TTT	-.043	.010	16.982	1	.000	.958	.939	.978
TTC	-.371	.111	11.146	1	.001	.690	.555	.858
TAXI	-.246	.379	.423	1	.516	.782	.372	1.642
PC	9.884	2.853	12.002	1	.001	19607.180	73.109	5.3E6

II. Model 2

Categorical Variable Codings[a]

		Frequency	(1)	(2)	(3)
INC_TAXI[b]	.00	465	1	0	0
	1.00	92	0	1	0
	2.00	175	0	0	1
	3.00	108	0	0	0

a. Category variable: INC_TAXI

b. Indicator Parameter Coding

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
399.240	174.686	6	.000	193.604	6	.000	193.604	6	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
TTT	-.043	.010	17.29	1	.000	.958	.938	.977
TTC	-.376	.113	11.05	1	.001	.686	.550	.857
TAXI	-.069	.487	.020	1	.888	.934	.360	2.424
PC	10.081	2.909	12.01	1	.001	23878.669	79.827	7142794.36
INC_TAXI			5.391	2[a]	.068			
INC_TAXI(2)	-.608	.387	2.467	1	.116	.544	.255	1.163
INC_TAXI(3)	.023	.354	.004	1	.949	1.023	.511	2.047

a. Degree of freedom reduced because of constant or linearly dependent covariates

III. Model 3

Categorical Variable Codings[a]

		Frequency	(1)	(2)
INC_PC[b]	.00	750	1	0
	2.00	22	0	1
	3.00	68	0	0

a. Category variable: INC_PC
b. Indicator Parameter Coding

Omnibus Tests of Model Coefficients

	Overall (score)			Change From Previous Step			Change From Previous Block		
-2 Log Likelihood	Chi-square	df	Sig..	Chi-square	df	Sig.	Chi-square	df	Sig.
397.765	170.003	5	.000	195.079	5	.000	195.079	5	.000

Variables in the Equation

							95.0% CI for Exp(B)	
	B	SE	Wald	df	Sig.	Exp(B)	Lower	Upper
TTT	-.041	.010	16.029	1	.000	.959	.940	.979
TTC	-.410	.120	11.738	1	.001	.664	.525	.839
TAXI	-.131	.388	.113	1	.736	.878	.411	1.876
PC	13.281	3.929	11.425	1	.001	585951.229	265.089	1.3E9
INC_PC			4.014	1[a]	.045			
INC_PC(2)	-3.604	1.799	4.014	1	.045	.027	.001	.925

a. Degree of freedom reduced because of constant or linearly dependent covariates

IV. Model 4

Categorical Variable Codings[a]

		Frequency	(1)	(2)
INC_PC[b]	.00	750	1	0
	2.00	22	0	1
	3.00	68	0	0

a. Category variable: INC_PC
b. Indicator Parameter Coding

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
396.208	170.320	7	.000	196.636	7	.000	196.636	7	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	Upper
TTT	-.041	.010	15.526	1	.000	.960	.941	.980
TTC	-.443	.126	12.283	1	.000	.642	.501	.823
TAXI	.054	.430	.016	1	.900	1.056	.455	2.450
PC	13.085	3.925	11.117	1	.001	481867.671	219.94	1.1E9
INC_PC			3.543	1[a]	.060			
INC_PC(2)	-3.464	1.840	3.543	1	.060	.031	.001	1.153
MALE_TAXI	-.170	.246	.479	1	.489	.844	.521	1.366
MALE_PC	1.356	1.391	.950	1	.330	3.879	.254	59.266

a. Degree of freedom reduced because of constant or linearly dependent covariates

V. *Model 5*

Categorical Variable Codings[a,c]

		Frequency	(1)	(2)
AGE_TAXI[b]	.00	465	1	0
	1.00	340	0	1
	2.00	35	0	0
INC_PC[b]	.00	750	1	0
	2.00	22	0	1
	3.00	68	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: INC_PC

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score) Chi-square	df	Sig.	Change From Previous Step Chi-square	df	Sig.	Change From Previous Block Chi-square	df	Sig.
392.171	173.897	6	.000	200.673	6	.000	200.673	6	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	Upper
TTT	-.044	.011	17.115	1	.000	.957	.938	.977
TTC	-.420	.120	12.343	1	.000	.657	.519	.830
TAXI	-1.045	.565	3.415	1	.065	.352	.116	1.065
PC	13.547	3.961	11.700	1	.001	764570.734	325.205	1.8E9
INC_PC			4.009	1[a]	.045			
INC_PC(2)	-3.682	1.839	4.009	1	.045	.025	.001	.925
AGE_TAXI			5.159	1[a]	.023			
AGE_TAXI(2)	.985	.434	5.159	1	.023	2.679	1.145	6.268

a. Degree of freedom reduced because of constant or linearly dependent covariates

VI. Model 6

Categorical Variable Codings[a,c,d]

		Frequency	(1)	(2)
AGE_TAXI[b]	.00	465	1	0
	1.00	340	0	1
	2.00	35	0	0
AGE_PC[b]	.00	750	1	0
	1.00	84	0	1
	2.00	6	0	0
INC_PC[b]	.00	750	1	0
	2.00	22	0	1
	3.00	68	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: AGE_PC
d. Category variable: INC_PC

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
392.156	174.092	7	.000	200.689	7	.000	200.689	7	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
TTT	-.044	.011	17.083	1	.000	.957	.938	.977
TTC	-.422	.121	12.291	1	.000	.655	.518	.830
TAXI	-1.036	.569	3.316	1	.069	.355	.116	1.082
PC	13.881	4.836	8.237	1	.004	1067541.347	81.580	1.4E10
INC_PC			3.969	1[a]	.046			
INC_PC(2)	-3.696	1.855	3.969	1	.046	.025	.001	.942
AGE_TAXI			5.093	1[a]	.024			
AGE_TAXI(2)	.982	.435	5.093	1	.024	2.668	1.138	6.259
AGE_PC			.016	1[a]	.901			
AGE_PC(2)	-.293	2.352	.016	1	.901	.746	.007	74.928

a. Degree of freedom reduced because of constant or linearly dependent covariates

VII. Model 7

Categorical Variable Codings[a,c]

		Frequency	(1)	(2)
AGE_TAXI[b]	.00	465	1	0
	1.00	340	0	1
	2.00	35	0	0
INC_PC[b]	.00	750	1	0
	2.00	22	0	1
	3.00	68	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: INC_PC

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	95.0% CI for Exp(B) Upper
TTT	-.043	.011	15.945	1	.000	.958	.939	.979
TTC	-.451	.124	13.161	1	.000	.637	.499	.813
TAXI	-1.101	.581	3.595	1	.058	.332	.107	1.038
PC	15.191	4.217	12.978	1	.000	3955741.813	1018.562	1.5E10
INC_PC			3.630	1[a]	.057			
INC_PC(2)	-3.735	1.961	3.630	1	.057	.024	.001	1.113
AGE_TAXI			5.425	1[a]	.020			
AGE_TAXI(2)	1.038	.446	5.425	1	.020	2.823	1.179	6.760
CAR_TAXI	1.382	.500	7.640	1	.006	3.984	1.495	10.618

a. Degree of freedom reduced because of constant or linearly dependent covariates

VIII. Model 8

Categorical Variable Codings[a,c]

		Frequency	(1)	(2)
AGE_TAXI[b]	.00	465	1	0
	1.00	340	0	1
	2.00	35	0	0
INC_PC[b]	.00	750	1	0
	2.00	22	0	1
	3.00	68	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: INC_PC

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	Upper
TTT	-.043	.011	15.967	1	.000	.958	.939	.979
TTC	-.438	.122	12.987	1	.000	.645	.508	.819
TAXI	-1.129	.578	3.819	1	.051	.323	.104	1.003
PC	16.816	25.086	.449	1	.503	2.0E8	.000	4.5E28
INC_PC			3.458	1[a]	.063			
INC_PC(2)	-3.361	1.807	3.458	1	.063	.035	.001	1.199
AGE_TAXI			5.412	1[a]	.020			
AGE_TAXI(2)	1.035	.445	5.412	1	.020	2.816	1.177	6.738
CAR_TAXI	1.374	.498	7.613	1	.006	3.950	1.489	10.479
CAR_PC	-2.324	24.934	.009	1	.926	.098	.000	1.6E20

a. Degree of freedom reduced because of constant or linearly dependent covariates

IX. Model 9

Categorical Variable Codings[a,c,d]

		Frequency	(1)	(2)	(3)	(4)
AGE_TAXI[b]	.00	465	1	0		
	1.00	340	0	1		
	2.00	35	0	0		
INC_PC[b]	.00	750	1	0		
	2.00	22	0	1		
	3.00	68	0	0		
TIME_TAXI[b]	.00	465	1	0	0	0
	1.00	168	0	1	0	0
	2.00	94	0	0	1	0
	3.00	83	0	0	0	1
	4.00	30	0	0	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: INC_PC
d. Category variable: TIME_TAXI

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score) Chi-square	df	Sig.	Change From Previous Step Chi-square	df	Sig.	Change From Previous Block Chi-square	df	Sig.
373.796	185.463	10	.000	219.048	10	.000	219.048	10	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	95.0% CI for Exp(B) Upper
TTT	-.046	.011	18.033	1	.000	.955	.935	.975
TTC	-.460	.126	13.366	1	.000	.631	.493	.808
TAXI	.056	.823	.005	1	.945	1.058	.211	5.306
PC	15.370	4.313	12.699	1	.000	4731312.075	1008.747	2.2E10
INC_PC			3.369	1[a]	.066			
INC_PC(2)	-3.820	2.081	3.369	1	.066	.022	.000	1.296
AGE_TAXI			4.130	1[a]	.042			
AGE_TAXI(2)	.904	.445	4.130	1	.042	2.470	1.033	5.909
CAR_TAXI	1.277	.507	6.335	1	.012	3.586	1.327	9.693
TIME_TAXI			9.321	3[a]	.025			
TIME_TAXI(2)	-1.377	.593	5.394	1	.020	.252	.079	.807
TIME_TAXI(3)	-.655	.614	1.140	1	.286	.519	.156	1.730
TIME_TAXI(4)	-1.211	.634	3.652	1	.056	.298	.086	1.032

a. Degree of freedom reduced because of constant or linearly dependent covariables

X. Model 10

Categorical Variable Codings[a,c,d,e]

		Frequency	(1)	(2)	(3)	(4)
AGE_TAXI[b]	.00	465	1	0		
	1.00	340	0	1		
	2.00	35	0	0		
INC_PC[b]	.00	750	1	0		
	2.00	22	0	1		
	3.00	68	0	0		
TIME_TAXI[b]	.00	465	1	0	0	0
	1.00	168	0	1	0	0
	2.00	94	0	0	1	0
	3.00	83	0	0	0	1
	4.00	30	0	0	0	0
TIME_PC[b]	.00	750	1	0	0	0
	1.00	22	0	1	0	0
	2.00	23	0	0	1	0
	3.00	30	0	0	0	1
	4.00	15	0	0	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: INC_PC
d. Category variable: TIME_TAXI
e. Category variable: TIME_PC

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
372.085	186.667	13	.000	220.759	13	.000	220.759	13	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	95.0% CI for Exp(B) Upper
TTT	-.046	.011	17.784	1	.000	.955	.935	.976
TTC	-.480	.132	13.322	1	.000	.619	.478	.801
TAXI	.074	.842	.008	1	.930	1.076	.207	5.602
PC	15.030	4.624	10.563	1	.001	3367858.145	389.922	2.9E10
INC_PC			4.265	1[a]	.039			
INC_PC(2)	-3.898	1.887	4.265	1	.039	.020	.001	.820
AGE_TAXI			4.088	1[a]	.043			
AGE_TAXI(2)	.902	.446	4.088	1	.043	2.464	1.028	5.906
CAR_TAXI	1.297	.510	6.474	1	.011	3.659	1.347	9.941
TIME_TAXI			9.385	3[a]	.025			
TIME_TAXI(2)	-1.333	.601	4.923	1	.026	.264	.081	.856
TIME_TAXI(3)	-.581	.625	.865	1	.352	.559	.164	1.904
TIME_TAXI(4)	-1.214	.643	3.571	1	.059	.297	.084	1.046
TIME_PC			1.553	3[a]	.670			
TIME_PC(2)	1.208	2.733	.195	1	.658	3.347	.016	709.001
TIME_PC(3)	2.013	2.738	.540	1	.462	7.483	.035	1601.624
TIME_PC(4)	-.203	2.634	.006	1	.939	.816	.005	142.531

b. Degree of freedom reduced because of constant or linearly dependent covariates

XI. Model 11

Categorical Variable Codings[a,c,d,e]

		Frequency	(1)	(2)	(3)	(4)
AGE_TAXI[b]	.00	465	1	0		
	1.00	340	0	1		
	2.00	35	0	0		
INC_PC[b]	.00	750	1	0		
	2.00	22	0	1		
	3.00	68	0	0		
TIME_TAXI[b]	.00	465	1	0	0	0
	1.00	168	0	1	0	0
	2.00	94	0	0	1	0
	3.00	83	0	0	0	1
	4.00	30	0	0	0	0

EDU_TAXI[b]	.00	465	1	0	0	0
	1.00	116	0	1	0	0
	2.00	69	0	0	1	0
	3.00	140	0	0	0	1
	4.00	50	0	0	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: INC_PC
d. Category variable: TIME_TAXI
e. Category variable: EDU_TAXI

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
355.659	197.946	13	.000	237.185	13	.000	237.185	13	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	95.0% CI for Exp(B) Upper
TTT	-.050	.011	19.285	1	.000	.951	.930	.973
TTC	-.462	.128	13.003	1	.000	.630	.490	.810
TAXI	.228	.957	.057	1	.812	1.256	.192	8.201
PC	15.494	4.468	12.024	1	.001	5.4E6	842.327	3.4E10
INC_PC			2.832	1[a]	.092			
INC_PC(2)	-3.766	2.238	2.832	1	.092	.023	.000	1.860
AGE_TAXI			2.485	1[a]	.115			
AGE_TAXI(2)	.711	.451	2.485	1	.115	2.037	.841	4.931
CAR_TAXI	1.215	.523	5.405	1	.020	3.371	1.210	9.388
TIME_TAXI			9.490	3[a]	.023			
TIME_TAXI(2)	-1.638	.627	6.827	1	.009	.194	.057	.664
TIME_TAXI(3)	-.998	.648	2.367	1	.124	.369	.103	1.314
TIME_TAXI(4)	-1.511	.668	5.110	1	.024	.221	.060	.818
EDU_TAXI			17.268	3[a]	.001			
EDU_TAXI(2)	-.361	.471	.587	1	.444	.697	.277	1.754
EDU_TAXI(3)	-.072	.492	.022	1	.883	.930	.355	2.441
EDU_TAXI(4)	.920	.461	3.983	1	.046	2.509	1.017	6.194

a. Degree of freedom reduced because of constant or linearly dependent covariates

XII. Model 12

Categorical Variable Codings[a,c,d,e,f]

		Frequency	(1)	(2)	(3)	(4)
AGE_TAXI[b]	.00	465	1	0		
	1.00	340	0	1		
	2.00	35	0	0		
INC_PC[b]	.00	750	1	0		
	2.00	22	0	1		
	3.00	68	0	0		
TIME_TAXI[b]	.00	465	1	0	0	0
	1.00	168	0	1	0	0
	2.00	94	0	0	1	0
	3.00	83	0	0	0	1
	4.00	30	0	0	0	0
EDU_TAXI[b]	.00	465	1	0	0	0
	1.00	116	0	1	0	0
	2.00	69	0	0	1	0
	3.00	140	0	0	0	1
	4.00	50	0	0	0	0
EDU_PC[b]	.00	750	1	0	0	0
	1.00	22	0	1	0	0
	2.00	10	0	0	1	0
	3.00	37	0	0	0	1
	4.00	21	0	0	0	0

a. Category variable: AGE_TAXI

b. Indicator Parameter Coding

c. Category variable: INC_PC

d. Category variable: TIME_TAXI

e. Category variable: EDU_TAXI

f. Category variable: EDU_PC

Omnibus Tests of Model Coefficients

	Overall (score)			Change From Previous Step			Change From Previous Block		
-2 Log Likelihood	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
355.494	198.592	16	.000	237.350	16	.000	237.350	16	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	Upper
TTT	-.050	.011	19.210	1	.000	.951	.930	.973
TTC	-.462	.129	12.796	1	.000	.630	.489	.812
TAXI	.217	.962	.051	1	.822	1.242	.188	8.195
PC	19.034	31.929	.355	1	.551	1.8E8	.000	2.8E25
INC_PC			2.582	1[a]	.108			
INC_PC(2)	-3.698	2.301	2.582	1	.108	.025	.000	2.254
AGE_TAXI			2.454	1[a]	.117			
AGE_TAXI(2)	.706	.451	2.454	1	.117	2.027	.837	4.904
CAR_TAXI	1.208	.525	5.309	1	.021	3.348	1.198	9.360
TIME_TAXI			9.165	3[a]	.027			
TIME_TAXI(2)	-1.616	.633	6.529	1	.011	.199	.058	.686
TIME_TAXI(3)	-.978	.653	2.245	1	.134	.376	.105	1.352
TIME_TAXI(4)	-1.488	.675	4.864	1	.027	.226	.060	.847
EDU_TAXI			16.877	3[a]	.001			
EDU_TAXI(2)	-.360	.472	.584	1	.445	.697	.277	1.758
EDU_TAXI(3)	-.084	.494	.029	1	.865	.919	.349	2.420
EDU_TAXI(4)	.913	.463	3.892	1	.049	2.492	1.006	6.174
EDU_PC			.105	3[a]	.991			
EDU_PC(2)	-3.204	31.803	.010	1	.920	.041	.000	4.8E25
EDU_PC(3)	-3.947	31.786	.015	1	.901	.019	.000	2.2E25
EDU_PC(4)	-3.641	31.753	.013	1	.909	.026	.000	2.8E25

a. Degree of freedom reduced because of constant or linearly dependent covariates

XIII. Model 13

Categorical Variable Codings[a,c,d,e]

		Frequency	(1)	(2)	(3)	(4)
AGE_TAXI[b]	.00	465	1	0		
	1.00	340	0	1		
	2.00	35	0	0		
INC_PC[b]	.00	750	1	0		
	2.00	22	0	1		
	3.00	68	0	0		
TIME_TAXI[b]	.00	465	1	0	0	0
	1.00	168	0	1	0	0
	2.00	94	0	0	1	0
	3.00	83	0	0	0	1
	4.00	30	0	0	0	0
EDU_TAXI[b]	.00	465	1	0	0	0
	1.00	116	0	1	0	0
	2.00	69	0	0	1	0
	3.00	140	0	0	0	1
	4.00	50	0	0	0	0

a. Category variable: AGE_TAXI
b. Indicator Parameter Coding
c. Category variable: INC_PC
d. Category variable: TIME_TAXI
e. Category variable: EDU_TAXI

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
352.584	198.631	14	.000	240.260	14	.000	240.260	14	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
IVTT	-.036	.014	6.812	1	.009	.965	.939	.991
OVTT	-.078	.021	14.371	1	.000	.925	.888	.963
TTC	-.454	.125	13.093	1	.000	.635	.497	.812
TAXI	.342	.968	.125	1	.724	1.408	.211	9.396
PC	14.836	4.386	11.443	1	.001	2.8E6	512.867	1.5E10
INC_PC			2.986	1[a]	.084			
INC_PC(2)	-3.785	2.190	2.986	1	.084	.023	.000	1.662
AGE_TAXI			2.532	1[a]	.112			
AGE_TAXI(2)	.721	.453	2.532	1	.112	2.056	.846	4.994
CAR_TAXI	1.213	.521	5.426	1	.020	3.364	1.212	9.337
TIME_TAXI			10.922	3[a]	.012			
TIME_TAXI(2)	-1.798	.648	7.691	1	.006	.166	.047	.590
TIME_TAXI(3)	-1.063	.661	2.584	1	.108	.345	.095	1.262
TIME_TAXI(4)	-1.621	.686	5.585	1	.018	.198	.052	.758
EDU_TAXI			18.665	3[a]	.000			
EDU_TAXI(2)	-.455	.480	.898	1	.343	.634	.248	1.626
EDU_TAXI(3)	-.097	.497	.038	1	.845	.907	.342	2.404
EDU_TAXI(4)	.918	.465	3.905	1	.048	2.504	1.008	6.225

a. Degree of freedom reduced because of constant or linearly dependent covariates

II. Stated preference modal choice SPSS parameter estimation results

Case Processing Summary

		N	Percent
Cases available in analysis	Event	375	30.9%
	Censored	840	69.1%
	Total	1215	100.0%
Cases dropped	Cases with missing values	0	0.0%
	Cases with negative time	0	0.0%
	Censored cases before the earliest event in a stratum	0	0.0%
	Total	0	0.0%
Total		1215	100.0%

Base model

Omnibus Tests of Model Coefficients

-2 Log Likelihood
875.742

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
333.511	509.792	3	.000	542.231	3	.000	542.231	3	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
LBUS	-3.912	.412	90.024	1	.000	.020	.009	.045
TAXI	-2.708	.231	137.488	1	.000	.067	.042	.105
PC	.261	.212	1.516	1	.218	1.299	.857	1.969

I. Model 1

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
292.603	523.487	5	.000	583.139	5	.000	583.139	5	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
LBUS	-1.771	.511	12.031	1	.001	.170	.063	.463
TAXI	-1.946	.266	53.347	1	.000	.143	.085	.241
PC	1.376	.887	2.408	1	.121	3.959	.696	22.513
TTT	-.084	.016	28.951	1	.000	.920	.892	.948
TTC	-.100	.047	4.597	1	.032	.905	.826	.991

II. Model 2

Categorical Variable Codings[a,c,d]

		Frequency	(1)	(2)	(3)
INC_LBUS[b]	.00	840	1	0	0
	1.00	92	0	1	0
	2.00	174	0	0	1
	3.00	109	0	0	0
INC_TAXI[b]	.00	840	1	0	0
	1.00	92	0	1	0
	2.00	176	0	0	1
	3.00	107	0	0	0
INC_PC[b]	.00	1125	1	0	
	2.00	22	0	1	
	3.00	68	0	0	

a. Category variable: INC_LBUS
b. Indicator Parameter Coding
c. Category variable: INC_TAXI
d. Category variable: INC_PC

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
291.706	527.660	10	.000	584.036	10	.000	584.036	10	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	Upper
LBUS	-1.842	1.077	2.929	1	.087	.158	.019	1.307
TAXI	-2.217	.615	12.978	1	.000	.109	.033	.364
PC	1.395	.891	2.448	1	.118	4.035	.703	23.157
TTT	-.084	.016	28.590	1	.000	.919	.891	.948
TTC	-.106	.048	4.824	1	.028	.899	.818	.989
INC_LBUS			.069	2[a]	.966			
INC_LBUS(2)	.248	1.253	.039	1	.843	1.281	.110	14.926
INC_LBUS(3)	.018	1.179	.000	1	.988	1.019	.101	10.269
INC_TAXI			.507	2[a]	.776			
INC_TAXI(2)	.172	.764	.051	1	.822	1.188	.266	5.312
INC_TAXI(3)	.433	.676	.410	1	.522	1.542	.410	5.803
INC_PC			.326	1[a]	.568			
INC_PC(2)	.327	.572	.326	1	.568	1.386	.452	4.256

a. Degree of freedom reduced because of constant or linearly dependent covariates

III. Model 3

Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score) Chi-square	df	Sig.	Change From Previous Step Chi-square	df	Sig.	Change From Previous Block Chi-square	df	Sig.
289.409	525.954	6	.000	586.333	6	.000	586.333	6	.000

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B) Lower	Upper
IVTT	-.100	.019	26.684	1	.000	.905	.872	.940
OVTT	-.045	.034	1.784	1	.182	.956	.895	1.021
TTC	-.104	.048	4.813	1	.028	.901	.821	.989
LBUS	-1.919	.541	12.576	1	.000	.147	.051	.424
TAXI	-2.040	.282	52.473	1	.000	.130	.075	.226
PC	2.073	1.046	3.927	1	.048	7.951	1.023	61.795