MEKELLE UNIVERSITY

ETHIOPIAN INSTITUTE OF TECHNOLOGY–MEKELLE UNIVERSITY (EIT-M)

SCHOOL OF CIVIL ENGINEERING

Developing Trip Production and Mode Choice Model for Home-Based Trips: The Case of Hawelti Sub-City Residents, Mekelle - Ethiopia

By:

Gerezgiher Gebrewahd Hagos

A Thesis Research submitted to the School of Civil Engineering the Ethiopian Institute of Technology-Mekelle University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Civil Engineering (Road and Transport Engineering).

Advisor:

Ashenafi Aregawi (PhD)

March, 2020

Mekelle, Ethiopia

ACKNOWLEDGMENT

First and foremost I would like to thank God for giving me inspiration, ability, and discipline to make it through.

I would like to extend my sincere thanks to my advisor Dr. Ashenafi Aregawi, for all his constructive, critical and valuable comments and professional support and excellent guidance throughout my thesis work and that made this thesis work possible.

I also express my deepest thanks and appreciation goes to the members of the field workers they really had played a significant role during field data collection. Without their devoted assistance, the field data collection in such a harsh environment would have not been possible.

I would like to express my gratitude to ERA, Mekelle University and Road and transportation engineering chair for all contributions provided to my thesis in various ways.

My special thanks go to Mekelle City transport office, Taxi Associations, Hawelti sub-city municipality office for their assistance and providing necessary documents and sharing ideas, all my friends and colleagues for their unlimited support and encouragement. I would like also to thank my family for their strong support, love, prayers and tolerating the time I spend working with my research.

Finally, I would be more pleased listing all those who lent their hands during my study. However, I extend my gratitude for all who were on my side without whom this effort would have been a rough path.

ABSTRACT

Hawelti sub City considered one of the most rapid increases in population and crowded sub-city in Mekelle city. Hawelti sub-city does not have a transportation model; hence, developing a model is considered very important for predicting the average daily trips and use of each mode. The aim of this study is to develop trip production and mode choice model for home-based trips in Hawelti sub-city residents. The trip production includes a trip for work, education, shopping, social and recreation whereas the mode choice focuses on work and education trips. A household interview survey was used as a data source in addition to secondary data obtained from the municipality of Hawelti Sub-City. Trip production was analyzed using multiple linear regressions and multinomial logistic regression was utilized in the analyses of mode choice with statistical data processing software SPSS. The modes of transport considered in the model are Walking, Bajaj, Private car, Minibus Taxi, and Service (Company or school provided transport). Minibus taxi is found to dominate with 44% share followed by the company provided service with a 22.8% *share for all trips. In the developed model the factors that affect the trip production are the number of students, employed persons, monthly income, the number of females and the age above 45 in the household. Moreover, the factors that significantly affect the choice of transport modes for employees were total travel time, work type and family size. The mode choice model developed for the case of students the significant factors were found to be travel time and gender. The developed models were significant at 95% confidence level and exhibits a good fit for the data with the coefficient of determination (R^2) of 0.889 for the general trip production and the adjusted goodness of fit measure statistics rho square ($\rho 2$) of 0.823 and 0.824 for the work and student models respectively for the whole study area. The models are validated to be capable of predicting 90.97% and 87.2% of the preference of trip makers for the work and education mode choice models respectively.*

Keywords: *Trip, production, mode choice, multiple linear regression, multinomial logit, model.*

TABLE OF CONTNTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

HTM	Handbook of transport modeling
ITE	Institute of Transportation Engineers
LRTS	Likelihood Ratio Test
MNL	Multinomial logit model
O-D	Origin –Destination
SPSS	Statistical Package for Social Science
TAZ	Traffic analysis zone
TP	Trip production

CHAPTER ONE

INTRODUCTION

1.1. Background

The complex nature of urban phenomenon needs to develop a model and, in specific, urban transportation models. The development of models for urban transport is the most important priority task and is an analytical tool used to support the urban transportation planning process and decision making on transportation policies. Transportation modeling plays a vital role in the development of the transportation network. One of the key components in transportation planning is travel demand forecasting. The transportation planning process has trip generation, trip distribution, mode choice, and travel assignment as the four basic stages in transportation modeling. Although the four stages of the transportation planning process are clearly inter-related, the traditional approach has been to treat each stage more or less as a separate entity (Qrtuzar and Willumsen, 2011). The classical transport modeling starts with the process of development of trip generation models and ends with the trip assignment, which helps in understanding the pattern of trips developed in the study area and used to estimate the traffic for future years. On the other hand, the two phases of the traditional four-step demand estimation method, trip production, and mode choice can be analyzed; this thesis focuses on trip production and mode choice models. Modeling of trip production and mode choice thus implies a procedure for predicting what travel decisions individuals would like to make given the generalized trips.

Trip production and mode choice modeling are aims to establish the spatial distribution of travel explicitly by means of an appropriate system of zones. The trip generation is the first phase in the travel demand forecasting process and involves the estimation of the total number of trips entering or leaving a particular area of any specific type of land use. The trip generation manual published by the Institute of Transportation Engineers (ITE) provides general guidelines and values of trip generation for different land-use developments (ITE, 2012). However, several studies reported that person trips estimated based on vehicular trips using the ITE manual resulted in an overestimation of such trips. Mustafa (2016) indicated that the trip generation rates for residential land use in Palestine are lower than the ITE rates. Thus, values provided in the manual should be calibrated for usage elsewhere outside the USA. Trip generation models are classified into two types as production models and attraction models. In this thesis, the trip production part of the trip generation model was studied. A trip that ends at home is called home-based trips or trip production. The trip production from the generation stage of the classical transport model is the model building phase starts with the first step commonly used to process and estimate the total number of trips produced by each area unit (zone) in conjunction with the land use and the socio-economic characteristics of each zone. This can be achieved in a number of ways: starting with the trips of the individuals or households who reside in each zone or directly with some of the properties of the zones: population, employment, number of cars, income, etc.

The trip production modeling is an appropriate analytical technique in the process of relating the trips produced by households or individuals to the factors that influencing trip production of the specific traffic analysis zone. Trip production techniques try to establish a relationship between the demographic and socio-economic characteristics of the population of an analysis unit and its trip production.

Although the individuals are usually the trip makers, the number of trips per household is usually estimated. The household is a one person or group of persons with or without family relationships, who normally live in the same housing unit, share meals, and make joint provision of food and other essentials of living. Some factors that can influence trip production are related to the household characteristics because the household is a major unit of trip production which are household size and composition, where household size is number of persons in the household, and composition like the age of persons in the household, the distribution of the sex of the individuals, number of employed persons, number of students, household income, vehicle ownership, etc. (Arasan, 2012).

Similarly, the locations of different land uses are related to the trip making of the analysis units. These procedures are based on the hypothesis of a causal relationship between population characteristics, land use, and the trip making behavior of people. The most common methods of analysis for trip production models are growth factor modeling and regression models.

Mode choice analysis is performed as one of the steps in the four-step transportation forecasting model following trip production. Mode choice analysis allows the modeler to replicate what mode of transport would be used, and what is the resulting model share. Mode choice models model the traveler’s choice of mode, which mode of transport to take, e.g. private car, taxi or bus. Most of the models for mode choice are based on random utility maximization principle derived from econometric theory. Cheng, et al. (2014) indicated that the developed multinomial logit model in the 1970s has become the most widely used tool. This, the model used to analyze and predict the traveler decisions is a mathematical function that estimates the probability of selecting individual travel choice based on the utility maximization principle. That is, random utility theory still assumes that an individual will choose that alternative which appears to maximize his or her utility at the time at which the choice is being made.

Different types of models that can be developed to produce existing demand conditions and actual travel patterns of people. At the individual level, urban transportation can be characterized by a trip purpose. Personal trips are made commonly based on their main purposes; work trips, shopping trips, social trips, recreational trips, school trips, and other trips. Each trip is made using one or multiple transportation modes for a defined purpose at a given time. However, analyses of travel behavior have traditionally focused on work trips as this behavior is easier to understand due to the more regular nature of the trips. Further, the trip purposes for work, education, shopping trips, social and recreational of individuals are the focused area of study in this research.

1.2. Statement of the Problem

Nowadays, transport becomes a challenge to the ever-increasing demand with the limited resource available especially for cities in developing countries like Mekelle city in Ethiopia. Transport study for Mekelle city including hawelti sub-city was undertaken as part of the master plan preparation process by a project office established under the regional Bureau of Works and Urban development in 2005 and Mekelle Structural Plan Revision was prepared by Mekelle university business and consultancy office in 2015 concerning all components of the transport system. These studies showed that Mekelle city suffers by transport problems including congestion, pollution, long waiting time at taxi station and mismanagement of existing transport services and the types of transport in the city due to lack of proper transportation planning. The studies excluded the characteristics of the trip maker and the four-stage travel demand model was not adopted. Without estimation of the number of trips, understanding the travel maker needs and travel mode's characteristics is difficult to make solutions for transport problems in the city. The development of these models is vital for planners, developers, and engineers.

Hawelti sub city is one of the critical sub-cities that have a rapid growth of population and the number of cars that generate a number of subsequent trips which cause problematic transportation. The increasing need for mobility according to population growth and motor vehicles, combined with the limited development of the transport networks within the sub-city, leads to congestion, delays, pollution, and accidents in the existing transport networks.

In order to adopt suitable transport policies for solving the expected transport problem resulting from population density and the increasing number of motor vehicles, there is a need for improving the transport planning process for Hawelti sub-city residents. One step should be improved is transportation modeling. Proper modeling is lacking describing the four analytical steps; trip generation, trip distribution, modal split, and route assignment. For Hawelti sub-city residential areas, trip production, and mode choice models are not yet developed. To provide a strong basis for the transportation planning process, this research considers the need for studies in the Hawelti sub-city, through modeling the trip production part of the trip generation and modal choice by considering the socio-economic, mode and trip-related data.

1.3. The objective of the thesis

The need to understand the trips produced and travel behavior of people on selection of the mode of transport and to be able to model the trip and behavior of people is important in order to give an emphasis by policymakers and planners in transportation systems for the benefit of the society and the economy, and the need of the people that maximize traveler's satisfaction.

In general, the objective of this thesis is to develop a trip production and mode choice model for Hawelti sub-city residents, thus studying and modeling trips produced and mode choice at the individual level according to their characteristics, depend on the principles of the regression analysis techniques.

1.3.1. The specific objective of the study

- To determine the factors that affect the trip production and mode choice behavior of people who make trips in the study area;
- To develop a trip production model based on trip purpose
- To develop a model of travel mode choice over the currently available modes of transport for the purpose of work and school trips.

1.4. Significance of the Study

This research is mainly concerned with the development of trip production and the mode choice model for Hawelti sub-city residents. Studies made yet in Ethiopia with the objective of trips for different purposes with identifying the factors affecting trip production and mode choice behavior are very small in number. Therefore, this study makes contributions towards extended research in the area of trip production by purpose and mode choice behavior of people of Ethiopia as a general and Mekelle city in particular.

Based on emphasis is given to prediction by modeling, identifying and analyzing the travel behavior of the residents of the study area, this study is used as an input to predict the future levels of different passengers' preference to the different attributes of a public transport mode which could indicate how a new or improved model would perform. In addition to that, the research can help for launching initiations, the government, municipal authorities and the community in the sub-city to determine the need for the travelers, planning inspections and improvement or new development of the model. Hence, the findings of the study could serve as an important base for transport planners, transportation professionals and policymakers to make a decision on policy and investment in the road transport system. And also the study is beneficial for researchers who conduct future detailed studies on similar researches like improving urban transport mobility with different models and other related issues.

1.5. Scope of the Study

To limit the scope of this research, from Mekelle city, Hawelti sub-city was taken as a study area and limited for trip production and mode choice model development. The reason for this limitation is time and financial constraints. The data were collected from the households of Hawelti sub-city residents and the administrative office of Mekelle city and municipality of Hawelti sub-city. The models were developed from the data collected using multiple linear regression analyses for the trip production model and multinomial logistic regression analysis for the mode choice model. In addition, Out of the different purposes of trips, the study in the trip production is delimited to the work, school, shopping, recreational and social trip purpose while the mode choice model is focused only on work and education trips made by the study area residents. The developed models were checked using statistical tests and by comparing model results with that observed data. The mode choice model is scoped to service, taxi, private vehicle, Bajaj and walk modes of transport.

1.6. Organization of the thesis

This thesis was arranged into six chapters. The first chapter contains the introduction part, which comprises the background of the study, statement of the problem, objectives of the study, significance of the study, scope, and limitation of the study. The second chapter covers related works of literature like model definitions, transport planning, factors that affect trip production and mode choice and techniques of modeling. Chapter three introduces research methodologies; in this chapter study design, data sources, data analysis, methods of data collection and model estimation techniques. In the fourth chapter the study area selection, sample size computation, data collection instrument and conducting field survey are involved. Results, discussions, and analyses are presented in chapter five. Finally, Chapter six concludes the study with main findings from the model developed and how the objectives of this study have been addressed and recommendations are forwarded.

CHAPTER TWO

LITERATURE REVIEW

2.1. Transportation Planning and Modeling

Transportation planning is a process of identifying opportunities to improve transportation systems and involved with the evaluation, design, and operation of transport facilities so that the access and mobility of people and businesses are accelerated. The transportation improvements should be based on the understanding of future demand and describing the effects of proposed transportation alternatives and explaining the benefits to the traveler of a new transportation system and its effect on the community. Transportation planning processes have been intensively used to estimate the demand for travel encountered in the future. The estimated travel demand is utilized as a basis to plan for future transportation facilities and services.

According to Meyer and Miller (2001), there are money transportation planning process at any given time in an urban area, every defined at a different level of purpose; while traffic engineers identify congestion reducing alternatives for the highway networks, regional planners look at urban development patterns and the provision of public services: individual employers consider employee transportation programs: and social service agencies examine transportation options to improve the delivery of their services to targeted population groups. Generally, transportation planning is the process that leads to decisions on transportation policies, strategies, and programs.

Modeling is one important part of the most decision-making process. Transport models can be defined as a simplified representation of the real world usually implemented in computer programs like (SPSS) which describes the impact of transport decisions. It is concerned with the methods, there may be quantitative or qualitative which allows us to study the relationships that underlie the decision-making process (Hensher and Button, 2000).

Transportation modeling plays an important role in supporting transportation planning and policymaking. One of the most important aspects of transportation modeling is to predict the travel choice behavior which is the most frequently modeled travel decision. It involves specific aspects of human behavior dedicated to choice decisions.

The urban transport model system often referred to as the 4-step model is commonly used to predict the flows on the links of a particular transportation network as a function of the land use activity system and socioeconomic activities that generate the travel. The model comprises four sub-models that are employed in a sequential process. Since then the interest of this field, as well as the growing complexity, has led to the further development of various travel demand models. The classic four-step method has trip generation, trip distribution, modal split, and trip assignment steps. These four stages are inter-related with that the traditional approach has been to treat each stage more or less as a separate entity. The first step from the classic four-step method the trip generation estimates the number of potential trips starting or ending in a given area. Analysis of the function of trip generation is used to establish a relationship between land use and trip making activity so that changes in land use can be used to predict subsequent changes in transport demand. It has two components in the form of trip production and trip attraction. Trip distribution associates origin and destination to each trip generated. Mode choice analyzes how the trips are split between modes of transportation. Mode choice is one of the most critical parts and the third steps in the sequential travel demand modeling process. It is the step where trips between a given origin and destination are split into trips using available modes of transportation. Mode choice analysis tells the traveler's preference for a mode of transport. The fourth step in travel demand modeling following trip generation, trip distribution, and mode choice is a trip assignment choice that concerns the selection of routes between origins and destinations and estimates volumes on different links of the transportation network.

On the other hand, although it is possible that two-step models of traditional four-step demand estimation method, trip generation, and mode choice, are separately analyzed, this thesis focuses on trip generation step (trip production) and mode choice. Since the purpose of the mode choice is to predict the probability of the transportation mode's selection, the final output is the number of people producing trips per day with respect to each mode. Both, the trip production and transport mode choice provide an understanding of why people travel, identify the essential factors influencing their travel decisions and provide information for the evaluation of alternative transportation policies by predicting the travel consequences of alternative policies or planning.

2.2. Trip Production

In transport modeling, a trip is a one-way movement of a point of origin to a point of destination (OrtUzar and Willumsen, 2011).

A Trip Production is defined as the home end of a home-based trip or as the origin of none home-based trip. A home-based trip is one where the home of the trip maker is either the origin or the destination of the trip and a non-home-based trip is, conversely, no one of the trips ends in the home of the trip maker. The trip production process is used to estimate the number of person-trips that will begin from or end at each travel analysis zone contained in the area of study. The process considers some socio-economic data as input producing trip production values as the output. Trip production models on this research estimate the number of home-based trips to and from zones of the trip makers resides.

Trips are classified by different purposes to obtain better and accurate trip production models into; work trips, shopping trips, social trips, recreational trips, school trips, home trips, and business trips. Based on other studies, trips can be also classified as trips for work, trips for education, trips for shopping, trips for recreation and others. Among these, the work and education trips are often referred to as obligatory trips and the rest as optional trips (OrtUzar and Willumsen, 2011). When transport policies are introduced, it would mostly impact on discretionary trips than compulsory trips. Trip production models for different types of trips can vary either by the factors in the equations or by the value of the coefficients of the same factor. In most modeling, the trip productions are estimated as typically stratified into three trip purposes: home-based work, home-based education and home-based other. In this research, the home-based trip purpose classifications are; work trips, school trips, shopping trips, social trips, and recreational trips to better predict the travel patterns in the study area. Since trip production is associated with the home, if the home is either origin or destination, the trips should be considered in the trip production model.

2.2.1. Aggregate and disaggregate approaches

There are two approaches in terms of data aggregation in trip production models: aggregate and disaggregate trip production models. According to the level of analysis, an aggregation level area (zonal) is the unit of analysis. In disaggregate models, the household or individuals are used. Dodeen (2014) analysis on the level of household is adopted, because of this level; more accurate results can be obtained, through the study of movements for each household. The advantage of disaggregate over aggregate person-level models the small amount of data is required for model estimation.

According to Jittrapirom and Emberger (2013), the Household is the preferable analysis unit to use in trip production modeling when the model estimation data are collected in a household travel survey. The empirical test in the Chiang Mai mobility and transport Survey indicates that household trip production models yield prediction of trips at the household levels are more accurate than those trips produced by person trip production models.

2.2.2. Factors affecting trip production

The factors influencing trip production reviewed studies highlight the relationship between trip frequency and the following measures: household size, employment, gender, age, household income, and mobility tools. In addition, some studies emphasize the importance of separating the generated trips based on the purpose of the trip. For instance, previous studies show that employment has a positive impact on the total number of generated all-purpose trips (Roorda et al., 2010), work trips (Chang et al., 2014; Huntsinger et al., 2013). Gender also shows the impact of production. Females have a positive impact for the purpose of shopping and social trips (Fox & Patruni, 2015). This is expected as in a multi-person household, females are typically more likely to engage in shopping and social trips.

Explanatory variables that influence trip production are income, car ownership, family size, household structure, the value of land, and residential density (Ortuzar and Willumsen 2011). In most of the studies, the approach was used more often; a range of socio-economic and demography variables influences the trip production. These variables include gender, age, and household income, and employment status, number of workers in the household, education status, vehicle ownership, and family size. The explanatory variables used in trip production models will differ depending on the type of trip being modeled. Different numbers of trip production studies have shown that the following explanatory variables that can affect personal trip production include individual and household characteristics of the trip makers with respect to the major trip purpose are listed as follows:

Individual characteristics

- Age
- Gender
- Occupation
- Education receiving status

Household characteristics

- Household size
- Household income
- Car ownership

2.2.3. Trip Production Modeling

In terms of the methods of analysis that are used in developing trip production models, the primary tool used is the regression analysis methods. Trip production models were commonly developed by regression analysis because of its power and simplicity. Trip production equations developed by the regression are used by planning agencies. Obtaining more detailed data for individual households requires a survey of trip production. Regression models are relatively straight forward to implement, thus cost-effective and the data needs are moderate in size (Chang et al., 2014). A frequently used regression model is the multiple linear regression model. In this technique, the dependent output variable is assumed to have a linear dependence on the independent input variables, which may or may not influence the trip production.

The general form of multiple regressions is shown in Equation 2.1.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots + \beta_n X_n \quad (2.1)$$

Where,

Y = dependent variable (i.e., the number of trips produced),

X_i = independent variables (i = 1, 2, 3……n)

$\dot{\alpha}$ = constant term

β_i = coefficient of independent variables (i = 1, 2,3……n)

n = number of independent variables

Definitions of the input and output variables vary with the type of linear regression approach used in the research. In this research, household-based multiple linear regression techniques are used to produce the travel patterns on a household level. The household-based regression tends to utilize various parameters associated with a household, in order to estimate the regression coefficients, such as household size, gender; the number of workers in a household; number of receiving education, car ownership, age groups and average monthly income in the household.

Assumptions in multiple linear regression models

Multiple linear regression analysis is a statistical method. The following are the several assumptions used in the analysis.

- Variable values, especially independent variables have a specific value or a value derived from the survey results with no significant errors.
- The dependent variable must have a linear correlation with independent variables. If the relationship between these variables is not linear, then linear transformations must do although this restriction will have other implications in the analysis of residuals.

- The independent variable's effect on the dependent variable is the sum, does not have to have a strong correlation among independent variables.
- Variations of dependent variables towards the regression line the same for all values of independent variables.
- Variations of dependent variables have to scattered normally or near-normal minimum.
- The correlation coefficient between the dependent variable and independent variables or among independent variables of positive one means that the correlation between the dependent variables and independent variables is positive. Conversely, if the value of the negative one, the correlation between dependent variables and independent variables is negative (the increase of independent value will affect the decrease of dependent value). The value of the correlation coefficient zero indicates no correlation between the variables.
- Any variables that have a regression coefficient is not statistically significant should be removed from the model.

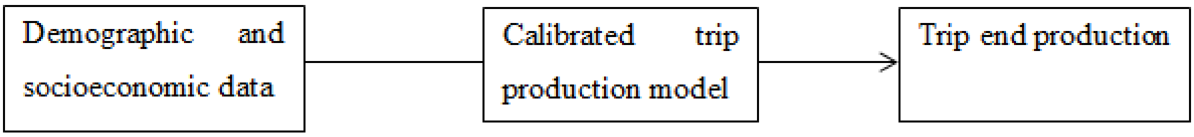

Figure2. 1. Trip production model inputs and output

2.3. Mode Choice

The modal split attempts the assignment of person-trips to the different alternative modes available in the study area. The development of transport mode choice models representing the behavior of travelers in the study area. The choice of transport mode is probably one of the most important classic four-step models in transport planning. This is because of the key role played by public transport in policymaking. Develop and use models are important which are sensitive to those travel attributes that influence individual choices of mode. The modal split step is used to determine the share of trips for each mode. There is a number of modal split models. Usually, mode choice models are classified into two groups according to the type of mathematical abstraction used; aggregate models (trip-end model, trip interchange model), and disaggregate behavioral models. Significant factors in the determination of the model used are grouped into three characteristics: trip, trip makers, and transportation system.

Mode choice can be performed before trip distribution (trip-end mode choice model) or after trip distribution (trip-interchange mode choice model).

The trip-end mode choice model is processed just after the trip generation step. In this way, the different characteristics of the person could be preserved and used to estimate the modal split. Trip-end mode choice modal split the total demand for travel for each transport zone by the available travel modes. The mode choice, in this case, is based on the attributes of the trip origin (that is, ease of access to each mode and the ability or inclination to use a particular mode). The advantage is that these models could be very accurate in the short run in the availability of public transport and there is little congestion. The limitation is that they are insensitive to policy decisions in improving public transport, restricting parking, etc. would have no effect on modal split according to these trip-end models (Ortuzar, Willumsen, 2011).

2.3.1. Aggregate and disaggregate approaches

There are two basic ways of modeling mode choice namely, aggregate and disaggregate approaches. The aggregate share of all or a segment of decision-makers choosing each alternative as a function of the characteristics of the alternatives and socio-demographic attributes of the group is directly models in the aggregate approach. Aggregate modeling primarily focuses on the mode choices made by average individual trips.

The disaggregate approach is to recognize that aggregate behavior is the result of numerous individual decisions and to model individual choice responses as a function of the characteristics of the alternatives available to and socio-demographic attributes of each individual. Disaggregate mode choice models have substantial advantages over the aggregate models for predicting the consequences of transportation policy measures that affect mode choice (Minal and Ch. Ravi, 2014).

The disaggregate approach explains why an individual makes a particular choice given her/his circumstances and is, therefore, better able to reflect changes in choice behavior due to changes in individual characteristics and attributes of alternatives while the aggregate approach, on the other hand, rests primarily on statistical associations among relevant variables at a level other than that of the decision-maker. As a result, the aggregate approach is unable to provide accurate and reliable estimates of the change in choice behavior due to changes in service or in the population. Disaggregate models more efficient than aggregate ones in terms of data and computational requirements (Abdulhaq, 2016).

Disaggregate approach models are being increasingly used to understand individual behavior so that those modify the attributes of alternatives which are important to individual decision-makers. In addition to that, the disaggregate approach is more suited for analysis since it is causal, less tied to the estimation data and more likely to include a range of relevant variables. The coefficients of the explanatory variables have a direct marginal utility interpretation (Ortuzar and Willumsen, 2011).

Furthermore, with present knowledge and computerized program (SPSS), it is possible to model the mode choice for more than two alternatives by considering both the travel maker's and the travel mode's characteristics. In addition, 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. Therefore, in this research, a Multinomial Logit model is developed, which is based on disaggregate socio-economic, mode and trip-related data, for the trips based on purpose made by Hawelti Sub-City residents. The different modes of road passenger transport considered in the model are walking, Bajaj, Minibus taxi, Private vehicle as driver and passenger and Company provided transport.

2.3.2. Factors Influencing Mode Choice

The factors that are assumed to influence the mode preference of trip makers are used to estimate the parameters of a discrete mode choice model. Trip maker mode choice is generally a function of three basic factors: characteristics of the trip, e.g., purpose, the socio-economic characteristics of the traveler, and the modal attributes a good mode choice model should include the most important of these factors (Ortuzar, Willumsen, 2011). According to Almasri and Alraee (2013) study for work trips in Gaza city and investigating the factors that affect the employed people's choice for transport modes. The factors that significantly affect the choice of transport modes are total travel time, the total cost divided by personal income, ownership of means of transport, distance, age, and average family monthly income. Earlier research hand clearly showed that individual and household socio-demographics exert a strong influence on travel mode choice decisions. The commonly used explanatory variables influencing mode choice models according to (Bhat and Sardesai, 2006) are presented as follows:

I. Characteristics of the trip maker

The variables which are generally believed to be important for the mode choice model can be household car ownership, household structure, household income; decisions made elsewhere, for example, the need to use a car at work, for shopping and recreation, for school to take children, etc. and residential density.

II. Characteristics of the trip/ journey

Mode choice is influenced by the trip purpose and time of the day when the journey is undertaken.

III. Characteristics of the transport facility

These can be the relative travel time by each mode and walking times and relative monetary costs and availability.

2.3.3. Utility maximization theory

The utility is an indicator of value to an individual. The utility maximization rule states that an individual will select the alternative from a set of available alternatives that maximize his/ her utility. Additionally, the rule implies that there is a function containing attributes of alternatives and characteristics of individuals that describes an individual‘s utility valuation for each alternative. 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).

The utility maximization rule, which states that an individual chooses the alternative with the highest utility, implies no uncertainty in the individual‘s decision process; that is, the individual is certain to prefer the highest ranked alternative under the observed choice conditions. 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. The probabilistic choice theory arises 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. Therefore the utility of the alternative from the perspective of the decision-maker will be decomposed into two components. One component of the utility function represents the portion of the utility observed by the analyst, which is often called the deterministic (or observable) portion of the utility and the other component is the difference between the unknown utility used by the individual and the utility estimated by the analyst known as random error. The utility, U is modeled as a random variable in order to reflect the uncertainty.

Formally, this is represented as:

$$U_{ij}=V_{ij}+\varepsilon_{ij} \qquad 2.2$$

Where; U_{ij} is the utility of the alternative mode i for individual j;

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

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

2.3.4. Mode choice Modeling

The experiences of some countries in modeling the mode choice for different types of trips that were reviewed are illustrated that the factors related to transport policies and the factors related to the characteristics of the trip maker are the most important factors in guiding the choice behavior of travelers. From the development of the mode choice model the logit model is the most widely used technique because it is simple in terms of formulation and estimation of the model in addition to its accuracy compared with the others

The mode choice is used to estimate future travel volumes by passengers choosing a mode. This step is based on the concept that the mode choice behavior of trip makers can be explained generally by three categories of factors: the characteristics of available modes; the socioeconomic status of the trip maker; and the characteristics of the trip (Ortuzar, Willumsen, 2011). The different types of models depend upon the application and the functional form of the error term distribution logit model was used in this research.

2.3.5. Logit Model

Logit models have been the most widely used modal split models in the area of transportation planning since they possess the ability to model complex travel behaviors of any population with simple mathematical techniques and thus proves to be the most widely used tool for mode choice modeling the framework of logit models is based on the theory of utility maximization. Due to the fact that the formula for the choice probabilities takes a closed-form and is readily interpretable all these made as popularity (Ben-Akiva and. Lerman, 1985). The logit model allocates person trips to alternative modes by comparing the utilities of all alternative modes. The mathematical framework of logit models is based on the theory of utility maximization (Almasri and Alraee, 2013). There are three basic types of logit models depend on whether the data or coefficients are chooser-specific or choice-specific. The multinomial logit model has chooser-specific data where coefficients vary over the choices. The conditional logit model has choice-specific data where the coefficients are equal for all choices. The logit model involves both types of data and coefficients. Briefly presenting the probability of an individual i selecting a mode n, out of M number of total available modes, is given as,

$$P_{in} = \frac{e^{V_{in}}}{\sum_{m=1}^{M} e^{V_{im}}} \quad (2.3)$$

Where,

V_{in} is the utility function of mode n for individual i;

V_{im} is the utility function of any mode, m in the choice set for individual i;

P_{in} is the probability of individual i selecting mode n; and

M is the total number of available traveling modes in the choice set for individual i.

All these three assumptions serve as the main postulates of the structure of logit models. The first assumption of the random component being Gumbel distributed indicates that all the utilities associated with the traveling modes should be considered as a linear sum of attributes and has the same scale parameter. The last two assumptions are normally grouped together to be referred to as a property of Independence of Irrelevant Alternatives (IIA property), simply meaning that all the travel modes used in modeling the travel behavior are independent of each other. Logit models are generally classified into two main categories namely binary and multinomial logit models. Binary choice models are capable of modeling with two discrete choices only, i.e. the individual having only two possible alternatives for selection, whereas the multinomial logit models imply a larger set of alternatives.

I. Binary Logit Models

The mathematical framework of a binary logit model can be represented by simplifying equation 2.2 with the total number of available alternatives limited to two, i.e. M = 2. An example of a binary logit model where the choice set contains car and public transport as two competing alternatives. By simplifying equation 2.2, the probability of individual i selecting the mode, m out of two available traveling modes m and n is given as,

$$P_{im} = \frac{e^{V_{im}}}{e^{V_{im}} + e^{V_{in}}} \quad (2.4)$$

Where,

V_{im} is the utility function associated with alternative m for individual i;

V_{in} is the utility function associated with alternative n for individual i;

P_{im} is the probability that alternative m will be selected by individual i;

II. Multinomial Logit Models

The multinomial logit model is widely used in a variety of transport choice contexts. Compared with the other choice models, the multinomial logit model is particularly attractive in many model developments due to the nature that is linked to the decision-making behavior through maximizing the utility (Li, 2011). A multinomial logit model is a regression model that generalizes logistic regression by allowing more than two discrete outcomes. The multinomial logit model is a model that is used to predict the probabilities of the different possible outcomes of a categorical dependent variable, given a set of independent variables. The multinomial logit model can be derived from the binary logit model to deal with more than two modes. The multinomial logit model is categorized into simple and nested multinomial logit models based on the characteristics of available traveling alternatives in the choice set. Possible levels of the dependent variable (mode choice) used for this study are shown below. The independent variables are those factors used to explain or predict the mode choice (e.g. trip purpose, demographics of travel, etc...). In this research, the multinomial logit models use the same mathematical framework and generally estimated using the maximum likelihood method. The simple multinomial logit models are presented below in Figures 2.2.

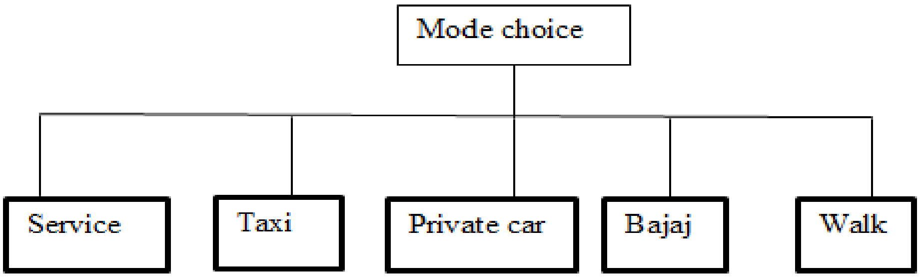

Figure2. 2.Multinomial logit modal choice alternatives

The mathematical form of the multinomial logit model is as follows. Suppose there are m total travel modes of interest (1, 2, 3… M) and that there are k factors (1, 2, 3… K) that is being used to predict the probability of a particular mode choice. These k factors, in general, may include continuous, binomial, or categorical data. To construct the logits in the multinomial case, one of the modes is considered the reference level and all other logits are constructed relative to the base. Here mode M is taken as the reference level. The probability of an individual i selecting a travel mode, m, out of M number of total available modes, is represented as p_{im}. The relationship between this probability and the K factors is given by the following multinomial logistic regression model

$$\text{Log}\left(\frac{pim}{\text{piM}}\right) = \beta_{0m}+\beta_{1m}X_{1i}+\beta_{2m}X_{2i}+\cdots\cdots+\beta_{km}X_{ki};\ \ m = 1, 2\ldots M\text{-}1,\ i = 1, 2 \ldots n. \qquad (2.5)$$

Where;

X1i… Xki are the k number of factors of mode, m for individual i;

β_{0m} is the mode-specific constant for mode, m;

$\beta_{1m} \ldots \beta_{km}$ are k number of coefficients of mode, m which need to be estimated from the data;

M is the set of all available traveling modes, and n is the number of individual/trip combinations in the dataset.

The above equation can be solved to yield the probability of an individual *i* selecting a travel mode, *m*, out of *M* number of total available modes as

$$p_{im} = \frac{\exp(\beta_{0m}+\beta_{1m}x_{1i}+\beta_{2m}x_{2i}+\cdots+\beta_{km}x_{ki})}{1+\sum_{l=1}^{M-1}\exp(\beta_{0l}+\beta_{1l}x_{1i}+\beta_{2l}x_{2i}+\cdots+\beta_{kl}x_{ki})}$$

and for the reference category,

$$p_{iM} = \frac{1}{1+\sum_{l=1}^{M-1}\exp(\beta_{0l}+\beta_{1l}x_{1i}+\beta_{2l}x_{2i}+\cdots+\beta_{kl}x_{ki})}$$

For this research, the model in equation 2.4 was fitted separately to the two different trip purposes (work, education, and education). The number of modes (M) was equal to five (walking, Bajaj, private vehicle, taxi, and service) where the taxi was considered the base level. The predictive factors included in each model are summarized in Table 4.3.

The form of the discrete choice multinomial logit model used in this research is based on the assumption that the choice of mode is a function of the characteristics of the traveler and/or the trip.

2.4. Review of similar studies

2.4.1. Previous studies in the trip production model

Different Studies have been applied around the world in the area of trip production modeling. The studies aimed at developing a trip production model that fits with the study area. In this research, it presents some studies in cities of developed and developing countries in order to benefit in developing a suitable trip production model for the study area.

Several research papers studied the variables that affect the trip production model and found some significant relationships among the variables. Most of the empirical researches has been conducted in developed countries and less empirical researches also conducted in developing countries from these, the researcher has summarized the results as follows.

Al-Masaeid and Fayyad's (2018) study were to develop trip production for residential areas, where Irbid city was selected as a case study. The city consists of seven regions and from that household survey were randomly selected and interviewed at their homes to collect data on trips and their purposes on typical workdays and holidays. Socio-economic characteristics of households were also obtained through field interviews. Both regression analysis and cross-classification approaches were used to model trip production rates and also a random sample was selected from all regions to check the performance of the developed regression and cross-classification approaches. The analysis carried out in the study indicated that the number of generated trips is influenced by family size, car ownership, and income level. On workdays, the analysis indicated that the number of home-based work trips constitutes about one-third of the total home-based trips. Also, it was found that the number of trips on holidays represents nearly one-third of the number of trips generated on workdays. Although the outcomes of both approaches were comparable and consistent with the actual field data, the rate approach provides better estimates. Thus, both models are applicable to estimate residential trips, provided that the data are within the limits and context of Irbid city. The performance of both approaches was very well, compared with developed countries, trip production rates for residential areas in Jordan were found to be substantially low.

Ashengar (2010) compares empirical trip rates modeled by way of multivariate regression and that of category analysis in Skudai town in the state of Johor in Malaysia which is selected as a study area in the study. Origin and Destination interview surveys were conducted to determine the method used for traveling. From the data collection, correlation analyses were carried out with trips as the dependent variable and independent variables to test the ability of those variables in estimating the observed trip production. Then the result was used to determine how well the trip production equation matches the surveyed data. Category and multivariate regression household trip production analysis techniques have been compared and contrasted. Although according to having a fixed formula predicting future travel, using the regression model seems much better and easier than Category Analysis. The study also recommends that the regression approach should be further investigated since it has advantages as a model building procedure and makes better use in a minimum sample data.

Chang, et al. (2014) investigated home-based work trips per household in Seoul metropolitan area using six models regression, Tobit, Poisson, ordered logit, category, and multiple classifications and compared the results. They considered regional characteristics such as demographics, regional economy, and transport system also in addition to household characteristics. The data for household characteristics and trip production rates were taken from a household travel dairy survey of the Seoul metropolitan area and regional characteristics were based on the data from the Korea Regional Development Total Information System (REDIS). The Regression model gave an R2 value of 0.4. The comparison of six models based on RMSE and NRMSE, measures of difference between observed value and model value revealed category-type models to be superior of all. Further, the regression model also produced acceptable performance.

Dodeen (2014) develop trip production models to predict the number of trips generated by households in the Palestinian areas considering Jericho City as the case study. The models are developed using multiple linear regression analysis techniques. The developed model by the study include three types of models, which was a general trip production model (i.e., a general model regardless of trip purpose and trip time), trip production models by trip purpose (work, educational, shopping, social, and the recreational) and five trip production models by trip time. The data consists of primary data, which was collected by conducting the representative household survey randomly selected from each of the 14 TAZ's. of Jericho City, the study area.

The results indicated that the estimated general trip production model has a good explanatory power and the variables that mostly affect trip production are found to be the number of persons receiving education in the household, the number of employed persons in the household, as well as the household monthly income. The number of employed persons in the household and the number of age between 31 to 50 years are the variables that mostly affect the work trip model. In the educational trip production model, the number of persons who are receiving education in the household is the main factor in this model. The shopping trip production model depends on the number of persons in the household and the monthly household income. The social trip production model depends mainly on the number of females in the household and the number of employed persons in the household. Finally, the recreational trip production model depends mainly on the number of persons receiving education in the household, the number of persons between 51 and 64 years old, and the monthly household income. For the temporal trip production models, the morning peak period and the evening peak period have the largest R-square values among all temporal trip production models.

Hu (2010) calibrated trip production model for work trips per household using National Travel Survey data, a household survey of travel covering residents of Great Britain. In addition, the author developed the trip production model for shopping trips per person using Edinburgh household survey data. In the shopping trip production model, parking costs reflecting the impact of transport policies were also included along with other attributes. Linear regression analysis and logistic analysis (binary, MNL, and NL models) were employed for both investigations. The linear regression trip production model of work trips yielded an R-square value of 0.326 and shopping trips yielded 0.135, which is very low.

Hunt and Broadstock (2010) developed a trip production model for a cross-section of residential developments in the United Kingdom. The empirical model tested whether trip making patterns for residential developments are independent of car ownership. The result was that trip production is dependent on car ownership, socio-economic factors, and site-specific characteristics, in particular, land-zone type. However, public transport services are not found to have a significant relationship with trip production. Consequently, a policy implication of the results was that increasing bus services to residential developments are not associated with a reduction in generated trips.

According to the research findings, there is a linear relationship between trip production and car ownership, household size and household income. These research outcomes are compatible with those of previous studies (Sowgat, 2012; Daniel and Ituen, 2013). And also these studies revealed that residential trips are mostly related to socio-economic variables.

Moussa (2013) conducted a study to develop a trip production model for Gaza city, Palestine to determine the household travel characteristics pattern in the study area. A household interview survey was conducted to collect primary data. The survey was distributed to 425 households in different districts of Gaza city. The results indicated that vehicle ownership, household size, income level, and the number of licensed drivers in the household are the main factors that affected trip production in Gaza city.

Pattersson and Schmockee (2010) conducted a study to analyze travel patterns by those aged 60 or over in Metro Manila, the Philippines. Trip frequency and tour complexity were analyzed with ordered probit regression, separating the effects of socio-demographic characteristics as well as land-use patterns. The results were compared to observations made for cities in developed countries, in particular, London as an example for a city in a first-world country. The research showed that there is a pronounced decrease in total trips made with increasing age in Manila. However, analyzing for specific trip purposes, the authors found similarly the number of recreational trips is fairly constant in all age groups to trends in developed countries.

Pokhrel (2017) conducted a comparative analysis of person-category and household-category trip production models. The data considered for the analysis excluded the weekend survey data. The analyses of travel survey data were performed using R software for statistical computing and graphics. Trips were categorized as home-based work, home-based education, home-based shopping, home-based other (for remaining home-based trips) and non-home-based trips. Two conventional techniques namely, multiple regression and cross-classification were used to model each of the five trip categories. Later, the benefits of these techniques were also analyzed. At the household level, socio-economic and demographic attributes of the households were included as independent variables while both personal and household attributes were included at a personal level. However, no household attributes were found to impact person-category models.

Later, household-category and person-category trip production models were compared based on the coefficient of determination (R^2), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE) values. In the case of multiple regression, except home-based shopping trips, all other trips yielded higher R^2 values at the household level. In the same way, all person-category trip production models yielded lower RMSE and only home-based shopping and home-based education yielded lower NRMSE at a personal level. At both levels, home-based work and home-based education performed well as they managed to explain more than 45% of the total variation in the model. In addition, this trend from the multiple regressions was replicated by cross-classification. Thus, with these results, the thesis concluded that neither household nor person-category could be claimed to be superior in trip production modeling. The results of the research concluded that neither household nor person-category could be claimed to be superior in trip production modeling.

Priyanto and friandi (2010) developed a trip production model for public transport passengers in Yogyakarta, Indonesia, using multiple linear regression analysis. The researcher established a relationship between the number of trips and socio-economic attributes. The data consisted of primary and secondary data were collected by conducting household surveys, which are randomly selected. The developed model showed that public transportation trips seem to have a negative correlation with income, motorcycle ownership, and car ownership. This means that the number of trips made by people decreases as income, the number of motorcycles, and cars owned increases. The model also showed that the number of public transport trips increases as the family size increases. Generally, the higher the number of family members, the more public transport trips will be made.

Sarsam and Al-Hassani (2011) developed statistical models to predict trip volumes for a proper target year in the Al-Karkah side of Baghdad city, Iraq. The traditional method to predict the trip production volume according to trip rate based on the family type was used in the study. The total study was divided into 45 zones based on the administrative divisions. A questionnaire was designed and interviewed at the selected zones from the study area. The developed models were total trips per household, work trips per household, education trips per household, shopping trips per household and social trips per household. These models were developed by SPSS software using a stepwise regression technique. The results showed that total trips per household are related to the family size and the structure variables such as the number of workers, the total number of students in the household, and the number of private vehicles. The model had a coefficient of determination was 0.67 for the whole study area. The results also showed that the home-based work trips are related to the number of workers in the household, a number of male and female workers in the household, and the number of persons between 25 and 60 years of age and the home-based education trips are strongly related to the number of students in the household. These models have a coefficient of determination equals to 0.82 and 0.90 to home-based work and home-based education trips respectively for the whole study area.

Sofia et al. (2012) developed a relationship between the daily household trips and socioeconomic characteristics for Al-Diwaniyah city, Iraq. The authors used the technique multiple linear regression techniques after the collected data had been fed to the SPSS software. The city was divided into five sectors with seventy zones covering an area of fifty-two square kilometers. The results showed that the trip production model mainly depends on family size, gender, the number of workers, and the number of students in the family.

Urquiza et al. (2010) developed a trip production model based on households through crossed classification techniques in the municipality of Palmira, Colombia. This paper proposed a trip production model, for the municipality of Palmira, based on the Crossed Classification technique, with a special interest put on the trips that are started or ended in the household, or household based-trip.

The classification of the information for the definition of Traffic Analysis Zones, the characterization of the same according to the household survey and the production model developed were spatially represented by the development of thematic maps as an analysis tool allowing quickly understanding the behavior of each zone on different temporal boundaries, recognizing trends, concentrations, and variations. This cartographical tool helps the interpretation and explanation of the behavior of variables that participate in the building of the model.

The work methodology was started with the zoning of areas relatively homogeneous Traffic Analysis Zones. Subsequently, with the collected data from the municipality through of Origin-Destination Household Survey, the average income for each Traffic Analysis Zones and the percentage of household on each income range, on this stage were estimated, in the making of graph relating the average income by Traffic Analysis Zones and the percentage of families by income range, the interpolation method of Cubical Tracer was used, the vehicle holding by the family for each income range, the number of the trip by household by day for each income range with vehicle holding, the number of the total trip by day generated on each Traffic Analysis Zones were calculated, and finally, the trip purpose categories: work, studies, and others; were generated. The technique of crossed classification was found that some difficulty on the incorporation and interrelation of new variables, limits the precision of estimation of future demands but the stability of supposed relations was acceptable.

2.4.2. Previous studies in the mode choice model

Many studies have been applied in the world in the area of mode choice modeling. The studies aimed at developing a mode choice model that fits with the study area. This research presents or reviewed some studies in order to help in developing a suitable mode choice model for the study area.

Abdulhaq (2016) developed a transportation mode choice model to predict the utility function for transportation modes used in Palestinian universities. The developed model covers three modes of transportation, Private car, Shared taxi, and Bus. The variables that mostly affected the mode choice are gender, car ownership, per capita family income, shared taxi travel cost, and bus travel time. As expected, those who own private cars prefer going to university by their cars more than using public transportation modes. The results also show that the use of private cars increases with the increase of the monthly family income divided by the family size (i.e., with the per capita income).

Almasri (2011) investigated the factors that affect travel choice of shared taxi versus bus for Palestinian university student trips. The results of this study indicated that factors that are significantly affecting the mode choice of students are: family income divided by family size, weighted travel time, out of vehicle travel time divided by distance. The results also show that the age and gender variables are statistically insignificant and it could be dropped from the model.

Al Raee (2012) developed a mode choice model for work trips in Gaza city and the results of his research showed that the factors that the total travel time, the total cost divided by personal income, ownership of means of transport, distance, age, and average family monthly income are the main factors that affect the choice of transportation modes. The results also indicated that the travel time, fare divided by personal income, frequency of service, age, average family monthly income and distance are the factors that affect the work mode choice model.

Amit, (2017) studied modal preference for school trips of college-aged students by using multinomial logit regression methods to model in Panipat city, India. A case study covers the age group of 16-22 years of different college students of Panipat institute of engineering and technology was carried out. In addition to safety; age, department, residential area, gender, and family class are the factors considered while developing the modeling. In the survey observed that 60% travel through college buses, 19 % through cars, 10 % walk to the college, 8% travel through motorbikes and rest through other modes of travel to colleges. Choices for trips different from colleges were also studied such as trips to picnic, trips to the market or other recreational areas. In this study, safety is the major concern for female students as compared to male students. Students with a good family background in terms of economy prefer personal vehicles more frequently in traveling. College bus is the most preferred mode due to the fact that the college campus located in rural areas away from the city. Students residing in the urban and sub-urban areas prefer personal vehicles more frequently as compare to students of rural areas. Through the study, other trips by the college-going students such as trips to the market or recreational places observe that they choose personal vehicles such as cars and motorbikes more frequently as compare to other modes. The study observes that students like to drive themselves to college but not possible due to family pressure and long-distance between college and their homes due to this they have to choose a college bus in most of the cases. The researcher concludes that in the future, the “model equation” model generated above in this article would be very helpful in future planning which provides the probability of choosing mode other than bus by school pupils. Parameters such as department, age, residence, gender, and family class are considered to be the contributing factors in estimating the overall pattern of the college students.

Danaf et al. (2014) developed a discrete mode choice model to predict the utility function for the choice among car, bus and shared taxi for students of the American University of Beirut and the general population of the greater Beirut area. It was found that travel time, cost, income, auto ownership, gender, and residence location are the main factors affecting mode choice.

Lin and Yu (2011) empirically analyzed the effects of the built environment on leisure travel among children. A multinomial logit model was used to analyze travel mode. This study reached the following empirical findings: intersection density, building density, employment density, and walkway quality encouraged a child to use transit systems or non-motorized travel modes for leisure travel; and vehicle density and leisure travel distance discouraged walking and biking but encouraged the use of transit systems for leisure travel involving children.

R Ashalatha, et al. (2013) studied the mode choice behavior of commuters in Thiruvanthpuram city and factors influence the commuters to use public transport. The multinomial logit model is used. A pilot questionnaire survey was conducted. Factors considering for the study are mode of transport (Bus, Two-wheelers, car), age group, gender, monthly income, vehicle ownership, distance, time/distance, and cost/distance. That is variables like socioeconomic variables, transport system variables, and attitudinal variables are discussed. MNL analysis revealed that those who own both car and two-wheeler prefer two-wheeler for shorter trips and car for longer trips. Age, gender and income have a massive role in the mode of choice. The study concluded that as the age of the commuter people prefer cars than a two-wheeler. In this paper, Modelling is helping to make policies and to introduce various schemes to improve the government-owned bus system.so it helps to be more attractive to public transportation.

Eluru, et al. (2012) explained that travel mode choice results clearly highlight the role of travel time, walking time, and initial waiting time on the propensity to choose transit. The results indicated that individuals find travel time on the bus mode the most onerous. 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 were also reported.

Thushara T et al (2013) worked on modeling of mode choice for work trips in Calicut city. The study considered different categories of work trips as government, private and self. The data was collected from interviews and random sampling methods are followed. A multinomial logit method was used. The preliminary analysis was done and identified the different characteristics. The study considered the traveler's gender, age, profession, and travel modes consider are car, two-wheeler, bus, and auto and the most influenced mode on the Calicut city is two-wheeler. Male commuters are dominant in the usage of total modes as well as a two-wheeler. In this study, the major concern is how employee‘s mode usage would change as they are getting older. Therefore, age is selected as an explanatory variable. Software SPSS 16.0 is used which is common in the development of models. Model fit is checked by pseudo-R-square. The prediction of the prepared sample is 86% and validated successfully. The study concluded that traveling time and traveling cost for work trips are the main factors can be used to identify the characteristics that influence mode choice behavior of employees.

Ulfarsson (2016) examined not only the effects of trip characteristics but also personal and household characteristics on the mode choice of older people, using data from the Household Travel Survey in the United States. This study is conducted due to experiencing a rapid increase in the number and proportion of the older population aged 65 and older. It was tested if the effects of characteristics were the same for the 65+ and the under 65 groups, and the analysis showed the effect of personal and household characteristics on the trip making varied with age based on the null hypotheses that were tested has no difference in effects between those under 65 and those 65plus. The main variables that had a statistically significant effect on increasing the probability of more trips were: being a householder, having a driver's license and automobile ownership, not being employed, and higher monthly household income. The effects of personal and household characteristics on the trip making of older adults were also different for the mandatory and discretionary trip purposes. Full-time work reduces mandatory trips and increases discretionary trips. High income has a greater effect on increasing mandatory trips than discretionary trips. Men make fewer trips than women, especially fewer mandatory trips. Trips were also classified into mandatory and discretionary trips and models estimated for the age groups greater than 65 for those two trip purposes. Specific results from the investigation were done showed that men 65 or older have a lower probability of making more trips than women. Higher-income was linked to more trips for older adults, especially mandatory trips. The results show the development of transport policies needs to consider the variance among the aged, in the case of gender, income, driving licensure and automobile ownership.

2.4.3. Review of previous studies in Ethiopia

In the context of Ethiopia, there may have more studies but the published studies which are uploaded on the internet there appear to be very limited work on the transport modeling especially on multinomial logit models. The studies that are previously done in the country are reviewed their methodology and findings briefly for each case.

The transport demand forecasting model developed for Addis Ababa by (Torrieri, 1985) had four sub-models (Generation model, Distribution model, model split model, and Assignment model). The modal split model, at first, splits between pedestrian and motorized mobility then the percentage of a pedestrian for each trip from a certain origin to a certain destination is computed. By assuming that the only determining factor in the choice between private and public transport modes is the economy a model is developed depicting the relation between household income and the share of private vehicle modes for various trip purposes. The first model ended by dividing the motorized trips into the motorized car and public transport but the second model has gone further developing a logit model for the two competing modes of public transport Minibus taxi and Anbessa bus by considering the relative cost and service properties of the trip and ignoring the socio-economic characteristics of the trip makers.

The Ethiopian Roads Authority (2005), with the assistance from the World Bank, has undertaken an urban transport study as a project entitled, "Urban Transport Study & Preparation of Pilot Project for Addis Ababa". In this study, the first mode choice between walk and vehicle trips was determined using income levels. For this, the zone-wise trips by walk mode and by purpose were also estimated. The trip that ends with vehicular modes has been obtained by deducting the walk trip share from the total trip ends. The analysis was undertaken upon the influence of income on the choice of mode in terms of the contribution of private modes to the total vehicular trip in order to understand the dynamics of mode choices. In the best fit curve from scattering data, the result shows that as income increases the share of private vehicle trips increases. Also, the rate of change per unit change in income decreases. The relation that explains the share of private vehicle trips for work purposes has a power function. A linear relation has also been observed to explain the variation in the share of private vehicle trips with respect to income in case of education and other purpose trips and a power function in case of non-home based trips.

For public transport users, the study considered two facilities; Anbessa Bus service and Minibus as available competing modes and a Logit model was developed in order to assess the choice behavior between Anbessa City Bus service and Minibus service based on the generalized cost of travel by these services. The study assumes that the first choice between walk and vehicular modes is influenced by the economic status of the household. Moreover, mathematical models were developed in order to assess the relationship between the choice of vehicular modes/walk and the economic status (i.e. income) of households. And finally, the model was calibrated for the observed data for the two competing modes Anbesa city bus and Minibus for work, education and other trip purposes.

Ali (2018) conducted a study on the development of a modal choice model for home-based work trips-the case of Jemo area residents in Addis Ababa. The study also aims to contribute to the improvement of the service quality of public transport systems by determining the basic factors that affect the preference of modes and how each factor affects the mode choice of trip makers to travel to the workplace in the study area. 475 questionnaires were collected in which 4/5th of the data were used for calibration and 1/5th for validation of the revealed and stated preference models. The collected data of variables related to the trip maker and characteristics of the model are fitted in a special type of multinomial logit model known as the conditional logit model using SPSS. Both revealed and stated preference model was developed. Preferences among large buses, minibus (taxi) and private cars were revealed, whereas the stated preference model was developed for preferences among the existing modes (large bus, minibus taxi, and private car) and the bus rapid transit 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 variables by taxi. The stated preference models indicated that the preference towards the bus rapid transit increases with the decrease in total travel cost, in-vehicle travel time and out of vehicle travel time.

Etsub, (2015) developed the multinomial logit model for work and school work trips along the railway route corridor in Addis Ababa, 183 and 100 trip makers were considered for the work and school trips respectively. A Multinomial Logit with statistical data processing software SPSS 20 (Statistical analysis in social science) was used for explaining travel patterns and mode choice of both employees and students in the study area. 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. It was also found that both workers and students who travel long distances choose Higer and Anbesa than a taxi to decrease the transition time and energy wastage during transferring from one taxi to another. The main modal shift to Light rail transit is found that from Anbesa, Higer, taxi and the least interest is shown from private car users. In the study, both revealed and stated preference models which consider the light rail service was included.

Gebeyehu and Takano (2005) studied the diagnostic Evaluation of Public Transportation Mode Choice in Addis Ababa. The research study attempts to develop an ordered logit model and a binary logit model to examine citizens' perceptions of the bus condition, with a determining factor for their choice of bus transportation and to analyze traveler choice behavior respectively. The diagnostic analysis is undertaken based on the two models. The result of the study using logit models shows that citizens' perceptions of the three chosen bus condition aspects of these; fare, convenience, and frequency have a significant influence on public transport mode choice. The sample survey of residents was undertaken in a door-to-door survey (interview) methods. The analysis results showed that socioeconomic as well as public transport mode related parameters have a significant negative or positive influence on the mode choice and the respondents' perceptions of the bus service condition. Based on the results and further analysis to improve the existing public transport services the strategic plan, policy indications, and physical solutions could be drawn as an effort. The spatial expansion of bus service, improving bus frequency conditions, and strengthening the bus linkage with respect to the land use are among the policy intervention areas

Teshome, (2007) developed a logit model of work trips for Addis Ababa Bole-sub city residents. The methodology adopted for this study is using a multinomial logit model by collecting both primary and secondary data. The primary data used for the research was generated from the household interview survey and the secondary data obtained from different authorities of the Addis Ababa City Administration. This study is done using disaggregates data and Statistical software STATA was used to develop the Multinomial Logit Model. The independent variables in the final optimal model include total travel time, travel cost, gender, age, monthly income, gender, child in the house or not and peak hour dummy variable.

CHAPTER THREE

ANALYTICAL FRAMEWORK AND METHODOLOGY

3.1. General Steps of the Methodology

In order to achieve the objective of the study the methodology adopted in this study is composed of existing information and field observation surveys. This chapter will discuss in details the methodology and procedures applied to achieve the main objective of this study and include the following phases:

- **Formulation of Objectives and Literature review:** In the first phase, objectives and goals were selected followed by the literature review on trip production and mode choice modeling.
- **Selection of study area:** This involves selecting the boundaries of the study area and dividing it into small traffic zones, according to the boundary of the study area governorates boundaries.
- **Sample size and designing questionnaire:** Identifying needed information and designing a questionnaire for proper data collection of relevant socio-economic variables and trip data are involved in this part. The researcher conducted reconnaissance visits to all traffic zones in order to capture appropriate samples in the study area. Collecting the required data is directly forward to the collection of data on the field from the sample size designed to the different traffic zones.
- **Analyzing of data**: analysis of data can be done through relevant computerized program SPSS and Excel software.
- **Building models:** This phase is concerned with the calibration and estimation of the model. This involves the trip production and mode choice model by trip purpose in the study area using multiple linear regression and multinomial logit model techniques.
- **Validation of results:** This is done by comparing the outputs of models with the actual numbers to verify the models, and make calibration of the variables used in model building. The details of the steps of the methodology are discussed in the following sections and chapters.
- **Results and Conclusions:** This phase includes results analysis, conclusions, and recommendations that present the main finding of the research.

3.2. Data sources

The data used for the purpose of accomplishing this research was both primary and secondary data. The primary data was collected from the household survey and secondary data were obtained from different authorities of the Mekelle City and Hawelti sub-city Administrations. The respondents of the household survey were people who are currently employed or are self-employed, others and students within the age of 10 and above years old. According to Dodeen (2014), trip details of children below 5 years are normally ignored.

3.3. Pilot study

A pilot study was distributed before collecting the final data from the whole sample. The objective of the pilot study was to test for the questionnaire, which involved testing the wording of the questions; identifying confusing questions, data coding, and data recording and testing the technique for collecting the data. Richardson et al. (1995) described various uses of conducting a pilot survey in detail as, determine the adequacy of the sampling frame; observe the variability of the parameters within the survey population; scrutinize the method used for data collection; check the word of question wording and layout of the questionnaire; and study the procedures of data entry, editing, and analysis. For conducting the pilot study, the preliminary questionnaire was prepared and interviewed to 40 households' respondents representing 10% of the sample size.

3.4. Sample Size of the study area

Sample generation is considered as a vital step in transportation modeling since the trip production and mode choice models are generally estimated using the data collected by surveying a sample of respondents from the targeted population. Therefore, it is essential that the sample generated for the thesis is representative of the characteristics of the population of the study area. Inappropriate sample generation can lead to erroneous modeling results involving biased estimated coefficients and non-representative travel behavior forecasts. 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 increase 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 remain within budgetary and time constraints (Koppleman & Chandra, 2006).

The methods used to determine the sample size for variables that are polychromous or continuous are proportion and mean. Cochran (1977) provides a formula to calculate sample size based on proportion. The way of estimating population variances for sample size determinations by using pilot study results.

The Cochran's sample size formula for the infinite population was calculated as follows;

$$n' = Z^2 * P * q / e^2 \qquad (3.1)$$

Where: n' is the required sample size, Z is the critical standard score, e is the desired level of precision (0.05), p is the population proportion, and q (q = 1 – p). The Z value is found from the Z table. To calculate the final sample size of finite population Cochran's correction formula should be used as follows:

$$n = \frac{n\prime}{1+(n\prime-1)/N} \qquad (3.2.)$$

Where: n is the required sample size, n' is an old sample size and N is the total size of the population.

In order to calculate the sample size based on statistical formulas, a random sample of 40 household respondents is initially taken to calculate the percentage of the population (P) that make different types of trips. Then, these percentages are used in the equation to calculate the actual sample size as shown in. The total population and the total number of households in the study area were 100,651 people and 29599 households respectively. The average household size in the study area was calculated to be 3.4 persons (100,651 /29599).

Trip type	P	1-P	A	Z	e	Old(n')	Modified (n)
Work	0.32	0.68	0.05	1.96	0.05	334	333
Education	0.29	0.72	0.05	1.96	0.05	316	315
Shopping	0.20	0.80	0.05	1.96	0.05	246	245
Recreational	0.12	0.88	0.05	1.96	0.05	162	162
Social	0.07	0.93	0.05	1.96	0.05	100	100
Total	1.00					1,158	1,155

The sample size based on Cochran's formula is 340 households (1,155/3.4).

In a simple random sampling technique to draw samples from the household population with a given number of households and confidence levels, the number of the sample can be determined. Yamane (1967) provides a simplified statistical formula for proportions to calculate sample sizes as shown below;

$$\mathrm{n} = \frac{\mathrm{N}}{1+\mathrm{N}(\mathrm{e})2} \qquad (3.4.)$$

Where: n is the sample size, e is the desired level of precision (0.05) and N is the total household population of the study area (100,651), the minimum sample size is calculated to be 395. This higher sample size is adopted to get a higher degree of accuracy.

Green (1991) recommended that the minimum sample size for regression analysis is,

$N >= 50 + 8 * m$ (3.5.)

Where; N and m are the sample size and number of factors considered in the model respectively. The total sample size required will not include the data for validation of the model which will collect another data one-fourth of the minimum sample size.

3.5. Sampling Method

In this research, proportionate stratified random sampling methods were used to collect the necessary data from the study area. This type of random sampling was selected because of having different population groups; a population that will produce trips from the study area, population within different levels of demography and a variety of socio-economic characteristics in the study area. To ensure sufficient inclusion of all categories of the population, it needs to identify the different characteristics and their actual representation (i.e. proportion) in the population. In stratified random sampling, the targeted population is split into distinct subpopulations (strata). The households sampled from each stratum (Tabiya) are selected randomly from the respective Tabiya address list whereby first calculate the proportion of each stratum within the population is shown in table 4.3.

After that, randomly the household is identified and all peoples in the household that made trips were interviewed until the estimated population/respondents in each Tabiya were obtained. These strata are classified on the basis of various factors of relevant interest to the survey and obtained by the simple random sampling within each stratum. For example, for a mode choice survey, the strata can be categorized on the basis of the users of various traveling modes, i.e. the individuals using private cars and public transport. Similarly, the classification can also be done on the basis of various socioeconomic conditions of the households such as age groups and occupations. The study area is divided into five Tabiya; Adishumdhin, Memona, Hdase, Hayelom, and Selam; this is used as a measure to classify the residents into five geographical strata. It is necessary actually to visit every residence in an area to establish a reliable sampling frame. For this study, Tabiya address lists were taken to be the sampling frame.

3.6. Sampling Unit

Variables can generally be inferred from the unit of analysis. For travel modeling, two units are normally considered: the household and the individual. Household is the preferable sampling unit to use in trip production and mode choice modeling because of various reasons like in a trip making point of, in this case, the home is the basis where most of the trips start and end; from an economic point of view, income or car ownership are usually shared by all members of the household; and from needs were best addressed by considering the total travel by the household, Thus, even though an individual may not travel greatly, the needs of that individual may be serviced if another member of that household was able to travel on their behalf, for their needs like shopping. Therefore, the households in the Hawelti sub-city are used as the unit of analysis/sampling in this thesis.

3.7. Method of survey

In order to conduct the actual travel, the survey revealed a preference survey was undertaken for this thesis. Among the various methods of revealed preference surveys, Household or Home Interview Survey was used as survey methods.

3.7.1. Interview Survey

In this method of the survey, the researcher or data collector visits the home of respondents to ask and fill up the prepared questionnaire. This method of survey used for travel and transportation surveys that provide information and data about the demographic, socio-economic and trip making characteristics of the entire area household's population including travel choice, and location. And also, the purpose of conducting an interview is to obtain the response of individuals with detail information about what the researcher needs. The other advantage of this method is that the interviewers can persuade (through motivation) respondents to answer questions and can explain the objectives of the survey. A door to door (household) interview survey involves going directly to the house of respondents and conducting the interview (Sincero, 2012). The method is the most effective and accurate method of data collection but time-consuming and costly. In addition to the above mentioned, the target population is comprised of all kinds of people. For these reasons, the household survey was adopted for the study.

As such, the survey was designed to collect data that characterize demographics of household, individual and travel patterns of developing models for trip production and modal choice selection.

3.7.2. Document Analysis

The necessary data used for this thesis were obtained from different authorities of the Hawelti sub-city municipality office, Mekelle city administration offices, unpublished and published literature and back-office reports in the Hawelti sub-city. The data were prepared to make suitable for analysis. Since the data that collected from different sources and period are converting to the same format.

3.8. Model specification and Estimation technique

3.8.1. Specification and Estimation techniques of the trip production model

The multiple linear regression is one of the popular forms of model structure, which can be applied for trip production. The most common form of the trip production model is shown in the above Equation 2.1 which is the linear function of the regression equation. The coefficients of the regression equation can be obtained by doing a regression analysis using statistical analysis software.

The trip production model would be estimated using the multiple linear regression techniques by regressing the dependent variables on each of the explanatory variables. In this research, the trip production models are divided into two categories. The first includes the general trip production model and the second includes the trip production models according to the trip purpose.

3.8.2. Assumptions and estimation techniques of the mode choice model

The search for a suitable model specification involves selecting the structure of the model, the explanatory variables to consider, and the form in which they enter the utility function and the identification of the individual's mode choice (Orrtuzar.J.d.D & Willumsen. LG.2004). This means that in the search for model specification, the following four assumptions must be addressed and discussed in detail below:

A. The decision-maker

In this research, the individual or the decision-making unit is a randomly selected individual from Hawelti Sub-City residents. To ensure that the individuals are a resident of Hawelti Sub-City, households in the Sub-city were selected randomly in each household were questioned. Moreover, the person must be made a trip since the model aims to find out what variables influence the trip mode choice behavior of Hawelti sub-city residents.

B. The Alternatives

A study reported on the structural plan of Mekelle city disclosed that different modes of road passenger transport in the city are available. Since Hawelti Sub-city is one part of the city, these modes of transport can be assumed to be available to Hawelti sub-city residents for their different purpose of trips. Therefore, the modes of transport considered in the universal mode choice for the model are:

Walking: - no cost

Bajaj: - there are blue painted with a capacity of three seats. The urban transport study for Mekelle has revealed that short distance or Bajaj provides more personalized services (door-to-door) on a contract basis.

However, during peak periods they also act as “Minibuses taxi” and transport commuters, but at a higher fare than the Minibuses taxi. Fares of short-distance Bajaj’s contract are “negotiable” on time and distance basis. There is a regulated tariff based on kilometer by the transport office but the Bajaj holders did not apply the regulated tariff.

Minibus taxis: - these are the blue-painted colors and providing the regular service in the sub-city with a seating capacity of twelve passengers. The fare of minibus taxi is based on distance (km) which is regulated by the government. Current tariff for Minibuses are 1.70 Birr for a trip less than 2.5km in length, 3.10 Birr from 2.6 km to 4.5 km, 4.15 Birr from 4.6 km to 6.0 km, 5.15 Birr from 6.1km to 10.0 km and 6.20 Birr from 10.1 km to 13.0 km (Mekelle transport office). Taxi is the dominant public transportation mode in Hawelti sub-city.

Private vehicle as driver and passenger: - these include all privately owned vehicles. 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 a passenger as an independent mode, it was found that trip maker who travels by private cars as a passenger were excluded from the study because the data obtained are small in number as compared to other modes. Therefore, in this study, the alternative ‘private car’ refers to a private car user driving by him/herself. The travel cost of a private car includes only fuel cost and maintenance cost. In summary, the total cost of driving is estimated as:

Travel Cost = Fuel Cost + Total Maintenance Cost (3.6)

Company or school provided transport/service: - this includes all company cars, which are not directly given to the individual. A person using this mode of transport should be waiting for the company car at a certain designated place to be picked up and travel to his/her working/education place.

C. The attributes

The choice of mode is affected by the characteristics of the trip maker, the characteristics of the journey and the characteristics of the transport facility (Orrtuzar & Willumsen, 2004). The variables considered under each category are defined as follows.

Characteristics of the Trip Maker

Socioeconomic attributes to be considered in specifying utility functions of travel choice alternatives. For this reason, it is assumed that transport mode choices are formed in response to mobility needs, which vary with individual and household socioeconomic characteristics. Therefore, the variables considered in the model development process are age, gender, car ownership, occupation, education level, monthly income, and household family size.

Characteristics of the Journey

In this study, the journey characteristics like trip purpose are considered and outlined below:

Trip Purpose: - in the trip production model the purpose of the trip for work, education, shopping, recreational and social purpose trips are under-considered whereas the mode choice only for work and education purpose trip model of Hawelti sub-city residents considered in this research.

Characteristics of the Transport Facility

Total travel time: - The specification for total travel time is different for motorized and non-motorized modes based upon the assumption that the utility value of time is not equal for motorized and non-motorized modes of transport. It is expected that the travelers in non-motorized modes to be more sensitive to travel time than travelers in motorized modes. Since the non-motorized modes like walking are physically more demanding than traveling in a car. In addition to that, the total travel time is classified in to out of vehicle travel time (access time, waiting time, egress time) and in-vehicle travel time. The waiting time for walking, car drive, and car passenger is zero, for the case of other modes of transport it is calculated through averaging the waiting time of respondents who have chosen their means of transport. Access time is a time used to reflect walk access/egress to/from the motorized vehicle elapsed by a traveler. It is calculated through averaging access/ egress time from/to home from/to station or location of taking transport.

Total travel cost: - To take account of the expectation that low-income travelers will be more sensitive to travel costs than high-income travelers the cost divided by income is used in place of cost as an explanatory variable. Such a specification implies that the importance of cost in mode choice reduces with increasing individual income.

D. The Decision Rule

The transport mode choice model for the trips in the study area is specified as a Multinomial Logit Model (MNL) with linear in parameter utility, assuming that the random component of utility functions is independently and identically distributed, having a double exponential distribution. This is the simplest and most popular choice model and the probability can be expressed as;

$$p_i = \frac{exp(\beta \mathrm{Xi})}{\sum \exp(\beta xj)} \qquad (3.7)$$

Where; $X_{i,\,j}$ Explanatory variables specific to each alternative

β Parameter estimated coefficients. These coefficients are assumed to be constant for all individuals but may vary across alternatives.

The estimation of the modal choice model the maximum likelihood method is applied. This method is used in order to infer the values of the unknown coefficients $\beta_1 \ldots \beta_k$ showed in the above equation.

The method of maximum likelihood is the most common procedure used for determining the estimators in logit models. Stated simply as, "The maximum likelihood estimators are the values of the parameters for which the observed sample is most likely to have occurred" (Ben-Akiva and Lerman, 1985).

3.9. Data Analysis Method

Statistical software, SPSS, was used to develop the required multiple linear regression and multinomial Logistic regression for trip production and mode choice model respectively. The SPSS program provides a wide range of procedures and tests used in statistics. The best and common method used to determine the parameter of the prediction model is a stepwise method (Abed, 2010). The trip production modeling using multiple linear regression analysis was applied by the method of stepwise while the modal choice modeling using multinomial logistic analysis was undertaken using a forward stepwise likelihood ratio method. A stepwise variable selection procedure, at each step, the independent variable that entered the equation has the smallest probability of F. This step continues until no more variables are eligible for inclusion or removal.

Forward stepwise methods used for the multinomial logistic regression start with a model that doesn't include any of the predictors. At each step, the predictor with the largest score statistic whose significance value is less than 0.05 is added to the model. The coefficient of determination(R-squared) measures the proportion of variance in the dependent variable associated with the predictor variables in a linear regression model. In a categorical dependent variable for the regression model is not possible to compute a single R-squared. Instead of that, Cox and Snell's R^2, Nagelkerke's R^2 and McFadden's R^2 statistics are computed. The model with the largest value of R-squared statistic is best. An adjusted version of the Cox & Snell R square that adjusts the scale of the statistic to cover the full range from 0 to 1 said to be the Nagelkerke's R-squared is and also used in this study.

All the responses of participants for each question were placed into the SPSS software and the percent of each response was calculated. Wide data was reduced to smaller analyzable units through the creation of categories and concepts that were derived from the data using such a coding system some of the qualitative data was quantified. As a result parameter with high percentage means it can affect the person trip production highly. To analyze data descriptive statistics frequency and percentage were computed. Descriptive statics computes to determine associations between variables. The major analysis, however, was set up right away after all the essential information was collected through the stated tools earlier.

3.10. Statistical Tests

The most common statistical tests that are used in the model selection process in multiple linear regression methods are Variance Inflation factor (VIF), R-squared (R^2), F-test and t-test. The Variance Inflation Factor (VIF) to detect multicollinearity more formally. The VIF shows how the variance of an estimator is inflated by the presence of multicollinearity. The R-squared (R^2), measures the goodness of fit of the regression model. The value of R-squared lies between zero and one. A value of R-squared close to one indicates that the model has a good fit, whereas a value closer to 0 indicates that the model has a poor fit. As R-squared, the coefficient of determination of a given explanatory variable with other remaining explanatory variables in the model increases toward one, the VIF also increases. If the value of VIF is larger, that will show the greater the degree of multicollinearity of one explanatory variable with other explanatory variables. As the VIF of a variable exceeds 10, that variable is considered to be highly collinear.

The ANOVA test results are used to show the analysis of the total variance in the dependent variable. The F-test is used to test whether the regression confidence is jointly equal to zero or not. The null hypothesis for testing the overall significance of the model is that the regression coefficients for the explanatory variables are all equal to zero. The alternative hypothesis is that at least one of these coefficients is not equal to zero. Usually, a 95% level of significance for the F-value is accepted.

The T-test is used to test the significance of individual regression coefficients. The t-test simply provides the significance level for rejecting the null hypothesis of having the true parameter value equals zero (rejecting that the parameter has no influence on the utility). If the calculated t-statistic is greater than two in absolute value, it is concluded that the estimate is statistically different from zero at the 95% level of significance.

The most common statistical tests used in multinomial logit analysis are log-likelihood ratio tests, the goodness of fit and sign of parameter value. Likelihood Ratio Test Checks the contribution of each independent variable to the dependent variable. The likelihood ratio test is used to compare two models provided that one is the restricted version of the other by formulating test statistics. The Likelihood Ratio Test provides the significance level for rejecting the null hypothesis that the parameter has no influence on the utility. The goodness of fit (asymptotic $\rho 2$) analysis used to determine the ability of the model to adequately describe the data. The asymptotic ρ2 (rho-squared), the sign of parameter is similar in concept to that of the coefficient of determination R-squared obtained for linear regression models. The difference between asymptotic ρ2 and R-squared is that asymptotic ρ2 is obtained for logistic regression models and R-squared obtained for linear regression models.

The goodness of fit index (ρ2) that measures the fraction of an initial likelihood value explained by the model, which can be calculated due to Mcfadden as:

$$\rho^2 = 1 - \frac{LL(\beta)}{LL(0)} \qquad (3.8)$$

Where:

ρ2 is the goodness of fit index

LL (β) is a log of likelihood function value at its maximum

LL (0) is a log of likelihood function value when all parameters are zero

The test of reasonableness of sign of parameter value allows judging if the variable conforms to a priori notation or theory about its overall behavior with respect to the utility. Variables with wrong signs should always be dropped from the model, even of the parameters pass the t-test. The explanation of parameter signs, parameters with significant negative coefficients was interpreted to decrease the likelihood of that response category with respect to the reference category while Parameters with positive coefficients increase the likelihood of that response category.

3.11. Model validation

After the calibration process is completed and the models have been compared, validation of trip production and mode choice model is checked. In this research, the 1/4th of the reserved data sets were used for validation purposes. The validity of the trip production model was checked by observed values with the value obtained from the estimated regression model and the validation of the mode choice model is done in two stages.

One of the methods of mode choice model validation is the test of reasonableness of the signs of parameters in the model, at this phase; models with opposite signs of parameters from the expected sign are to be dropped. Usually, this phase is completed during the comparison of models in the calibration process. The second one is the computation of the ratio of prediction which is the most important method of validation in the assessment of the prediction capability of the model (Essam & Sadi, 2013). The data that was reserved for validation will be used to predict the mode choice of the trip makers using the calibrated model. By computing the utility of all the modes in the developed model the predicted chosen mode can be determined. Finally, the prediction ratio is computed by dividing the number of modes that are correctly predicted by the total number of validation data.

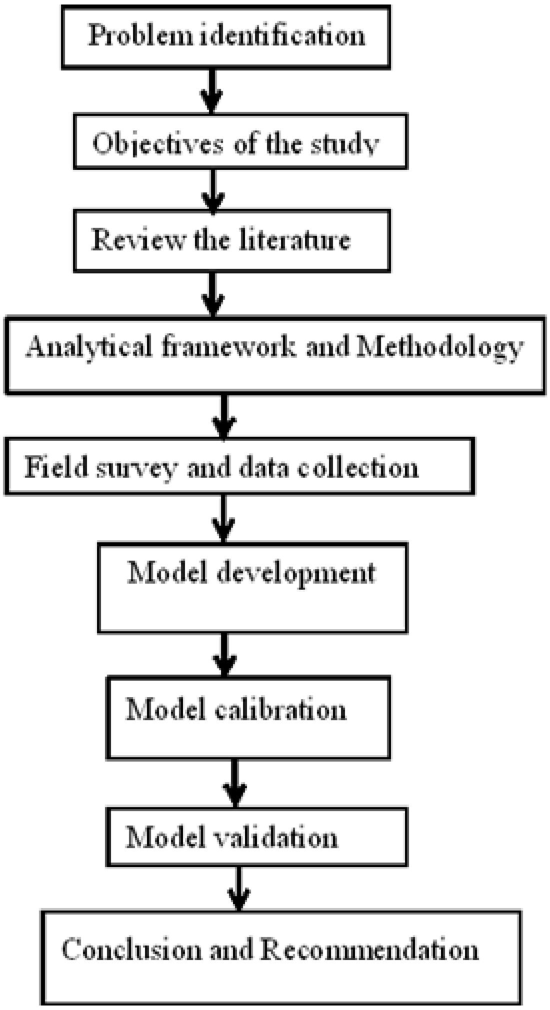

Figure3. 1.Methodology flow chart

CHAPTER FOUR

FIELD SURVEY AND DATA COLLECTION

4.1. Selection and characteristics of the study area

For transportation research purposes, a study area is generally regarded as a geographical region in which transport planning needs to be done, for reasons such as estimating and forecasting the travel behavior of the population by developing a model. It is essential for trip production and mode choice modeler to have accurate information and statistics on the boundaries, population growth, and mode of transport in the study area. The Hawelti sub city from Mekelle city was selected as the study area for this thesis. The main reason for specifically selecting the study area was dictated by the purpose of the study. In addition to that the high population growth rate and the increasing number of vehicles, the high trip origins and destinations, due to the narrow and cobblestone road in the main route, people walk on foot for more than 500-meter distance from or to the housing settlements in one location of the Minibus taxi routes which is a maximum distance from the international standard recommended (Mekelle City structural plan Revision, 2016). Moreover, there is a demand for transportation in the sub-city in which, there are long queues of people waiting for Minibus taxis during morning and evening peak periods. Under this chapter, the various demographics and statistical profiles of the study area were presents in detail.

4.2. Description of the Study Area

Mekelle city is serving as the capital of Tigray Regional State which is located some 783 kilometers from Addis Ababa in the northern part of Ethiopia. Mekelle city has been structured into seven sub-cities each empowered with municipal authority in managing its sub city municipal tasks. The study area, Hawelti sub city is one of the highly populated sub-city and is located in the western part of Mekelle city with a total area of 1732.5 hectares and it holds about 100,651 population (Hawelti sub-city, 2018). The total population in Hawelt Sub-city is increasing rapidly compared to the other seven sub-cities of Mekelle. According to the Mekelle city administration office: the sub-city, at present is undergoing a long process of transition in economic development and future expansion of the city is to the western part of Mekelle city. It is a relatively well developed and populated sub-city.

Now a day, the Hawelti sub-city is organized into five Tabiya's (Adishumdhin, Memona, Hdase, Hayelom, and Selam). Within the sub-city, the highest population is found in Tabiya Hdase and the lowest in Tabiya Selam. Tabiya is the small administrative division in the Hawelti sub-city.

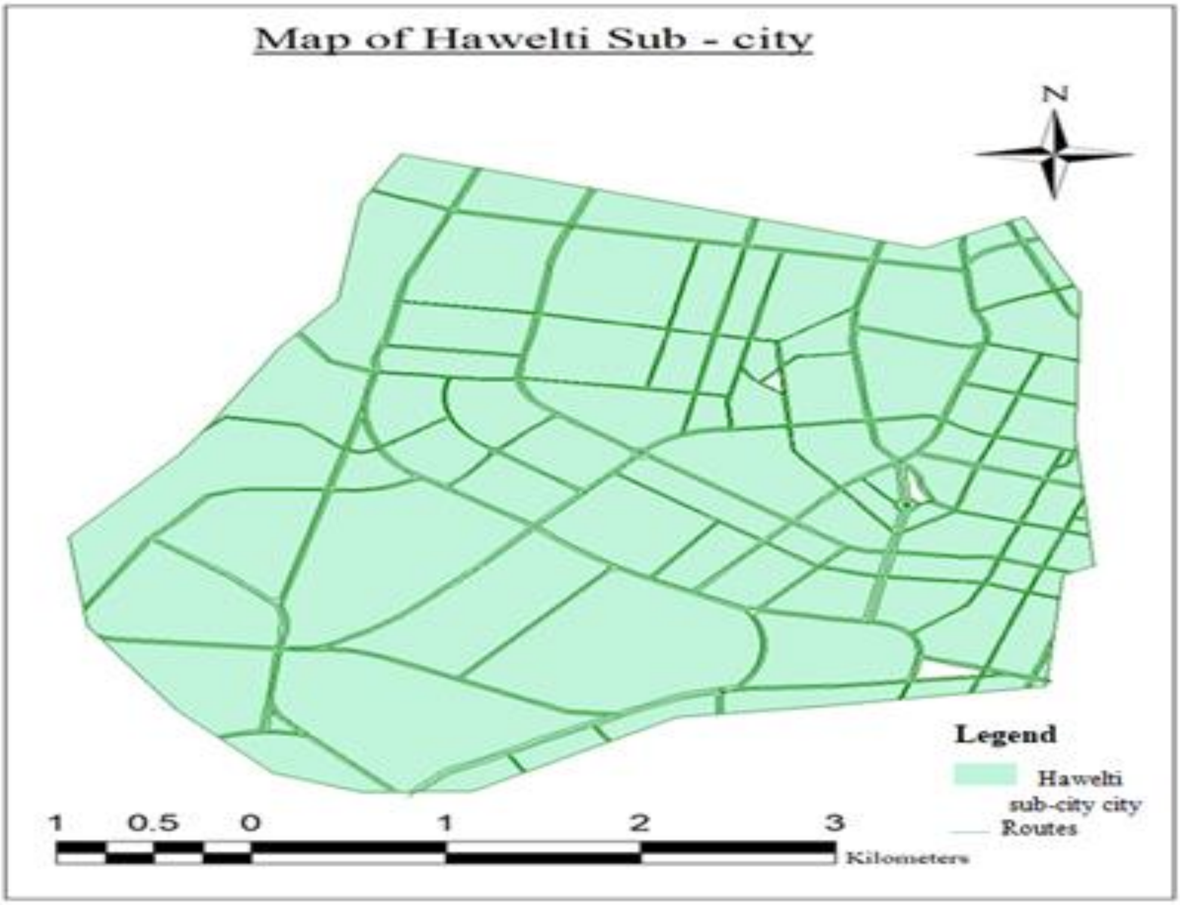

Figure4. 1.Map of Hawelti sub-city.

4.3. The population of the study area

The population of Hawelti sub-city residents is considered as the focus of the survey. The estimated total population of the sub-city for the year 2018 is 100,651 among which 51,955 are males and the remaining 59,348 are expected to be females. Table 4.1 shows the distribution of the population of Hawelti Sub-City by Tabiyas and gender.

Table4. 1.Distribution of Population by Tabiya's and gender

No.	Name of the Tabiya's in Hawelti sub-city	Estimated total population		
		Male	Female	Total
1	Selam	4989	5373	10362
2	Hayelom	5595	8477	14072
3	Adishumdhin	12492	14326	26818
4	Memona	10889	10892	21781
5	Hdase	13458	14160	27618
Total		47423	53228	100651

(Source; Hawelti sub-city municipal office, 2018)

According to the finance and economy development bureau, the sex distribution in Hawelti sub-city for the year 2018 is 47423 male and 53228 female residents. Based on this data the sex ratio of the Hawelti sub-city is 89 males per 100 females. This means the number of males in the sub-city is less by 11% compared to that of females. The sex distribution of all Tabiya's for the year of 2018 has shown that there are more females than males in all administrative Tabiya's. According to the data obtained from Mekelle structural plan report, the population density of the sub-city in the year 2010 was 18949. Out of the seven sub-cities of Mekelle, the Hawelti sub-city was the smallest size of the population with 8.18% of the total population of the city. The population of the sub-city has increased from the smallest size of the population to the second populated sub-city with a 100651 in 20118 (Hawelti sub-city, 2018). Furthermore, the population trends of the study area are presented in table 4.2 for the last few years and population projections for the year 2019.

Table4. 2. Population trends of the five Tabiya's of the study area

Hawelti sub-city Tabiya's	Population (2015)	Population (2016)	Population (2017)	Population (2018)	Projected Population (2019*)
Selam	7477	8336	9294	10362	11553
Hayelom	11947	11320	12621	14072	15690
Adishumdhin	19349	21573	24053	26818	29902
Memona	15715	17521	19535	21781	24286
Hdase	19927	22217	24770	27618	29868
Total	74415	80967	90273	100651	111303

(Source; Hawelti sub-city municipal office, 2018)

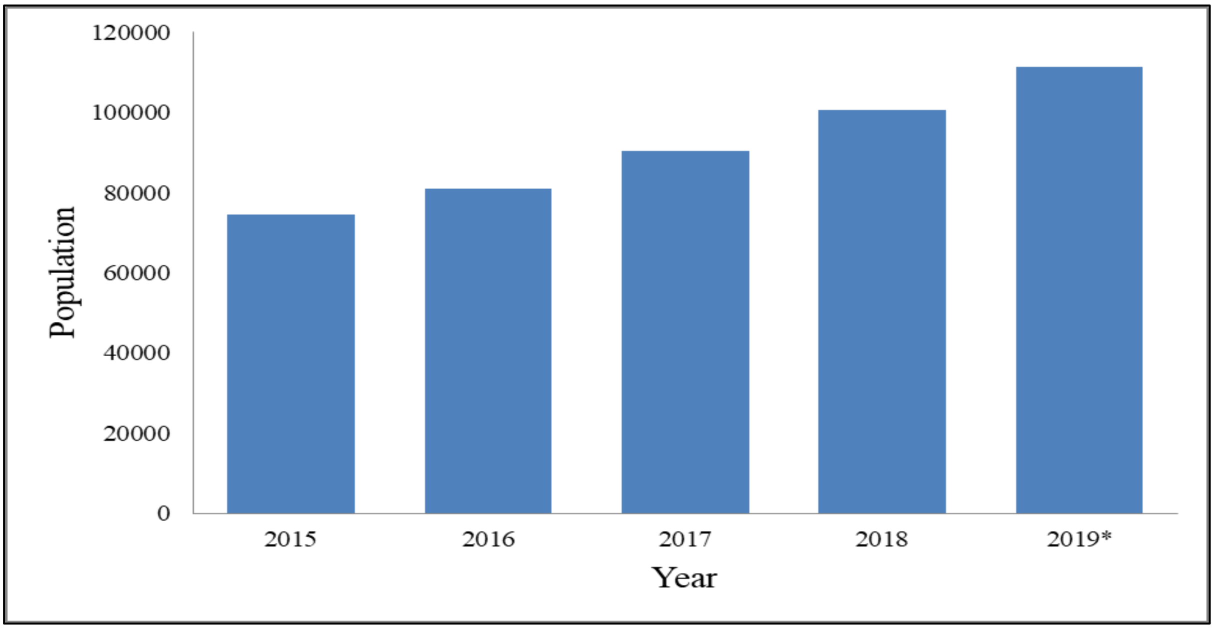

Figure4. 2.Population trends

4.4. Sample Size of the Study Area

The total population and household size of the study area were 100,651 people and 29599 households respectively. The average household size was calculated to be 3.4 persons (100,651 /29599). In order to calculate the sample size that will be used in this study, first, the total population and household size are specified from the Hawelti sub-city. The sample size for the trip production model was calculated based on equation 3.2 is 395. The socio-demographic of employees and students' data collected for trip production model development is also used for mode choice modeling. The sample size was checked by using equation 3.3 for the recommended minimum sample size for logistic regression analysis. In the student's case, there were eight independent variables and in the case of employees, there were 11 explanatory variables. The sample size required for work and school purpose is above the recommended minimum sample size, the indicated numbers were taken as a sample. Therefore the sample size for employees taken for the study is 501 and 574 for the case of students. An additional data from the study area were collected for validation of the model which is one-fourth of the minimum sample size. The total collected data for the study area was the validation and calibration data that are shown in the table below 4.3.

Table4. 3. Sample size from the statistical formulas

Hawelti Sub-city Tabiya's	Selam	Hayelom	Adishumdhin	Memona	Hdase	Total
Population	10362	14072	26818	21781	27618	100651
# of Housing units	3047	4138	7887	6405	8122	29599
% of Housing units	10.29	13.98	26.65	21.64	27.44	100
Sample size for trip production model (calibration + validation)	51	69	132	107	135	494
Sample size for work mode choice model (calibration +validation)	74	100	191	156	197	718
Sample size for school mode choice model (calibration +validation)	64	88	167	135	172	626

The actual sample size computed for this study satisfies both equations (3.2., 3.3) that were discussed in the subtitle 3.5 sample size determination for this study. In the implementation of this study, respondents were interviewed at their homes/residents. The sample that is drawn from the household population is stratified random sampling. Stratification reflects geographic locations defined as traffic analysis zone. From these interviews, relevant data on the trip characteristics for each member and detailed individual and household socioeconomic information and mode choice selection were obtained. The study area was divided into five traffic analysis zones. The samples are randomly selected from each of the five traffic analysis zones. The sample sizes for each traffic analysis zone are shown in table 4.3.

4.5. Questionnaire Design

Once the survey objectives and the tabulation plan have been determined, the relevant questionnaire can be developed. The questionnaire plays a central role in the survey process in which information is transferred from the respondents to the user/researcher.

Questionnaires are best used for collecting factual data and appropriate design is essential to ensure that responses are obtained valid to the raised questions. The questionnaire was developed in order to collect the data required for the development of trip production and mode choice models. The questions were arranged, summarize and formulated as follows:

- What is the number of family size and composition, average family monthly income?

- What is the type of occupation and educational level?
- Do you own a private car by the household?
- What is the purpose of the trip?
- What are the origin and destination of the trip?
- Which transport mode do you usually use to travel?
- What are the total travel time and total travel cost with respect to the mode of transport?
- Is the time of leaving your home to travel at the morning peak hour?

The purpose of this data is to analyze the socio-economic status of the subjects in order to know the basic issues play in the decision making with regards to trip production and mode choice. Both open and closed choice questionnaires were used here in order to collect the quantitative data regarding trip production and mode choice which can be expressed in number or percent based on the socio-economic characteristic of the subjects. And also the open and closed questionnaire was used to collect qualitative and quantitative research data which used to get depth information on characteristics of travel and number of trips originating or destined for particular traffic analysis zone.

4.6. Data Collection Instruments

Instruments used for data collection in this study were an interview, direct field survey observations, and document analysis. Questionnaires were prepared to collect the primary data and in general, it had two parts (See Appendix A).

In this study, a household interview survey was carried out in 395 households for the development of the models and 99 households for validation to the Hawelti sub-city residents in five wards. The survey was conducted for a period of seven consecutive days. Travel behavior of individuals was analyzed with the survey results from the socio-economic profile, trip production, and trip purpose and transport mode standpoints. The first part of this survey deals with the personal and household socio-economic characteristics of the subjects of the study. The designed data for this section of the questionnaire comprises gender, occupation, age, average monthly income, household size, car ownership, and educational background. The ultimate purpose of this data is to analyze the socio-economic status of the subjects, in order to know the basic issue that can play a role in the decision making with regards to trip production and mode choice. Closed choice questionnaires were used here in order to collect the quantitative data regarding trip production and mode choice which can be expressed in number or percent based on the socio-economic characteristic of the subjects.

The second part deals with the travel behavior of the subjects, the design data used for this part include mode of transport used, the purpose of travel, origin, and destination of the trip, time of travel and travel cost of mode choice. Both closed and open-ended questionnaire was used to collect qualitative and quantitative research data which used to get depth information on characteristics of travel and number of trips originating or destined for particular traffic analysis zone.

4.6.1. Individual and household characteristics

The household travel surveys involve contacting respondents at their home and collecting information regarding their household characteristics, their personal characteristics and the travel decisions made in the recent past (e.g., number of trips, mode of travel for each trip, etc.). The survey data collected consisted of the socioeconomic variables used to reduce the possibility of a subjective classification. The traditional approaches were extensively adopted for travel surveys, i.e. face-to-face interviews.

Socioeconomic details contain age, gender, monthly income, car ownership, number of family members traveling, job and educational background. The total trips produced by the household and the number of trips for each type of trip purposes are divided into five groups (work trips, education trips, shopping trips, social trips, and recreational trips). Usually, most of the shopping, social, and recreational trips are made irregular trips. The obligatory trips which are made for the work and educational purpose trips are the two that are going to be developed for the mode choice model. The explanatory variables that have been considered for possible use in the modeling process are summarized in table 4.4.

4.6.2. Trip data

The movement outside a building or premise with a given purpose is called a trip. The information sought considers trips by the change of mode (including walking). change of mode is characterized on the basis of variables such as origin and destination (normally expressed by their nearest road junction), purpose, start and end times, the mode of transport used, amount of money paid for the trip, and so on (Ortuzar and Willumsen,2011). Data's regarding trips purpose are an important step in mode choice analysis because different trip maker behaviors are expected in selecting a mode for different trip purposes.

According to the points mentioned in the questionnaire design and the required information, the questionnaire form used in this thesis is shown in the appendix. The data collected by the questionnaire are considered as explanatory variables and are used to build the models, which will be examined statistically and logically to adopt the most appropriate model that can predict the trips produced and the mode choice.

Table4. 4. Explanatory variables used in trip production and mode choice model

For trip production	
Variable	Description
FS	Number of persons in the household
Gender	Gender of respondents
Age	Age of respondents
Job	Job of respondents
CPH	Car ownership by a household
AMI	Monthly household income
For the mode choice model	
Variable	Description
Constant	Mode specific constant
FS	Number of persons in the household
Gender	Gender of respondents(1 if female and 0 otherwise)
Age	Age of respondents in years
Job	Job of respondents
CPH	Car ownership (1 for yes and 0 otherwise)
AMI	Monthly household income in birr
PH	Peak hour (1 for the individual time of leaving home between 7:30 am to 8:30 am and 0 otherwise)
TT	Related total travel time in minutes (generic variable)
TC	Related total travel cost in birr (generic variable)
OVTT	Related out of vehicle travel time in minutes (generic variable)
IVTT	Related in-vehicle travel time in minutes (generic variable)

Table4. 4. Dependent variables used in trip production and mode choice model

For trip production	
TPR	Number of daily trips made by household
TPW	Number of daily work trips made by household
TPED	Number of daily educational trips made by household
TPSHO	Number of daily shopping trips made by household

TPSO	Number of daily social trips made by household
TPRE	Number of daily recreational trips made household
For the mode choice model	
$U_{walking}$	The utility function of walking
U_{bajaj}	The utility function of Bajaj
$U_{private\ car}$	The utility function of private car
U_{taxi}	The utility function of taxi
$U_{service}$	The utility function of service

4.7. Conducting Field Survey

After determining the sample size, designing the questionnaire, and choosing the methods of survey, the field survey was conducted according to the following steps;

- Enumerators were trained on how they will perform the interview and they were provided with support letters from the University to facilitate the response rate and to make the respondents understand the purpose of the research and encourage them to give accurate information.
- The enumerators will be given random household addresses. The random sample of households in different traffic zones will be numbered on a printed map and the same numbers are put on the questionnaire layout, where each number on the map and the questionnaire represents one household.
- The enumerators have to first get permission to be surveyed from the household. The heads of households and each member of the households were visited and interviewed at their homes by the enumerators using the structured questionnaire to record socioeconomic and travel details for seven consecutive days. Below five years aged children are not included in the trip details.
- Since the actual survey may take place at any time during the day, the respondents are required to answer the questions about the travel details according to the previous day.
- Finally, the necessary data was collected under the close supervision of the researcher.

CHAPTER FIVE
RESULTS AND DISCUSSION

5.1. Data analysis

This section is also intended to give an overview of the distribution of various trips according to their purpose and transport mode used, this shows which trips and mode of alternatives are the most frequent. Initially, the frequencies were obtained on a whole dataset to determine the distribution of travelers for various travel and socioeconomic characteristics. The descriptive statistical analysis is conducted by categorizing the collected data by users, presenting the findings in the form of tables and figures, using SPSS, in order to infer the basic characteristics related to the number of daily trips and traffic behavior of the mode users with each of the explanatory variables and transportation mode used from home to the work and education trip purposes of the Hawelti sub-city residents.

Moreover, this section includes figures that relate the number of daily trips with each of the explanatory variables so as to show the relationship between the variables graphically. This also shows as the relationships are linear or not.

Next, SPSS (statistical package for the social science) software has been used to analyze the data. The multiple linear regression model and multinomial logistic regression model have been used for the trip production model and modal split model developments respectively. In this research which assumes that the trip production and all of the alternatives are independent and each should be analyzed separately. Relevant statistical tests' results will be presented and discussed in detail in the below sections.

5.1.1. General analysis of socio-demographic data

The data used in this analysis were obtained from the questionnaires collected from a sample of Hawelti sub-city residents. In addition to that independent variables were grouped initially to be analyzed using the SPSS program and there were some ungrouped variables that have been analyzed and remain in its original categories. The results of this analysis are summarized in the form of frequency, tables, and figures. Generally, descriptive analysis was used to get an overall picture of the survey results. The distribution of travelers for various travel and socio-economic characteristics are presented in the following subsections.

Table5. 1. Demography and socio-economic profile of the trip makers in the study area

Parameters	Category	Frequency	Percentage
Gender	Number of males in the household	595	44.5
	Number of females in the household	742	55.5
Age	Below18 in the household	276	20.7
	18-45 in the household	539	40.3
	Above 45 in the household	522	39.0
Education level	Illiterate	57	4.3
	Primary school	262	19.6
	Secondary school	397	29.7
	Certificate and Diploma	228	17.0
	Degree and Above	393	29.4
Occupation type	Number of Students in the household	501	37.5
	Number of employees in the household	574	42.9
	Others in the household	262	19.6
Household income	5000 and Below	58	14.7
	5001-10000	229	58.0
	10001-15000	98	24.8
	Above 15000	10	2.5
Household Car ownership	No	337	85.3
	Yes	58	14.7
Household family size	Below Two	98	24.8
	Three–Four	234	59.2
	Above Four	63	16.0

The gender distributions of trip maker survey respondents shown above in the table about 44.5% of the respondents are male and 55.5% are female. Females are the dominant respondents in the study area. The age of respondents was categorized into three. The first age group of ages below18 years includes persons who are assumed to be in school have 20.7 % and 40.3% of the sample obtained from the second age groups of age 18 to 45 years includes a person who is either continuing education or part of the labor force. Age above 45 years grouped under third includes a person who is head of families and part of the labor force which represents about 39.0 % of the whole sample. The distributions of respondents based on the educational background were 4.3% Illiterate and 29.7% were secondary school these are the lowest and highest percentage of respondents respectively. The secondary school, and degree and above have the same percentage of respondents. The respondents' job category results reported in the table above show that the majority of respondents in the household were employees about 42.9% and the least one are other categories.

The household average family monthly income results reported in the table show that the majority and least of respondents have a monthly income between birr 5001-10,000 and above birr 15,000 which represents about 58.0% and 2.5% of the whole sample respectively. The car ownership in the household indicates that only 14.7% of the respondents have a private car, while 85.3% of the respondents don't own private cars. Based on the analysis of the collected data, the average family size is 3.4 persons. The family size was categorized into three categories. The result shows that the dominant percent of the household size has a family size between 3-4 persons which represents about 59.2% of the whole sample. While the minimum family size represents 16% of the sample.

5.1.2. The relationship between trip production and socioeconomic characteristics

Independent variables were grouped initially to be analyzed using the SPSS program based on the results obtained from the questionnaire and also based on previous studies (Etsub, 2015 and Ali, 2018). However, there were some ungrouped variables that have been analyzed and remain in their original categories. In the categorical analysis, trip observations are aggregated in terms of the common socioeconomic characteristics. Comparison graphs were established between the grouped explanatory variables to the average daily trips per household and that is used in the process of prediction trip production was presented as below.Figure 5.1 shows that the relationship between the average daily trip production per household made by the gender difference in the household and it is remarkable that trip making increases with an increase in both the number of females and males in the household. The average daily trip per household made by the number of males higher than these trips made by the number of females in the household until it reaches three males per household and then slightly less than at three and four or more females per household.

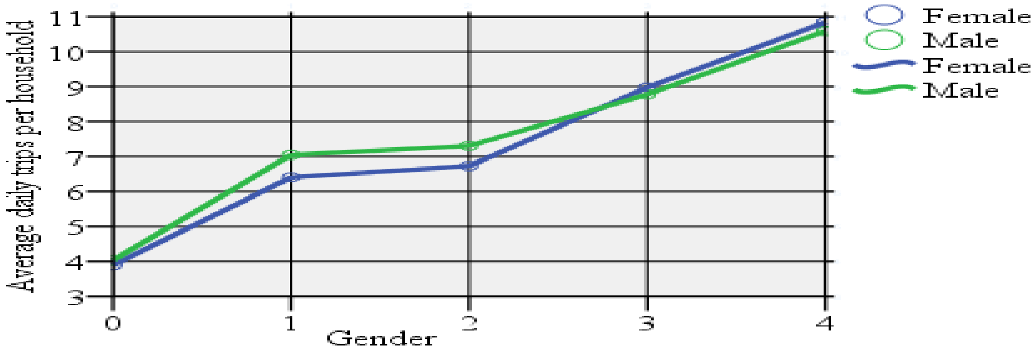

Figure5. 1.Average daily trip production and gender

The average trips made by the occupation groups in figure 5. 2 show that the trip making increases with increasing the number of employed persons and the number of students. The average trips from zero to two made by students were greater than the average trips made by employed persons in the household. At three or greater employed persons in the household average trips per household were exceed those made by students.

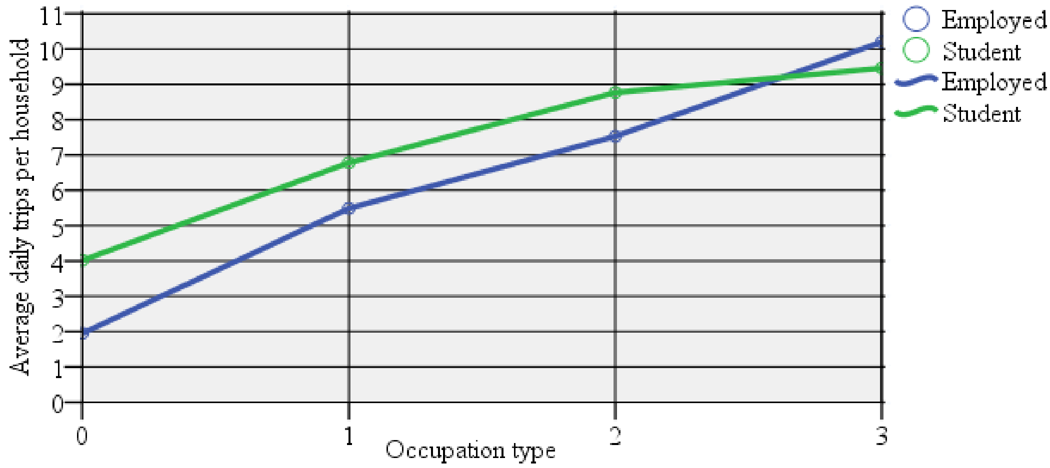

Figure5. 2.Average Daily Household Trips and Occupation type

Figure 5.3 Results show that the average daily trip made by family size by vehicle ownership and it is remarkable that the average daily trip for families with the vehicle is higher for all households' sizes except for households with one and two persons. It could be explained by that the ratio between the average daily trips to the number of households is higher for families with the vehicle than those with no vehicle in the household.

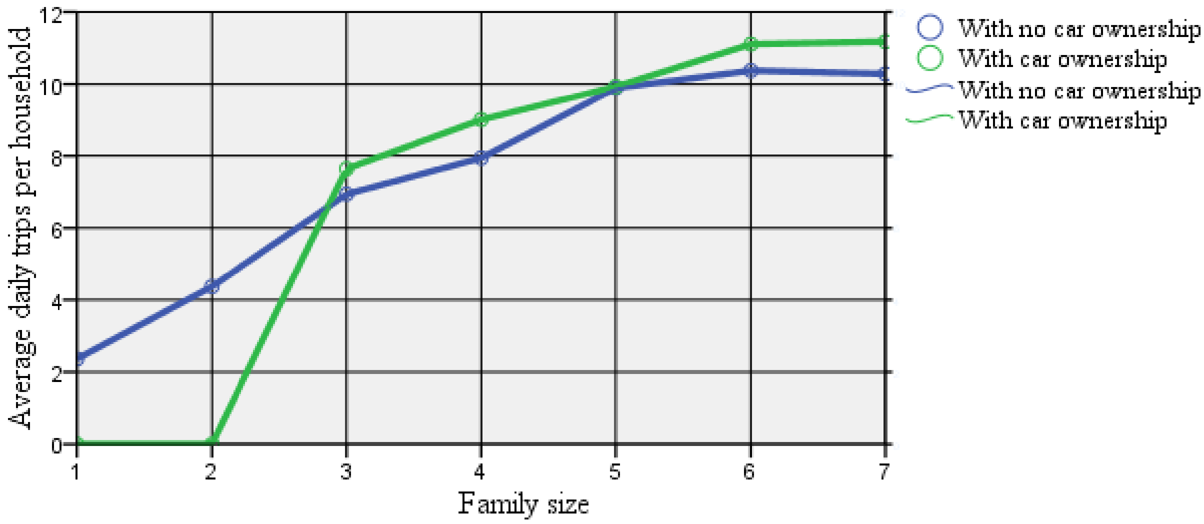

Figure5. 3.The average daily trip made by family size by vehicle ownership

The average trips by age group shown in figure 5.4 indicate that age less than 18 and age between 18 and 45 persons made almost similar average trips whereas age above 45 shows more different from the other age groups. However, the trip making increases with increasing the members in the household for all age groups.

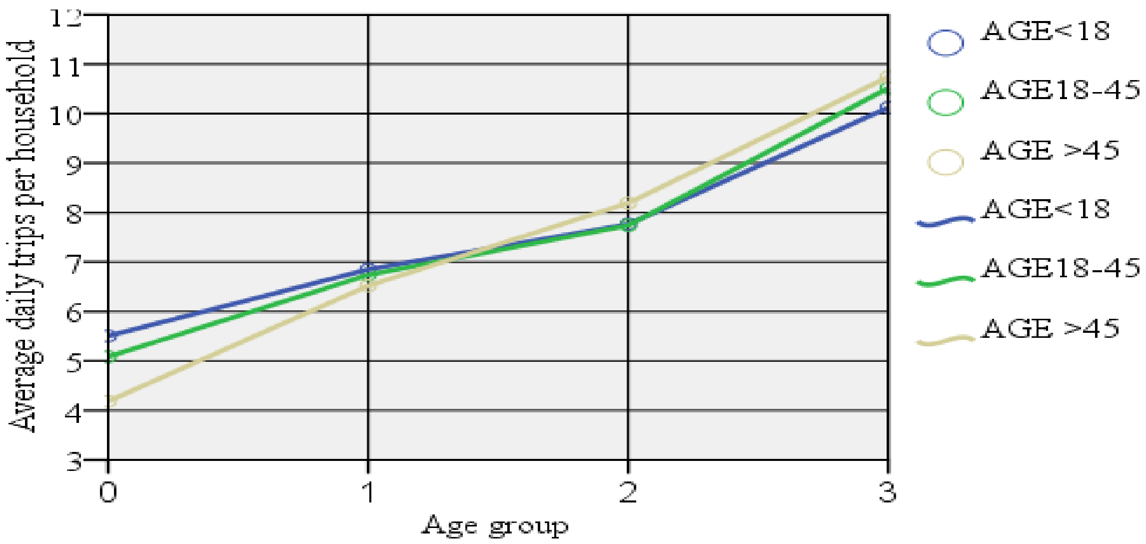

Figure5. 4.Average Daily Household Trips and Age group

The relationship between the average trips per day per household and the explanatory variables used in the model are depicted as follows.

From figure 5.5. Below it is shown that there is a direct relationship between the total average daily trips made by a household and the number of males in the household which means as the number of males increases in a household, the average daily trips made by that household also increases.

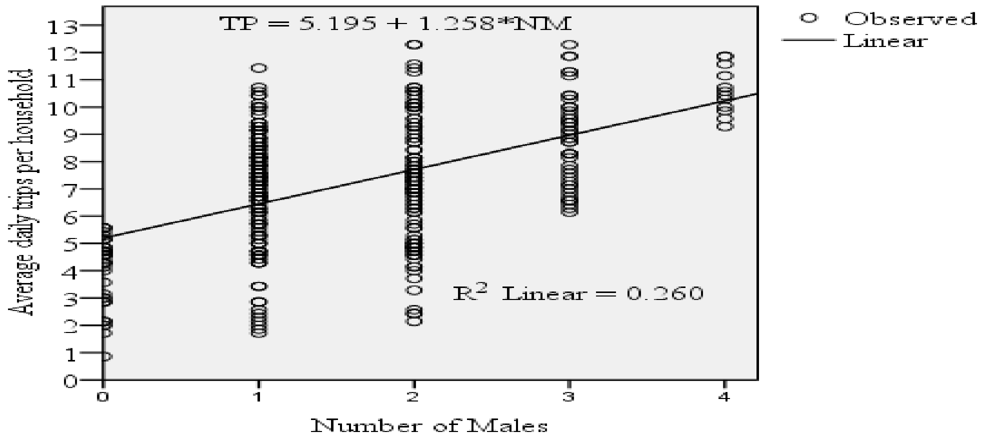

Figure5. 5.Average Daily Household Trips and Number of Males

The number of females in households and the total average daily trip production given in Figure 5.6 shows a positive relationship between these two variables. The result shows that, as the number of female's increases in the household, the total average daily trips made by the household also increases.

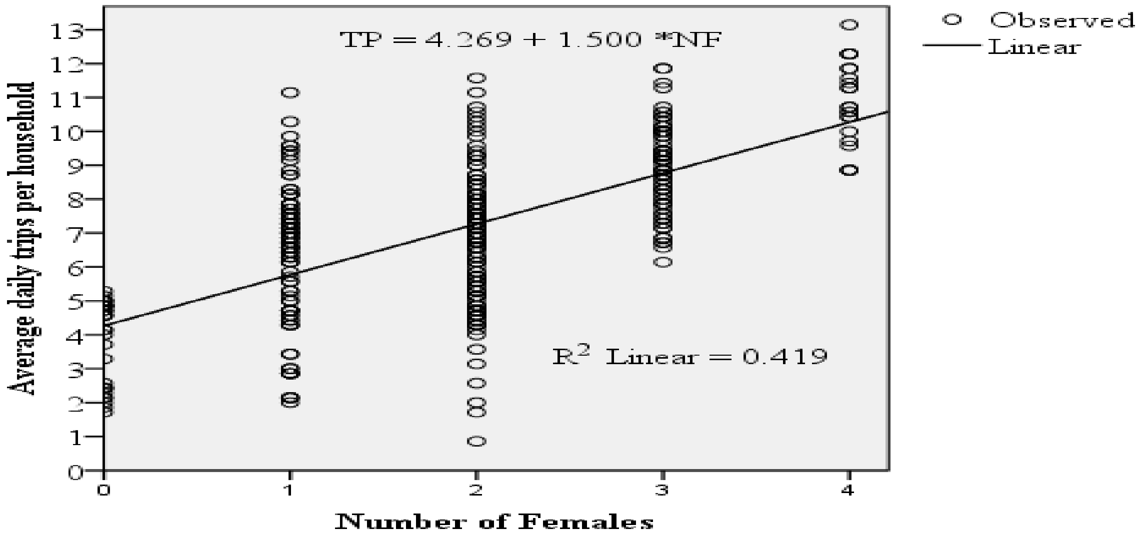

Figure5. 6.Average Daily Household Trips and Number of Females

The relationship between the total number of average daily trips made by a household and the number of employed persons in the household is shown in Figure 5.7. The figure shows a relatively positive relationship between these two variables.

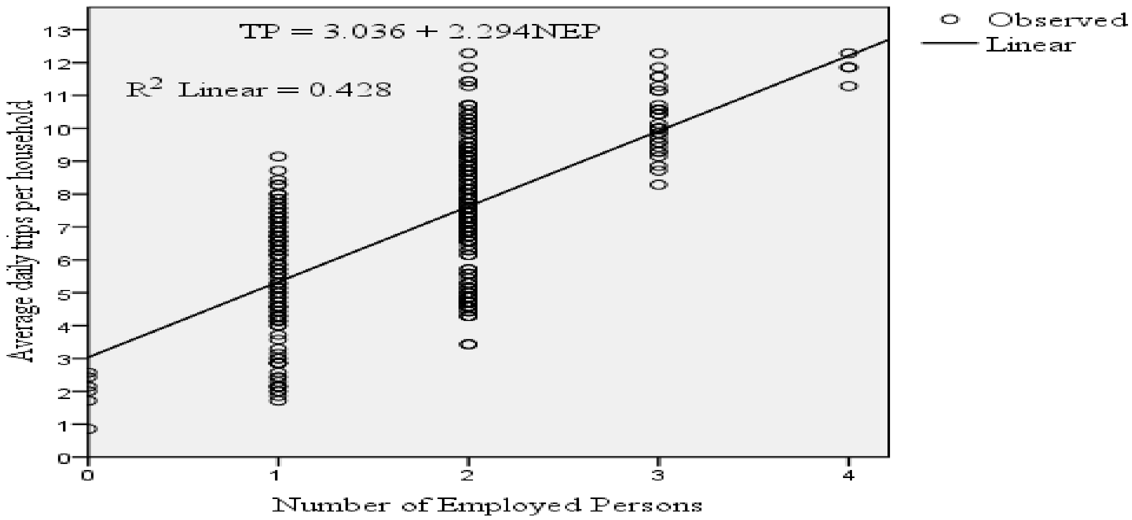

Figure5. 7.Average Daily Household Trips and Number of Employed Persons

The average daily trips that are produced from a household and the number of students in the household are presented in figure 5.8 which shows that the relationship between the average daily trips per household and the number of persons receiving an education.

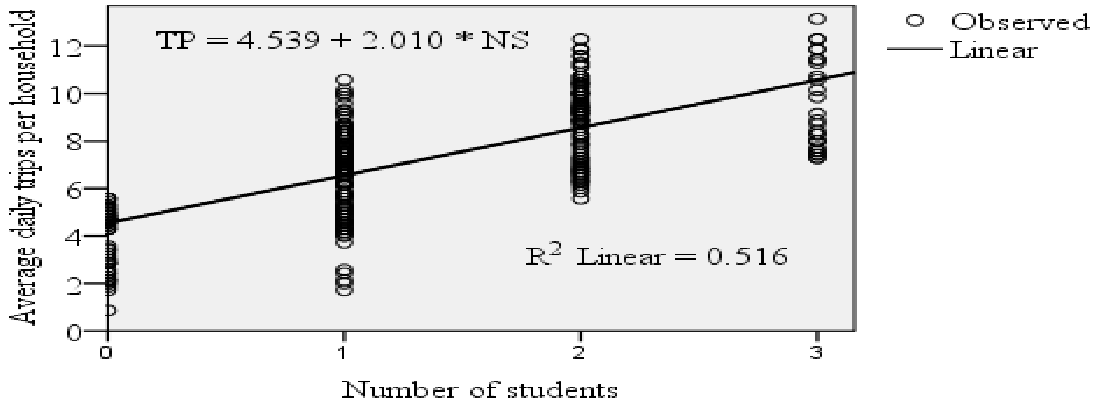

Figure5. 8.Average Daily Household Trips and number of Students

Fig. 5.9. Demonstrates the number of cars owned by the household versus average daily trips per household. It is apparent from the figure that there is no apparent relationship between these two variables.

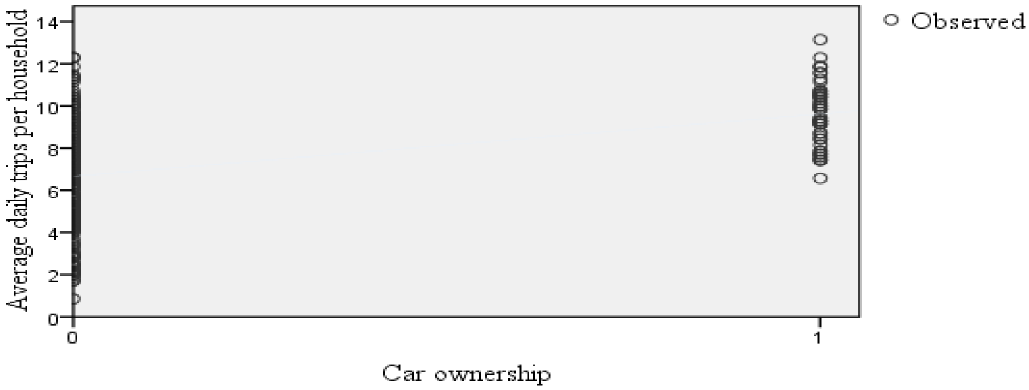

Figure5. 9.Average Daily Household Trips and Cars ownership

Figure 5.10. illustrates the relationship between the average daily trips per household and the house household size as measured by the number of persons in the household is represented. It is remarkable that the trip making increase with increasing the number of family size.

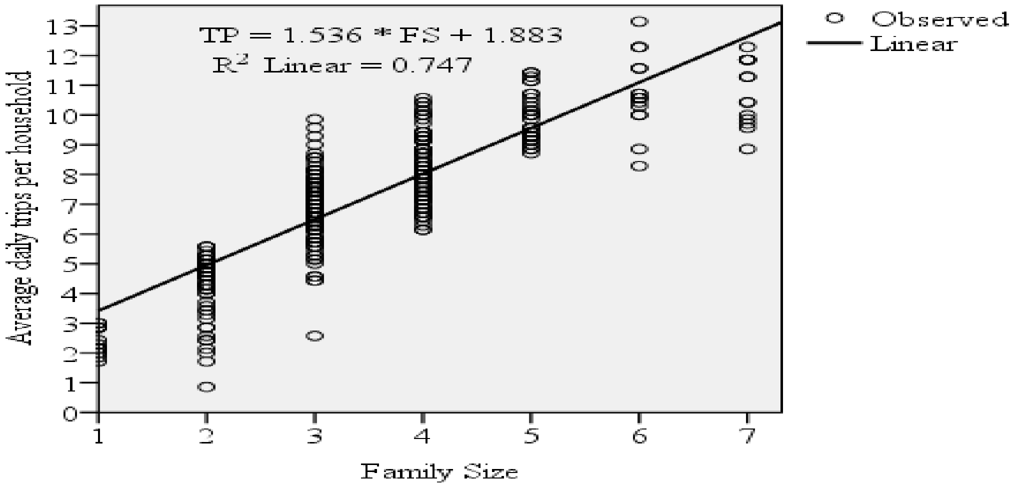

Figure5. 10.Average Daily Household Trips and Household Size

The relationship between household monthly income and the average daily trips made by a household is shown in figure 5.11. The figure shows a relatively positive relationship between these two variables. This positive relation explains that, as the household income increases, the average daily trips also increases.

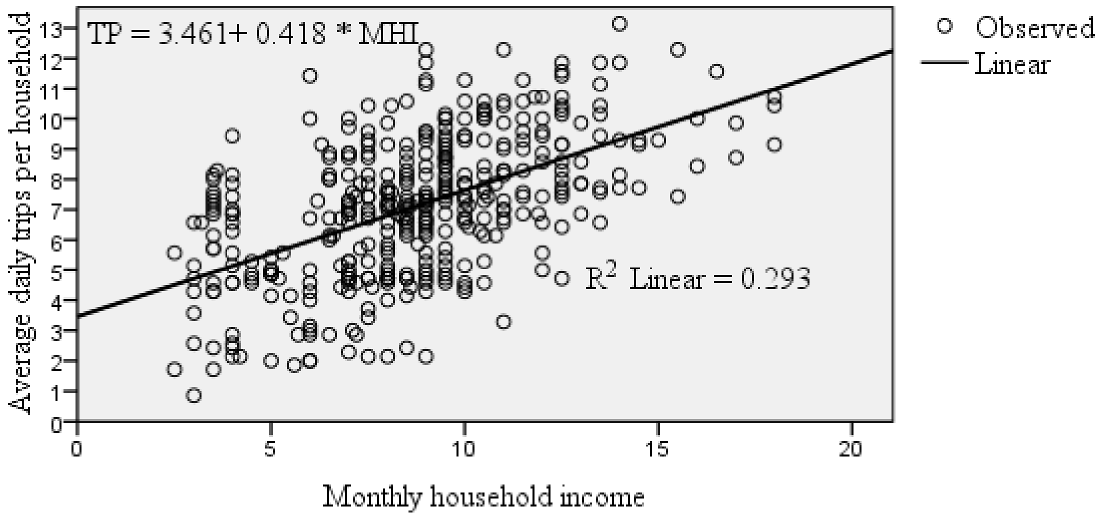

Figure5. 11.Average Daily Household Trips and Monthly Household Income

The modes of transport used by the respondents were Service, Minibus taxi, Private cars, Bajaj and Walk is shown in table5.2. The Minibus taxi average trips represent 44.05% of all trips. The company provided a mode of transport (service), Private cars and Bajaj constitute 22.81%, 7.60%, and 5.12% respectively. The fraction of trips using non-motorized modes (walk) constitutes 20.42 %.

Table5. 2.The average trips by transportation mode

Mode of transport	Average trip per Household	Percent (%)
Service	637.650	22.81
Minibus taxi	1231.416	44.05
Private cars	212.438	7.60
Bajaj	142.977	5.12
Walk	570.760	20.42
Total	2795.241	100

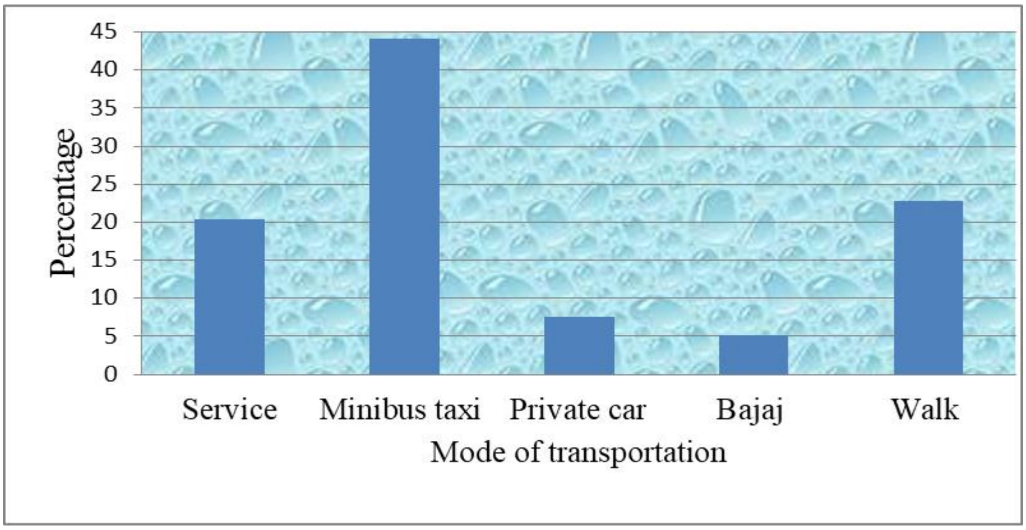

Figure5. 12.Percentage of Trips by Mode of Transportation

Graphically in figure 5.12. Shows that the majority of households mode of transport used for trip purpose were 44% and that indicates they use Minibus taxi, whereas the only 5.12% they use bajaj which is the list percent.

5.1.3. Trip purpose

Classification of trips according to purpose is important to be carried out since people make their trips for various reasons. The Hawelti sub-city survey data for this research is a household travel survey carried out in 2018/19. The survey captured trip data that represents 2,795.241 trips made by 1,337persons in 395 households within Hawelti sub-city residents.

The trips can be classified as trips for work, trips for education, trips for shopping, trips for recreation and social trips. Among these, the work and education trips are often referred to as mandatory trips and the rest as discretionary trips. In this research, among 2,795.241 trips produced from 395 surveyed households, most of the trips (48.85%) are work trips. Education trips account for 27.42% of total average daily trips produced. The distribution of household trips according to purpose is shown in table 5.2 and graphically in figure 5.12.

Table5. 3.Distribution of household trips according to the purpose of the trip

Trip production based on the purpose	Average trip per Household	Percent (%)
Home-based work trips	1365.532	48.852
Home-based education trips	766.546	27.423
Home-based shopping trips	228.000	8.157
Home-based recreational trips	227.569	8.141
Home-based social trips	207.594	7.427
Total trips	2795.241	100.000

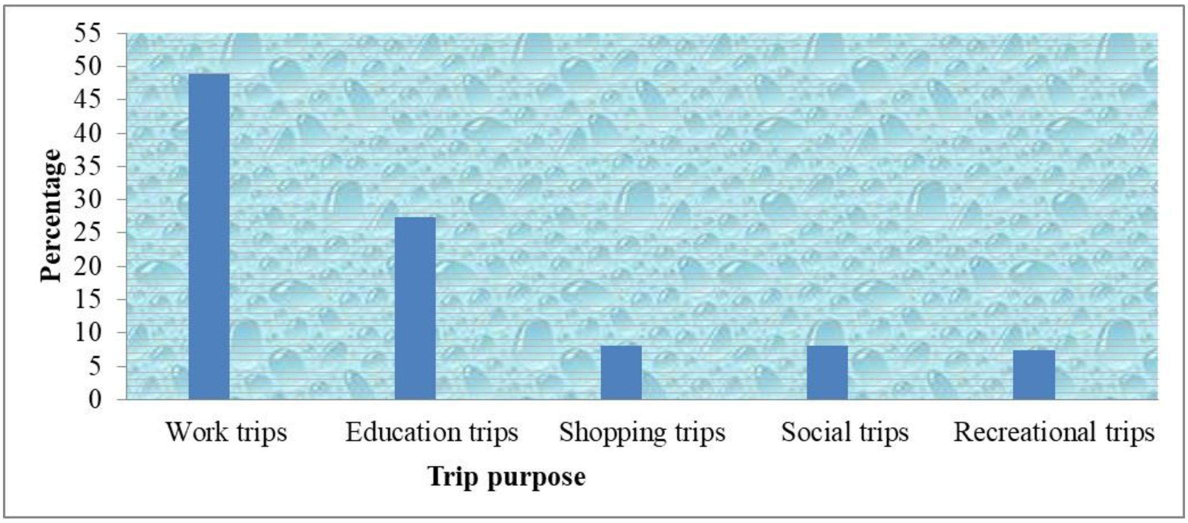

Figure5. 13.Distribution of Daily Household Trips by Purpose

To evaluate the assumption of normality for the dependent variable, the SPSS package help in this way by several methods, statistics tables (as Kolmogorov- Smirnov test or Shapiro-Wilk test and skewness or kurtosis values) or by visual figures. Normality is an important assumption in multiple linear regression and states that all residuals are normally distributed. In this study a visual check is adopted using a normal probability plot of residuals figures See Figure 5.14 The figures give an indication of the smooth normality of the regression standardized residuals for all trip purposes; therefore this assumption appears to be satisfied. In addition to that results of the statistics tables for Kolmogorov- Smirnov test or Shapiro-Wilk test and skewness or kurtosis values also show that all dependent variables satisfied the criteria for a normal distribution.

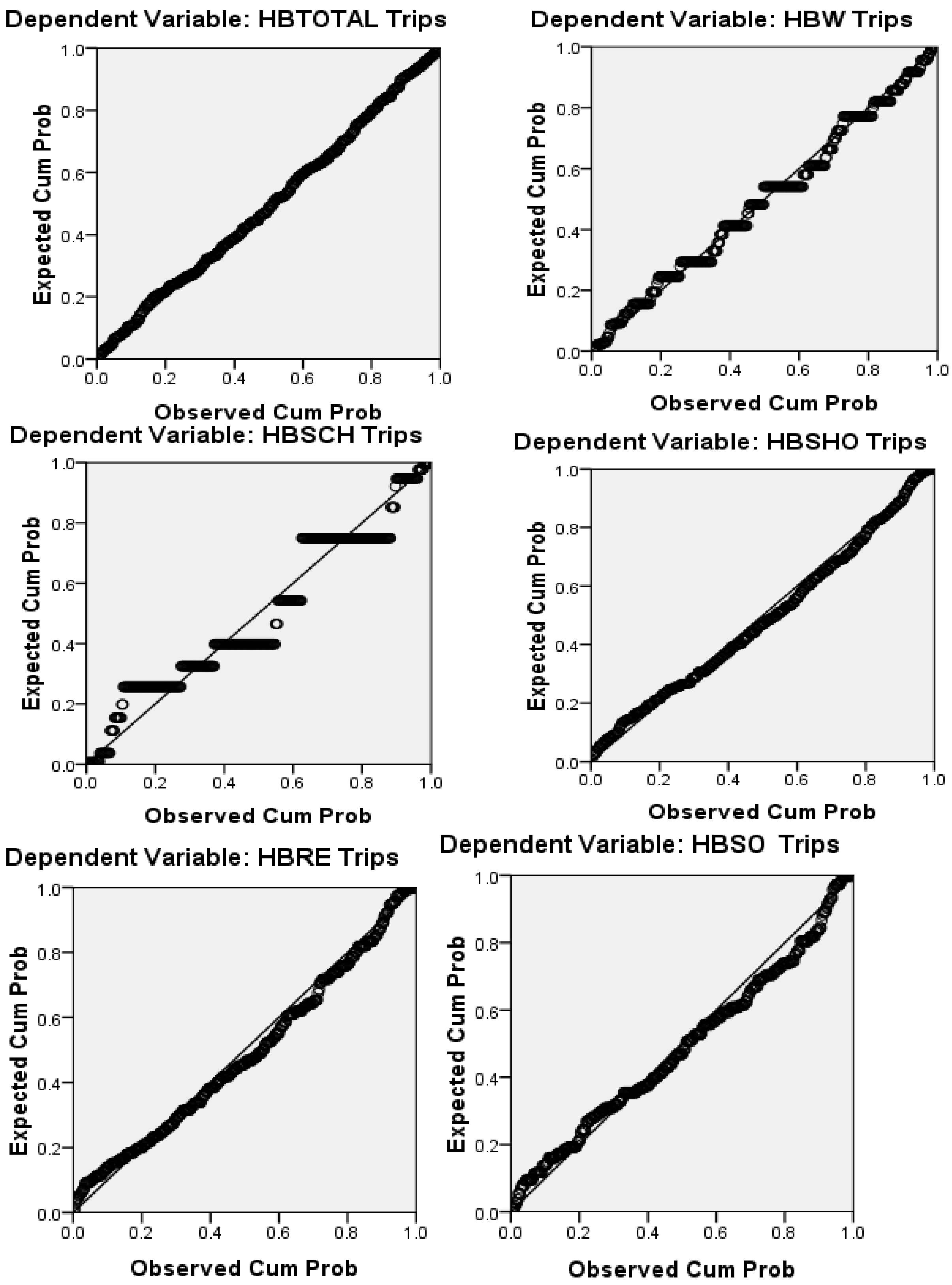

Figure5. 14.Normal P-P Plots of Regression Standardized Residual by Trip Purpose

5.2. Development of Trip Production Model

The trip production in this thesis is used to estimate the average daily trips per household, based on purpose in the study area. Trip production is performed by relating the average daily trips to the characteristics of the individuals and households. The variables used in the development of the model were the first analyzed by using Pearson correlation among the independent variables to determine how to relate the strength of association between the variables since a high correlation would indicate that one variable could explain the other and any one of them could represent others in the model. From the result of the correlation matrix gives the family size (FS) is considered as strongly correlated with the NF, NM, NS, $AGE_{under17}$, and $AGE_{above45}$, while most of the other variables result in the pairwise correlations matrix are not high, suggesting that there is no severing multicollinearity problem**(Appendix B)**.

5.2.1. General Trip Production Model

This section estimates the general trip production model using the collected data from the household respondents and analyzed using multiple linear regression with the stepwise method. During regression, all variables were entered in the regression equation in a stepwise procedure. In each step, the regression model is evaluated according to the statistical tests. The variables which are statistically significant at the 95% level were included in the development of the model. The final developed model to estimate general trip production is shown in equation 5.1.

$$TP=0.572+ 1.629NS + 1.855NEP + 0.079AFMI+0.150\,NF+0.143AGE_{above45} \quad (5.1)$$

Table5. 4. Parameter estimation for general trip production model

Constant and Variables	Coefficients	Std. Error	t- Value	Significance	VIF
Constant	0.572	0.142	4.041	0.000	
NS	1.629	0.066	24.848	0.000	1.919
NEP	1.855	0.075	24.693	0.000	1.600
AFMI	0.079	0.015	5.276	0.000	1.328
NF	0.150	0.052	2.883	0.004	1.769
$AGE_{above45}$	0.143	0.070	2.056	0.040	1.619
R^2			0.889		
F-value			620.662		
Sample			395		

The developed regression equation for trip production in the above and all the parameters was significant at the 95% probability level, with a coefficient of determination of 0.889. The R squared value 0.889 yielded by the model denotes 88.9% of the trip produced is influenced by the variation in the variables included in the model. Such a value of R squared indicates that the model is considered as a good fit.

All the coefficient of NS, NEP, AFMI, NF, and $AGE_{above45}$ (number of students, number of employed persons, the monthly household income, number of females and age above 45 in the household) in the model has a positive sign, which means that the variables are positively related to daily trips made by the household. The t-value for the coefficient of all of the regression coefficients have above two. Thus, the hypothesis that the number of students, number of employed persons, the monthly household income, number of females and age above 45 in the household has no effect on daily household trips (TP) is rejected and the alternative hypothesis that the number of students, number of employed persons, the monthly household income, number of females and age above 45 in the household and the number of daily household trips are positively correlated is accepted. The correlation matrix was usually obtained to test for the multicollinearity of variables used in the model development. In the appendix at the end of this paper gives the correlation matrix, which has been used in the model building process. In the correlation matrix, the entries on the man diagonal (those running from the upper left- hand corner to the lower right- hand corner) give the correlation of one variable with itself. The entries off the main diagonal are the pair-wise correlation among the explanatory variables. From the correlation matrix table, most of these pairwise correlations are not high; suggesting that there is no severing multicollinearity problem.

The variance inflation factors (VIF) for each of the explanatory variables that are included in the general trip production model have a value of less than 10. Therefore, the problem of multicollinearity does not exist in the developed general trip production model. Table 5.5 an ANOVA result that presents the analysis of the total variance in the dependent variable. This variance is used to test the regression model at a 95% level of significance.

Table5. 5.ANOVA table for the general trip production model

Model	Sum of Squares	Degree of freedom	Mean Square	F-value	Significance
Regression	1903.953	5	380.791	620.662	0.000
Residual	238.047	388	0.614		
Total	2142.000	393			

The F-statistics tests the hypothesis that all the slope confidence of the multiple regression model is simultaneously zero; that is all the independent variables jointly have no impact on the dependent variable. From the ANOVA table shown in the above, the F- statistics for the general trip production model is 620.662. Since this is a highly significant value, the hypothesis that all the independent variables entered into the general trip production model (number of daily trips per household) is rejected at the 95% level of significance and the alternative hypothesis that all these variables jointly affect the dependent variable is thus accepted. In summary, it is concluded that NS, NEP, AFMI, NF, and $AGE_{above45}$ (number of students, number of employed persons, the monthly household income, number of females and age above 45 in the household) do explain the variation in the number of daily trips per household (TP).

5.2.2. Trip Production Models Based on Purpose

In this section, trip production based on trip purpose will be developed. Similar to the general trip production model, modeling of all the trip categories based on purpose also analyzed using multiple linear regression techniques. In this research, the trip production model categorized as work, education, and shopping, social, and recreational purpose trips.

5.2.3. Work trip production model

The work trips in this thesis refer to all trips starting at the home and ending at the workplace of the trip makers or vice versa. Based on multiple regression analysis, the following regression models were developed to estimate trips for work purpose;

$$TP_W = 0.513 + 1.649NEP + 0.241CPH \quad (5.2)$$

The model shows that work trip production is produced by the NEP and CPH (the number of employed persons in the household and car per household, respectively). The coefficients are significant at a 95% level of significance. These coefficients have positive signs as expected. This means that the positive signs of the variables positively affect the number of daily work trips. The explanatory variables included in the model have a variance inflation factor less than ten (10) indicates that there is no multicollinearity problem in the developed model. The R^2 value was 0.867 indicates that the independent variables in the model explain about 86.7% of the variation in the dependent variable (number of daily work trips per household). The model developed by NEP and CPH explains the variation in the number of daily work trips per household (TP_W). Table 5.5 summarizes the estimated work trip, production model. The full results of the work trip production model using the SPSS package are included in the appendix.

Table5. 6. Parameter estimation for the work trip production model

Constant and Variables	Coefficients	Std. Error	T- Value	Significance	VIF
Constant	0.513	0.065	7.905	0.000	
NEP	1.649	0.036	45.752	0.000	1.149
CPH	0.241	0.068	3.535	0.000	1.149
R^2			0.867		
F-value			1276.636		
Sample			395		

5.2.4. Education trip production model

The trip production model for education is estimated using multiple linear regression analysis. The estimated education trip production model is

$$TP_{SCH} = 0.066 + 1.478NS \quad (5.3)$$

The coefficient of NS (number of persons receiving education in the household) is significant at the 95% level of significance. A positive sign in the coefficient shows that as the number of persons receiving education in the household increases, the average number of daily education trips per household will also increase. The sign is also as expected. The R squared value 0.960 which indicates that about 96% of the trip produced for education is influenced by the variety of variables in the model. Based on this excellent R squared value, the education trip production model has strong explanatory power.

The F-statistic for the education trip production model is 9369. since this is a significant value, the hypothesis that the independent variable entered into the education trip production model has an impact on the dependent variable is accepted at a 95% level of confidence (number of daily education trips per household). Table 5.7 shows the summarize regression results for the estimated educational trip production model. The full results of the educational trip production model using the SPSS package are included in the appendix.

Table5. 7.Parameter estimation for education trip production model

Intercept and Variables	Coefficients	Std. Error	t- Value	Significance	VIF
Constant	.066	.023	2.833	0.005	
NS	1.478	.015	96.797	0.000	1.000
R^2			0.960		
F-value			9369.699		
Sample			395		

5.2.5. Shopping trip production model

The estimated shopping trip production model using the multiple linear regression analysis is

TP_{shoP}= 0.035AFMI + 0.149NEP (5.4)

The coefficients of AFMI and NEP (the monthly household income and the number of employed persons in the household) are all significant at the 95% levels of significance. These coefficients have positive signs. The model indicates that shopping trips increase with the increase in household income and the number of employed persons in the household. Each of the explanatory variables in the model has a VIF of about 6.839. This value indicates that there is no multicollinearity problem in the developed model since the value is less than 10. The R squared value of the shopping model was 0.798 which shows that the independent variables in the model are able to describe 79.8% of the total variation in the dependent variable (number of daily shopping trips per household).

The F-statistic for the shopping trip production model is 768.911. Since this is a significant value, the hypothesis that all the independent variables jointly affect the dependent variable is accepted. Table 5.8 presents the summarized results for the estimated shopping trip production model. The full results of a shopping trip production model using the SPSS package are included in the appendix.

Table5. 8. Parameter estimation for a shopping trip production model

Intercept and Variables	Coefficients	Std. Error	t- Value	Significance	VIF
NEP	0.149	0.021	7.095	0.000	6.839
AFMI	0.035	0.004	8.192	0.000	6.839
R^2			0.798		
F-value			768.911		
Sample			395		

5.2.6. Social trip production model

The social trip production model is estimated using the multiple linear regression analysis as

TP_{SO} = 0.110NEP + 0.104NF + 0.096$AGE_{above45}$ (5.5)

The coefficients of NEP, $AGE_{above45,}$ and NF (the number of employed persons in the household, age above 45 persons in the household and number of females in the household, respectively) have t-values above two in absolute value. Thus, these coefficients are significant at the 95% level of significance. Table 5.9 summarizes the most important regression results for the estimated social trip production model.

Table5. 9.Parameter estimation for the social trip production model

Intercept and Variables	Coefficients	Std. Error	t- Value	Significance	VIF
NEP	0.110	0.016	6.768	0.000	6.874
NF	0.104	0.013	7.911	0.000	5.745
$AGE_{above45}$	0.096	0.019	5.122	0.000	6.143
R^2			0.853		
F-value			754.616		
Sample			395		

As the number of employed persons and the age above 45 persons in the household increase, the average number of daily social trips per household will also increase. This is indicated by the positive coefficients of the number of employed persons in the household.

The coefficient of NF indicates that as the number of females increases in the household, the average number of daily social trips per household will increase. This result shows that females have more social activities than males.

None of the explanatory variables included in the developed model has a VIF that exceeds 10. Therefore, the developed model does not have a multicollinearity problem.

The social model has an R-squared value of about 0.853 indicating that the independent variables entered into the model explained 85.3% of the variation in the number of daily social trips per household. This value is considered a good model fit.

The F-statistic for the social trip production model is 754.616. Since this is a significant value, the hypothesis that all the independent variables jointly affect the dependent variable is accepted at the 95% level of significance.

5.2.7. Recreational trip production model

The statistically significant model for recreational trip production is estimated using the multiple linear regression analysis as

$$TP_{RE} = 0.189NS + 0.038AFMI \quad (5.6)$$

The coefficients of independent variables at the 95% level of significance in the recreational trip production model are NS and AFMI (number of persons receiving education in the household and the average family monthly income in the household). The model coefficient shows that the household with a higher number of receiving education and monthly income in the household makes more recreational trips. Both the coefficients in the number of persons receiving education in the household and the average family monthly income in the household are positively related to the number of daily recreational trips per household, which indicates that as the number of persons receiving education in the household and the average family monthly income in the household increases, the average number of daily recreational trips per household will also increase.

The variance inflation rate factor for each of the explanatory variables included in the model is less than 10. Hence, there is no multicollinearity problem in the developed model.

The R-squared value of the final model was 0.794 which means the independent variables in the model describes 79.4% of the total variation in the dependent variable (number of daily recreational trips per household).

The F-statistic for the recreational model is 756.503. Thus, the model as a whole is significant at the 95% level of significance. Therefore, the hypothesis that all the independent variables that are included in the model mutually affect the dependent variables is accepted.

The model result is presented in Table 5.10. The full results of the recreational trip production model using the SPSS package are included in the appendix.

Table5. 10. Parameter estimation for a recreational trip production model

Variables	Coefficients	Std. Error	t- Value	Significance	VIF
NS	0.189	0.020	11.364	0.000	3.811
AFMI	0.038	0.003	9.297	0.000	3.811
R^2			0.794		
F-value			756.503		
Sample			395		

5.3. Estimation of Mode Choice Model

As discussed earlier, the trips made by residents of the study area were modeled on the basis of trip purposes. In this case presents the specifications of the mode choice models, developed for trips taken within residents of the study area. Based on the collected data, service (company provided and school bus), private car, minibus taxi, bajaj, and walking are the modes that people of the study area used to travel from home to destination of trips in Hawelti sub-city residents. The travel behavior of the residents of the study area, for this study, was modeled on the basis of trip purposes. The trips based on purpose were categorized into two namely work and school trips and which are on a number of mode choice response were attained for each trip purpose. Therefore, two unique sets of mode choice models were developed for the study area. To accomplish this model development multinomial logistic regression model has been developed. Multinomial logistic regression models are used to predict the probabilities of the different possible outcomes of the dependent variable (mode choices), given a set of independent variables. Different specifications for the models have been evaluated to determine which specifications best build the model for the trip purposes. The variables that have been used in this research include the variables that have been found in the literature review to influence the trip maker choice such as in-vehicle travel time (IVTT), out of vehicle travel time (OVTT), total travel time (TT), total travel cost (TC), age of respondent (AGE), Gender of the respondent (GENDER), household car ownership (CPH), family size(FS), family monthly income (AFMI), Job type (JOB), peak hour (PH), and educational level (EDL). The travel time and travel cost variables are included as a generic variable meaning that an increase of one unit of this variable has the same impact on all the transport modes. The socio-economic independent variables are considered alternative specific; with different parameters for all the modes or an increase of one unit of these variables has different impacts for all the modes.

In this research, minibus taxi was the base or reference alternative in SPSS for all the parameter estimation of the independent variables except for the generic variables (like TT and TC) and for variables with categories more than one, always the last category is the reference category by default in the SPSS software. Different models were developed and analyzed step by step to determine which one of the developed models is the fit or true model that shows the data better. Estimation and calibration of the multinomial model is the process by which the multinomial model is filtered to the data through the use of the maximum likelihood principle and methods. The inclusion potential of variables is tested and the parameters of the utility functions are computed using SPSS software. In this research, two models were estimated for the different purposes of trips for Hawelti sub-city residents as shown in the table below. These Models are: By comparing and evaluating the models using informal judgment based on tests, the goodness of fit measures and statistical tests the optimal model will choose.

5.3.1. Mode choice model for work trips

I. Socio-demographic characteristics

The work trips in this research refer to all trips starting at the home and ending at the workplace of the trip makers. In the development of the model, there were 574 work trip makers data collected from the study area with interviews at their households. The socio-economic collected data of employees include age, gender, education level, occupation type, household monthly income, household size, household car ownership and start time of trip maker and the set of alternatives chosen by employees were Walk, Bajaj, Private car, Taxi and company provided service. These data can be summarized as shown in table 5.11. below.

Table5. 11. Summary of mode choice and socio-economic profile of employees

	Attributes	Label	Category	Frequency	Percentage
Individual characteristics	Age	1	18-45	324	56.4%
		2	Above 45	250	43.6%
	Gender	0	Male	253	44.1%
		1	Female	321	55.9%
	Education level	1	Primary school and below	84	14.6%
		2	Secondary school	161	28.0%
		3	Diploma or certificate	111	19.3%
		4	Degree and Above	218	38.0%
	Occupation type	1	Government employee	248	43.2%
		2	Self-employed	132	23.0%
		3	Private employee	194	33.8%
Household characteristics	Household income	1	Under 5000	69	12.0%
		2	5001-15000	483	84.1%
		3	Above 15001	22	3.8%
	Car ownership	0	No	453	78.9%
		1	Yes	121	21.1%
	Family size	1	Under 2	84	14.6%
		2	3-4	341	59.4%
		3	Above 5	149	26.0%
Travel attributes	Peak hour	0	No	44	7.7%
		1	Yes	530	92.3%
	Mode of transport chosen	1	Walk	73	12.7%
		2	Bajaj	16	2.8%
		3	Private	48	8.4%
		4	Minibus Taxi	281	49.0%
		5	Service	156	27.2%

Source:(Field survey, 2018)

From the 574 work trips, 44.1% of trips are made by men and 55.9% trips are by women. 56.4% of the trips are made by persons between the age groups of 18 to 45 and 43.6% were in the age group of above 45 years. The larger and lowest percentage of the trip makers based on the education level was a degree and above (38.0%) and primary school (14.60%) respectively. The distribution of job respondents in the table shows that about (43.2%) of the respondents are governmental employees, followed by (33.8%) private-sector employees and (23.0%) are under the self-employed category.

Only 3.8% of total work trips are made by persons with monthly household income levels higher than birr 15,000. The highest percentage (84.1%) of the respondents have a monthly household income between birr 5,000 and 15,000 and the other 12.0% was fall into to below birr 5,000 income group. Car ownership in the household indicates that about 81.7% do not own private cars and 18.3% have a private car. The distribution of respondents' family size was categorized into three categories. From the family categories, the largest percentage (59.4%) of the workers household have between three to four family members, 14.6% have family size of two and below, and the remaining 26.0% have above five members. Most of the trip makers (92.3%) in the study area leave their homes between 7:30 AM and 8:30 AM in the morning and only 7.7% of workers leave their homes for work before7:30 AM or after8:30 AM.

The distributions of the modes of transport that are usually used by the respondents in the study area are presented in the above table. Minibus taxi is the largest percent (49.0%) of the respondents used mode for work trips followed by the company provided service, walk, private car and Bajaj respectively. The use of Bajaj presented on the last order by 2.8%. 12.7% of the work trips are made by walking which shows the nearness of some residential areas to the work area.

II. Model result

A discrete choice data was analyzed using the MNL program of SPSS software. The program runs with different models using various attributes to ascertain the essential attributes of the model. In the analysis of the data multinomial logistic regression analysis was undertaken using the forward stepwise likelihood ratio method. Forward stepwise methods begin with a model that doesn't include any of the predictors. At each step, the predictor with the largest score statistic or likelihood ratio statistic whose significance value is less than a specified value (0.05) is added to the model. Different models were developed step by step to determine the best fit model. The step by step addition of variables is tested and coefficients of the utility functions are computed in all procedures. After performing numerous model estimation runs, the finalized Estimation results of the final best developed multinomial logit model are displayed in table 5.15 below. From the total collected sample of work trip makers due to missing data and other reasons 574 samples were used for the development of the model. According to the results obtained, the relationship between the modal choice variable and the set of independent variables is statistically significant. The result from the SPSS software for the development of a model for modal choice can be discussed as follows;

Model Fitting Information

The likelihood ratio test is performed by estimating two models and comparing the fit of one model to the fit of the other. In this case, the likelihood ratio test shows whether the model fits the data better than a null model. It is a likelihood ratio test of the final model against the one in which all the parameter coefficients are 0 (Null). The Chi-Square statistic in the model fitting is used for testing relationships between the dependent variable and the independent variables. ; They are independent. The chi-square statistic is the difference between the -2 log-likelihoods of the Null and Final models. In the model fitting table below the developed model have the chi-square value of 815.545 with significance (0.000) and is less than 0.001, so the null hypothesis of the Chi-Square test is that no relationship exists on the dependent and the independent variables is rejected, there is a significant relationship between the dependent variable and the entered independent variables. The multinomial logit model is much better than only a constant model.

Table5. 12. Model Fitting Information for work mode choice model

Model Fitting Information				
Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	1461.743			
Final	646.197	815.545	16	.000

Goodness-of-Fit

The goodness-of-fit table presents the null hypothesis for two tests that the model adequately fits the data. In multinomial logistic regression, Pearson and Deviance goodness-of-fit statistics are reported. In the case of the null is true; the Pearson and deviance statistics have chi-square distributions with the displayed degrees of freedom. If the significance value is less than 0.05, then the model does not adequately fit the data. The significance value for this case is greater than 0.10, so the data are consistent with the model assumptions.

Table5. 13. The goodness of fit for work mode choice model

Goodness-of-Fit			
	Chi-Square	Df	Sig.
Pearson	1457.662	1736	1.000
Deviance	646.197	1736	1.000

Pseudo R-Square

The coefficient of determination (R^2), in the linear regression model, summarizes the proportion of variance in the dependent variable associated with the predictor variables. The larger coefficient of determination (R^2) values indicating that more of the variation is explained by the model, to a maximum of 1, cannot be computed for multinomial logistic regression models. The Cox and Snell R^2 measure functions like R^2, with higher values indicating greater model fit. This cannot reach the maximum value of 1. The Nagelkerke proposed a modification that had the value range from 0 to 1and its measure is relied on indicating the strength of the relationship. Further, it has a relatively high pseudo-R^2 value of 0.823, indicating that approximately 82.3% of the variation in the dependent variable (modal choice) can be explained by the estimated MNL regression model. Thus, it can be concluded that the MNL model fits the sample data well.

Table5. 14. Pseudo R-Square value for mode choice model

Pseudo R-Square	
Cox and Snell	.758
Nagelkerke	.823
McFadden	.558

Likelihood Ratio Tests

The likelihood ratio tests in this research used to check the contribution of each effect to the work mode choice model. The result from the likelihood ratio test in Table 5.14 indicates that, the contribution of the variable to the overall relationship between the dependent and individual independent variables in differentiating between the groups specified by the dependent variable. The likelihood ratio test is a hypothesis test that the variable contributes to the reduction in error measured by the -2 log-likelihood statistic. If the significance value of the test is less than 0.05, then the effect contributes to the work mode choice model. In this case, the variables; family size, job type and total travel time contribute significantly to explaining the mode choice.

Table5. 15. Likelihood Ratio Tests for work mode choice model

Likelihood Ratio Tests				
Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	Df	Sig.
Intercept	646.197	.000	0	.
TT	1331.877	685.680	4	.000
FS	690.119	43.922	8	.000
JOB	706.741	60.544	4	.000

Parameter Estimates

All variable coefficients estimated using the multinomial logit model for work trips were found to be statistically significant at the 0.05 confidence level. If the independence variables significance level is less than 0.05 that variable is selected for this method. The most basic test of the estimation results is to examine the signs of the estimated parameters with theory and judgment regarding the expected impact of the corresponding variables. The parameter estimates of the minimum acceptable model are reported in table 5.16.

Table5. 16. Parameter Estimates for work mode choice model

Parameter Estimates							
Travel mode	Intercept and variables	B	Std. Error	Wald	Df	Sig.	Exp(B)
Walk	Intercept	10.241	1.385	54.704	1	0.000	
	TT	-0.446	0.051	76.437	1	0.000	0.64
Bajaj	Intercept	17.467	2.124	67.615	1	0.000	
	TT	-0.77	0.087	77.61	1	0.000	0.463
	[JOB=0]	-2.485	0.873	8.096	1	0.004	0.083
Private	Intercept	26.561	2.285	135.128	1	0.000	
	TT	-1.282	0.12	115.084	1	0.000	0.277
	[FS=0]	-3.048	1.161	6.888	1	0.009	0.047
	[FS=1]	-3.521	0.774	20.684	1	0.000	0.03
Service	Intercept	16.853	1.65	104.357	1	0.000	
	TT	-0.738	0.065	128.729	1	0.000	0.478
	[JOB=0]	1.564	0.467	11.214	1	0.001	4.776
The travel mode reference category is Taxi.							

III. Discussion on the estimated coefficients

In the interpretation of the sign of the parameter, parameters with significant negative coefficients are interpreted to result in a decrease in the likelihood of that response category with respect to the reference category and parameters with significant positive coefficients are increasing the likelihood of that response category with respect to the reference category.

From the final model estimation results for work trips, using the multinomial logit model, shown that all the estimated confidence was found to have negative signs, other than (JOB=0) government employee for service. The negative estimated coefficient of the walk, Bajaj, private car and service total travel time variable is -0.446, -0.77, -1.282 and -0.738 respectively. This implies that the utility of the walk, Bajaj, private car and service decrease as the total travel time increases. The results also show that the family size variable for private cars has negative coefficients which imply as the family size increases the employees choosing the private car relative to the reference category (Minibus taxi) decrease. The other variable with a negative coefficient in the model is the job for Bajaj (JOB_{Bajaj}). This job category (JOB=0) government employee for Bajaj indicates that the individuals who travel for work prefer Bajaj relative to Minibus taxi tend to decreases. On the other hand for the case of (JOB=0) government employee for service parameter sign is positive which means government employees are more likely to prefer service over taxi as their work trip purpose.

The results also show that total travel time (TT), the job for Bajaj and service ($JOB_{Bajaj\ and\ service}$), and family size for private car ($FS_{private\ car}$) all these parameters have a P-value of less than 0.05 at 95% confidence level, the regression coefficient for all the total travel time, job category and family size group has been found to be statistically significant at the significance level of 0.05.

By comparing the formulated models it is obtained that the empirical utility functions for different mode can be written as follows:

$$U_{walk} = 10.241 - 0.446\ (TT)$$
$$U_{Bajaj} = 17.467 - 0.77\ (TT) - 2.485\ (JOB=0) \qquad (5.7)$$
$$U_{private\ car} = 26.561 - 1.282\ (TT) - 3.048\ (FS=0) - 3.521\ (FS=1)$$
$$U_{service} = 16.853 - 0.738\ (TT) + 1.564\ (JOB=0)$$

Based on the developed models; total travel time of the modes, family size interacted with Private car, and job category interacted with Bajaj and company provided service of the trip maker are the factors that statistically affect the preference of trip makers among walk, bajaj, private car, and company provided service.

5.3.2. Mode choice model for education trips

I. Socio-demographic characteristics

The total number of revealed mode choice response sample achieved for educational trips were 501. The descriptive statistics on the socio-demographic characteristics of students for the study area are presented below in table 5.17.

Table5. 17. Descriptive statistics on the socio-demographic characteristics of students

Socio-demographic characteristics					
Dimension	Attributes	Label	Category	Frequency	Percentage
Individual characteristics	Gender	0	Male	240	47.9%
		1	Female	261	52.1%
	Age	1	Age 10-13	153	30.5%
		2	Age13-15	186	37.1%
		3	Above15-18	162	32.3%
	EDL	1	Primary school	239	47.7%
		2	Secondary school	262	52.3%
Household characteristics	AFMI	1	Below5000	45	9.0%
		2	Between5001-15000	437	87.2%
		3	Above15000	19	3.8%
	CPH	0	No	414	82.6%
		1	Yes	87	17.4%
Total sample				501	100.0%

As to the socio-demographic characteristics of the respondents, from the total sample size female respondents has a greater percent for educational trips (52.1%) and the remaining percent (47.9%) are males. The age level of students was from 10 to 18. The group of ages from 13 to 15 takes the highest proportion (37.1%) while from 15 to 18 placed in the second with (32.3%) and age students 10 to 13 fall in the last order about (30.5%). There is no much difference in the case of age group percentage. In terms of income, the majority of the respondents are in the ranges of between birr 500 to 15000 household income level which is (87.2%) and the other range fails in birr Below5000 and birr above 1500 thus, with (9.0%) and (3.8%) respectively. About (82.6%) of the respondents do not have a private car in their households while (17.4%) have a private car. From the total sample for educational trips, 47.7% of trips are made by primary school and 52.3% trips are by secondary school students.

II. Mode of transport chosen by Students

The mode of transport was chosen by students for educational trips are; walk, Bajaj, Minibus taxi, private car, and School service. The table below illustrates the percentage of modal split chosen by students.

Table5. 18. Modal choice chosen by students

Mode of Transport	Frequency	Percentage
Walk	112	22.4%
Bajaj	18	3.6%
Private car	55	11.0%
Taxi	241	48.1%
Service	75	15.0%
Total	501	100%

According to the general analysis results of the survey, 48.1% of the respondents use Minibus taxi, 22.4% use walk, 15.0% use Service 11.0% use Private car, and 3.6% use Bajaj as their major mode of transportation. From the table above the respondents usually, use the Minibus taxi is the dominant mode of transportation to go their education. The use of Bajaj presents as the smallest one from the sample.

III. Modeling results

Same as of employees the data were analyzed using the SPSS software Multinomial logistic regression. A sample of 501 was collected and taken for model formulation. The program ran with different models using various attributes to establish the essential attributes of the model. In general, the maximum likelihood estimation method is used for the regression of the model utility function of the probability of mode selection for education. After many iterations, through the multinomial logistic regression analysis of the above data, the final model was obtained as shown in table 5.22. All the parameters in the final model are statistically significant at the 95% confidence level, with a good fit and a high pseudo-R-squared value. In addition, the goodness-of-fit measure (i.e., pseudo-R-squared) is 0.824, indicating that the independent variable could explain 82.4% of the variation in the dependent variable (modal choice) and which represents a good fit.

Table5. 19. Pseudo R-Square for school mode choice model

Pseudo R-Square	
Cox and Snell	.767
Nagelkerke	.824
McFadden	.546

Table5. 20.Goodness-of-Fit for school mode choice model

Goodness-of-Fit			
	Chi-Square	Df	Sig.
Pearson	722.965	664	.056
Deviance	606.517	664	.946

Model fitting information

The model fitting information shown in Table 5.20 below is the Likelihood Ratio Chi-Square test that at least one of the predictors' regression coefficients is not equal to zero in the model. The analysis of Multinomial Logistic Regression is used to describe the overall test of the relationship between the dependent and independent variables. This relationship is based on the statistical significance of the final model Chi-Square. The difference between the initial Log Likelihood value (with no independent variable) and the final Log Likelihood value (including the independent variable) is a Chi-Square value of 729.423 and its significance value is 0.000 (which is less than 0.05). This shows that there is a significant relationship between the dependent variable and the independent variables. So, the null model is rejected by the final model.

Table5. 21. Model fitting information for school mode choice model

Model Fitting Information				
Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	1335.940			
Final	606.517	729.423	8	.000

Likelihood Ratio

The contribution of the individual independent variable to the dependent variable Checks by the Likelihood Ratio Test. In the test effect contributes significantly to the model has to be a significance value of less than 0.05. In the Likelihood Ratio Test table the travel time and gender variables have a significance value less than 0.05 then they have an effect that contributes to the model.

Table5. 22. Likelihood Ratio Tests for school mode choice model

Likelihood Ratio Tests				
Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	Df	Sig.
Intercept	606.517	.000	0	
TT	1327.722	721.205	4	.000
GENDER	621.338	14.821	4	.005

IV. Parameter Estimates

The Multinomial Logistic Regression Model used in the estimates of the effect of the variables on the probability of choosing the mode of transport to the educational trip purpose. A multinomial logistic regression model is designed to define four categories such as walk, Bajaj, private car and school service. In the interpretation of the Multinomial Logit, for a unit change in the predictor variable, the logit of outcome relative to the reference group is expected to be changed by its respective parameter estimate. If the significance of the statistic is small (i.e., less than 0.05) then the parameter is beneficial to the model. The coefficients were estimated by fitting the data into the model were shown in table 5.23.

Table5. 23. Parameter Estimates for school mode choice model

Parameter Estimates							
Travel Mode	Intercept and variables	B	Std. Error	Wald	Df	Sig.	Exp(B)
Walk	Intercept	61.409	12.142	25.578	1	.000	
	TT	-2.505	.477	27.556	1	.000	.082
Bajaj	Intercept	68.553	12.219	31.478	1	.000	
	TT	-3.256	.493	43.639	1	.000	.039
Private	Intercept	66.275	12.176	29.629	1	.000	
	TT	-2.938	.482	37.106	1	.000	.053
	[GENDER=0]	1.497	.635	5.559	1	.018	4.467
Service	Intercept	59.188	12.136	23.788	1	.000	
	TT	-2.393	.477	25.189	1	.000	.091
The travel mode reference category is Taxi.							

V. Discussion on the estimated coefficients

Parameters having significant negative coefficients decrease the likelihood of that response while Parameters with positive coefficients increase the likelihood of that response category. The travel time variables in the utility functions of all modes appeared with negative signs indicating a declined utility with travel time (TT) increase. The estimation results of the model presented in table5.22 show that the travel time coefficients have a correct negative sign. The results also show that the negative sign of parameters of total travel time (TT) in the model indicates that the decrease in total travel time of a mode will improve the utility of the mode choice for a trip maker. The coefficient of the gender (GENDER=0) males appeared with a positive sign in the utility function of private cars, which indicates that males are expected to prefer private car mode to taxi mode. Based on the developed models, total travel time of the modes and gender of the trip maker interacted with private cars are the factors that statistically significantly affect the reveled preference of trip makers among school service, taxi (minibus), private car, Bajaj and walking. From the entire model specifications tested, the most satisfactory models for educational trip purposes are presented at last.

The empirical utility functions for the multinomial logit model based on the final model in the table above for different mode can be written as follows:

$U_{walk} = 61.409 - 2.505\ (TT)$

$U_{Bajaj} = 68.553 - 3.256\ (TT)$ (5.8)

$U_{private\ car} = 66.275 - 2.938\ (TT) + 1.497\ (GENDER)$

$U_{service} = 59.188 - 2.393\ (TT)$

5.4. Trip production and Mode choice model validation

5.4.1. Validation of Trip Production Model

The validation of the model is an important step to make sure that the model functions behave consistently in future estimates. Since validation is basically represented by the comparison between the observed values and the modeled values. Specifically, in order to validate the estimated general trip production model, a comparison of the estimated average daily trips per household (as calculated from the estimated regression model) with the actual average daily trips per household (from observed values) is carried out using 99 samples of data. If the estimated model results and the survey observations are in acceptable agreement, the model can be considered validated. The whole result of the observed trip production compared to estimated trips from the model is given in the appendix (see Appendix C).

Table5. 24. General trip production model validation

observation no.	Observed Trip P production	Estimated Trip P production	Difference
1	8.57	8.98	-0.41
2	6.29	6.65	-0.36
3	7.14	7.09	0.05
4	8.29	8.78	-0.50
5	10.71	10.65	0.06
6	5.72	6.57	-0.86
7	11.71	10.73	0.98

The results in above Table 5.18 show that there is no sever difference between the estimated and observed value. A good match between the estimated average daily trips per household and the actual average daily trips per household for most of the randomly selected observations. Thus, the results support the estimated multiple linear regression model as a good model for daily trip production per household.

5.4.2. Mode choice model validation

I. Work Model Validation

Model validation is considered a very important process to evaluate the performance of the calibrated mode choice model and the model's ability to predict individuals' mode choice behavior. The validation process is tested on two different phases. The first phase is the test of reasonableness validation process which was tested during the calibration process depending on the expected sign of estimators. So, all the variables included in the developed model have correct signs of coefficients.

The second phase for the validation process is to calculate the prediction capability of the calibrated model. To calculate the prediction ratio, the utility for each trip maker was calculated then the probability of each alternative (mode) was estimated. The alternative with the highest probability is predicted to be the chosen mode for that particular individual. The number of travelers correctly predicted was summed up to each alternative and compared to yield the prediction value. For the purpose of model validation, 144 observations have been used. The calculated prediction ratio of the model is 0.9097 which means that the model is capable to predict about 90.97% of the choices of the trip makers' correctly.

Table5. 25. Sample of work mode choice model validation

No.	Observed mode choice	Probability of choosing a mode					Predicted mode choice	Correct prediction
		Walk	Bajaj	Private Car	Service	Minibus Taxi		
1	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
2	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
3	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
4	Private car	0.00034	0.00011	0.71741	0.28213	1.26E-06	Private car	Yes
5	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
6	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes

The main diagonal cells give the match number between observed and predicted mode choice and off-diagonal provides the erroneous one, which is shown in the table below. The overall correctly predicted percentage indicates the percent of cases where the actual choices and the predicted choices of the individuals match.

Table5. 26.Observed and predicted for work modal split of validation data

		Predicted Mode Choice					Total
		Walk	Bajaj	Private car	Minibus Taxi	Service	
Observed Mode Choice	Walk	0	0	0	0	5	5
	Bajaj	0	0	0	0	4	4
	Private car	0	0	16	0	2	18
	Taxi	0	0	0	70	2	72
	Service	0	0	0	0	45	45
Total		0	0	16	70	58	144

Prediction accuracy $= \frac{0+0+16+70+45}{144} = 90.97\%$

II. Education Model Validation

To test the validation of the education model, the same approach was used in the work model validation. The informal test for reasonableness of basic parameters and computation of the ratio of prediction for the validation of the education model is also done. The reasonableness test of the parameters validation process was tested during the calibration process as expected by the estimators. For example, the generic variable, total travel time in the developed model is negative which is the same as the expected signs since the utility of modes towards the trip makers decreases as the travel time increase. To calculate the prediction ratio, first, 125 randomly observed data and the utility for each trip maker is computed then the probability of each alternative is computed. The alternative with the highest probability is predicted to be chosen the mode for that particular individual. The prediction ratio was then calculated by dividing the correct choices by the total number of individuals' choices. The estimated prediction value is 0.872. In other words, the calibrated model is capable of correctly predicting about 87.2% of the choices of the trip makers, as computed below. The result of the samples that were correctly predicted using the developed models and the samples that the developed models predicted other than the observed chosen modes are shown below in table 5.27. The resulting probability of choosing the modes for each education trip in the observed preference and the predicted preferences are as shown in table below5.27, partially.

Table5. 27. Sample of school mode choice model validation

No.	Observed mode choice	Probability of choosing a mode					Predicted mode choice	Correct prediction
		Walk	Bajaj	Private car	Service	Minibus Taxi		
1	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
2	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
3	Walk	0.209082	0.032289	0.671644	0.086984	5.08E-15	Walk	Yes
4	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
5	Service	0.394244	0.007324	0.373497	0.224935	1.12E-11	Walk	No
6	Private car	0.015497	0.251856	0.729427	0.00322	6.77E-23	Private car	Yes

probabilities for each mode of transportation were calculated for each trip in the validation dataset which is shown in Table 5.27 in the above shows the probability of mode choice and that was used for the for prediction ratio computation.

Table5. 28.Observed and predicted for school modal split of validation data

		Predicted Mode Choice					Total
		Walk	Bajaj	Private car	Minibus Taxi	Service	
Observed Mode Choice	Walk	17	0	0	0	2	26
	Bajaj	3	3	0	0	0	6
	Private car	0	0	11	0	0	11
	Taxi	0	0	0	50	0	50
	Service	0	0	0	5	31	39
Total		23	3	11	55	33	125

Prediction accuracy $= \frac{17+3+11+50+31}{125} = 87.2\%$

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

The thesis performed the development of the home-based trip production model and mode choice model. Trip production and mode choice is a critical step of the conventional four steps model estimates in the urban transportation modeling system. The trip production models were developed for general trip production and five different trip production model categories using multiple regression at the household level. The five-trip production models ware categorized according to the trip purpose. The trip productions based on purpose such as work, education, shopping, social and recreation home-based trips. In order to estimate the generated trips, the household questionnaire survey was done based on stratified random sampling, with the sample size distributed to each Tabiya in the study area. The sample size collected from Home interview surveys for the trip production model was 395 households. This study identified that the significant variables affecting the trip production (TP) of the Hawelti sub-city residents are the monthly household income, number of persons receiving education, the number of employed persons, the number of females and age above 45 persons in the household. The general trip production model was checked by comparing its results with the field data.

Based on the trip purpose, the variables that are significant in the work trips production model are the number of employed persons in the household and car ownership per household. The most important variable that affects the education trips production model is the number of persons receiving education in the household. The most relevant explanatory variables for the shopping trip production model are found to be the number of employed persons in the household and the monthly household income. The variables that are the most relevant in the social trip production model are age above 45 persons in the household, the number of females in the household, and the number of employed persons in the household. The most important explanatory variables in the recreational trip production model are the monthly household income and number of persons receiving education in the household. All the developed trip production models have good explanatory power with an R^2 value of greater than 0.7. This high value of goodness of fit indicates that the model is well defined and the trips are a function of the variables chosen for this study.

A mode choice model was developed for Hawelti sub-city residents. The mode choice model consists of multinomial logit models developed for two-trip purpose categories namely home-based work and home-based education trips. The analysis is conducted based on household travel survey data. The data used for the model were socioeconomic data, travel characteristics and travel modes. The five available modes (walking, Bajaj, private vehicle, taxi, and service) were taken into account for mode choice analysis according to the survey data. All input data were transformed into the proper format for analysis in the SPSS and a multinomial logistic regression model has been adopted to reveal the factors affecting the travel mode choice. To reach the final model, a number of models have been developed and tested for the goodness of fit, the sign of parameters and the significance of coefficients.

The mode choice model developed for work showed that the total travel time of the mode, job type interacted with Bajaj and company provided service and family size interacted with Bajaj, private car and company provided service is proved the significant factors that influence the mode usage of the work trips. The negative sign in total travel time in the developed models showed that decreasing a unit travel time of a transport mode to increase its utility. The effect of job on mode choice is found to be positive for the service and negative for traveling using Bajaj relative to using a taxi. The positive sign associated with the company provided service transport shows that people tend to prefer a service relative to using a taxi and the negative sign associated with Bajaj indicates that, the preference of using the Bajaj mode decrease than using a taxi. The effect of family size on mode choice is found to be negative for traveling using private cars, Bajaj and service relative to using a taxi. In addition, Bajaj is found to have a more negative parameter than private cars and services indicating that people choose Bajaj less than to use private cars and services as family members increase. The developed models for the revealed preference data for work trip purposes are validated to be capable of predicting 80% of the preference of modes of trip makers of the study area.

The mode choice model for the case of a school trip to the study area travel time and gender are proved to be the significant factors that influence the mode usage of the students. The coefficients of gender were positive indicates male students are more likely to use private car transport than all other modes. The developed models for the revealed preference data for educational purpose trips are validated to be capable of predicting 80% of the preference of modes of trip makers of the study area. The developed models for trip production and mode choice are tested and compared at a 95% level of confidence level. Both, the trip production and mode choice models were analyzed, calibrated and validated based on the collected data from trip makers of Hawelti sub-city residents.

6.2. Recommendations

In this study, the developed model for trip production and mode choice models is beneficial to transport planners and policymakers in the estimation of the number of trips and modal attributes' and the socio-economic characteristics of the trip makers in the study area. And also it is recommended that, using the results found in this thesis to compare with research conducting a trip production model based on time in the sub-city.

Developing the mode choice model by considering both the travel maker's and the travel mode's characteristics were studied. The socio-economic and travel mode variables are used for a better understanding of the people's behavior in choosing among the modes of transport. The results found in this research as a base, researchers are encouraged to develop modal choice models for non-home based trips in Mekelle city and including the variables related to comfort, safety, availability and reliability of transport modes. Furthermore, by using this study as a starting point, recommends the development of the trip production and modal choice models considering all types of trips and the whole combinations of modes for the practical applications of the Hawelti sub-city. Finally, the results are encouraged to check model validation by collecting a new data whether that the produced average daily trips base estimation model's formula used in this study can be used to estimates generated trip and mode choice characteristics in other areas of Mekelle city by comparing the estimated trips from the developed models with the observed trips.

REFERENCE

Abdulhaq, D. M. N. (2016). Transportation Mode Choice Model for Palestinian Universities Students: A Case Study on An-Najah New Campus (Doctoral dissertation).

Ali, F. (2018). Development of a modal choice model for home-based work trips: the case of Jemo area residents Addis Ababa: Addis Ababa University.

Al-Masaeid, H. R., and Fayyad, S. S. (2018). Estimation of Trip Generation Rates for Residential Areas in Jordan. Jordan Journal of Civil Engineering, 12(1).

Almasri, E.H. (2011). Factors affecting travel choice of shared taxi versus bus for Palestinian university student trips. International Review of Civil Engineering, 2(1), 35-45.

Almasri, E., & Alraee, S. (2013). Factors Affecting Mode Choice of Work Trips in Developing Cities—Gaza as a Case Study. Journal of Transportation Technologies, 3(04), 247.

Alraee, S. I. S. (2012). Development of the Mode Choice Model for Gaza City. Development of the Mode Choice Model for Gaza City.

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

Arasan, T. V. (2012). Urban Transportation systems planning. Unpublished Hand Book presented at Short Term Course organized by Kwame Nkrumah University of Science and Technology and Indian Institute of Technology Madras, Accra.

Chang, J. S., Jung, D., Kim, J., and Kang, T. (2014). Comparative analysis of trip generation models: results using home-based work trips in the Seoul metropolitan area. Transportation Letters, 6(2), 78-88.

Cheng, L., Chen, X., Wei, M., Wu, J., & Hou, X. (2014). Modeling mode choice behavior incorporating household and individual sociodemographic and travel attributes based on rough sets theory. Computational intelligence and neuroscience, 2014, 26.

Danaf, M., Abou-Zeid, M., and Kaysi, I. (2014). Modeling travel choices of students at a private, urban university: insights and policy implications. Case studies on transport policy, 2(3), 142-152.

Daniel, E., and Ituen, U. (2013). Inter-urban trip generation models for the urban centers in Akwa Ibom State, Nigeria. Civil and Environmental Research, 3(4), 55-62.

Dodeen, A. M. Y. (2014). Developing trip generation models utilizing linear regression analysis: Jericho city as a case study.

Eluru, N., Chakour, V. and El-Geneidy, A. (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.

Fox, J., and Patruni, B. (2015). STM3 2011 base frequency, mode-destination and car ownership models. No. RR-1125-BTS. Cambridge, UK: RAND Corporation.

Gebeyehu, M., & Takano, S. E. (2007). Diagnostic evaluation of public transportation mode choice in Addis Ababa. Journal of Public Transportation, 10(4), 2.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? Multivariate behavioral research, 26(3), 499-510.

Hahn, J. S., Kim, H. C., Kim, J. K., and Ulfarsson, G. F. (2016). Trip making of older adults in Seoul: Differences in the effects of personal and household characteristics by age group and trip purpose. Journal of transport geography, 57, 55-62.

Hensher, D. A., & Button, K. J. (2000). Handbook of Transport Modelling (Handbooks in Transport). εκδ. Pergamon.

Hunt, L., and Broadstock, D. (2010). Modeling Car Trip Generations for UK Residential Developments Using Data from TRICS, Transportation Planning, and Technology, 33 (8), PP. 671-678.

Huntsinger, L. F., Rouphail, N. M., & Bloomfield, P. (2013). Trip generation models using cumulative logistic regression. Journal of Urban Planning and Development, 139(3), 176-184.

Hutchinson, B.G. (1974). Principles of Urban Transport System Planning, McGrew-Hill Book Company, New York

Institute of Transportation Engineers, ITE. (2012). "Trip generation manual". 9th Edition, ISBN No: 0935403-79- 5, Washington, D.C., USA.

Jittrapirom, P., Emberger, G., Brezina, T., & Shibayama, T. (2013). Chiang Mai City Mobility and Transport Survey (CM-MTS) 2012 (No. 1).

Juan de Dios Ortuzar and Luis G. Willumsen. (2011). "Modelling transport". 4th Edition. John Wiley and Sons, Ltd., United Kingdom.

Koppelman, F. S., & Bhat, C. (2006). A self-instructing a course in mode choice modeling: multinomial and nested logit models.

Li, B. (2011). The multinomial logit model revisited: A semi-parametric approach in discrete choice analysis. Transportation Research Part B: Methodological, 45(3), 461-473.

Lin, K. Y., and Lu, H. P. (2011). Why people use social networking sites: An empirical study integrating network externalities and motivation theory. Computers in human behavior, 27(3), 1152-1161.

Mengstu. E. H. (2015). Logit Model for School and Work Trip Mode Choice on Railway Route Corridor. Addis Ababa.

Meyer, M. D., and Miller, E. J. (2001). Urban transportation planning, McGraw-Hill Book Company, New York.

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

Mousa, H. (2013), Development of a Trip Generation Model for Gaza City, M.Sc. The Islamic University, Gaza.

Mustafa, A. (2016). Trip generation models for selected land use in the West Bank. M.Sc. Thesis, An-Najah University, Palestine.

Ortuzar, J. d. D. & Willumsen, L. G., (2011). Modeling transport. 4th edition. Southern Gate: A John Wiley and Sons, Ltd., New York.

Pettersson, P., and Schmöcker, J. D. (2010). Active aging in developing countries?–trip generation and tour complexity of older people in Metro Manila. Journal of Transport Geography, 18(5), 613-623.

Priyanto, S., and Friandi, E. P. (2010). Analyzing of Public Transport Trip Generation in Developing Countries; A Case Study in Yogyakarta, Indonesia. World Academy of Science, Engineering and Technology, 42, 6-24.

R Ashalatha, V S Manju, Arun Baby Zacharia, (2013). Mode choice behavior of commuters in Thiruvananthapuram city, American Society of Civil Engineers.

Roorda, M. J., Páez, A., Morency, C., Mercado, R., and Farber, S. (2010). Trip generation of vulnerable populations in three Canadian cities: a spatial ordered probit approach. Transportation, 37(3), 525-548.

Sarsam, S. I. (2011). Modeling household trip generation for selected zones at the AL-Karkh side of Baghdad city. Journal of Engineering, 17(6), 1461-1472.

Sincero, S. M. (2012). Personal Interview Survey. Retrieved June 19, 2018.

Sofia, G. G., Ali, A. H. A., & Al-Zubaidy, H. A. (2012). Trip Generation Modeling for Selected Zone in Al-Diwaniyah City. Journal of Engineering and Sustainable Development, 16(4), 167-180.

Sowgat, T. (2012). "Pro-poor planning in Bangladesh: a case study of Khulna city". Doctoral Dissertation, Heriot-Watt University, Edinburgh, UK.

Torrieri, V. (1985). Mobility and Traffic Forecasting, Urban Transportation Sector in AAMP, People's Democratic Republic of Ethiopia, National Urban Planning Institute.

Tushara, T., Rajalakshmi, P., & Bino, I. K. (2013). Mode Choice Modelling For Work Trips in Calicut City. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN, 2278-3075.

Yamane, Taro. (1967). Statistics, an introductory analysis, 2nd edition, New York: Harper and Row.

APPENDICES

Mekelle University

Ethiopian Institute of Technology – Mekelle

School of Civil Engineering

Master of Science in Road and Transportation Engineering

Developing Trip Production and Mode Choice Model for Home-Based Trips: The Case of Hawelti Sub-City Residents

Appendix A: Questionnaire

General Information

This research survey is designed to fulfill an academic requirement for the M.Sc. degree program in Road and Transport Engineering at Mekelle University. I can assure you that, the research data will be used only for academic purposes. Your open and prompt response is highly appreciated. This questionnaire for interview aims to collect socioeconomic data of the individuals/households of Hawelti sub-city, the situation of the transport system and the factors that affect the people's choice for transportation modes these data will be used to study their effect on the purpose of trips and to build a mathematical trip production and mode choice model for Hawelti sub-city residents.

Thanks for your invaluable cooperation.

Prepared by:

Gerezgiher Gebrewahd Hagos

Supervised by:

Dr. Ashenafi Aregawi

Part I. Individual and household socioeconomic characteristics of the traveler

1. Gender

□ Male □ Female

2. What is your age in years…………………………

3. What is your occupation?

□ Student

□ Government employee

□ Private employee

□self-employed

□ Other

4. Education level?

□ Other (Illiterate)

□ Primary school (8th grade and below)

□ High School (9th -12th grade)

□ Certificate or Diploma

□ Degree and above

5. What is your family average income per month in birr..................?

6. What is your family size?....................person

7. Do you have a car in your household?

□ Yes □ No

Part II. Trip Characteristics

1. Which transport mode do you usually use to travel?

□ Walking

□ Bajaj

□ Private car

□ Minibus Taxi

□ Service (company provided, school service/bus)

2. What is the purpose of your travel?

□ Work trips

□ Education trips

□ Shopping trips

□ Recreational trips

□ Social trips

3. At what time of the day do you usually leave your home to travel? ....
4. What is your trip origin? --
5. What is your trip destination? --
6. Please give your estimation for travel time and cost for the modes you can use on your trip? (Fill only for the modes you can use).

A. Walking mode

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

B. Bajaj

- Travel time in bajaj in minutes ……………………………………………
- Travel cost/fare in birr …………………………………………

C. Private car as a driver

- Travel time in minutes **…………………………………**
- Average monthly fuel consumption in liters ……………
- How many kilometers do you travel with 1-liter fuel? ....
- How often do you give your car for maintenance and how much do you pay in Birr….…

D. Minibus taxi

- Access time from home to nearby station in minutes …………………………
- Time of waiting for a taxi in minutes ……………………………………………………
- Travel time in the taxi in minutes ……………………………………………………
- Travel cost/fare in birr ……………………………………….................................

E. Service(company provided and school bus)

- Access time from home to nearby station in minutes ………………………
- Travel time on the bus in minutes …………………………………………………………
- Travel cost/fare in birr ……………………………………….................................

APPENDIX B: correlation matrix and SPSS results for the trip production model

Correlation matrix

	FS	NF	NM	NS	NEP	OTHER	$Age_{Under17}$	$Age_{18and45}$	$Age_{above45}$	CPH	AFMI
FS	1	.696	.646	.768	.553	.556	.659	.581	.677	.432	.489
NF	.696	1	-.097	.549	.400	.347	.478	.414	.447	.327	.302
NM	.646	-.097	1	.481	.341	.403	.405	.366	.464	.251	.357
NS	.768	.549	.481	1	.077	.224	.657	.376	.466	.255	.383
NEP	.553	.400	.341	.077	1	-.006	.260	.345	.436	.360	.344
OTHER	.556	.347	.403	.224	-.006	1	.261	.401	.384	.210	.171
$Age_{Under17}$	.659	.478	.405	.657	.260	.261	1	.065	.302	.279	.305
$Age_{18and45}$	.581	.414	.366	.376	.345	.401	.065	1	-.020	.259	.318
$Age_{above45}$	.677	.447	.464	.466	.436	.384	.302	-.020	1	.290	.311
CPH	.432	.327	.251	.255	.360	.210	.279	.259	.290	1	.552
AFMI	.489	.302	.357	.383	.344	.171	.305	.318	.311	.552	1

I. General trip production model

The regression results for the estimated general trip production model from the SPSS software are shown below

Model Summary			
R	R Square	Adjusted R Square	Std.error of the Estimate
0.943	0.889	0.887	0.783277

ANOVA					
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1903.953	5	380.791	620.662	.000
Residual	238.047	388	.614		
Total	2142.000	393			

Coefficients					
Constant and variables	Coefficient	Std. Error	T – value	Sig.	VIF
Constant	.572	.142	4.041	.000	
NS	1.629	.066	24.848	.000	1.919
NEP	1.855	.075	24.693	.000	1.600
AFMI	.079	.015	5.276	.000	1.328
NF	.150	.052	2.883	.004	1.769
$Age_{above45}$	.143	.070	2.056	.040	1.619

II. Work Trip Production Model

The results obtained from the SPSS software for the work trip production model are shown below:

Model Summary			
R	R Square	Adjusted R Square	Std.Error of the Estimate
.931	.867	.867	.443643

ANOVA					
Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Regression	502.533	2	251.267	1.277E3	.000
Residual	76.956	391	.197		
Total	579.490	393			

Coefficients					
Variables	Coefficients	Std. Error	t- value	Sig.	VIF
Constant	.513	.065	7.905	.000	
NEP	1.649	.036	45.752	.000	1.149
CPH	.241	.068	3.535	.000	1.149

III. Educational Trip Production Model

Model Summary			
R	R Square	Adjusted R Square	Std.Error of the Estimate
.980	.960	.960	.252771

ANOVA Table					
Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Regression	598.659	1	598.659	9.370E3	.000
Residual	25.046	392	.064		
Total	623.706	393			

Coefficients					
Variables	Coefficients	Standard Error	t –value	Sig.	VIF
Constant	.066	.023	2.833	.005	
NS	1.478	.015	96.797	.000	1.000

IV. Shopping Trip Production Model

The final regression analysis results obtained from the SPSS software for shopping trip production model are given below:

Model Summary			
R	R Square	Adjusted R Square	Std.Error of the Estimate
.893	.797	.796	.301798

ANOVA Table					
Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Regression	140.068	2	70.034	768.911	.000
Residual	35.704	392	.091		
Total	175.772	395			

Coefficients					
Variables	Coefficients	Standard Error	t –value	Sig.	VIF
NEP	.149	.021	27.095	.000	6.839
AFMI	.035	.004	8.19	.000	6.839

V. Social Trip Production Model

The regression results obtained from the SPSS software for the estimated social trip production model are shown below:

Model Summary

R	R Square	Adjusted R Square	Std.Error of the Estimate
0.923	0.853	0.852	.231952

ANOVA Table

Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Regression	121.799	3	40.600	754.616	.000
Residual	21.036	391	.054		
Total	142.835	395			

Coefficients

Variables	Coefficients	Standard Error	t –value	Sig.	VIF
NEP	.110	.016	6.768	.000	6.874
NF	.104	.013	7.911	.000	5.745
$Age_{above45}$	.096	.019	5.122	.000	6.143

VI. Recreational Trip Production Model

The regression results obtained from the SPSS software for the estimated recreational trip production model are shown below:

Model Summary

R	R Square	Adjusted R Square	Std.Error of the Estimate
.891	.794	.793	.314360

ANOVA Table

Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Regression	149.519	2	74.759	756.503	.000
Residual	38.738	392	.099		
Total	188.257	395			

Coefficients

Variables	Coefficients	Standard Error	t –value	Sig.	VIF
AFMI	.038	.003	11.364	.000	3.811
NS	.189	.020	9.297	.000	3.811

APPENDIX C: Validation results of the trip production model

No.	Observed TP	Estimated TP	Difference
1	8.57	8.98	-0.41
2	6.29	6.65	-0.36
3	7.14	7.09	0.05
4	8.29	8.78	-0.50
5	10.71	10.65	0.06
6	5.72	6.57	-0.86
7	11.71	10.73	0.98
8	7.86	9.06	-1.20
9	9.00	11.04	-2.04
10	7.86	6.87	0.99
11	11.43	11.37	0.06

12	6.286	6.87	-0.58
13	4.72	5.17	-0.46
14	7.43	6.88	0.55
15	7.57	8.69	-1.12
16	10.14	10.51	-0.36
17	8.43	8.96	-0.53
18	6.29	6.77	-0.48
19	5.00	4.75	0.25
20	8.29	7.48	0.80
21	9.72	8.80	0.91
22	11.00	10.96	0.04
23	5.14	5.13	0.01
24	3.29	4.82	-1.53
25	6.72	6.72	-0.01
26	7.43	6.67	0.76
27	5.86	6.94	-1.08
28	10.29	11.26	-0.98
29	5.86	5.09	0.77
30	5.86	7.40	-1.54
31	9.43	9.00	0.43
32	7.29	7.17	0.12
33	5.29	6.76	-1.47
34	4.14	4.55	-0.41
35	5.86	6.96	-1.10
36	11.28	11.15	0.13
37	5.86	5.19	0.67
38	6.29	6.84	-0.55
39	7.00	7.21	-0.21
40	9.71	11.04	-1.33
41	4.86	5.01	-0.16
42	12.29	9.11	3.18
43	7.29	9.15	-1.87
44	9.57	10.89	-1.32
45	5.57	6.88	-1.31
46	6.43	5.33	1.10
47	1.86	2.82	-0.96
48	8.86	8.43	0.43
49	6.14	7.01	-0.87

50	4.86	4.93	-0.08
51	6.29	6.98	-0.69
52	4.29	5.10	-0.82
53	2.86	3.02	-0.16
54	4.43	5.29	-0.86
55	8.72	8.73	-0.01
56	5.14	5.17	-0.03
57	4.29	5.52	-1.24
58	8.00	8.66	-0.66
59	7.00	6.86	0.14
60	4.72	5.25	-0.53
61	6.00	6.34	-0.34
62	8.00	10.77	-2.77
63	3.57	4.71	-1.14
64	8.14	7.20	0.94
65	7.86	6.81	1.05
66	4.57	5.21	-0.64
67	10.14	8.74	1.40
68	3.14	3.15	0.00
69	8.57	8.95	-0.38
70	4.71	5.34	-0.63
71	11.57	10.73	0.84
72	7.86	6.89	0.97
73	2.43	2.44	-0.01
74	10.86	10.74	0.12
75	6.43	6.62	-0.19
76	7.00	8.69	-1.69
77	4.86	6.84	-1.98
78	4.71	5.33	-0.61
79	10.29	11.37	-1.09
80	3.86	5.09	-1.23
81	9.86	11.17	-1.31
82	8.00	8.88	-0.88
83	7.00	6.69	0.31
84	8.29	6.77	1.52
85	9.14	10.36	-1.21
86	7.57	8.66	-1.09
87	8.57	6.98	1.59

88	8.43	9.52	-1.09	95	6.43	6.88	-0.45
89	4.86	6.80	-1.94	96	3.71	5.33	-1.62
90	6.14	6.87	-0.72	97	6.86	6.94	-0.08
91	2.14	2.62	-0.48	98	4.71	5.33	-0.61
92	5.14	4.81	0.34	99	6.00	7.37	-1.37
93	4.71	5.67	-0.96	Total	681.86	662.06	19.80
94	4.29	4.83	-0.54				

APPENDIX D: Validation results of work mode choice model

No.	Observed mode choice	Probability of a mode					Predicted mode choice	Correct prediction
		walk	Bajaj	private car	Service	minibus Taxi		
1	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
2	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
3	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
4	Private car	0.00034	0.00011	0.71741	0.28213	1.26E-06	Private car	Yes
5	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
6	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
7	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
8	Private car	0.01108	0.00984	0.48547	0.4933	0.00032	Service	No
9	Bajaj	0.12479	0.01276	5.72E-11	0.79232	0.07014	Service	No
10	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
11	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
12	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
13	Private car	0.00034	0.00011	0.71741	0.28213	1.26E-06	Private car	Yes
14	Private car	0.01108	0.00984	0.48547	0.4933	0.00032	Service	No
15	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
16	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
17	Walk	0.15809	0.01024	1.13E-05	0.66502	0.16664	Service	No
18	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
19	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
20	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
21	Private car	0.00034	0.00011	0.71741	0.28213	1.26E-06	Private car	Yes
22	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
23	Walk	0.15809	0.01024	1.13E-05	0.66502	0.16664	Service	No
24	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
25	Walk	0.15809	0.01024	1.13E-05	0.66502	0.16664	Service	No
26	Bajaj	0.00066	0.00043	0.01093	0.98799	9.49E-07	Service	No

27	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
28	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
29	Walk	0.15809	0.01024	2.23E-11	0.66502	0.16664	Service	No
30	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
31	Walk	0.15809	0.01024	1.13E-05	0.66502	0.16664	Service	No
32	Service	0.15569	0.01048	1.23E-05	0.67826	0.15556	Service	Yes
33	Minibus Taxi	0.07133	0.00018	9.48E-10	0.70723	0.22126	Service	No
34	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
35	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
36	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
37	Service	0.15569	0.01048	1.23E-05	0.67826	0.15556	Service	Yes
38	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
39	Private car	0.00034	0.00011	0.71741	0.28213	1.26E-06	Private car	Yes
40	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
41	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
42	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
43	Minibus Taxi	0.02154	9.12E-05	1.35E-09	0.00775	0.97062	Minibus Taxi	Yes
44	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
45	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
46	Minibus Taxi	0.07133	0.00018	9.48E-10	0.70723	0.22126	Service	No
47	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
48	Service	0.04382	0.00025	3.47E-06	0.91215	0.04378	Service	Yes
49	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
50	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
51	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
52	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
53	Service	0.04382	0.00025	4.87E-09	0.91215	0.04379	Service	Yes
54	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
55	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes

56	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
57	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
58	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
59	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
60	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
61	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
62	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
63	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
64	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
65	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
66	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
67	Service	0.15569	0.01048	1.23E-05	0.67826	0.15556	Service	Yes
68	Minibus Taxi	0.02093	7.38E-06	1.31E-09	0.03599	0.94308	Minibus Taxi	Yes
69	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
70	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
71	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
72	Minibus Taxi	0.02093	7.38E-06	1.31E-09	0.03599	0.94308	Minibus Taxi	Yes
73	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
74	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
75	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
76	Bajaj	0.00066	0.00043	1.55E-05	0.99889	9.60E-07	Service	No
77	Minibus Taxi	0.02093	7.38E-06	1.31E-09	0.03599	0.94308	Minibus Taxi	Yes
78	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
79	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
80	Minibus Taxi	0.16112	0.00476	1.53E-06	0.33434	0.49977	Minibus Taxi	Yes
81	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
82	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
83	Private car	0.00034	0.00011	0.71741	0.28213	1.26E-06	Private car	Yes
84	Minibus	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus	Yes

	Taxi						Taxi	
85	Service	0.04382	0.00025	3.47E-06	0.91215	0.04378	Service	Yes
86	Minibus Taxi	0.02154	9.12E-05	1.35E-09	0.00775	0.97062	Minibus Taxi	Yes
87	Service	0.15569	0.01048	1.23E-05	0.67826	0.15556	Service	Yes
88	Minibus Taxi	0.02154	9.12E-05	1.35E-09	0.00775	0.97062	Minibus Taxi	Yes
89	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
90	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
91	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
92	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
93	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
94	Service	0.15569	0.01048	1.23E-05	0.67826	0.15556	Service	Yes
95	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
96	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
97	Minibus Taxi	0.02154	9.12E-05	1.35E-09	0.00775	0.97062	Minibus Taxi	Yes
98	Minibus Taxi	0.16112	0.00476	1.53E-06	0.33434	0.49977	Minibus Taxi	Yes
99	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
100	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
101	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
102	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
103	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
104	Minibus Taxi	0.16112	0.00476	3.01E-12	0.33434	0.49978	Minibus Taxi	Yes
105	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
106	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
107	Bajaj	0.00309	0.02404	7.21E-05	0.9728	4.47E-06	Service	No
108	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
109	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
110	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes

111	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
112	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
113	Minibus Taxi	0.02154	9.12E-05	1.35E-09	0.00775	0.97062	Minibus Taxi	Yes
114	Service	0.15569	0.01048	1.23E-05	0.67826	0.15556	Service	Yes
115	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
116	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
117	Minibus Taxi	0.16112	0.00476	3.01E-12	0.33434	0.49978	Minibus Taxi	Yes
118	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
119	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
120	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
121	Minibus Taxi	0.02154	9.12E-05	1.35E-09	0.00775	0.97062	Minibus Taxi	Yes
122	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
123	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
124	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
125	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
126	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
127	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
128	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
129	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
130	Minibus Taxi	0.16112	0.00476	3.01E-12	0.33434	0.49978	Minibus Taxi	Yes
131	Service	0.15569	0.01048	2.43E-11	0.67827	0.15556	Service	Yes
132	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
133	Service	0.15569	0.01048	1.73E-08	0.67827	0.15556	Service	Yes
134	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
135	Minibus Taxi	0.16112	0.00476	2.14E-09	0.33434	0.49978	Minibus Taxi	Yes
136	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
137	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes

138	Service	0.04382	0.00025	3.47E-06	0.91215	0.04378	Service	Yes
139	Minibus Taxi	0.02154	9.12E-05	1.35E-09	0.00775	0.97062	Minibus Taxi	Yes
140	Minibus Taxi	0.02093	7.38E-06	1.84E-12	0.03599	0.94308	Minibus Taxi	Yes
141	Minibus Taxi	0.02154	9.12E-05	1.90E-12	0.00775	0.97062	Minibus Taxi	Yes
142	Minibus Taxi	0.02154	9.12E-05	2.66E-15	0.00775	0.97062	Minibus Taxi	Yes
143	Private car	0.00044	0.00175	0.92192	0.07588	1.62E-06	Private car	Yes
144	Private car	0.00034	0.00011	0.71741	0.28213	1.26E-06	Private car	Yes

APPENDIX E: Validation results of Education mode choice model

NO.		Probability of a mode					Predicted mode choice	Correct prediction
	Observed mode choice	Walk	Bajaj	Private car	Service	Minibus Taxi		
1	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
2	Walk	0.346193	6.52E-06	0.00616	0.552248	0.095392	Service	No
3	Service	0.041758	2.14E-07	3.51E-04	0.080855	0.877035	Minibus Taxi	No
4	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
5	Service	0.394244	0.007324	0.373497	0.224935	1.12E-11	Walk	No
6	Walk	0.209082	0.032289	0.671644	0.086984	5.08E-15	Walk	Yes
7	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
8	Service	0.394244	0.007324	0.373497	0.224935	1.12E-11	Walk	No
9	Private car	0.015497	0.251856	0.729427	0.00322	6.77E-23	Private car	Yes
10	Private car	0.302261	0.156398	0.43634	0.105001	1.30E-16	Private car	Yes
11	Service	0.555202	0.010314	0.117716	0.316769	1.58E-11	Walk	No
12	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
13	Walk	0.209082	0.032289	0.671644	0.086984	5.08E-15	Walk	Yes
14	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
15	Private car	0.302261	0.156398	0.43634	0.105001	1.30E-16	Private car	Yes
16	Walk	0.347857	6.55E-06	0.001385	0.554901	0.095851	Service	No
17	Walk	0.209082	0.032289	0.671644	0.086984	5.08E-15	Walk	Yes
18	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
19	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
20	Service	0.041758	2.14E-07	3.51E-04	0.080855	0.877035	Minibus Taxi	No
21	Walk	0.209082	0.032289	0.671644	0.086984	5.08E-15	Walk	Yes

22	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
23	Bajaj	0.267723	0.185667	0.457581	0.089028	4.34E-17	Walk	No
24	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
25	Walk	0.436799	0.067456	0.314025	0.181721	1.06E-14	Walk	Yes
26	Walk	0.209082	0.032289	0.671644	0.086984	5.08E-15	Walk	Yes
27	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
28	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
29	Service	0.041758	2.14E-07	3.51E-04	0.080855	0.877035	Minibus Taxi	No
30	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
31	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
32	Bajaj	0.035299	0.160028	0.795801	0.008872	1.09E-20	Baja	Yes
33	Minibus	0.068056	4.15E-07	1.42E-04	0.128423	0.80338	Minibus Taxi	Yes
34	Service	0.041758	2.14E-07	3.51E-04	0.080855	0.877035	Minibus Taxi	No
35	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
36	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
37	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
38	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
39	Bajaj	0.035299	0.160028	0.795801	0.008872	1.09E-20	Baja	Yes
40	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
41	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
42	Minibus	0.068056	4.15E-07	1.42E-04	0.128423	0.80338	Minibus Taxi	Yes
43	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
44	Walk	0.5348	9.12E-04	0.028614	0.435674	4.37E-08	Walk	Yes
45	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
46	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
47	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
48	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
49	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
50	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
51	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes

52	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
53	Minibus	0.068056	4.15E-07	1.42E-04	0.128423	0.80338	Minibus Taxi	Yes
54	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
55	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
56	Private car	0.120263	0.062227	0.775733	0.041778	5.17E-17	Private car	Yes
57	Service	0.041758	2.14E-07	3.51E-04	0.080855	0.877035	Minibus Taxi	No
58	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
59	Private car	0.120263	0.062227	0.775733	0.041778	5.17E-17	Private car	Yes
60	Private car	0.302261	0.156398	0.43634	0.105001	1.30E-16	Private car	Yes
61	Minibus	0.068056	4.15E-07	1.42E-04	0.128423	0.80338	Minibus Taxi	Yes
62	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
63	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
64	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
65	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
66	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
67	Bajaj	0.103487	0.071769	0.79033	0.034414	1.68E-17	Walk	Yes
68	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
69	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
70	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
71	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
72	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
73	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
74	Minibus	0.068056	4.15E-07	1.42E-04	0.128423	0.80338	Minibus Taxi	Yes
75	Private car	0.302261	0.156398	0.43634	0.105001	1.30E-16	Private car	Yes
76	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes

77	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
78	Bajaj	0.092333	0.418594	0.465867	0.023206	2.85E-20	Baja	Yes
79	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
80	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
81	Private car	0.120263	0.062227	0.775733	0.041778	5.17E-17	Private car	Yes
82	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
83	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
84	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
85	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
86	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
87	Walk	0.209082	0.032289	0.671644	0.086984	5.08E-15	Walk	Yes
88	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
89	Walk	0.436799	0.067456	0.314025	0.181721	1.06E-14	Walk	Yes
90	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
91	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
92	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
93	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
94	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
95	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
96	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
97	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
98	Minibus	0.068056	4.15E-07	1.42E-04	0.128423	0.80338	Minibus Taxi	Yes
99	Bajaj	0.557252	0.009043	0.109291	0.324413	2.48E-11	Walk	No
100	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
101	Private car	0.120263	0.062227	0.775733	0.041778	5.17E-17	Private car	Yes
102	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus	Yes

							Taxi	
103	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
104	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
105	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
106	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
107	Private car	0.302261	0.156398	0.43634	0.105001	1.30E-16	Private car	Yes
108	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
109	Private car	0.302261	0.156398	0.43634	0.105001	1.30E-16	Private car	Yes
110	Minibus	0.068022	4.15E-07	6.32E-04	0.12836	0.802985	Minibus Taxi	Yes
111	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
112	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
113	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
114	Minibus	0.068056	4.15E-07	1.42E-04	0.128423	0.80338	Minibus Taxi	Yes
115	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
116	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
117	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
118	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
119	Minibus	2.51E-08	1.70E-15	3.89E-12	9.29E-08	1	Minibus Taxi	Yes
120	Service	0.46723	1.45E-04	0.041801	0.490813	1.13E-05	Service	Yes
121	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
122	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
123	Minibus	2.51E-08	1.70E-15	1.74E-11	9.29E-08	1	Minibus Taxi	Yes
124	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes
125	Service	0.482898	1.50E-04	0.009669	0.507271	1.16E-05	Service	Yes