ANALYSIS AND MODELING OF TRAVEL BEHAVIOR FOR A SMALL SIZED INDIAN CITY

A Thesis Submitted In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

In

Civil Engineering

By

Partha Pratim Sarkar

Under the supervision of

Dr. Mallikarjuna C.

DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039, ASSAM, INDIA

MAY 2014

Department of Civil Engineering
Indian Institute of Technology Guwahati
Assam – 781039
India

CERTIFICATE

This is to certify that the work contained in this thesis entitled "Analysis and Modeling of Travel Behavior for A Small Sized Indian City" submitted by Mr. Partha Pratim Sarkar (09610410) to the Indian Institute of Technology Guwahati, for the award of the degree of Doctor of Philosophy in Civil Engineering, has been carried out under my supervision. This work has not been submitted elsewhere for the award of any other degree or diploma.

Dr. Mallikarjuna C.
Associate Professor
Department of Civil Engineering
Indian Institute of Technology Guwahati
Assam -781039, India.

Acknowledgements

I would like to express my deep gratitude and heartfelt thanks to my supervisor Dr. Mallikarjuna C. for his guidance, encouragement and gracious support throghout my reaserch work. His valuable suggestions, effusive co-operation and encouraging interactions were a great driving force for me to carry out this reasearch work. Conducting the research and writing this dissertation would not have been possible without his patience, guidance and tireless devotion to this work.

I deeply express my sincere thanks to the chairman of Doctoral Committee Prof. Anjan Dutta and committee members Dr. Bulu Pradhan and Dr. K.V. Krishna for their sincere advice and suggestions to improve the quality of my work. I would also like to extend my sincere gratitude to Head, Department of Civil Engineering, and the other faculty members for the various supports and encouragement extended to me directly or indirectly to carry out my research.

I would like to convey my thanks to Ms Rupjyoti Bordoloi, Mr. Amit. Mote, Mr. Dibyendu. Pal, and Mr. Syed. Omar, research scholars, Transportation Engineering, IIT Guwahati for their inspiration, interesting discussions and feedback they always provided. I am also thankful to Mr. Samik Saha, Mr. Tony Jamoh and Mr. Rumen Ghosh for their support in preparation of the land use map.

I also like to take this opportunity to thank the Director, NIT Agartala, Head, Department of Civil Engineering and all other faculty members of NIT Agartala, who has supported and encouraged me directly or indirectly in various aspects.

I would like acknowledge with gratitude and appreciation to AICTE for providing financial support for travel data collection as a part of the sponsored project “Analysis of dynamic travel behavior using panel data: A case study of Agartala city”, numbered 8023/BOR/RID/RPS-250.

I express my gratitude to my parents and all other family members for their constant support and encouragement. Last but not the least; I would like to appreciate my better half Ms. Rumela Datta Sarkar and my daughter, Prajakta for their support and sacrifices, without which I would not have reached to the status which I have achieved now.

Partha Pratim Sarkar

Abstract

Analysis and modeling of mode choice behavior is essential in formulating appropriate transport policies. Discrete choice models, based on the random utility theory, are widely used in modeling the individuals' choice behavior. Utilities of various modes depend mainly on the modes' characteristics, and the individual's socioeconomic characteristics. Many researchers have stressed the relevance of appropriate model specification on the parameter estimates. As pointed out in many of the past studies, some of the important aspects in the model specification are the choice set availability to the trip maker, specification of the observed component of the utility, and the specification of the unobserved component of the utility. Recently, several researchers have also pointed out the effects of land use mix on the mode choice behavior of the individuals. Researchers have also pointed out the effects of qualitative aspects of the modes on the mode choice behavior.

Mode choice behavior of the trip makers residing in the smaller Indian cities has not been studied much in the past. These cities are changing rapidly and in most of these cities trip makers are shifting towards the private motorized modes for making the work and shopping related trips. In this work, a study has been carried out to analyze and model the mode choice behavior of the trip makers residing in the city of Agartala, capital of Tripura, a north-east Indian state. Data required for the present study have been collected through a household survey as well as through some field studies. Both the revealed preference (RP) and the stated preference (SP) data have been collected through the household survey. Main objectives of the present study are to understand and overcome the issues related to the formulation and application of discrete choice models for smaller Indian cities; to find out various important determinants of mode choice for work trips; to understand and model the effect of mixed land use on mode choice for work and shopping trips as well as on other travel parameters; to understand the effects of psychological factors and the error in travel time data on choice modeling, using hybrid choice modeling approach.

Quantification of mixed land use observed in the smaller Indian cities is a major hindrance in studying its impact on travel behavior. In this study, various existing mixed land use indices, slightly modified indices, and a few new indices have been used to

quantify the mixed land use. A new index proposed in this study, termed as Area Index, is found to be significant in explaining the modal preferences as well as the trip lengths.

When the choice models were estimated using the RP data, it has been found that in case of many modes the modal attributes such as travel time and travel cost were either insignificant or were having wrong signs. Socioeconomic characteristics such as income, gender, vehicle ownership, years of education, possession of driving license, and the mixed land use parameter 'area index' were found to be significant in explaining the observed component of the preference heterogeneity towards various modes.

From the models estimated with the SP data, it has been observed that the coefficient of comfort (interacted with income and gender) is higher for bus than motorized three wheeler (MThW). This implies that the policy makers have to focus on this aspect of transit so that people can be motivated to use the transit mode. From the elasticity analysis of the parameters, it has been observed that cost parameter of bus is highly elastic indicating higher sensitivity of the bus users towards the travel cost. Same is the case with the users of motorized two wheeler (MTW) i.e., if any other competitive mode (in terms of cost) is available there is a scope for modal shift. Results from mode choice analysts reflect that modal share can be controlled by regulating fares. Further, regulating fare in combination of increase in land use mix can bring desired modal share.

When the models have been estimated with the combined SP-RP data, there was a significant improvement in the model performance when compared to the models estimated with only SP data. Compared to the value of travel times obtained from the models estimated with either RP or SP data, the value of travel times obtained from the MMNL model seems to be realistic. The coefficients obtained from the estimation of hybrid choice models are found to be statistically more significant than the parameters of the MNL model, both estimated with the RP data. When latent variables like comfort of car, flexibility of MTW, and flexibility of MThW were included in the model, the fitness of the model improved considerably, adjusted ρ^2 increased to 0.83 compared to the base MNL model (0.502). In hybrid choice model for travel time measurement error correction, there was a significant improvement in the coefficient of travel time for MThW and MTW due to which the value of travel time has increased from Rs 11.72/hour (from the MNL model estimated using the RP data) to Rs 27.77/hour.

Contents

List of Figures

List of Tables

Abbreviations

SP	Stated Preference
RP	Revealed Preference
MThW	Motorized Three Wheeler
MTW	Motorized Two Wheeler
LOS	Level of Service
GIS	Geographic Information Systems
VMT	Vehicle miles travelled
WTP	Willingness to pay
ICLV	Integrated choice and latent variable
NMT	Non-motorized transport
DI	Dissimilarity Index
MNL	Multinomial Logit Model
MMNL	Mixed Multinomial Logit model.
PWSE	Probability Weighted Sample Enumeration
WTP	Willingness to pay
IID	Independently and Identically Distributed

Chapter 1

Introduction

Travel behavior analysis and modeling is essential in planning and the provision of urban transport infrastructure. In most of the smaller Indian cities, with population less than five hundred thousand, there have been significant changes in the trip making behavior of the individuals. Factors such as growing geographical area, changing socioeconomic and land use patterns, increasing number of motorized personal vehicles, absence of planned public transportation system might be influencing the change in travel behavior observed in these cities. According to a report of Ministry of Urban Development (MoUD), Govt. of India (2008), in most of the small sized Indian cities there will be a significant increase in the modal shares of private transport modes such as cars and motorized two wheeler (MTW) (57% in 2007 to 72% in 2031). Also, the share of the public transport modes (5% in 2007) and non-motorized modes (NMT) (38% in 2007) is decreasing in many of these cities. According to this report, the percentage mode share of public transport for small cities was very low and also predicted it to be much lower in the future (2% in 2031). Government of India has recently started investing in improving the infrastructure, including the public transportation facilities, of many of these cities. At this juncture it is important to understand and model the travel behavior observed in these cities.

Disaggregate travel behavior analysis and modeling techniques are widely being used in understanding the individual's travel behavior. Discrete choice models, derived based on the random utility maximization behavior, are widely used for this purpose. In the context of smaller developing cities of India, there were no significant studies focused on understanding the disaggregate travel behavior. In this study, disaggregate travel behavior of Agartala's residents has been analyzed and modeled. As a part of this effort, a thorough study of existing data collection methodologies and the associated problems, as well as the state-of-art choice modeling approaches has been carried out. Finding out the important determinants of travel behavior in the context of smaller Indian cities being the main objective, this study has also tried to address several issues associated to primary and the secondary data used for choice modeling.

1.1 Need for the present study

Choice set determination is one of the important aspects of discrete choice modeling framework. In the context of smaller Indian cities, determining the choice set available to the trip makers, and getting the data corresponding to the attributes of various alternatives is a major task. In general, choice set available to any trip maker (for work related trips) is determined based on the socioeconomic characteristics such as the number of workers in the household, number of vehicles owned by the household, and the availability of public transport mode based on the details of the public transport network. But in case of the present study area, the category of public transport mode covers variety of modes such as cycle rickshaw, motorized three wheeler (MThW), and various types of buses. The former two modes of this category offer greater flexibility/convenience and the attributes of these modes are different from that of the buses. Also, the services offered by the MThW depend on its driver as there are no fixed regulations on its use. In this scenario, based on the socioeconomic characteristics and the network characteristics, it is difficult to ascertain the choice set of any individual.

Another important aspect of travel behavior observed in smaller Indian cities is the effect of mixed land use on travel behavior. In a majority of the small sized Indian cities, it can be said that the land use is mixed. The land use mix in these cities can be scaled at building level (such as the multi-functional buildings), street level (different buildings, located on a street, with different functionalities), and ward’s level (coexistence of residences, shops, schools, offices, recreational areas, and industries). Also, there exists heterogeneity within a particular component of the mixed land use. For example, within the residential component of the mixed land use people with various socioeconomic backgrounds live together. One other peculiarity is that the extent of land utilized by different land uses varies from very small areas such as an isolated shop of five square meters to bigger areas such a shopping complex of 3000 square meter. Major difficulty in handling such a kind of land use mix lies in developing the Geographic Information Systems (GIS) based land use data base and its quantification. Due to the above difficulties with the data, there were no past studies that analyze the land use based travel determinants in the context of Indian cities.

Also, in the smaller Indian cities road network databases are not available. In this scenario it would be difficult to get the mode related data such as travel time and travel cost of the non-chosen modes. So it is necessary to develop a transportation network and from the network it is necessary to collect the data on the level of service (LOS) variables. Complexities associated to the nature of various kinds of modes used as public transport modes make it more difficult to collect the LOS data from the network database. In this context, it is necessary to understand the complexities in collecting the LOS data from the network database and to understand the implications of such data on the mode choice behavior.

Many researchers have also pointed out that the attitudes and perceptions play an important role in the mode selection process of an individual. In Indian cities, people have certain perceptions about various modes and there seems to be significant variability in the perceptions on various modes by a similar set of trip makers. The presence of flexible as well as convenient public (more precisely Paratransit) transport services necessitates the collection and usage of the attitudinal and perception data. In this context, it is necessary to understand various discrete choice modeling approaches that also consider the attitudinal and perception data as well as to analyze and model their implications on the mode choice behavior.

1.2 Objectives and scope

Considering the need of the present study the following objectives are set for the present study;

1. To analyze the travel data collected from household survey and to address the difficulties related to the network data
2. To analyze the interaction of land use and travel behavior
3. To find important determinants of travel behavior in the context of small sized Indian cities
4. To understand the effect of attitude and perception on mode choice for work trips
5. To examine the sensitivity of different modeling approaches in finding the value placed by the individual on various important modal attributes

Coming to the scope, land use mixing at building level was not considered in this study while preparing the land use database. Internal to external trips are also not considered as the land use database is not available for the destinations of such trips. This study is mainly concentrating on analyzing and modeling the work related trips, going from home to work place. Trips made for the other purposes are considered only for analyzing the effects of land use mix on travel parameters.

1.3 Organization of the Thesis

The thesis has been organized into eight chapters. **Chapter 1** presents general introduction, objectives and scope of the work. A state of the art review on literature has been presented in **Chapter 2**. **Chapter 3** deals with the methodology related to the modeling of travel behavior, and hybrid choice models. **Chapter 4** provides details on data collected for the study which includes description and land use details of the study area, network preparation, questionnaire details, and exploratory analysis of the sample data. **Chapter 5** describes the effects of land use on trip length and mode choice behavior of the individual. In **Chapter 6,** various travel behavior models prepared with Revealed Preference (RP), Stated Preference (SP), and combined SP-RP data and the corresponding model results are discussed. **Chapter 7** deals with understanding the effect of latent variables in the mode choice behavior as well as the effects of measurement error in travel time data obtained from network using hybrid choice model. Finally, **Chapter 8** provides the summary of the work, important conlcusions, and future scope for research in this area.

Chapter 2

Literature Review

Travel demand models are the key elements in planning transportation facilities in the urban areas. Many cities in the developing countries are seeing rapid growth in the recent past and started realizing the importance of planning in the provision of urban transport infrastructure. In the developed countries, discrete choice models are widely being used in travel demand forecasting. But, there are many constraints in using these modeling approaches in the context of smaller Indian cities and requires thorough understanding of these approaches. As a part of the present study, an extensive literature review of mode choice modeling approaches has been carried out and discussed in the following sections. The specific objectives of the present review were to identify various factors affecting the mode choice of an individual and to understand the methodologies adopted by previous researchers for modeling the travel behavior. Literature review is organized into six sections. First section presents the literature related to the choice set determination in the multi-modal choice context. Second section presents the literature that deals with the effects of socio-economic parameters on mode choice behavior. Third section deals with the literature that highlights the effects of land use on travel behavior. Fourth section presents the literature that deals with the role of mode related attributes in mode choice. Fifth section presents the literature on various data used for estimating the choice models. In the final section a brief summary on the literature review is presented.

2.1 Choice-set availability

All the individuals in the sample may not be having same set of alternatives for making a trip. Determination of choice set availability is crucial in mode choice modeling. Assumptions on the choice set, such as all the modes are available to all the individuals may underestimate the probabilities of the available modes. Manski (1977) has used a probabilistic approach that models the choice set generation process as the joint probability of selecting a choice set and selecting an alternative from this choice set. In this method the number of choice sets to be considered was quite high even though the total number of alternatives available for an individual is limited. Swait (2001) has proposed a multivariate Generalized Extreme Value (GEV) model to implicitly model the choice set availability along with the choice probability. Martinez et al. (2009) have implicitly estimated the choice set through elimination of alternatives. Bierlaire et al. (2010) have compared these two techniques and found that the method proposed by Martinez et al. (2009) was a poor approximation of the method proposed by Manski (1977). Srinivasan et al. (2007) have analyzed the effects of choice set availability to the individuals in the context of developing countries. They have considered two different approaches, namely, explicit and implicit choice set availability to the sampled individuals. In case of implicit choice set availability, unavailability of modes was indirectly captured using the variables such as vehicle ownership and trip length. In case of explicit choice set representation, the utility values of unavailable modes (found based on vehicle ownership) were set to negative infinity. They have observed that the explicit choice set representation led to better parameter estimates. This approach can be considered as deterministic explicit rule based choice set generation approach. Enam and Choudhary (2011) have addressed the problem of choice set availability using a probabilistic choice set generation model. Availability of a set of alternatives to any individual has been modeled using the socioeconomic characteristics and the trip related data.

From this limited review, it can be said that though comprehensive approaches proposed by Manski (1977) and Swait (2001) could be used to model the choice set availability, several researchers have adopted much simpler and practical approaches to address the choice set related problems.

2.2 Socioeconomic variables

Socioeconomic characteristics of an individual, along with the mode's characteristics, play a significant role in the mode choice modeling. Several researchers have studied the implications of socioeconomic characteristics on mode choice modeling. Socioeconomic variables are generally added to the modal utilities to capture the observed component of taste and preference heterogeneities. Socioeconomic variables and trip characteristics are useful in accommodating the systematic preference and response heterogeneity within the logit modeling framework (Bhat, 1998; Train 2009). Bhat (1998) has also emphasized that the researchers must strive to attribute most of the preference and response heterogeneity to the systematic variations. Important socioeconomic characteristics include age, gender, education, household size, household composition, and household income.

There were many studies carried out to understand the effect of gender on travel behavior. Women usually tend to make trips less frequently than men (Hanson and Hanson, 1980; Kitamura and Kostynuik, 1986). Hanson and Johnston (1985) and Kostynuik and Cleveland (1978) have reported that the trips made by women were shorter as compared to men, and women tend to use public transit and walk modes and less car travel compared to men. Hanson and Hanson (1980), Vance and Hedel (2007), Srinivasan and Rogers (2005), and Simma and Axhausen (2004) have also obtained similar results. Best and Lanzendorf (2005) have reported that women make more trips for maintenance activities such as shopping and child care. Dissanayake et al. (2012) from a study carried out on Nagoya city reported that females do not prefer rail as a mode of transport. Dissanayake et al. (2012) from a study on four Asian metropolitan cities reported that male travelers in all the study areas have shown greater propensities for using cars and MTW. In Kuala Lumpur, young travelers have shown negative propensities for bus use. As was also found by Bhat (1998), all the above studies indicate that gender can be used to model the systematic component of preference heterogeneity. Gender is also found to be useful in explaining the intrinsic mode preferences and systematic taste heterogeneity (Bhat, 1998). The response to the out-of-vehicle travel time is significantly different across the respondents and this variation can be explained using the gender. Response to the frequency of service was also found to be varying across the respondents and the gender was found to be useful in explaining this variation.

Newbold et al. (2005) have studied the travel behavior of Canadians aged 65 years and above to find if their travel patterns were different from that of the younger Canadians. They concluded that older Canadians undertake fewer trips and their reliance on private automobile is also significant. Giuliano and Dargay (2006) reported that individuals in the age group of 18-34 years tend to travel more than people older than 65 years. Cao et al. (2009) from the data collected from Northern California reported that age was negatively associated with the usage of car and walking, and biking trip frequency. This may reflect the mobility limitations or possibility of safety concerns. According to Dissanayake et al. (2012), when people are older than 45 years, the corresponding parameter in the bus utility function was significantly negative indicating that the travelers in middle and old age may prefer to use taxi or private vehicles. In addition, Chintakayala and Maitra (2010) have found that the age of the trip maker is useful in explaining the systematic response or taste heterogeneity towards travel time.

Sun et al. (1998), in a study based on 1994 Portland activity based travel survey, reported that the household size, income, vehicle ownership were positively associated with the number of household trips and vehicle miles travelled (VMT). Simma and Axhausen (2004), in a study carried out for Upper Austria, reported that the car owners make fewer trips on foot and on public transport and more trips by car. Giuliano and Dargay (2006) have reported that people prefer less costly modes due to the higher transport costs in Great Britain. Per capita income is also lower in Great Britain compared to the United States. Lower income and higher cost of car use combine to promote alternative modes in Britain. Srinivasan et al. (2007) have found that the vehicle ownership was found to be significantly influencing the sensitivity of the decision maker towards travel time. They have modeled this effect by segmenting the decision makers based on the vehicle ownership and the vehicle availability for making the trip. Cao et al. (2009) reported that with the increase in income the utility for car increases and that with the presence of a driving license transit utility is reduced. Dissanayake and Morikawa (2010), from a study on Bangkok metropolitan area, identified household income, job status, and presence of school children in the households as the key considerations leading to household's decisions on vehicle ownership, mode choice, and trip sharing. Walker et al. (2010), from a study carried out on Chengdu, China, reported that car ownership increases the utility of car.

Cherchi and Ortuzar (2003) have suggested the use of socioeconomic variables interacting with the LOS variables. They have analyzed the significance of socioeconomic variables in capturing the taste heterogeneity (within the standard logit modeling framework) when compared to the mixed logit model. They have concluded that the inclusion of socioeconomic variables (in terms of linear in parameters) may help in only improving the overall fit. They have found that the nested logit models with interaction terms perform in a better way compared to mixed logit models though the mixed logit models explain the random taste heterogeneity as well. Pinjari and Bhat (2006) have also mentioned that the taste heterogeneity for LOS variables may partly can be explained through the socioeconomic characteristics of the individuals interacting with the LOS variables.

From the above review, it can be said that the socioeconomic variables have significant effect on travel parameters such as the total distance traveled, trip frequency, vehicle ownership, and mode choice. Since VMT and trip frequency influence the mode choice, it can be said that the socioeconomic characteristics have significant effect on the mode selection. Socioeconomic characteristics can enter the models linearly or through interaction with the LOS variables. Some of the socioeconomic characteristics are useful in explaining the systematic preference heterogeneity when they linearly enter the model in the form of alternative specific variables. Socioeconomic characteristics, when enter the model through interaction with the LOS variables, are quite useful in explaining the systematic taste heterogeneity.

2.3 Effect of land use on travel behavior

Land use pattern has been characterized using parameters such as the population/employment density, land use mix, distance to facilities for shopping, accessibility, street connectivity, centeredness, transit accessibility, and road way design. Land use mix is one of the important measures of land use development pattern and it refers to the diversity of land uses within an area. When diverse land use exists in a given area, it is expected that trips originating from that area may have trip ends in the same area. Land use mix is generally characterized using indices like entropy, dissimilarity index, Gini coefficient, and Herfindahl index.

Most of the earlier studies have defined the mixed land use as the co-existence of different land uses such as office space, retail space, residential area, institutional area, etc., in the same locality. According to Litman (2012), land use mix can occur at various scales, including mixing at a building level (such as ground-floor for retail, with offices and residential places in the above floors). Mixed land use can happen along the streets, within neighborhoods, and mixing at housing types based on different demographic and income classes. This type of mixing is commonly observed in urban areas of developing countries like India. In a mixed land use area, it is assumed that people are less likely to use car and more likely to walk to their destinations. In most of the past studies, researchers have attributed several transportation-related benefits to the mixed land use. As referred in Cervero and Kokelman (1997), Cervero (1989a) has first measured the land use mix using a factor called the 'entropy'.

$$\text{Entropy} = \sum_j P_j \times \frac{\ln(P_j)}{\ln(J)} \tag{2.1}$$

where, P_j is the proportion of developed land corresponding to j^{th} land use category found in the tract being analyzed. J is the total number of land uses considered in the proposed study.

Since the entropy value is normalized with respect to natural logarithm of number of distinct uses considered, the entropy value varies between zero and one. The value nearer to one means land use is more diverse.

Cervero (1989b) pointed out that balancing the jobs and residential locations may bring reduction in commuting times. Cervero (1991), in a study carried out on six urban activity centers in United States, evaluated the effect of land use parameters on the travel behavior of the employees of the suburban activity centers. He found that the transit share was greater in mixed land use and multi-story buildings, and also the average vehicle occupancy was higher in mixed use buildings. Land use mix was not significant in explaining the percentage of work trips made using private automobile vehicles.

Frank and Pivo (1994) have used entropy index to describe the evenness of the distribution of built square footage, among seven land use categories. It was found that the increase in land use mix increases the use of transit and walk mode. They have also suggested using land use mix parameters corresponding to the origin and the destination tracts to get better result rather considering only for the origin. Cervero (1995) concluded that the residential density has a greater influence on commuting mode choice (excluding walking/cycling) compared to the land use mix. He has found that the car is the main commuting mode for people who are living in the areas of low residential density than those living in areas of high residential density. Messenger and Ewing (1996) have used land use mix, overall density (residential + employment), job housing balance, commercial jobs at zonal level, in modeling the transit share for work trips. They found that except the overall density, other measures have no impact on transit share for work trips. Cervero and Radisch (1996) concluded that mix land use was a better predictor for mode choice for non-work trips. Sun et al. (1998) reported that the land use mix and population density do not affect the number of household trips but both of these variables have significant impact on VMT.

Cervero and kockelman (1997), in a study carried out on San Francisco, reported that land use parameters such as accessibility, land use mix, and land use balance have significant impact on travel parameters such as VMT, automobile ownership, and mode choice. They have proposed a new index called dissimilarity index for quantifying the land use mix and found it to be better compared to the entropy index.

$$\text{Dissimilarity Index} = \sum_k \frac{1}{K} \sum_i^8 \frac{X_{ik}}{8} \tag{2.2}$$

where,

K = number of actively developed cells in a census tract.

X_{ik} = 1 if central cell's land use differs from that of a neighboring cell and 0 otherwise.

Compared to the census tract used as the analysis unit for entropy measurements, finer cell of 100m x 100m size was used in computing the dissimilarity index. This helps in getting more information about the type or intensity of mixing. Bhat and Gossen (2004) have measured the land use mix using an empirical formula given below.

$$\text{Land use mix diversity} = 1 - \frac{\left|\frac{r}{T}-\frac{1}{3}\right| + \left|\frac{c}{T}-\frac{1}{3}\right| + \left|\frac{o}{T}-\frac{1}{3}\right|}{\left(\frac{4}{3}\right)} \tag{2.3}$$

where r is the residential area in acres, c is commercial/industrial area in acres, and o is the area of other land uses in acres. Land use diversity is 0 if the land use consists of only one category and the value is one when the land use is equally divided among the three land uses. Their study concluded that the mixed land use variable was not significant in explaining the propensity to participate in recreational trips and also have suggested for finer resolution in computing land use diversity.

Limtanakool et al. (2006) have concluded that the land use attributes like population density, land use diversity (in terms of entropy), distribution of employment, services and population across urban space, are important in explaining the variation in mode choice for medium and longer distance travel. Their analysis has confirmed that the spatial configuration of land use and transport infrastructure has a significant impact on mode choice. Furthermore, they found that most of the land use variables, included in the expanded models along with the travel time, have stronger effects on mode choice. Their findings suggested that the impact of land use may be underestimated if travel times are not simultaneously considered. Cervero and Duncan (2003) have found that the land use mix in and around the person's neighborhood, characterized using 'mixed use entropy', was the strongest predictor for walking trips. Mixed use entropy was nothing but the entropy measured within one mile radius of the origin of the trip. They have analyzed only the trips that involve less luggage or goods. They also found that the land use pattern, characterized using density, diversity, and design corresponding to the origin of the trip, significantly influence the bicycle trips. Rajamani et al. (2003) have found that the land use mix, characterized using 'land use mix diversity' (Bhat and Gossen, 2004), significantly influence the walk trips made for non-work purpose. They also found that the street design, measured in terms of cul-de-sac streets, results in more walking trips made for shopping. Chapman and Frank (2007) have measured the land use mix in terms of entropy index and found that the mixed land use decreases the VMT.

Vance and Hedel (2007), from a study carried out in Germany, have reported that the urban form variables such as the commercial density, road density, and walking time to public transit were significant in modeling the car usage and distance travelled. Zegras (2007) reported that the land devoted to commercial and service purposes, in the zone of trip origin, increases the likelihood of making home based non-work, non-school walk trips, while percentage of vacant land has the opposite effect. Pinjari et al. (2007) have characterized the built environment in terms of population density, employment density, and the block density. Any change in these characteristics of the built environment significantly affects the non-motorized commuting trips. Cao et al. (2009), in a study on eight neighborhoods in north California, used accessibility indicators to the number of establishments within specified distance to analyze the travel parameters. They found that the number of business type establishments within 400m radius of residence was negatively correlated with car travel and positively correlated with non-motorized trip frequency. They have also found that the number of business type establishments within 800m radius have positive effect on transit trip frequency. From this study, it can be inferred that mixing the land use tends to discourage the car trips and facilitates the use of transit and non-motorized modes. Ewing and Cervero (2010), using the meta analysis of several studies, reported that the land use mix measured in terms of distance to store has high elasticity value compared to the other diversity measures when used for analyzing walk trips. Among the other land use variables, intersection density has higher elasticity value.

Tracy et al. (2011) have used three different land use characteristics namely density, diversity, and design, and tried to find their effect on non-motorized mode choice, transit mode choice, personal vehicle choice, home based vehicle hours travelled, and home based VMT. They used dissimilarity index to measure the land use mix and found it to be more significant in non-motorized mode and transit choices. Tsai et al. (2012), based on a study conducted on Sydney, reported that the land use density, road length, and accessibility have positive impacts on public transport mode choice.

Matt et al. (2005) have suggested some general strategies which may reduce travel by car. One of the strategies was to reduce the trip distance. They have also suggested that reducing total number of trips may result in reduced car travel. From the literature review, it is clear that the land use pattern has significant impact on trip frequency and the trip distance. Reduced travel distance may increase the attractiveness of walking and cycling in place of using the car. Some studies have reported the effect of mixed land use is insignificant in explaining the personal vehicle mode choice (Cervero 1991; Frank and Pivo, 1994; Cervero and Kockelman, 1997; Tsai et al., 2012). On the other hand, some of the studies (Vance and Hedel, 2007; Cervero, 1995) have concluded that the residential density, compared to land use mix, has good impact on personal vehicle as a mode of travel, and the distance travelled. Limtanakool (2006) has emphasized the use of travel time as an explanatory variable while studying the impact of mixed land use on travel behavior.

Another important observation from the literature is the use of different indices in quantifying the land use mix. Hess et al. (2001) have pointed out specific drawbacks regarding the Entropy and dissimilarity index. They pointed out that land use parameters attributed to large tracts cannot capture the variations in land use mix that may affect travel as well as the spatial intersection of different land uses. They have recommended a new measure for land use mix in terms of edge-contrast. Bhat and Gossen (2004) have suggested for using finer resolution while measuring the land use mix.

From the above review it can be observed that in most of the past studies Dissimilarity Index was computed for census tract, using a cell of 100 m X 100 m, and entropy was also calculated for a census tract. In the context of smaller Indian cities, smaller size land uses associated to the shopping and commercial usages and their heterogeneity within a census tract (called as municipal wards), direct application of the existing land use parameters in explaining the travel behavior may not yield promising results.

2.4 Mode related attributes

Travel time and travel cost are the most commonly used mode related attributes in mode choice modeling. Many other subjective LOS attributes such as the comfort, reliability, frequency, safety, and convenience are also used in mode choice modeling (Hensher et al., 2005). Many of the subjective attributes cannot be observed and even the travel time and travel cost data are estimated from the network database. Some researchers have imputed the unobserved travel times based on either limited number of observed data or network skims (Hensher et al., 2005; Washington et al. 2014). McFadden (2000) reported that one of the most common issues in disaggregate behavioral forecasting is the accurate measurement of travel time and cost components at individual level. Steimetz and Brownstone (2005) have used a method to correct the measurement error that employs multiple imputations. This method can be used when one has accurate data for a sub-sample of the observations. Enam and Choudhary (2011) have estimated the travel time corresponding to the non-chosen modes based on the stated (by the decision makers) travel times corresponding to the chosen mode. They have formulated a linear regression model between the stated travel time of the chosen mode and the network O-D travel time. The relationships have been developed for various predominantly used modes. In this study, they have not estimated the travel costs corresponding to the non-chosen modes.

Some researchers have considered travel time to be a latent variable, and the latent travel time has been modeled using the travel time estimated from the network and the socioeconomic characteristics of the individual (Ben-Akiva et al., 2002; Bolduc et al., 2005; Walker et al. 2010). Walker et al. (2010) suggested using latent variable approach to correct for measurement errors introduced in travel demand modeling. They treated true travel time as latent variable and the model was estimated using hybrid choice modeling framework. They concluded that the estimates with measurement errors in travel time tend to significantly under estimate the value of travel time.

In addition, decision maker's sensitivity towards the LOS variables was found to be heterogeneous. Some of this heterogeneity can be explained using the socioeconomic characteristics but its random component was also found to be significant and needs to be considered to get meaningful estimates for the willingness to pay measures (Bhat, 2000; Phanikumar and Maitra, 2007; Das et. al., 2009; Chatterjee, 2011). In the context of developing countries, Srinivasan et al. (2007) reported significant difference in sensitivity to travel time, cost, and vehicle availability observed across different user segments. They modeled the sensitivity to travel time and cost for user groups segmented based on captivity status and work distance. Sensitivity to travel time and cost also varies across modes, with greater sensitivity to non-motorized and public transport modes. Effects of random heterogeneity in the response towards the LOS variables can be captured in the mixed logit framework.

Willingness to pay for travel time saving or value of travel time is of prime importance for any mode choice model. This can be estimated when the responses to travel time and travel cost are properly estimated. Value of travel time saving (VTTS) is an important willingness-to-pay indicator with respect to the cost-benefit analysis in the context of planning new transport system or for pricing (Hess et al., 2005). The ratio of estimated travel time and travel cost parameters provides an estimate of value of travel time. In case of models with fixed coefficient for travel time, researcher expects a negative sign for the travel time coefficient. Models producing positive travel time coefficient are rejected on the ground of model misspecification. In random parameter models or mixed logit models (used to capture the heterogeneity in the response) various probability distributions have been used to model the density of the response measure. The use of unbounded distributions like normal distribution may lead to positive and negative travel time coefficients. This means some share of population prefers higher travel time which is inappropriate. The choice of distribution for coefficient plays an important role in the modeling process. Normal distributions are relatively easy to implement in both classical and Bayesian methods and are widely used as the distribution for random parameter.

Lognormal distribution prevents the possibility of unexpected parameter sign since the distribution contains positive values only. However lognormal distribution has some disadvantages also. Many researchers have found that during simulation, the log normal distribution has problems such as slow convergence or even no convergence. (Hess et. al., 2005; Algers et al.,1998). Further, long right tail can lead to some unrealistically high standard deviations (Sikka 2012). Hess et al. (2005) suggested the use of bounded distributions such as triangular or Johnson's S_B where bounds are estimated from the data. Hensher et al. (2005) have recommended the use of constrained triangular distribution to get non-negative WTP measure. Sikka (2012) has pointed out that the constrained triangular distribution heavily depends on the mean. Besides the attributes of interest for which the WTP needs to be calculated, the cost coefficient can also be randomly distributed. But the resulting distribution of the WTP measure may have undefined population moments (Daly et. al., 2012). Assuming a fixed parameter for cost coefficient and letting the coefficients of interests to vary randomly for which the WTP measures needs to be calculated lead to satisfactory estimates (Daly et. al., 2012).

Few researchers have estimated the value of travel time in the context of developing countries. Phanikumar and Maitra (2006) have found out the value of in-vehicle travel time to be between Rs 4.4/hr to Rs 4.87/hr for transit users in Kolkata. The variations in the willingness-to-pay (WTP) have resulted from different specifications of the choice models. The upper limit for the WTP has resulted from the mixed logit specification and the lower limit has resulted from standard logit specification. In the mixed logit framework, the random component of the response heterogeneity has been modeled using constrained triangular distribution. The systematic heterogeneity in the response for in-vehicle travel time has been explained using household income. Chintakayala and Maitra (2010), from a study on Kolkata city, used stated preference data for modeling generalized cost of travel for taxi and found that the value of travel time for taxi to be Rs 66.0/hr. In this case, they have used mixed logit model to estimate the value of travel time.

Apart from the travel time, travel cost, subjective, and attitude related parameters are also significant and sufficiently improve the models (Ben-Akiva et al. 2002; Yanez et al. 2010; Paix et al. 2011). Collection of the data on subjective parameters has been a major constraint in utilizing these data in choice models. Many researchers have collected the data on subjective factors in the form of ranking on a Likert scale (Srinivasan et. al., 2007; Ahmed and Datta, 2006). Ahmed and Datta (2006) have collected the data in the form of rankings and found that the level of service variables such as comfort, safety, convenience, and reliability significantly influence the utility of paratransit mode.

Subjective and attitudinal variables were also analyzed using the structural equation models considering these variables to be latent. Latent variables, unlike observable variables, are not directly observed but rather inferred from variables that are observed. One significant advantage of latent variable approach is that a large number of indicator variables can be aggregated together to visualize an underlying concept. For including the latent variables in choice models sequential and simultaneous approaches are available. Ben-Akiva et al. (2002) have found that the parameter estimation by the later approach i.e., the simultaneous approach has been more efficient. Ben-Akiva et al. (2002) presented a general methodology and framework for incorporating the attitude and perception data in the choice model. Johanssen et al. (2006) studied the effect of attitude and personality traits on mode choice by constructing sequential integrated choice and latent variable modeling approach. They concluded that the flexibility, comfort, pro-environment friendly, influence individual's mode choice. Kitrinou et al. (2009) have developed integrated choice and latent variable (ICLV) models for residential relocation decision. They have found that the addition of latent variables significantly influence the model outcome and considerably improve the goodness-of-fit of the model. Raveau et al. (2010) have reported that empirical performance of sequential and simultaneous approaches was not statistically different. Paix et al. (2011) have used hybrid choice model to jointly estimate the choice and latent variable for finding relationship between urban environment and travel behavior. They have concluded that the hybrid choice models show a major improvement in the goodness-of-fit of the model as compared to the classical discrete choice models. From a study on Chengdu, Walker et al. (2010) considered the travel time to be latent and estimated a hybrid choice model. They have modeled the latent travel time in terms of the network based travel times. They found that the use of latent travel time had significant influence on the value of travel time.

From the above review, it can be said that the response towards the LOS variables such as travel time and travel cost is heterogeneous and part of this heterogeneity can be explained using the socioeconomic characteristics of the individual and the trip's characteristics. Mixed logit modeling framework can be used in estimating the random component of the heterogeneity. Collection of data on travel time and travel cost is prone to errors and there are different ways to correct these errors. Considering the travel time to be latent and estimating the resulting hybrid choice model may yield better estimates for the value of travel time. It can also be observed that the subjective and attitudinal factors also play major role in mode choice modeling and hybrid choice and latent variable model needs to be used for getting better estimates of the corresponding coefficients.

2.5 Data for model estimation

RP data has high reliability and face validity as these are the real choices made by the individuals. Hence, RP data are particularly well suited for short term forecasting of small departures from current state of affairs (Vonkova 2011). RP data collected on the attributes and their levels are likely to be ill conditioned (i.e., being largely invariant, problem of multi-collinearity) and the parameter estimates (other than constant terms) estimated using these data are likely to be biased (Hensher et. al., 2005). On the other hand, the SP data set attributes are likely to be of good condition and hence the associated parameter estimates from models estimated using the SP data are likely to unbiased (Hensher et. al., 2005). SP data has some specific advantages over RP data such as high variability, provision for analyzing new proposed alternatives, new attributes, and level of attributes.

Strengths of both the data can be exploited if the models are estimated using the combined SP-RP data. The process of pooling RP and SP data and estimating a model from the pooled data is called data enrichment (Louviere et. al., 2000). This process was originally proposed by Morikawa (1989) whose motivation was to use SP data to identify parameters that RP data cannot, thereby improving the model's efficiency. Later on many researchers (Ben-Akiva and Morikawa, 1990;, Hensher and Bradley, 1993; Bradley and Daly, 1994; Swait et al., 1994; Adamowicz et al., 1997; Brownstone et al., 1996; Bhat and Casteler, 2002; Hensher et al., 2008) have used the combined SP-RP data for estimating their models.

Combining data sources requires consideration of differences in variance (scale) associated with the unobserved influences on choice. Accommodating the scale difference between data sources has been carried out by assuming the scale homogeneity within each data source. Sequential estimation (Morikawa, 1989; Swait et al. 1994) or joint estimation methods (Morikawa, 1989; Bradley and Daly, 1991; Hensher and Bradley, 1993) can be used to find the differences in scale. Bhat and Castelar (2002) have reported that the joint SP-RP estimation can induce state dependency, defined as influence of actual (revealed) choice on the stated choice of the individual. They have examined the state dependence, defined as $\varphi_q(1-\delta_{qt,RP})$, where $\delta_{qt,RP}$ =1 for RP observation, otherwise 0, and φ is the parameter estimate of state dependence which can be fixed or random. State dependence effect may be positive or negative. A positive effect may result in due to habit persistence or inertia to explore another alternative. A negative effect could be the result of variety seeking or result of latent frustration with inconvenience associated to the currently used alternative. Bhat and Casteler (2002) and Hensher et al. (2008) have found that the model estimated with unobserved heterogeneity and state dependence found to improve the model parameters significantly.

From the above literature review, it can be seen that the scale heterogeneity must be considered while combining the SP-RP data and the combined data results in a better model. State dependency also needs to be considered to know the effect of RP choice on the SP choices of the same individual. Consideration of unobserved heterogeneity also sufficiently improves the model parameters estimated with the combined SP-RP data.

2.6 Summary of Literature review

Socioeconomic characteristics play significant role in capturing the systematic components of the response (taste) and preference heterogeneity. In the cities of developing countries like India it is difficult to collect the segmented sample so that the choice probabilities can be estimated for different segments, separately. In Indian scenario it can be hypothesized that the tastes vary across the individuals depending on the socioeconomic characteristics. Very few researchers have attempted to understand this influence of socioeconomic characteristics on travel behavior. Many researchers have also found that the mixed land use has significant impact on mode choice directly or indirectly (through VMT and trip frequency). Given the mixed land use in Indian cities significant impact of land use on travel behavior can be expected but the traditional indices may not be useful in quantifying the land use mix. LOS variables associated to various modes govern the mode choice behavior of the individuals. Researchers have resorted to different ways to collect/estimate the data. Some researchers have used hybrid mode choice and latent variable models.

But most of these data cannot be collected from the individuals or estimated from the network. LOS data, specifically the travel time and travel cost data on non-chosen modes, cannot be collected from the individuals since these data are to be recollected (if already that individual had used the mode in past) or to be guessed. Network (both road and transit/paratransit networks) complexities makes it to difficult to guess the LOS data. The same complexities make it to difficult to impute the LOS variable data from the digitized road and transit/paratransit networks. Most of the cases the digitized road and transit/paratransit networks are not available for many of the cities. Another problem associated with many of the Indian cities is the flexibility (no specific routes or schedules) of transit/paratransit networks and this flexibility makes it difficult to develop the digital database.

Chapter 3

Theory of Choice modeling

Evolution of discrete choice models, in terms of specification and estimation, has been thoroughly reviewed in the chapter on Literature Review. Some of the important aspects of formulation and estimation of choice models, mostly adopted from various previous studies, are discussed in this chapter. Methodology adopted in identifying the choice set available to the sampled individuals is briefly discussed in section 3.1. Basic formulation of standard logit model is briefly described in section 3.2. Basics of mixed logit models are briefly discussed in this section. Model specification and estimation of combined SP-RP model is also discussed in this section. Hybrid choice model (Integrated Choice and Latent Variable model) formulation and estimation is discussed in section 3.3.

3.1 Identification of choice set for an individual

Determination of choice set for an individual i.e., modes available to an individual for making a trip, may not be easy to determine in the context of smaller Indian cities. Difficulties associated to the determination of modes available to an individual have been discussed in Chapter 4. To avoid any inconsistency, choice set was determined based on predefined logical rules. Trips less than 100m were not considered in this study. MThW and bus were considered to be available to all the trip makers for work trips. Though buses run on some of the major routes, it has been assumed that the trip makers utilize access modes to reach the bus routes. Walk and cycle-rickshaw were considered to be non-available when trip length is greater than 3 kilometer. This has been decided based on the trip length data obtained in the household survey. Cycle rickshaw was considered in the choice set only for the individuals living within three km radius of the city center. Vehicle ownership was the sole criterion for including bicycle, MTW, and car in the choice set. In the households, where only one vehicle is available and if it was chosen for the work trip by any member of the household, then that mode is considered to be non-available for the other members. This is not a constraint for this study as most of the households have one worker only.

There can be many other criteria based on which the availability of public transport (run by private people) and intermediate public transport, for making work trips, can be determined. For example transport services run by private people are mostly governed by the demand and are available only during the peak periods. In this study, this criterion is not considered in determining the choice set availability. Choice set availability was also revealed by the individual in the RP survey and it was found that there is a difference between the revealed choice set and the rule based choice set.

3.2 Modeling of the travel behavior

Discrete choice models describe, explain, and predict the choices between two or more alternatives, such as choosing between the modes of transport or choosing from various brands of goods. Discrete choice models statistically relate the choice made by each person to the attributes of the person and the attributes of the alternatives available to the person. In the standard logit modeling framework, the probability of an individual choosing an alternative is given as below;

$$\Pr(\mathrm{i}) = \frac{e^{V_i}}{\sum_{j=1}^{J} e^{V_j}} \tag{3.1}$$

where, Pr(i) is the probability of decision maker choosing an alternative i; V_j is the systematic component of the utility of alternative j. MNL model has three main drawbacks, namely, inability to explain the unobserved component of the taste variation, unrealistic proportional substitution across alternatives, and the inability to explain the correlation of the unobserved factors which may result in case of panel data (Train, 2009).

Sample collected in the present study is not stratified based on the socioeconomic characteristics. Given the wide variations in the socioeconomic characteristics of the sampled individuals, it is expected to have unobserved taste heterogeneity among the sampled individuals. Since the choice set contains many similar alternatives, proportional substitution pattern may also not hold well. Also, in the present study, the choice model has been estimated with both the SP and RP data. In case of SP data, there can be significant correlations among the unobserved components of the modal utilities corresponding to various choice scenarios.

These drawbacks can be handled by either GEV class models such as the nested logit model or more flexible mixed logit models. Mixed logit model generalises a standard logit model by allowing a parameter (i.e., coefficient) associated with each observed attribute to vary randomly across the individuals. In case of random parameter logit model, the utility derived by person n from alternative j is specified as;

$$U_{nj} = \beta_n X_{nj} + \varepsilon_{nj} \tag{3.2}$$

where, X_{nj} are the observable attributes of the alternative and decision maker; β_n is a vector of coefficients of variables for person n representating the person's taste; ε_{nj} is the random term that is independently and identically distributed (IID) extreme value. Here the coefficients vary across the decision makers in the population with density f(β). The specification is same as for standard logit model except that β varies over decision makers rather than being fixed. The probability conditional on β is given by;

$$L_{ni}\ (\beta_n) = \frac{e^{\beta_n x_{ni}}}{\sum_j e^{\beta_n' x_{nj}}} \tag{3.3}$$

where, L_{ni} (β_n) is the logit probability evaluated at parameter β_n.

As the researcher does not observe β_n and, therefore, cannot condition on β. The unconditional choice probability over all values of β_n is expressed as

$$P_{ni} = \int \left(\frac{e^{\beta x_{ni})}}{\sum_j e^{\beta x_{nj}}} \right) f\ (\beta|\theta)\ \delta\beta \tag{3.4}$$

where, θ refers collectively to the parameters of chosen distribution (such as mean and variance of β).

A mixed logit model can also be used without a random-cofficient interpretation, as simply representing the error components that create correlations among the unobserved utilities for different alternatives (Train, 2009). The utility in this case is specified as;

$$U_{nj} = \alpha' X_{nj} + \mu_n' Z_{nj} + \varepsilon_{nj} \tag{3.5}$$

where,

X_{nj} and Z_{nj} are the vectors of observable variables related to alternative j,

α' is a vector of fixed coefficients,

μ_n' is a vector of random terms with zero mean, and,

ε_{nj} is error component which is iid extreme value.

The terms in Z_{nj} are error components that, along with ε_{nj} forms the stochastic portion of utility. The unobserved component of the utility is $\mu_n' Z_{nj} + \varepsilon_{nj}$, which can be correlated over alternatives depending upon the specification of Z_{nj}. In case of standard logit model Z_{nj} is zero and there is no correlation in the utility over the alternatives.

The unconditional choice probability as given in equation 3.7 cannot be calculated exactly as the integral is not in closed form. Probabilities are approximated through simulation for a given value of θ. For simulation, a value of β from $f(\beta|\theta)$ is drawn and logit formula L_{ni} (β_n) is calculated. These steps are repeated and the results are averaged. This average is the simulated probability and is given as (Train, 2009);

$$\hat{P}_{ni} = \frac{1}{R}\sum_{r=1}^{R} L_{ni}(\beta^r) \tag{3.6}$$

where, R is the number of draws. $\hat{P}_{ni}$ is the unbaised estimator of P_{ni}. Its variance decreases as R increases and is positive so that ln $\hat{P}_{ni}$ is defined.The simulated probalities are inserted into log likelihood function to get the simulated log likelihood.

3.2.1 Handling of SP data in mixed logit framework

Data from the SP survey contains repeated choice data from a single individual. The number of choice situations can vary over people, and choice sets can vary over people and choice situations. Mixed logit model with panel specification can handle the correlations among the unobserved error components, mainly associated with the SP data. In case of repeated choice data, as in case of panel data, utility from alternative j in choice situation t by an individual n is given as (Revelt and Train, 1998);

$$U_{njt} = \beta_n x_{njt} + \varepsilon_{njt} \tag{3.7}$$

where , ε_{njt} is iid extreme value over individual, alternative, and time.

Conditional on β, the probability that the person making the sequence of choices is the product of logit formulas given below;

$$L_{ni}(\beta) = \prod_{t=1}^{T}\left[\frac{e^{\beta_n x_{ni}}}{\sum_j e^{\beta_n' x_{nj}}}\right] \tag{3.8}$$

The unconditional probability is the integral of this product over all values of β.

$$P_{ni} = \int L_{ni}(\beta) f(\beta)\, \delta\beta. \tag{3.9}$$

The simulated probabilities and the simulated likelihood functions can be obtained following the procedure discussed in the previous sub-section.

3.2.2 Handling of combined SP-RP data

When two sources of data, SP and RP, are used to estimate the choice models there will be issues with the scales of the data. Combining two data sources involves imposing the restriction that the common attributes (atleast one) have the same parameters in both the data sources, i.e., $\beta^{RP} = \beta^{SP} = \beta$. Data enrichment includes pooling of two choice data sources under the restriction that common parameters are equal while controlling for the scale factors. Scales of the utilities, corresponding to the SP and the RP data would be different as the unobserved variance may not be same in the utilities corresponding to these data. Since it is difficult to identify both the scale parameters, it is conventional to fix the scale of RP data set ($\lambda^{RP} = 1$) and estimate the λ^{SP}, which represents the relative scale. The corresponding choice model may be written as follows (Louviere et al. 2000);

$$P_i^{RP} = \frac{exp\,[\lambda^{RP}(\alpha_i{}^{RP} + \beta^{RP}\, X_i^{RP} + \omega Z_i)]}{\sum_{j \in C^{RP}} exp[\lambda^{RP}\left(\alpha_j{}^{RP} + \beta^{RP}\, X_j^{RP} + \omega Z_j\right)]} \quad \forall\, i \in C^{RP} \tag{3.10}$$

$$P_i^{SP} = \frac{\exp[\lambda^{SP}(\alpha_i{}^{SP} + \beta^{SP}\, X_i^{SP} + \partial Z_i)]}{\sum_{j \in C^{RP}} exp\left[\lambda^{SP}\left(\alpha_j{}^{SP} + \beta^{SP}\, X_j^{SP} + \partial Z_j\right)\right]} \quad \forall\, i \in C^{SP} \tag{3.11}$$

where, i is an alternative in choice sets C^{SP} or C^{RP}, α's are the data source-specific alternative specific constants (ASCs) and β^{RP} and β^{SP} are the utility parameters for attributes common to both the data sets, and δ and ω are utility parameters for unique attributes in each data set. λ^{RP} and λ^{SP} are the scale factors.

So the deterministic utility portion for a mode for both the RP and SP parts can be written as

$$V_{in} = \alpha_i{}^{RP} + \beta^{RP}\, X_i^{RP} + \omega Z_i + \lambda^{SP}\ (\alpha_i^{SP} + \beta^{SP} X_i^{SP} + \delta W_i\ ) \tag{3.12}$$

3.3 Hybrid choice model

In case of Hybrid choice modeling, this model consists of two parts, a discrete choice model and a latent variable model. Each part consists of one or more measurement equations. The integrated choice and latent structure explicitly models the latent variables that influence the choice process. To specify both discrete choice and the latent variables, two types of equations are required: a) Measurement equation which links the latent variable to the indicator variables and b) a structural equation that links latent variable to the explanatory variables. The structural equation for the latent variable model (for Figure 3.1) can be written as;

$$(LV)\eta_{ilq} = \sum_r \lambda_{ilr}\, \xi_{irq} + \zeta_{ilq} \tag{3.13}$$

Measurement equations can be written as;

$$I_{ipq} = \sum_l \gamma_{ilr}\, \eta_{ilq} + \delta_{ipq} \tag{3.14}$$

where, ξ represents explanatory variables, ζ and δ is the error term normally distributed with zero mean and standard deviation σ, index i refers to an alternative, l to latent variable, p to indicator and r to explanatory variable. The equation 3.13 results into one equation for each latent variable and equation 3.14 results in one equation for each indicator (i.e., each survey question). These measurement equations usually contain only the latent variables on the right hand side. λ and γ are the coefficients to be estimated in the latent variable model.

Considering the error components (ε, ζ , δ) are independent and the joint probability of the observable variables y and I, conditional on the exogenous variable ξ is given as;

$$f(\mathrm{y}, \mathrm{I}|\xi; \lambda, \beta, \gamma \textstyle\sum_\zeta \sum_\varepsilon \sum_\delta) = \int_\eta P(y|\,\xi, \eta; \beta, \textstyle\sum_\varepsilon) \;\; \int_3 (I|\,\xi, \eta; \lambda, \textstyle\sum_\zeta) \;\; \int_1 (\eta|\,\xi, \eta; \gamma, \textstyle\sum_\delta) d\eta \tag{3.15}$$

The first term of the integrand corresponds to the choice model, the second term corresponds to the measurement equation from the latent variable model, and the third term corresponds to the structural equation of the latent variable model.

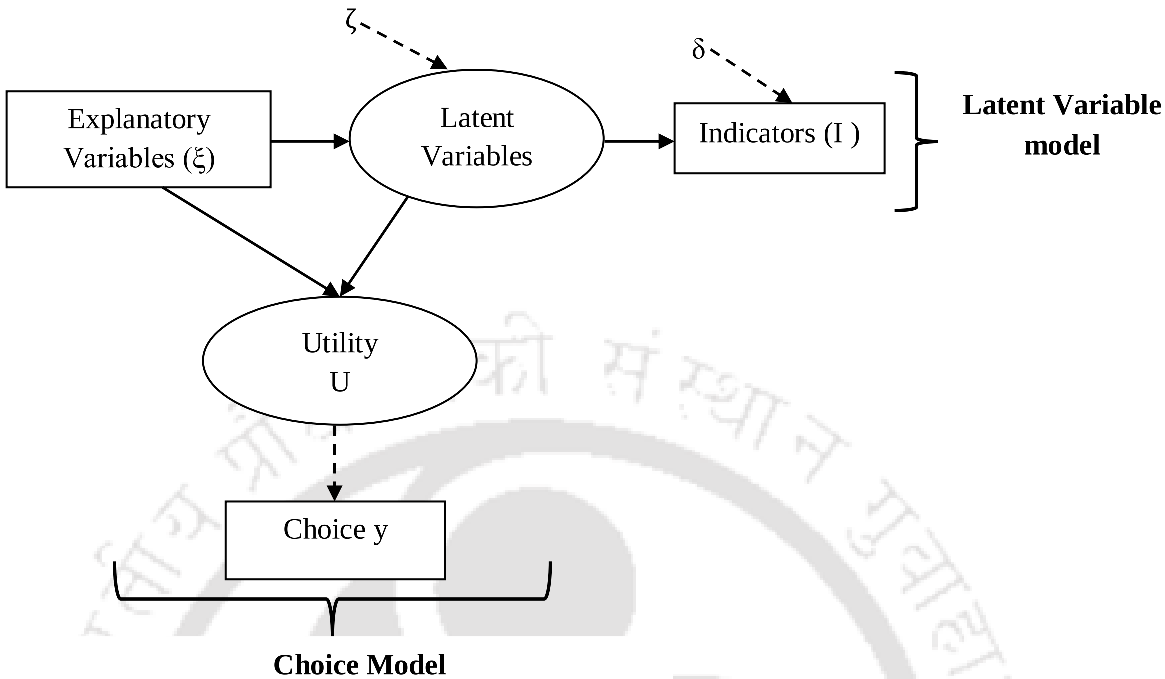


Figure 3.1: Integrated Choice and Latent Variable Model (Adapted from Ben-Akiva et al., 2002)

3.4 Summary

Various theoretical concepts, mainly related to the choice models developed in the present study, are discussed in the previous sections. The concepts covered in this chapter were selected based on the hypotheses related to the travel behavior observed in the case of smaller Indian cities.

Chapter 4

Data Collection

Methodology adopted in collecting various data including the household travel data, land use data, and various important statistics of the household data are presented in this chapter. Household travel data have been collected using RP and SP surveys. Study area description, RP questionnaire and the design of SP survey are explained in detail. Since the database on road network is not available, data collection on LOS variables of non-chosen mode is a challenge. Detailed methodology adopted in collecting these data is also briefly explained. Developing land use database is also a major task in the context of developing cities in India. Detailed methodology adopted in collecting the land use data has also been discussed in detail. This chapter is organized as follows: Section 4.1 presents an overview of the study area. Section 4.2 provides the land use details and land use data collection methodology. Section 4.3 discusses the importance and procedure for preparation of the present network data. Section 4.4 presents the questionnaire used for the collection of travel related household data, which includes questionnaire for RP survey, SP survey, and the questions related to attitude and perception. Section 4.5 provides exploratory analysis of the household survey data with reference to the socioeconomic, trip purpose, and mode choice related data of the sample.

4.1 Study area description

In this study, Agartala, the capital city of Tripura, a state located in north eastern part of India, has been chosen as the study area (Figure 4.1). Agartala municipality consists of 35 municipal wards, divided mainly for administrative purposes. Agartala is the second largest city in the north-east India, after Guwahati, in terms of municipal area. According to census data of the year 2011, the population of Agartala was 3,99,688. The average annual rainfall of Agartala city is 220 cm. The city is situated at $23^{0}45'$- $23^{0}55$N latitude and $91^{0}15'$ - $91^{0}20'$ E longitude, in the flood plain of Haora river. The climate of Agartala is of tropical monsoon type with moderate temperature and high humidity. From the report available with Govt. of Tripura, Department of Urban development, a socioeconomic survey shows that around 45% of the people have migrated to the city implying the presence of employment opportunities. As per the city development plan for Agartala (2006), in the year 2005, the expenditure and income levels in Agartala are much more than the other urban areas of the state. The average household income per month was Rs 8506, of which, the expenditure was Rs 6255 per month. Approximately, 10% of the income i.e., Rs. 850 per month was being spent by the household on transportation needs.

With respect to the road network and connectivity, the road network follows grid iron pattern in the central part of the Agartala city. Percentage of road area, including the local roads is about 6.2 % of the total area of the city. A majority of the roads in the city are about 6-10 m wide and some roads have footpath facility but no cycling facility. Recently there has been high growth in the vehicle ownership, which may be attributed to the recent economic boom seen in the country. Vehicle population growth data presented in Figure 4.2 shows that there is a significant growth in car and MTW population and a decrease in the growth of number of buses in Agartala city. This sharp increase in the percentage of private vehicle ownership may be one of the reasons behind people shifting to private modes from sustainable modes of transportation.

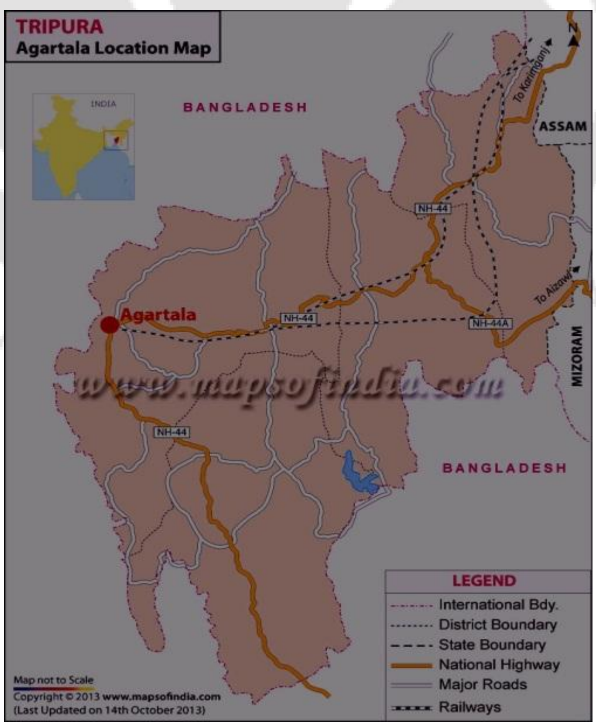
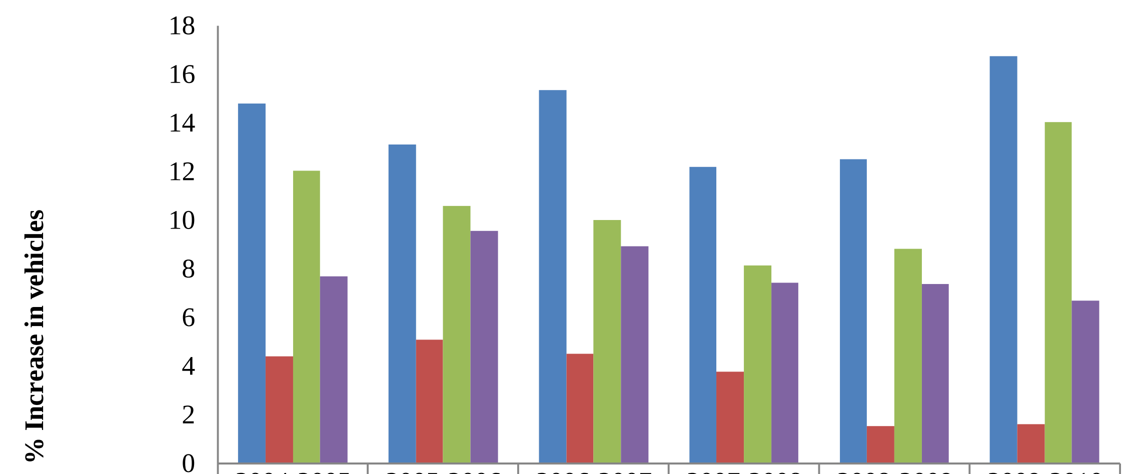


Figure 4.1: Location of the study area (Agartala)



	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010
% Increase in Cars	14.81	13.12	15.37	12.21	12.53	16.77
% Increase in Bus	4.4	5.09	4.51	3.78	1.54	1.62
% Increase in MTW	12.05	10.59	10.02	8.14	8.84	14.04
% Increase in MThW	7.71	9.57	8.94	7.43	7.39	6.7

Figure 4.2: Percentage growth of vehicles in the recent years in Agartala

Figure 4.3 shows the boundary of the Greater Agartala Planning Area (GAPA) as decided for the City development plan prepared for Urban Development Department, Govt. of Tripura, by LEA Associates South Asia pvt Ltd., in association with CEPT, Ahmedabad. Part of the GAPA bounded with the blue cordon line is considered as the study area for this work.

Table 4.1 shows the land use composition details of Greater Agartala Planning Area in the year 2001. The residential area is more as compared to the other land use types. Land uses that induce travel such as commercial, industrial, and public utilities are relatively smaller in terms of the land area occupied.

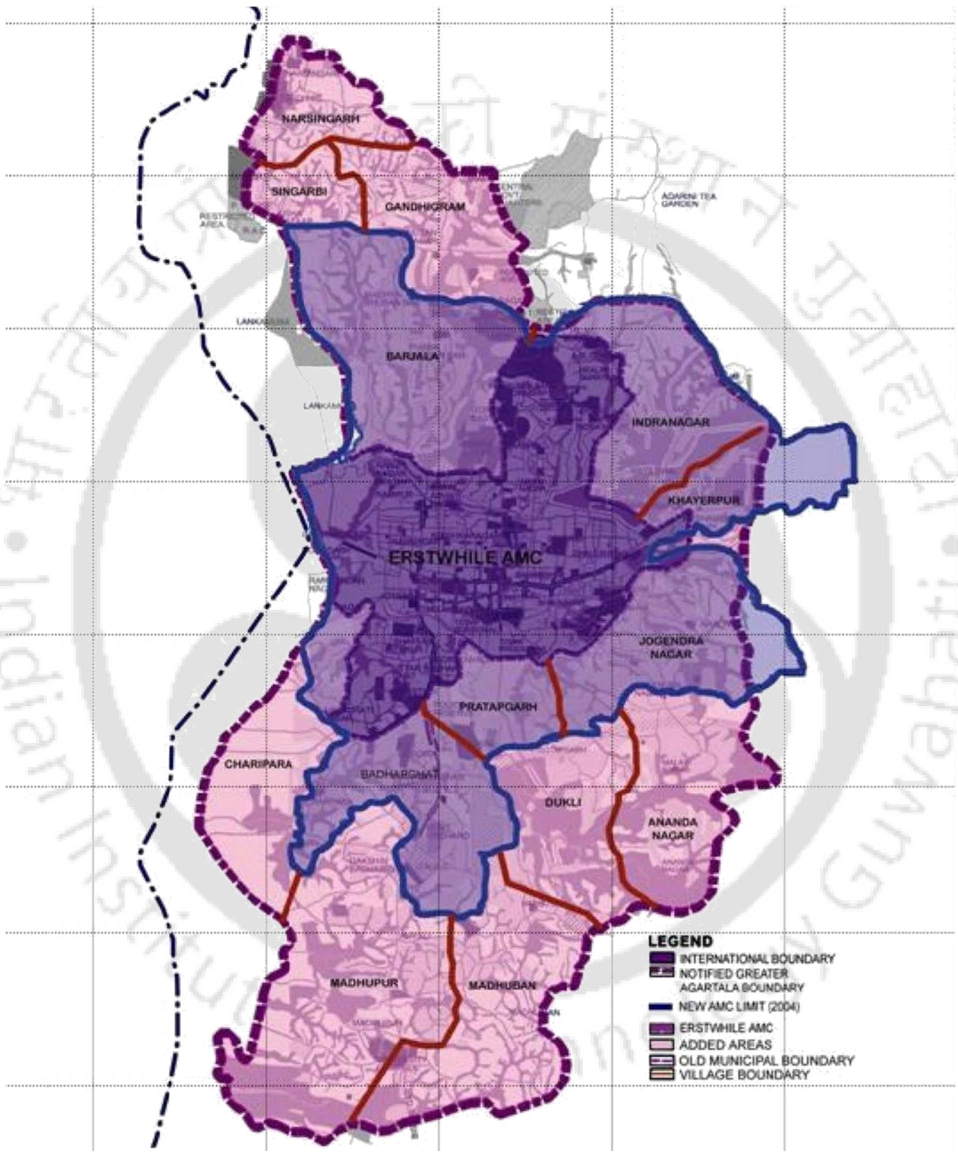


Figure 4.3: Greater Agartala Planning Area (Source: City development plan-Agartala, May 2006)

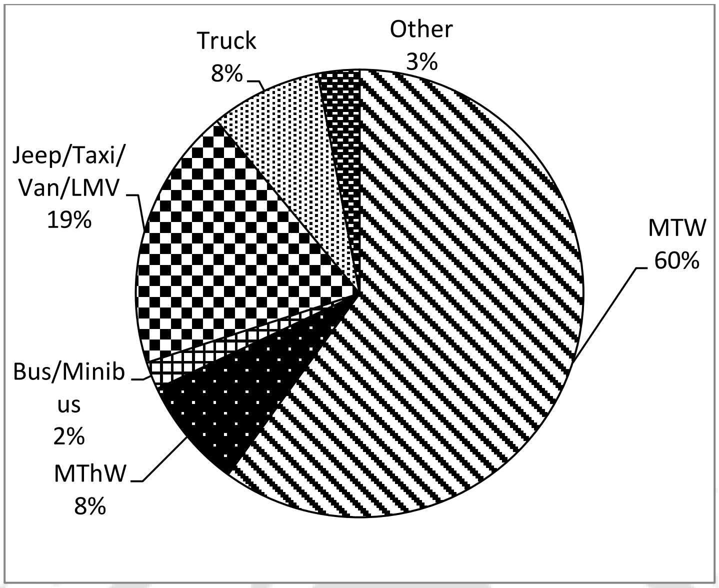


Figure 4.4: Composition of vehicles registered in Agartala city

Table 4.1: Land use details of Greater Agartala Planning Area in the year 2001

Land use	Area(Ha)	Percentage
Residential	2277.55	24.75
Commercial	50.75	0.55
Industrial	50.35	0.56
Public Utilities	18.5	0.2
Institutional	175.25	1.9
Government Function	70.5	0.76
Organized Open Space	24.55	0.26
Vacant Land	2205.75	23.97
Land under defense	79.5	0.86
Transport and communication	455.4	4.95
Agricultural	2257.75	24.54
Others (Including Forests, Water Bodies)	1532.15	16.7
Total	9200	100

* Source : Agartala City Development Plan, 2006

Table 4.2: Land use details of the study area considered for the present work

Land use	Area(Ha)	Percentage
Residential	2217.58	34.38
Commercial & Industrial	37.84	0.59
Educational	62.39	0.97
Service	21.43	0.33
Social and others misc.	6.95	0.11
Others (Agricultural, forest, drains, river and other water bodies etc.)	2048.33	31.75
Vacant Land	2056.192	31.88
Total	6450.71	100

The land use composition of the study area for the year 2010, given in Table 4.2, clearly shows the dominance of residential area in the city. Area corresponding to commercial, industrial, service, social and other land uses is considerably small. These details are obtained from the field studies conducted as part of the present work. Description of the land use data collection is given in the following section.

4.2 Land use details of the study area

Availability of the digitized land use map is a major problem in studying the effect of land use mix on travel behavior of smaller Indian cities. In this study an effort has been made to collect and digitize the land use details of Agartala city. Different land use data like residential, commercial, industrial, service, educational, forest, agricultural, river, ponds, social welfare centers, playgrounds and vacant land have been collected. Except forest, agricultural, vacant, and water bodies, all the remaining land uses have been categorized into five different types by suitably merging different land uses. Residential, commercial, educational, service and others are the five land use types considered in the present study. All the retail shops, including shopping complexes and the buildings with retail shops and industries have been considered as commercial land use. Buildings meant only for office use have been considered as service area. In order to avoid the complications in land use map preparation, land use mixing at building level was neglected. When a particular building is used for both commercial and residential purpose (for example, ground floor shops and other floor for residential), that building area was considered as commercial. If more than one floor is used for commercial purpose only, the area corresponding to ground floor was taken into consideration. During collection of land use data, GPS device (Handheld Trimble JUNO-SB) was used for measuring the area of the big markets, offices and shopping complexes. For isolated small shops, the location was digitized using the GPS and the area was calculated using a measuring tape. ArcGIS10 software was used for storing and analyzing the land use and travel data. Raster image of the study area, with municipal ward boundaries, has been digitized and shown in Figure 4.5. Figure 4.6 shows the comprehensive land use details of the study area. Figure 4.7 shows the land use details of the study area after categorizing various land uses into five types, as mentioned earlier. Chapter 5 deals with the methodologies adopted in extraction of the land use parameters from the digitized map shown in Figure 4.7.

4.3 Network Data

It is general practice to collect the information about the travel time and cost of different modes through network analysis (Bhat and Koppleman, 2006; Ortuzar and Willumsen, 2011). In case of Indian cities, this is a major problem as most of these cities do not have digitized road network with relevent attribute information. Public transport network details are also not available and most of the times public transport operates with flexible schedules as well as routes. Data collection on para transit or intermediate public transport modes is even more difficult. Parta transit modes include MThW and Jeeps, commonly used for ferrying the passengers in the smaller cities and are operated by the private people.

In this study entire road network of the study area has been digitised. Raster images of the road network (taken from google maps) have been digitized and the same was cross checked with a sample road network developed using GPS. The prepared network data was imported into TransCAD 5.0 software for network analysis. Average travel times of various types of vehicles have been manually collected using the stop-watches from all the major and minor road stretches of the network at different times of the day on different type of roads (morning off-peak from 7 AM - 9 AM, peak hour from 9:45 AM to 11 AM, noon off-peak 2 PM- 4 PM and evening peak 6 to 8 PM).

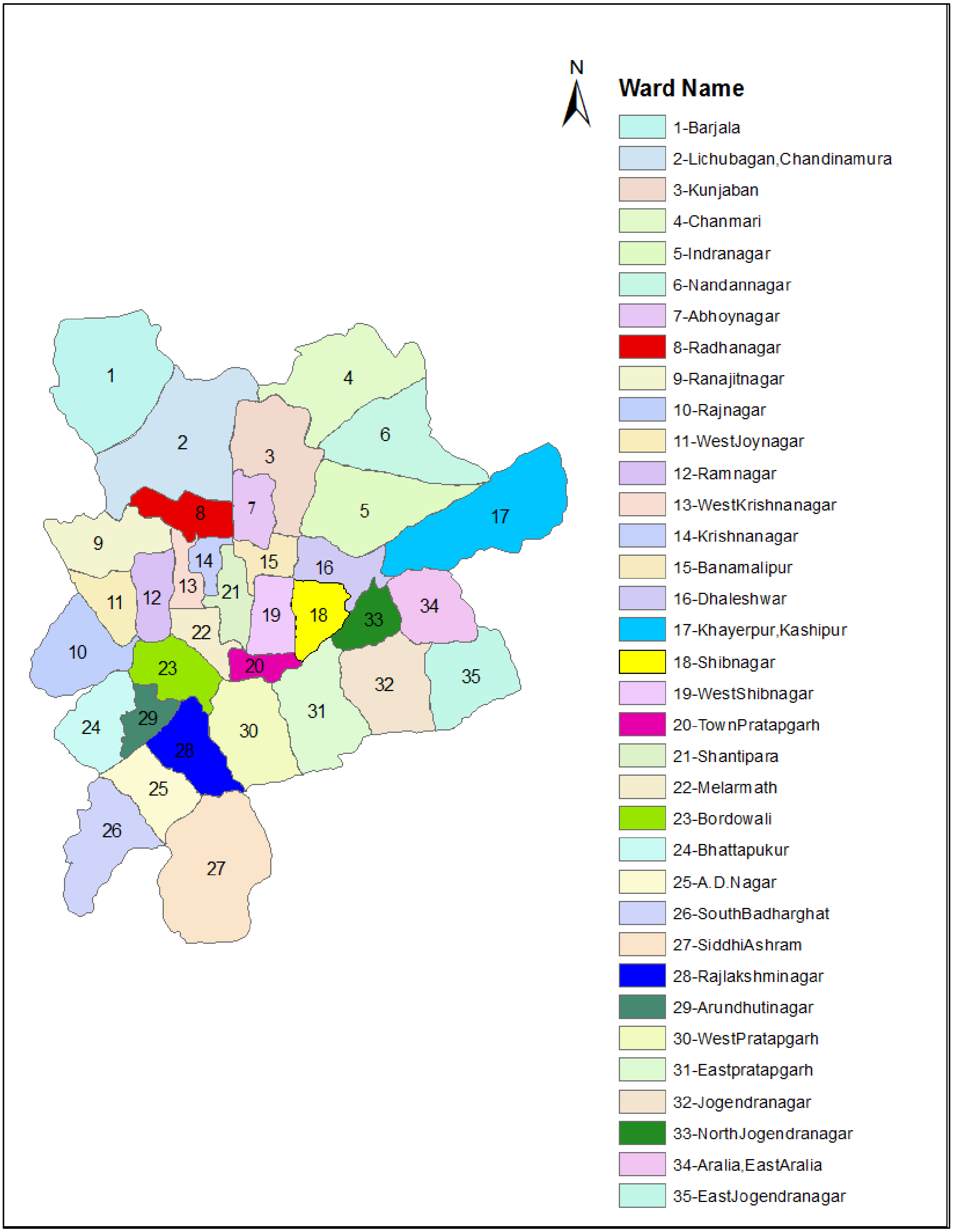


Figure 4.5: Ward/zonal details of the study area, Agartala city

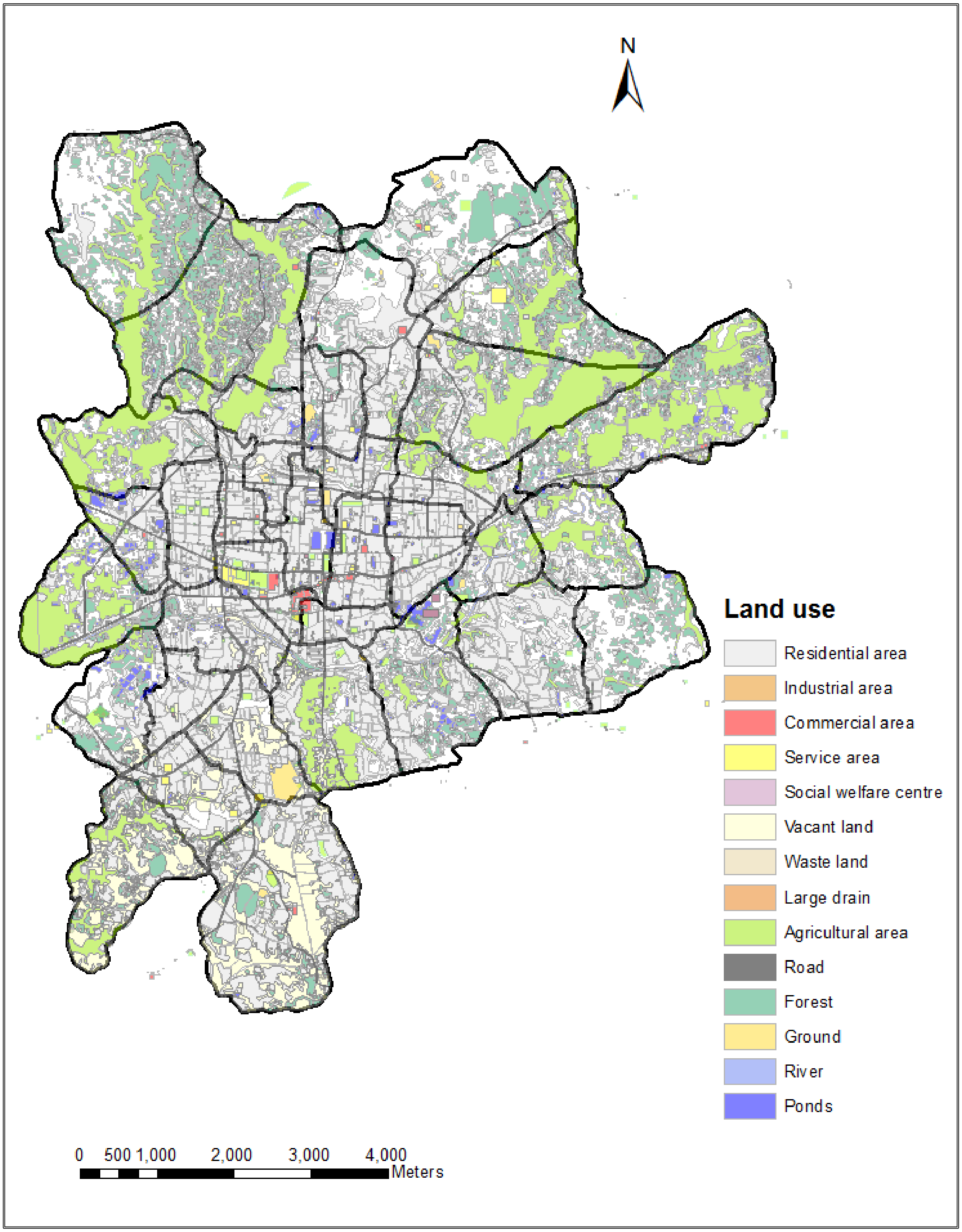


Figure 4.6: Digitized land use map of the study area.

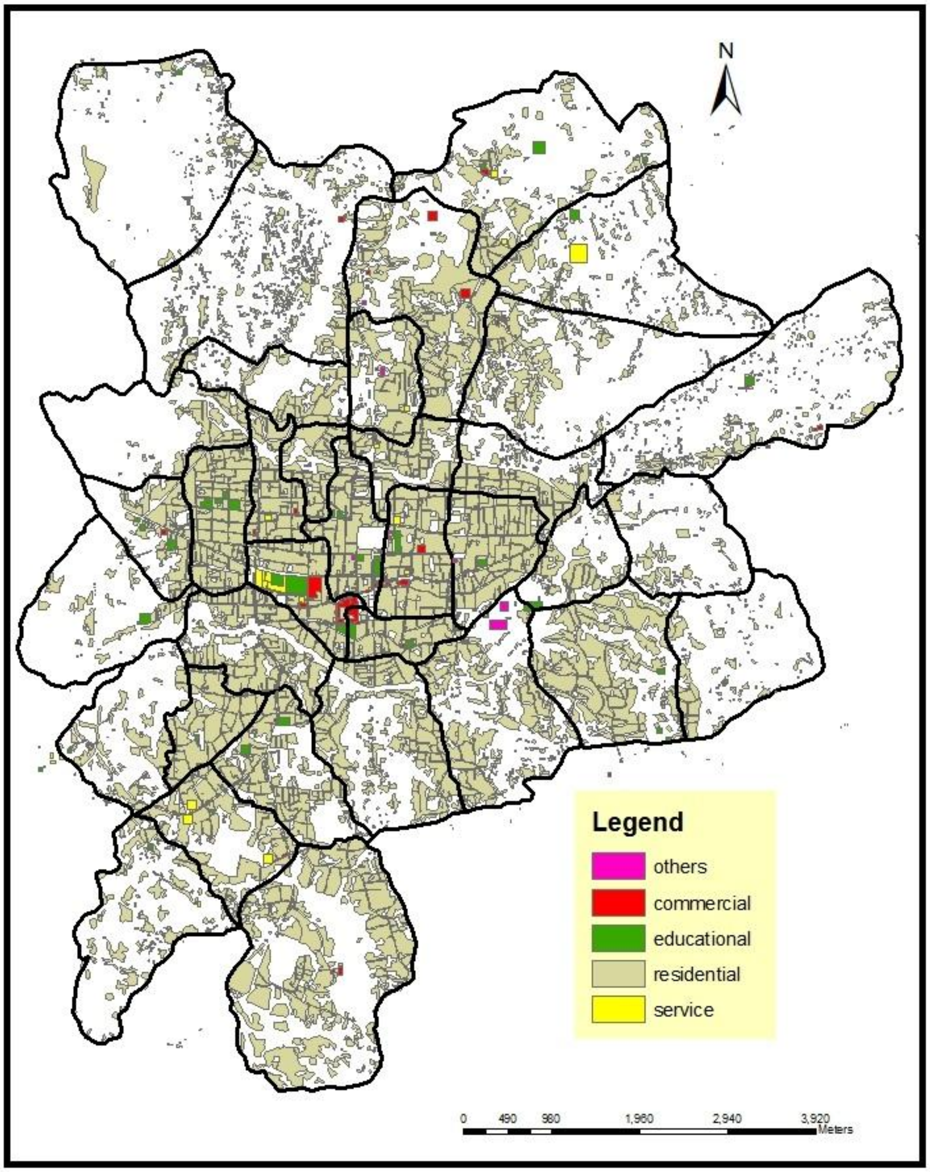


Figure 4.7: Land use map showing the land details considered in the study area

Speeds estimated based on the link length and the average travel time were loaded onto the links corresponding to the data collection stretches for different categories of vehicles and for different times of the day. The resulting network is useful in getting the travel related data for all the non-chosen alternatives. Mode related data like distance, travel time from the origin to the destination have been collected using the digitized road network. Taking travel time and distance data from the network essentially reduces the error as the travel time and trip distance given by the individual (for the chosen mode) found to have significant error as many individuals do not know the actual travel time and travel distance between their origins and destinations. For estimating the choice models, all the LOS data have been taken from the digitized network. Table 4.3 shows few trip related data taken from the network simulation. This table also shows the corresponding data of the chosen modes given by the respondents. Data related to various modes such as car, MThW, and MTW are provided in this table. From this table, it can be seen that there is a significant difference between the two sets of data, specifically in case of public and intermediate public transport modes. This difference is sufficient enough to have an impact on modeling. Data collected from the network simulation have been compared with the actual travel times (corresponding to few trips) and found that this data is more reliable. But, due to the nature of the services operated by the public and intermediate public transit services operated by the private people, there will be some differences in the actual and simulated data. For example, private operators wait for passengers, sometimes even at the non-designated stops, and this waiting times are completely random.

Figure 4.8 shows road network of Agartala city that was prepared for extracting the travel related data. Further, Figure 4.9 shows the screen shot of network prepared in TransCAD 5 showing the road network information.

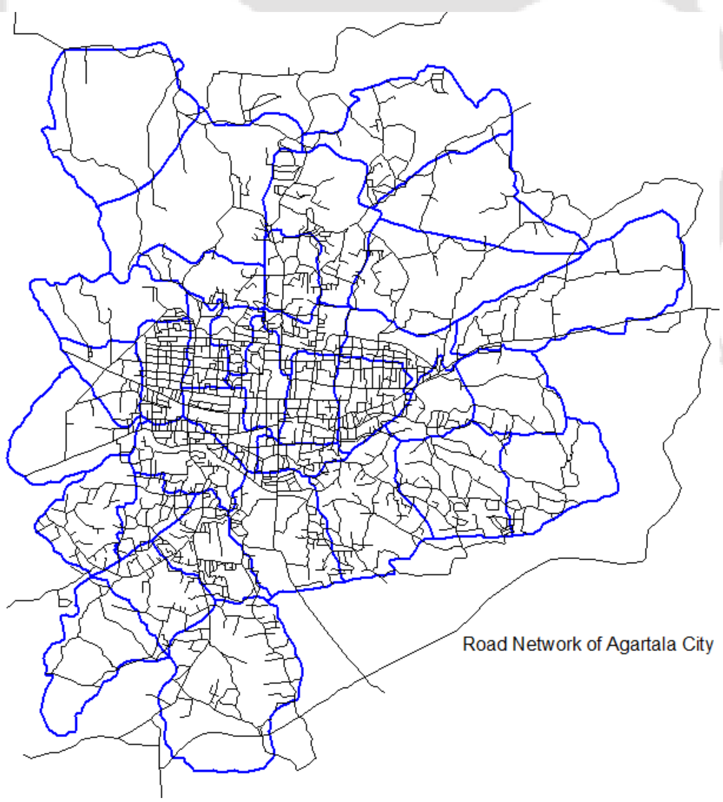


Figure 4.8: Road network of Agartala City

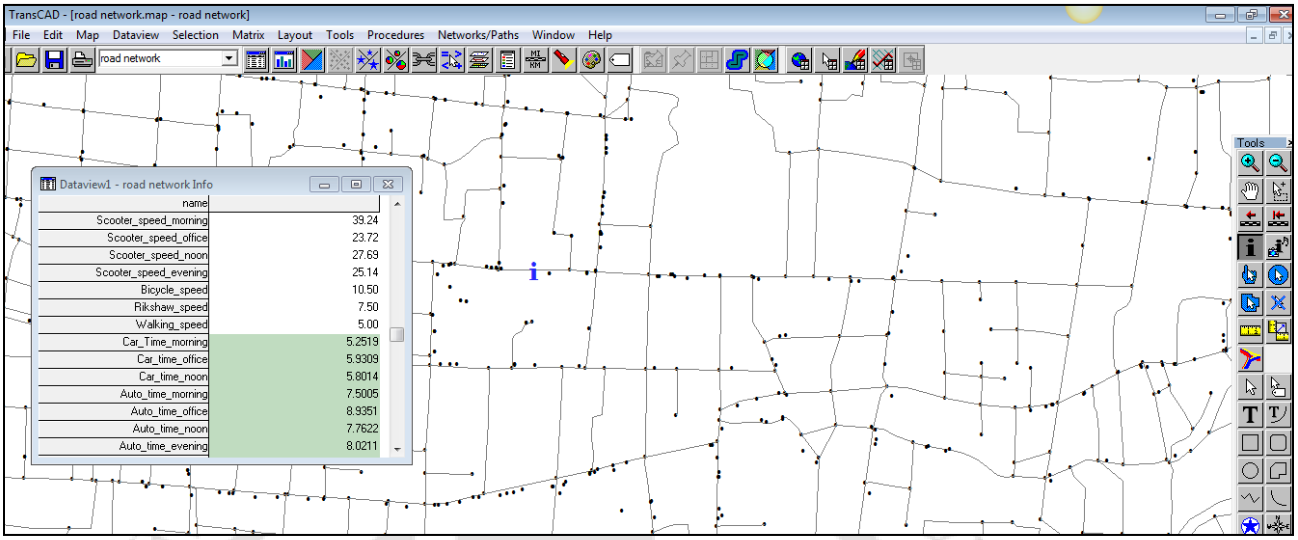


Figure 4.9: Screen-shot of road network prepared in TransCAD 5

Table 4.3: Comparison of mode related data obtained from the network analysis and stated by the individual for the chosen mode in a trip

SN	Mode	Mode related data given by respondent		Mode related data from Network Simulation		Percentage difference in travel time from network extracted time	Percentage difference in trip length from network extracted data
		Travel Time (min)	Distance (Km)	Travel Time (min)	Distance (Km)		
1	Bus	45	8	23.016	6.875	-95.516	-16.364
2	Bus	30	10	36.490	9.731	17.786	-2.764
3	Bus	40	20	29.538	18.764	-35.419	-6.587
4	Bus	25	12	40.988	10.930	39.007	-9.790
5	Bus	30	6	18.582	5.602	-61.447	-7.105
6	Auto	10	3	10.100	1.694	0.990	-77.096
7	Auto	15	3	8.450	1.210	-77.515	-147.934
8	Auto	20	4	15.320	3.602	-30.548	-11.049
9	Auto	10	6	12.120	3.512	17.492	-70.843
10	Auto	30	6	11.270	4.553	-166.193	-31.781
11	Auto	15	2	3.390	1.336	-342.478	-49.701
12	Auto	15	3	13.090	3.838	-14.591	21.834
13	Auto	30	6	27.100	3.033	-10.701	-97.824
14	Auto	15	3	9.430	2.175	-59.067	-37.931
15	Car	15	5	13.190	4.270	-13.723	-17.096
16	Car	10	5	7.530	5.902	-32.802	15.283
17	Car	5	2	2.530	1.213	-97.628	-64.880
18	Car	60	25	52.584	21.091	-14.103	-18.534
19	Car	15	8	12.310	5.500	-21.852	-45.455
20	MTH	10	2.5	6.260	1.989	-59.744	-25.691

4.4 Questionnaire for the household survey

Questionnaire used in the present study has been prepared with a view to gather comprehensive household travel related data. The travel survey has been prepared in line with the standard household surveys conducted in the developed countries. The questionnaire include all the essential requirements of a comprehensive household survey such as household details, personal details, vehicle ownership, travel diary, SP survey, perception and attitudinal survey (as per report on household datacollection,www.public.asu.edu/~rpendyala/HouseholdTravelSurveyDataCollectionPlan.pdf). Detailed questionnaire is given in the Annexure.

4.4.1 Questionnaire for the RP survey

RP data have high reliability and face validity as these data are related to the real choices made by the individuals. Moreover, in real market situation, choices made by the individuals are bounded by real constraints and this is captured in the RP data. These constraints that affect the choices are not observed in the SP data. SP survey is widely used for collecting travel behavior data, mainly for predicting performance of new alternatives and hypothetical scenarios. Face to face interview was conducted for household data collection. Surveyors were recruited for data collection and were given training on using the questionnaire. A few students from NIT Agartala have also helped in collecting the household data.

4.4.1.1 Household Details

This section contains information related to the household such as the address, family size, status of the household (rented/owned), number of employed and unemployed adults, monthly travel expenditure, and aggregated monthly income. Household location has also been digitized using the location data collected with GPS. Figure 4.10 shows the locations of households from where data have been collected. Also, detailed information about the vehicles owned by the household, like number of vehicles owned, type of vehicles, usage of vehicles, i.e., for private or public use, model and year of manufacture, odometer reading of the vehicle, has been obtained.

4.4.1.2 Personal Data Set

Personal data are obtained from the individual persons of the household. This contains information like gender, age, educational status, monthly income, distance between the workplace and home, most preferred mode, least preferred mode, monthly expenditure on travel, etc. This section of survey also includes some questions related to the use of bicycle, public transit, and walking mode, which helps the analyst to understand the reasons for motivating or discouraging the commuters in using a particular mode.

4.4.1.3 Travel Diary

The information related to the trips made by the individuals like origin, destination, purpose of the trip, mode of travel, length of the trip, travel time, fare/cost, waiting time, activity duration and other modes available for making the same trip is obtained using this part of the questionnaire.

4.4.2 Questionnaire for the Stated Preference (SP) Survey

From the past studies, it has been evident that the RP data have little variability i.e., the attribute values do not vary much. Moreover, the attributes are highly correlated. Finalizing the choice set for RP survey is very difficult, whereas the choice set ambiguity is not there in case of SP survey. The purpose of the SP survey is to collect the data required for efficient model estimation with as little bias as possible.

From the vehicle composition data (see Figure 4.1), it can be seen that the percentages of MTW, LMVs, and MThW are predominant while the share of bus is negligible. In the proposed SP survey, it was decided to study the trip maker's behavior towards the three major modes i.e., MThW, MTW, car and public transport (bus), which is being implemented as a part of JNNURM (Jawaharlal Nehru National Urban Renewal Mission) scheme, initiated by the Govt. of India.

A preliminary O-D survey was carried out to know the important factors influencing the mode choice. Ranking data for a number of attributes such as travel time, cost, comfort, accessibility (in terms of walking distance), convenience, availability, flexibility, traffic congestion, bad road condition, safety, and distance were collected as a part of the preliminary survey.

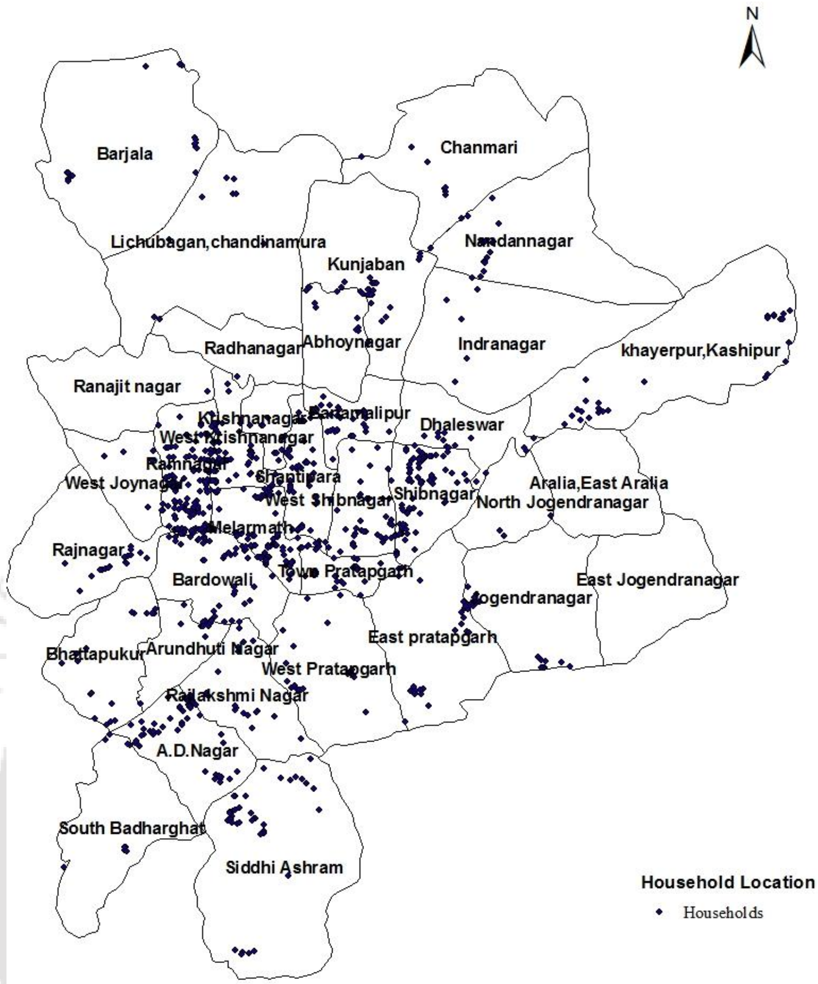


Figure 4.10: Location of the households from where the data have been collected

After analyzing the preliminary data, convenience, availability of modes, comfort, cost, travel time, and flexibility were found to be influencing the mode choice. The availability of mode can be represented by the choice set used for the SP survey and frequency of transit has been added in the SP survey. There were already some studies in which convenience and comfort were represented as a single attribute. Hence, cost/fare, comfort, travel time, and frequency (for only public transit) were selected as the attributes in the design of SP survey.

SP survey was carried out for four motorized modes car, bus, MThW and MTW. Four levels of attributes were considered in the SP survey, except for comfort which was measured as a binary variable where two levels (comfortable and not comfortable) were considered. For MThW, higher level of comfort represents the availability of seat, and less number of co-passengers. For bus, higher level of comfort represents the availability of seat, and no-standing scenario. Based on the fuel prices, travel cost is assumed to be changing in case of car and MTW. In case of the other two modes fares are regulated by the state government but more or less governed by the fuel prices in addition to the other factors. The price of petrol was around Rs 65 at the time of survey and a decrease or increase will sufficiently affect the cost of travel. Fuel prices are assumed to be Rs 50/liter, Rs65/liter, Rs80/liter and Rs 100/liter. Travel times may change due to the increasing number of personal vehicles which may lead to congested roads and this eventually results in increased travel times. For travel time, four levels were considered assuming that speed takes the values of 10/20/30/40 km/hr for bus/MThW. For MTW and car, speed was considered as 10/25/40/55 km/hr. From the field data collected on speed it was found that there is a difference in average speed between bus/MThW and MTW/car. For car, four levels of travel cost were taken; Rs 3/km, Rs 6.5/km, Rs 8.0/km and Rs 10/km. For MTW, the cost of the trip is considered as Rs1.33/km,Rs1.85/km, Rs 2.42/km and 2.85/km. Fares for bus and MThW are provided in Tables 4.5 to 4.8. For bus mode four levels for frequency (time between every buses) 5 min, 15 min, 30 min and 45 min have been considered. Further, different trip lengths have been considered in the SP survey as the average trip distance varies significantly for different modes as observed in the preliminary survey. This type of design was considered so that it can improve realism of the SP task, and individuals will not have face unrealistic scenario. Table 4.4 to Table 4.8 show the attribute levels and their values used in the SP survey.

Block design method was used in formulating the SP questionnaire i.e., the number of experiments are divided into sets (blocks), and the blocks were distributed over a number of respondents. The success of this test depends on the assumption that preference across the sample of respondents will be sufficiently homogenous such that respondent's data can be combined over the blocks of the experiment. Different SP questionnaires have been designed for different choice sets and for different trip lengths. Sample questionnaire for four modes with 3 km travel distance is given in the Annexure. Description of these questionnaires is given below:

1) **Choice set with Bus, MThW, Car and MTW**: Minimum number of experiments required for finding the main effects for four alternatives with 11 attributes are 45

(M (Number of alternatives) * A (Number of attributes) + 1 = 4 * 11 + 1 = 45). Two attributes, cost and travel time, were considered for MTW and car. For bus, four attributes cost, travel time, comfort and frequency, were considered. For MThW, three attributes, cost, travel time, and comfort were considered. The master design set for choice task is 4^9 x 2^2 i.e., 1048576 (Using L^{MA} L is number of levels,M the number of alternatives and A the number of attributes). The design used in the present study was carried out as fractional design with 64 runs for main effects only. The blocks were prepared in such a way that each individual can see all the levels of attributes, for which two extra attributes were added as blocking variables (Hensher et al., 2005). 16 blocks were prepared with 4 experiments in each block so that the respondent fatigue can be reduced. All the 16 blocks are to be combined to make the experiment orthogonal and to estimate the main effects.

2) **Choice set with alternatives (Bus, Car and MThW) and (Bus, MThW and MTW):** Minimum number of experiments required for finding the main effects, for three alternatives with 9 attributes, is 28. The master design set for choice task is 4^7 x 2^2 i.e., 65536. The design was carried out as fractional design with 32 runs (8 blocks), for main effects only. Eight blocks were prepared each with 4 experiments, for which an extra attribute was required as a blocking variable. 8 blocks need to be combined to make the experiment orthogonal for estimating the main effects.
3) **Choice sets with alternatives (Bus and MThW):** Minimum number of experiments required for finding the main effects, for two alternatives with 7 attributes, is 15. The master design set for choice task is 4^5 x 2^2 i.e., 4096. The design was carried out as fractional design with 32 experiments (8 blocks), for main effects only. Eight blocks were prepared, each with 4 experiments, for which an extra attribute was added as a blocking variable. All the eight blocks need to be combined to make the experiment orthogonal and to estimate the main effects.

In this way SP instruments, matching with the specific choice sets of the individuals, could be administered. The block is so generated that each individual is shown four choice experiments and each individual can see atleast once all the levels of attributes. Orthogonality for main effects and all the two-way interactions was verified.

Table 4.4: Orthogonal coding of the attributes considered for fractional factorial design

Modes\Attributes	Comfort		Cost				Time				Frequency			
Car	Constant		-2	-1	1	2	-2	-1	1	2				
Bus	1	-1	-2	-1	1	2	-2	-1	1	2	-2	-1	1	2
MThW	1	-1	-2	-1	1	2	-2	-1	1	2				
MTW	Constant		-2	-1	1	2	-2	-1	1	2				

Table 4.5: Levels and values of attributes for 3 km trip length

Modes\Attributes	Comfort		Cost (Rs)				Time (min)				Frequency(min)			
Car	Constant		9	20	25	30	3.5	4.5	7	18				
Bus	1	-1	3	5	10	15	4.5	6	9	18	5	15	30	45
MThW	1	-1	3	5	10	15	4.5	6	9	18				
MTW	Constant		4	5	7	9	3.5	4.5	7	18				

Table 4.6: Levels and values of attributes for 6 km trip length

Modes\Attributes	Comfort		Cost (Rs)				Time (min)				Frequency(min)			
Car	Constant		18	40	50	60	6.5	9	15	36				
Bus	1	-1	3	5	10	15	9	12	18	36	5	15	30	45
MThW	1	-1	5	8	15	20	9	12	18	36				
MTW	Constant		9	11	14	17	6.5	9	15	36				

Table 4.7: Levels and values of attributes for 9 km trip length

Modes\Attributes	Comfort		Cost (Rs)				Time (min)				Frequency(min)			
Car	Constant		27	60	70	90	10	13.5	22	54				
Bus	1	-1	5	10	15	20	13.5	18	27	54	5	15	30	45
MThW	1	-1	5	10	15	20	13.5	18	27	54				
MTW	Constant		13	17	21	26	10	13.5	22	54				

Table 4.8: Levels and values of attributes for 12 km trip length

Modes\Attributes	Comfort		Cost (Rs)				Time (min)				Frequency(min)			
Car	Constant		36	80	95	120	13	18	29	72				
Bus	1	-1	5	10	15	20	18	24	36	72	5	15	30	45
MThW	1	-1	10	15	20	25	18	24	36	72				
MTW	Constant		17	23	28	34	13	18	29	72				

4.4.3 Attitudinal and perception related questions

Many researchers in the past have suggested that consideration of psychological factors, generally considered as latent variables, may help in better understanding of the travel behavior of the individuals. In order to understand the effect of latent variables, individuals were asked some questions (agreement and disagreement statement) related to their attitude towards various modes. Further, individuals were also asked to rate different modes based on the attributes like flexibility, reliability, safety, and comfort. Table 4.23 provides the perception ratings of different travel modes, and Table 4.24 provides the mean ratings of agreement and disagreement statement related to the attitudes of the individuals.

4.5 Exploratory analysis of the sample data

Travel data have been collected through a household survey conducted in the study area during March-September, 2012. Sample size, in terms of households, is about 1% of the total number of households (1028 households) of the study area. In this section, statistics of socioeconomic characteristics like age, gender, years of education, household size, household income, vehicle ownership and driving license status of commuter is presented. These statistics are useful in examining the variability of socioeconomic information, which could be useful in formulation of the choice models.

4.5.1 Socioeconomic characteristics of the sample

The sample was found to be representing the overall travel patterns of Agartala residents. In the collected sample, 30% of the trips were non-motorized and more or less similar figures were given in a report (2008) of Ministry of Urban Development, Government of India as well as in a report on comprehensive mobility of Agartala, 2008. Basic composition of the sample data are given in Table 4.9.

Table 4.9: Socioeconomic composition of the sampled data

Socioeconomic characteristic	Value in Percentage
Gender	
Male	73.38
Female	26.62
% of individual in the age category	
Up to 20	3.95
20 -30	17.69
30 – 40	18.53
40 -50	24.08
50 -60	22.77
> 60	12.98
% of household having driving License	
Having	49.5
Not Having	50.5
% of individuals (Years of education)	
0	0.19
1 to 5	4.05
5 to 8	9.13
8-10	20.51
11 to 12	15.62
13-15	33.02
16 to 18	17.31
19-21	0.00
More	0.19
Car Ownership	13.83
MTW ownership	44.21
Monthly household income (in Indian rupees)	
0-2000	0.28
2001-10000	33.96
10000-20000	30.48
20000-50000	24.46
> 50000	10.82

Table 4.10: Descriptive statistics of the socioeconomic characteristics of the sampled data

Socio-economic parameters	Mean	Median	Mode	Standard Deviation	Minimum	Maximum
Income	28119.47	15000.00	10000.00	30689.93	2000.00	200000.00
Family size	4.17	4.00	4.00	1.99	1.00	18.00
Age	43.21	44.00	52.00	15.12	14.00	96.00
Years of Education	12.50	12.00	15.00	3.37	0.00	22.00
Vehicle ownership	0.65	0.00	0.00	0.80	0.00	6.00

From Table 4.9, it can be seen that the sample contains less female representation than males. Further, 50.5 % of households are captive to public/para-transit modes as they are not having driving skills, as revealed in the choice set availability to the individuals. Car ownership in the sample is 13.83% and MTW ownership is 44.21%. From Figure 4.11 and Table 4.10, it can be seen that the average household income of the sampled individuals is ranging between Rs 20000.00 to 30000.00 per month. The average income of the sampled households is slightly higher than the income reported (8,506.00 rupees per month in the year 2005) in a report of with Govt. of Tripura, Urban development department. Given the socioeconomic changes happened after the year 2005, the rise in the average household income seems to be meaningful.

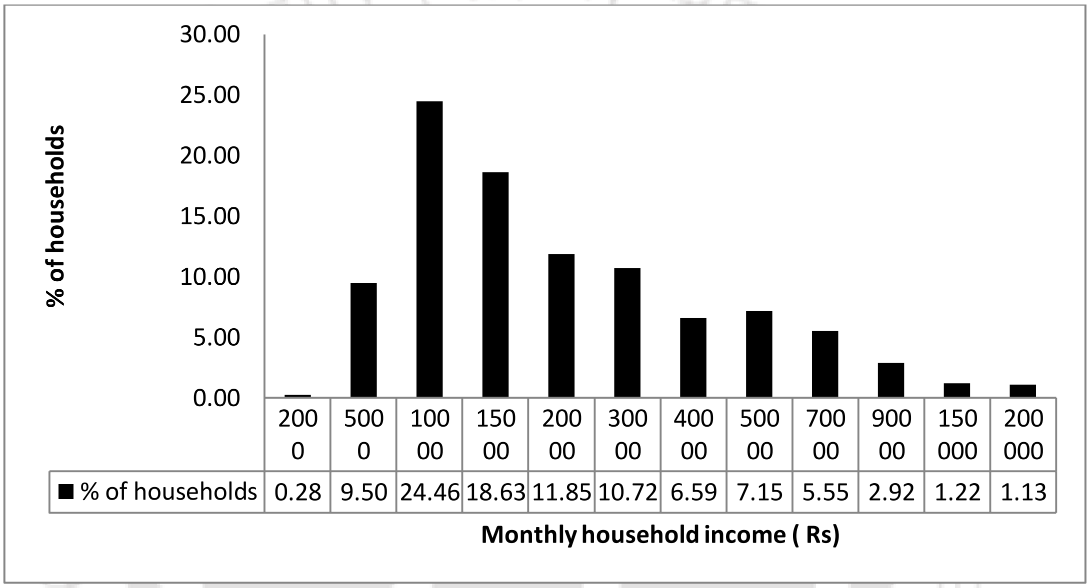


Figure 4.11: Distribution of the household income in the sampled households

In the collected sample, the mode of the family size is about four persons (Figure 4.12). In the sampled data, age of the persons is found to be distributed evenly between 30 to 60 years (Figure 4.13) with a mean age of 43 years (Table 4.10). From Figure 4.14 and Table 4.10, the average years of education were found to be 12.5 years and the sample contains significant percentage of individuals with 13 to 15 years of education. From Figure 4.15, it can be seen that about 50.6% of the households do not own private vehicles and the trip makers of these households may be captive to the other modes of transport i.e., public transit, intermediate public transit, and NMT modes.

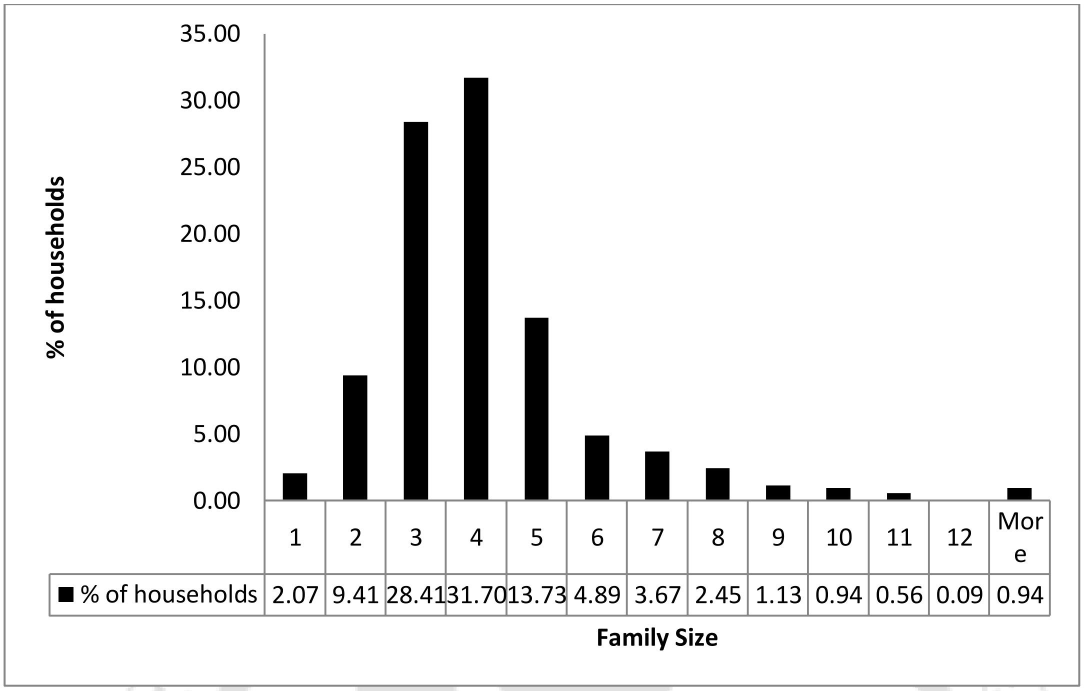


Figure 4.12: Distribution of family size in the sample data

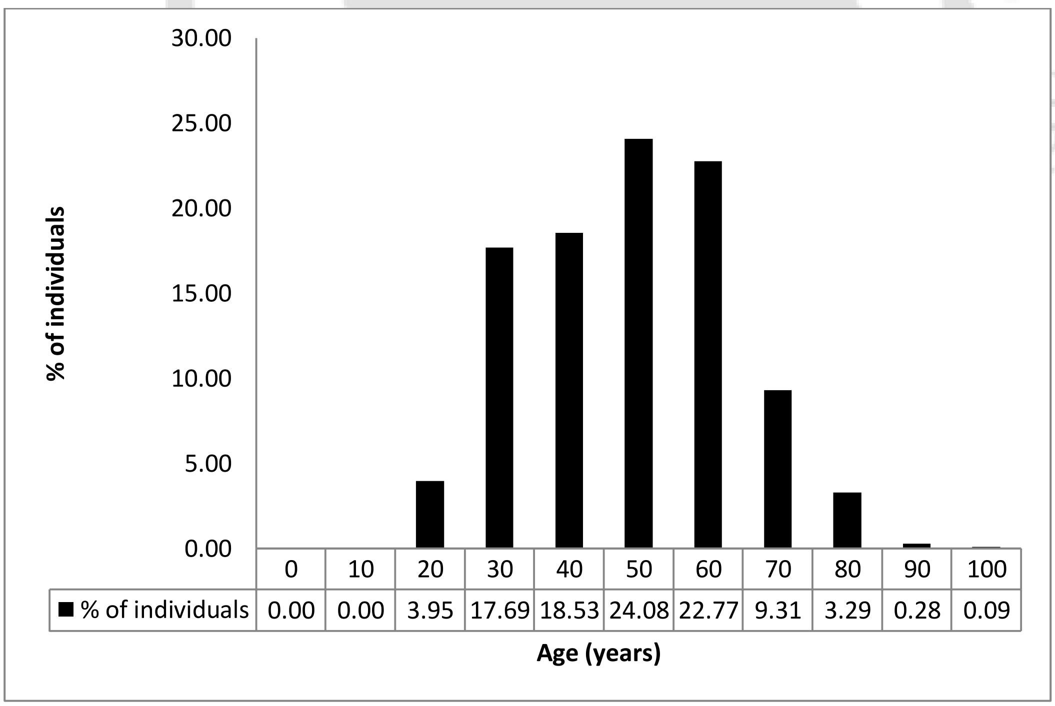


Figure 4.13: Distribution of age of the individuals in the sample

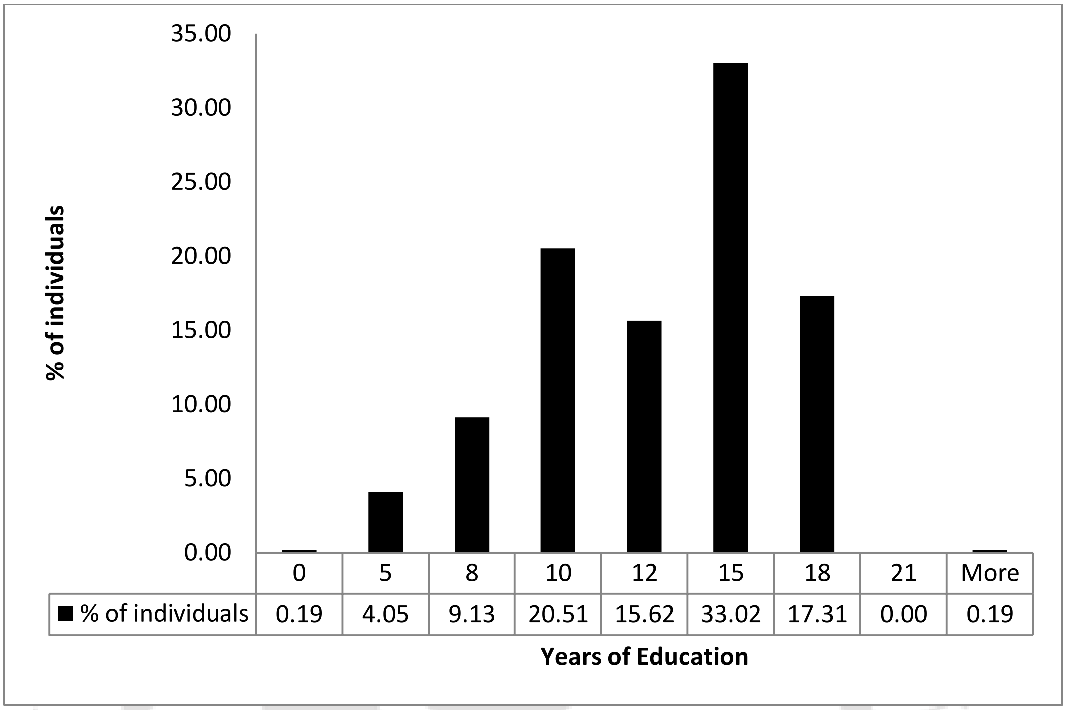


Figure 4.14: Distribution of years of education in the sampled individuals

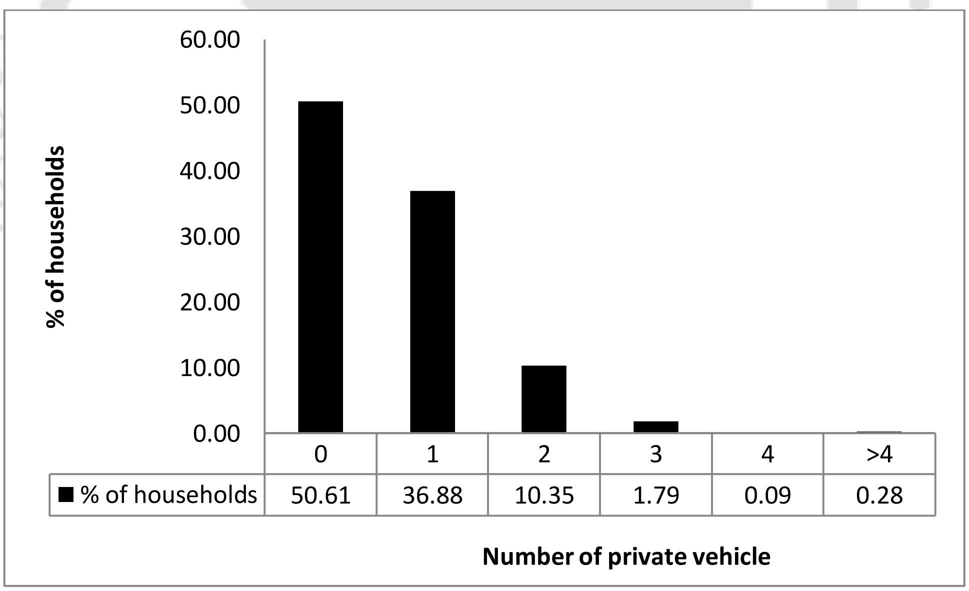


Figure 4.15: Distribution of vehicle ownership in the sample

4.5.2 Exploratory analysis of the travel related data

4.5.2.1 Purpose of the Trips

In the household survey, thirteen different trip purposes were mentioned and the data were collected accordingly. From the collected data a majority trip purposes were found to be work, shopping, social and educational trips. Hence, all the trips were categorized into four trip groups, work trip (41.46%), shopping trip (21.62%), educational trips (10.13%) and others trip (26.79%).Trips made for the purchase of grocery, vegetables and other items such as clothes, electronics goods and miscellaneous items were considered as shopping trips. Trips such as visiting doctors, social trips, trips to cultural events, and entertainment were categorized as other trips.

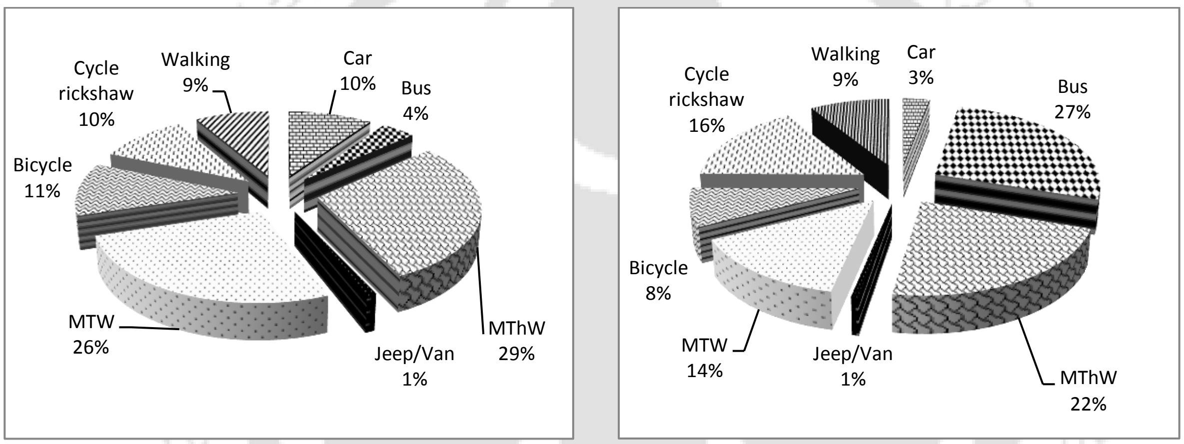


Figure 4.16: Modal composition for (a) Work trips (b) Educational trips

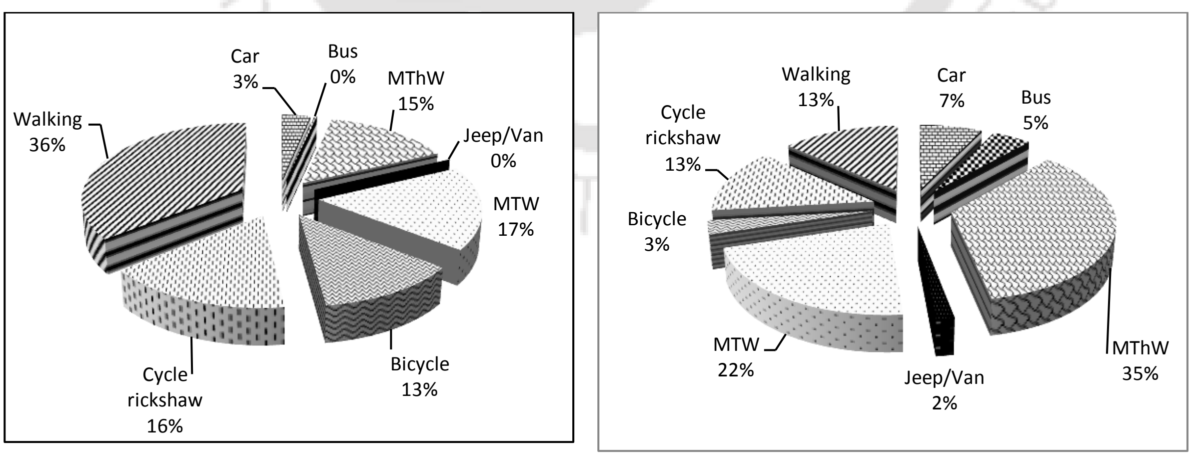


Figure 4.17: Modal composition for (a) Shopping trips (b) Other trips

From Figure 4.16 (a), it can be seen that MThW (29%) and MTW (26%) are the predominant modes used for making the work trips. Bus (4%) and Jeep/van (1%) are less frequently used as the commute modes. Car, bicycle, cycle rickshaw and walk modes have got equal share in the work related trips. In case of the shopping trips (Figure 4.17(a)), walk (36%) is the mostly chosen mode. MThW, MTW and cycle rickshaw were also commonly used for making the shopping trips.

From Figure 4.16 (b), it can be seen that a significant portion of education trips were made using NMT and transit (NMT 33% and transit 27%). Share of the transit mode (Bus) was less in trip purposes like work trips, shopping trip and other trips. In case of shopping trips, NMT mode was preferred (about 65%) compared to the 30% of work trips, 29% of social and other trips, and 33% of educational trips. This may be inferred shopping trips are shorter and involved with not much luggage. This may be due to the availability of the shops (of different nature) within the vicinity of the households.

4.5.2.2 Mode choice with respect to socioeconomic characteristics

Analysis of the effect of the socioeconomic characteristics on mode choice of the sampled individuals is analyzed in this section. This analysis has been carried out as it would be useful in finalizing the socioeconomic characteristics that would be useful for choice modeling. This analysis has been carried out for work and shopping trips only. Analysis of the effect of socioeconomic characteristics on work related mode choice is presented in the following sub-section. Similar analysis for shopping related trips is presented in the ensuing sub-section.

4.5.2.3 Mode choice for work trips

Data presented in Figure 4.18 and Table 4.11 clearly shows that higher income people mostly own a car. Low income people generally own/use bicycle and walking modes. In case of the work trips, share of MTW (25%) and MThW (33%) is high for households with income less than Rs 40000.00 per month. With increasing income, the patronage for car is found to be increasing. Among the households with more than Rs 50000.00 monthly income, 38% of the trip makers use car and its share is highest compared to the other modes.

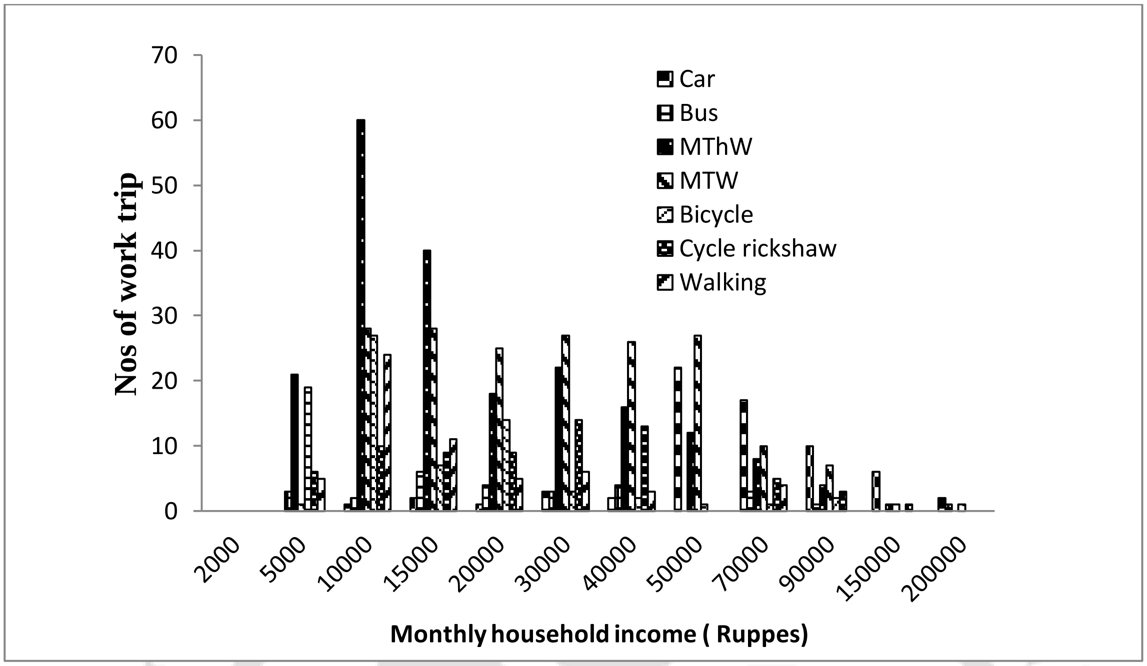


Figure 4.18: Distribution of work trips based on income

Table 4.11: Descriptive statistics of the income of the individuals using various modes for work trips

Modes	Mean	Median	Mode	Standard Deviation
Car	71515.15	70000.00	50000.00	39922.71
Bus	35370.37	20000.00	15000.00	39877.84
MThW	22945.54	15000.00	10000.00	20669.44
MTW	33618.78	30000.00	10000.00	25402.83
Bicycle	16052.63	10000.00	10000.00	16438.08
Cycle rickshaw	30785.71	30000.00	30000.00	25575.06
Walking	19137.93	12500.00	10000.00	16520.53

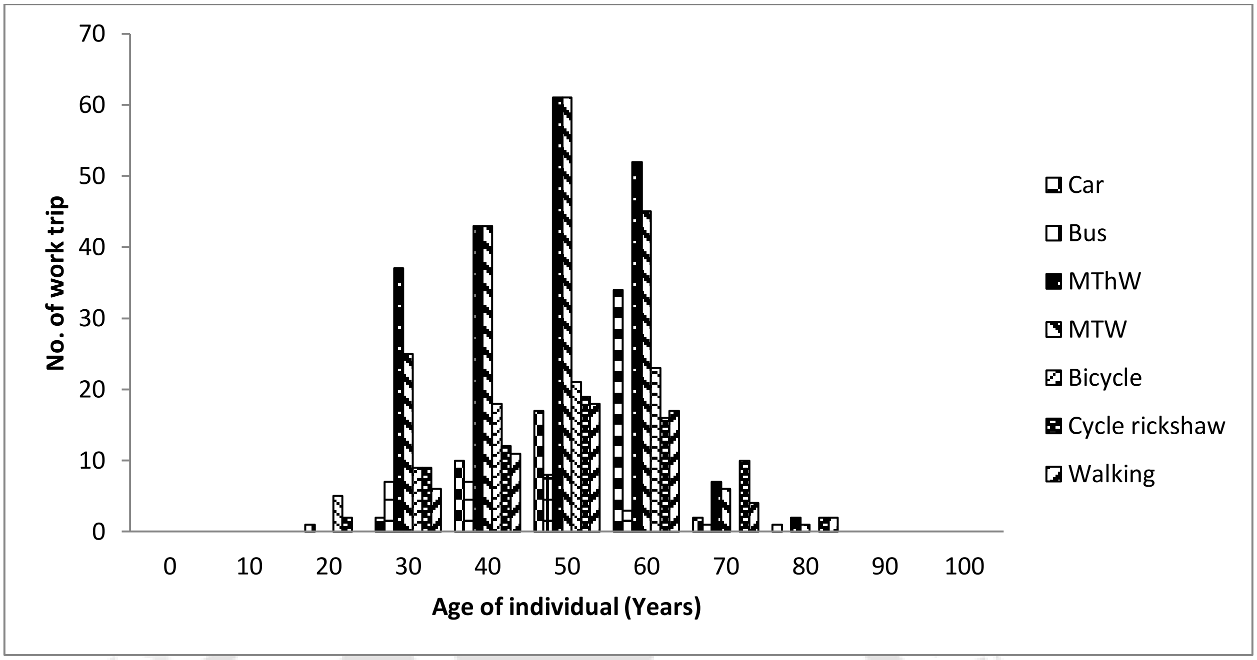


Figure 4.19: Distribution of work trips based on age

Table 4.12: Descriptive statistics of the age of the individuals using different modes, for work trips

Modes	Mean	Median	Mode	Standard Deviation
Car	48.73	51.00	51.00	8.15
Bus	38.93	37.00	32.00	12.65
MThW	43.50	45.00	50.00	11.78
MTW	43.86	45.00	52.00	10.66
Bicycle	42.09	42.00	42.00	11.69
Cycle rickshaw	45.79	46.00	62.00	13.62
Walking	47.29	47.00	52.00	11.61

From Figure 4.19, it can be inferred that MTW and MThW are the preferred modes for making work related trips irrespective of the age of the sampled individuals. From the sample, it has been observed that significant number of individuals aged between 50 to 60 years have made work trips using car.

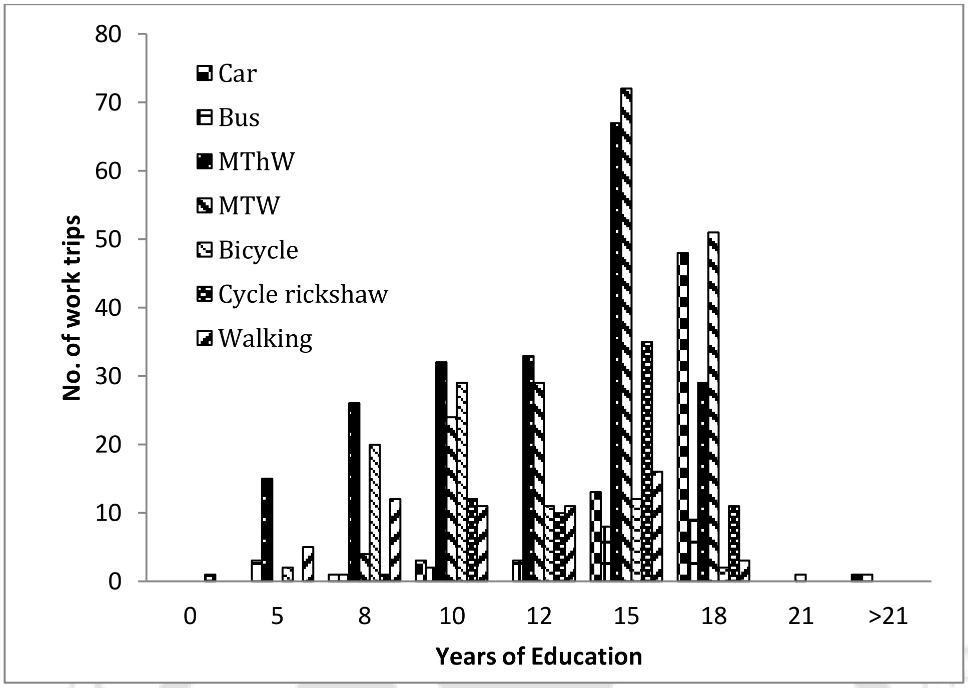


Figure 4.20: Distribution of work trips based on years of education

Table 4.13 Descriptive statistics of years of education of the individuals using different modes for work trips

Modes	Average years of education	Median	Mode	Standard Deviation
Car	15.71	16.00	16.00	2.02
Bus	13.63	15.00	15.00	4.35
MThW	12.24	12.00	15.00	3.61
MTW	14.06	15.00	15.00	2.58
Bicycle	10.36	10.00	10.00	2.85
Cycle rickshaw	13.63	15.00	15.00	2.85
Walking	11.07	12.00	15.00	3.49

From Figure 4.20, it can be seen that usage of car is increasing for making the work trips with increasing years of education. From Table 4.13, it can be seen that people using walk and bicycle as travel modes are less educated than the people choosing the other modes.

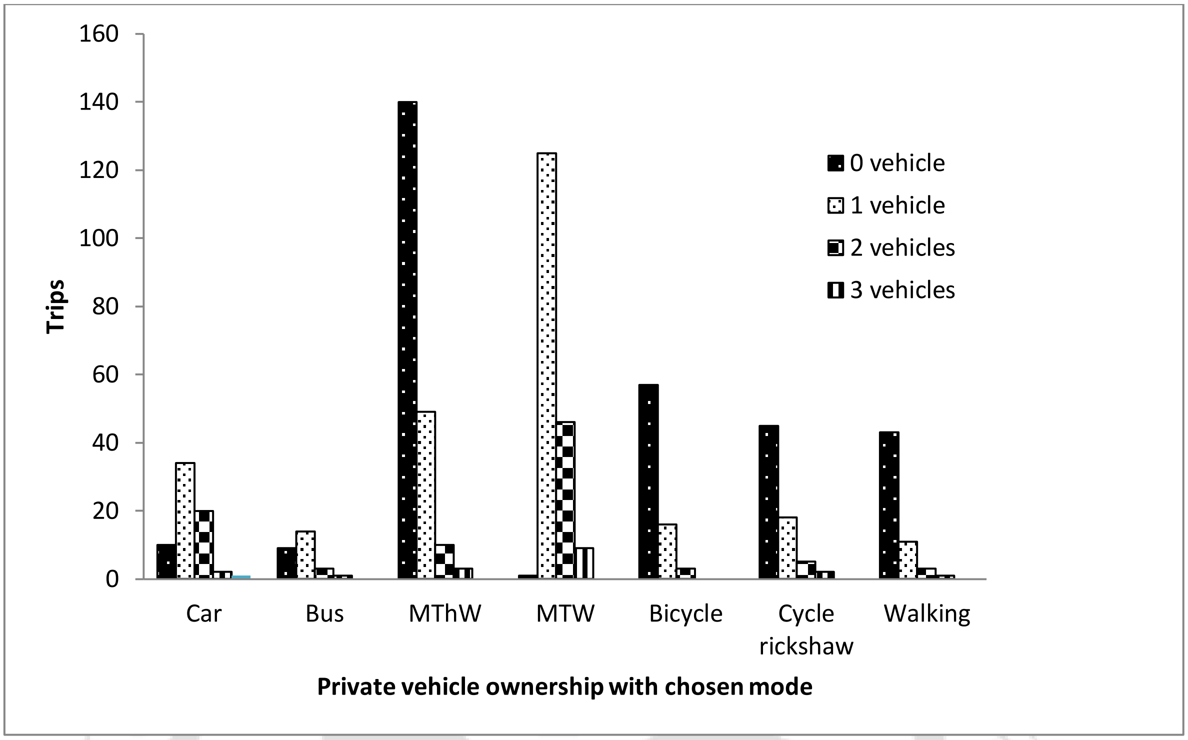


Figure 4.21: Distribution of work trips based on vehicle ownership (MTW + Car)

Table 4.14: Descriptive statistics of vehicle ownership in case of work trips

Modes	Average private ownership	Median	Mode	Standard Deviation
Car	1.21	1.00	1.00	0.73
Bus	0.85	1.00	1.00	0.77
MThW	0.39	0.00	0.00	0.65
MTW	1.35	1.00	1.00	0.58
Bicycle	0.29	0.00	0.00	0.54
Cycle rickshaw	0.49	0.00	0.00	0.76
Walking	0.34	0.00	0.00	0.66

From Figure 4.21, it can be seen that people without any private vehicle ownership prefer MThW rather than the other modes of transport. This may be due to the fact that people are not having reliable public transportation mode and people without private vehicle becoming virtually captive to MThW.

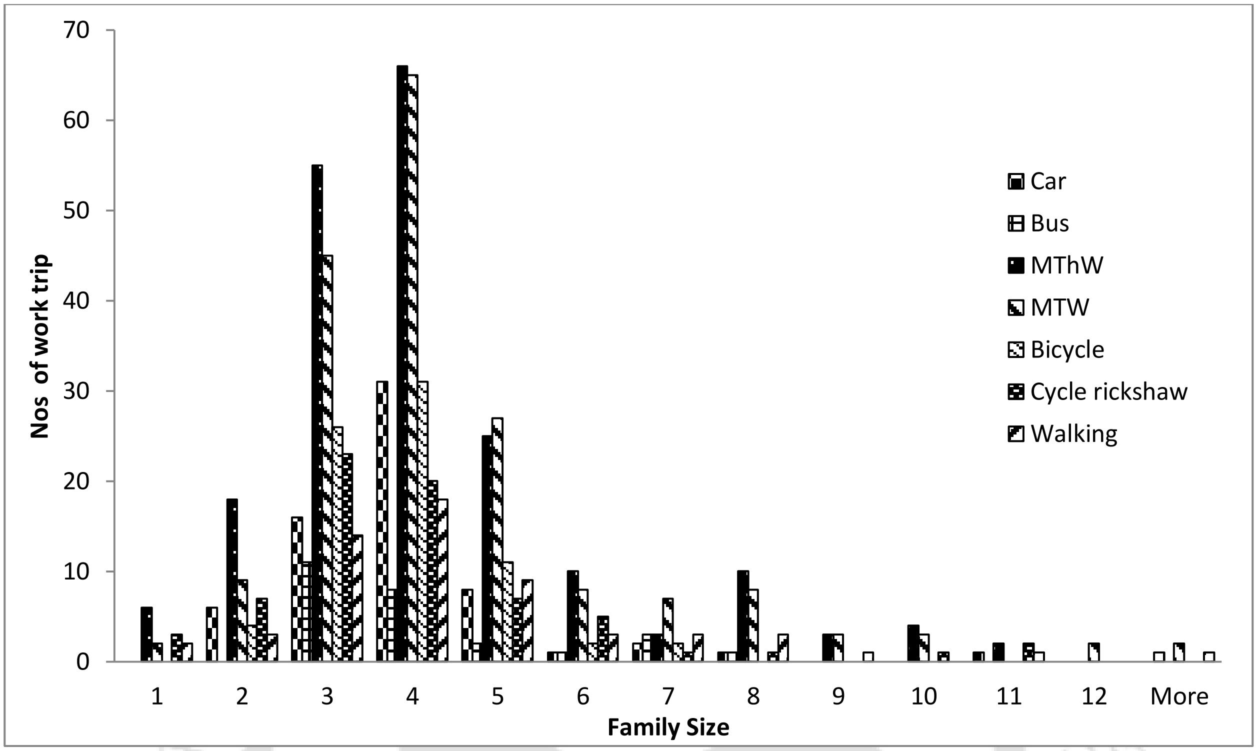


Figure 4.22: Distribution of work trips based on the family size

Table 4.15: Statistics of sampled household size for various chosen modes

Modes	Average family size	Median	Mode	Standard Deviation
Car	3.98	4.00	4.00	1.44
Bus	4.70	4.00	3.00	2.88
MThW	4.19	4.00	4.00	1.92
MTW	4.53	4.00	4.00	2.12
Bicycle	3.83	4.00	4.00	1.01
Cycle rickshaw	3.97	4.00	3.00	1.96
Walking	4.55	4.00	4.00	2.28

Figure 4.22 shows the mode wise frequency of trips with family size. The mean size of the family is around 4 people. From Table 4.15, it can be seen that the individuals from the smaller families are choosing bicycle as travel mode for work trips.

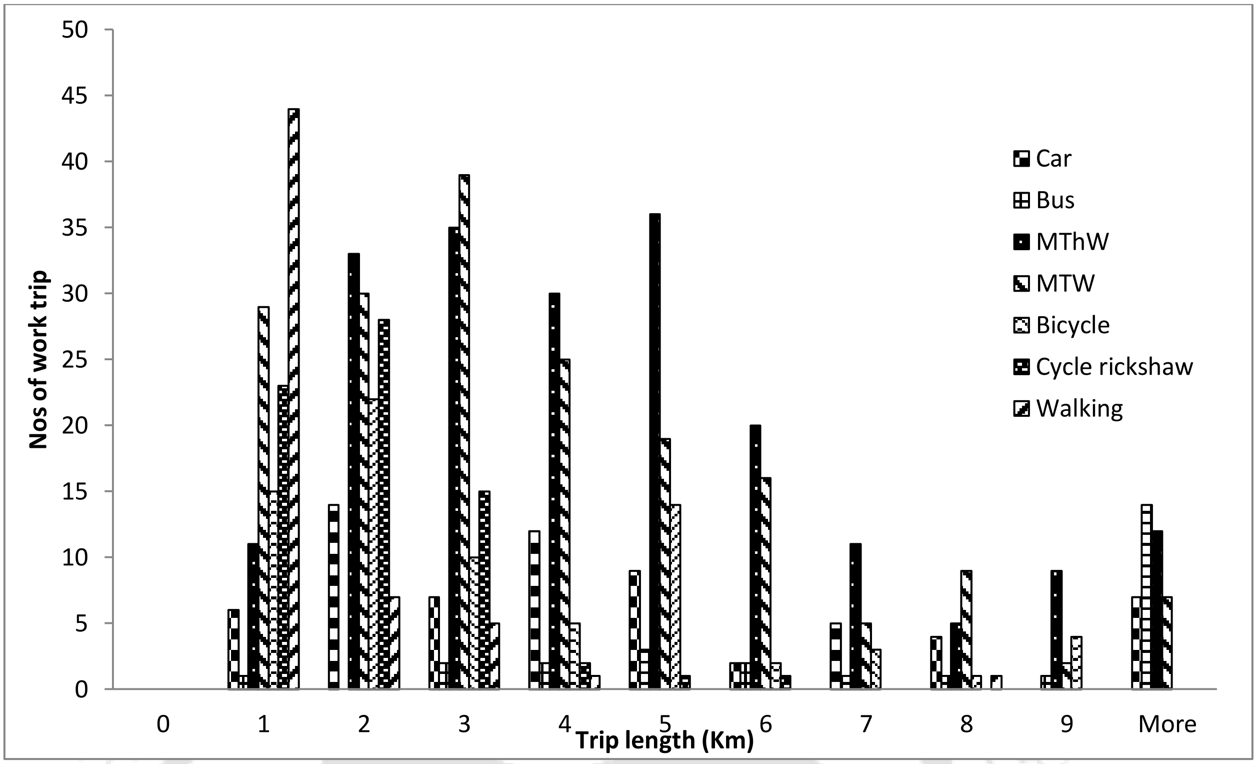


Figure 4.23: Distribution of work trips based on the trip length

Table 4.16: Statistics of the sampled individuals' trip length for various chosen modes

Modes	Mean	Median	Mode	Standard Deviation
Car	4.46	3.40	12.00	3.73
MThW	4.29	3.81	5.02	3.11
MTW	3.54	2.89	1.00	3.21
Bicycle	2.89	2.05	5.00	2.21
Cycle Rickshaw	1.56	1.41	0.51	0.92
Walking	0.90	0.55	0.19	1.12

As shown in Figure 4.23, individuals prefer bus as travel mode when trip length is more. The mean trip length for walk trips is 0.9 km and for cycle rickshaw and bicycle it is 1.56km and 2.89km, respectively. This implies that there is considerable number of workplaces near the residential locations within 1km buffer radius, which induces 9% of walk trips with mean distance of 0.9 km. People are using motorized modes when the mean trip lengths are around 3.5km. With respect to gender, it can be seen from the Figure 4.24 that female trip makers prefer cycle rickshaw and MThW as the modes of travel for work trips. MTW and bicycle are preferred modes for male trip makers.

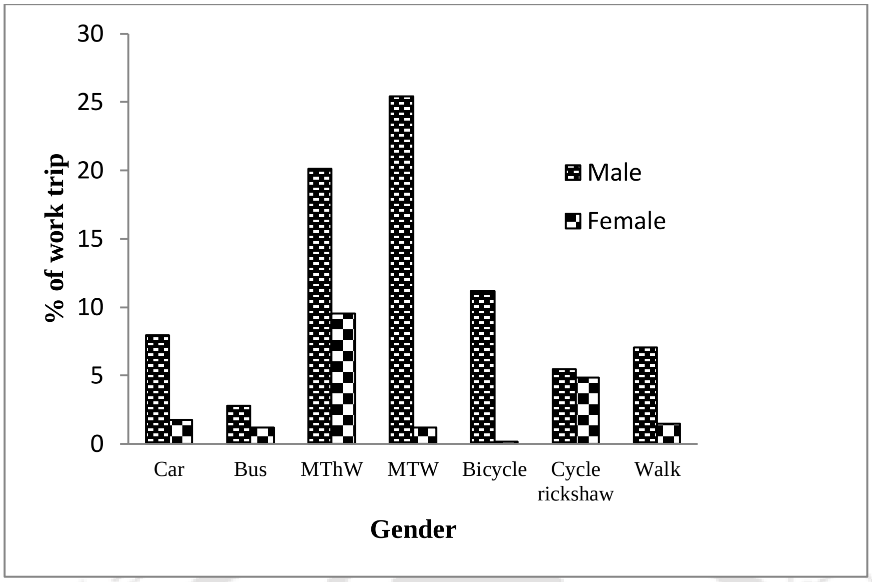


Figure 4.24: Distribution of work trips based on gender

4.5.2.4 Exploratory analysis of mode choice with respect to socioeconomic data for shopping trips

Figure 4.25 shows the frequency of trips versus income. People with income less than Rs40000.00 prefer walk mode compared to the other available modes. As per the data shown in Table 4.17, people using bicycle and MThW are having less monthly income. People above 50 years of age have predominant share of walk mode. Share of MTW is high when the age of the trip maker is less than 40 years (Figure 4.26).

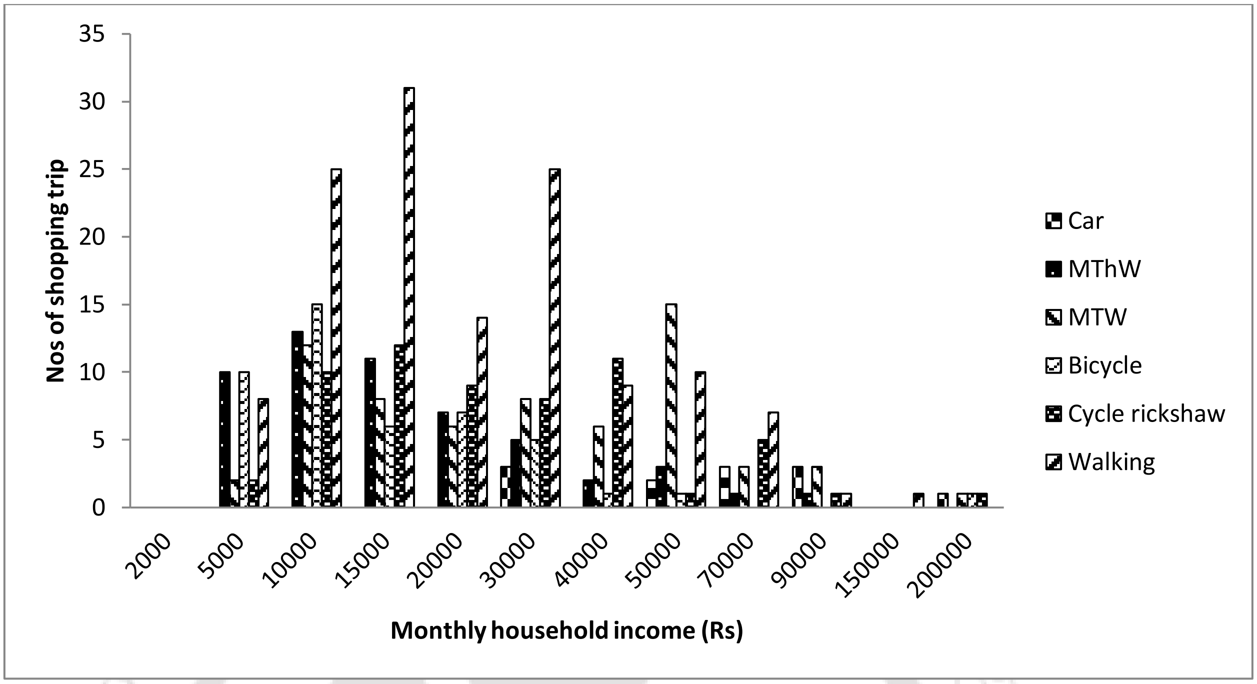


Figure 4.25: Distribution of shopping trips based on household income

Table 4.17: Statistics of the sampled individuals' income for various chosen mode

Modes	Average Income(Rs)	Median(Rs)	Mode(Rs)	Standard Deviation
Car	72500.00	70000.00	30000.00	46343.58
Auto-Rickshaw	19339.62	15000.00	10000.00	17098.07
MTW	36562.50	30000.00	50000.00	32755.41
Bicycle	18913.04	10000.00	10000.00	29076.78
Cycle Rickshaw	30666.67	20000.00	15000.00	29349.83
Walking	25763.36	20000.00	15000.00	424028185.55

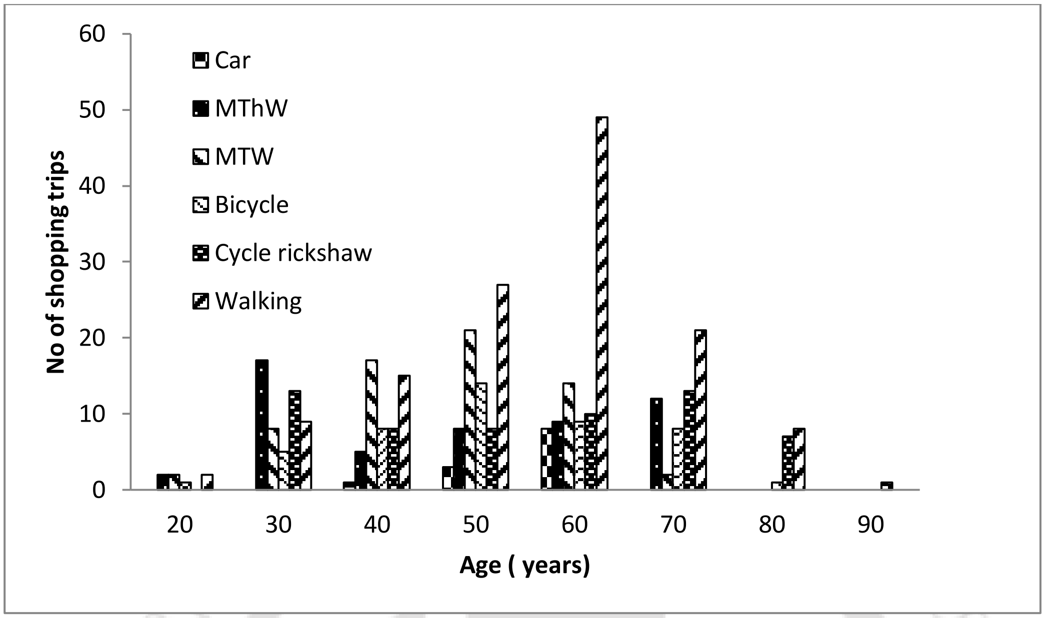


Figure 4.26: Distribution of shopping trips based on the age of the sampled individuals

Table 4.18: Statistics of the sampled individuals' age for various chosen modes

Modes	Average Age	Median	Mode	Standard Deviation
Car	50.08	52.00	52.00	5.09
Motorized Rickshaw	43.15	45.00	60.00	16.77
MTW	42.59	45.50	39.00	11.25
Bicycle	47.07	48.00	48.00	12.41
Cycle Rickshaw	49.50	52.00	62.00	17.53
Walking	51.34	53.00	55.00	13.16

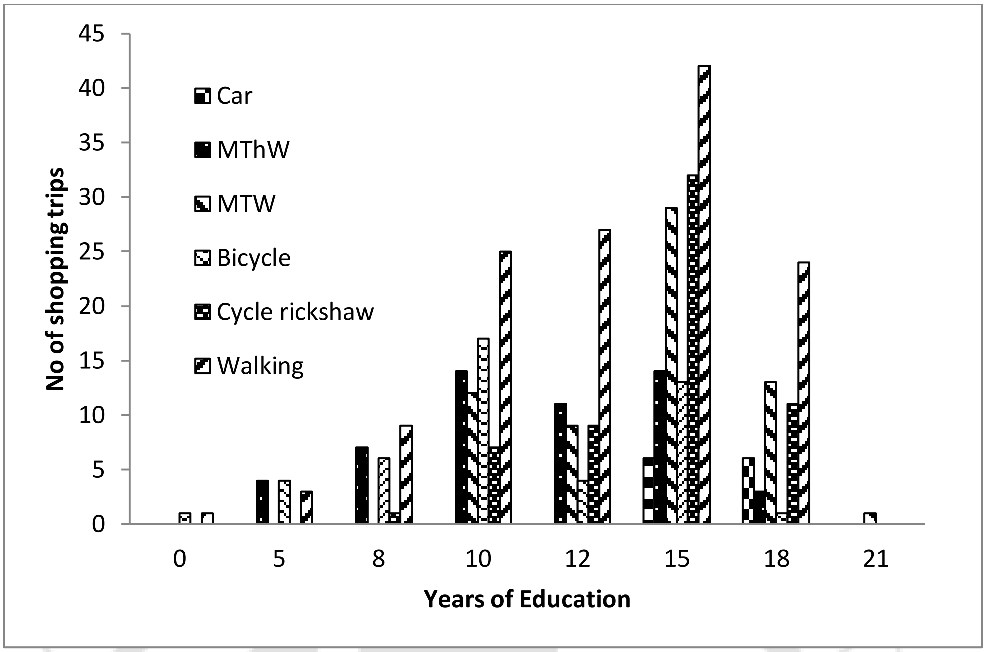


Figure 4.27: Distribution of shopping trips based on years of education

Table 4.19 Statistics of the sampled individuals' educational level for various chosen mode

Modes	Average year of education	Median	Mode	Standard Deviation
Car	15.75	15.50	15.00	0.97
MThW	11.30	12.00	15.00	3.46
MTW	13.83	15.00	15.00	2.43
Bicycle	10.61	10.00	15.00	3.56
Cycle Rickshaw	14.15	15.00	15.00	2.28
Walking	12.73	13.00	15.00	3.26

As opposed to the case of work trips, for shopping, the trip makers with 15 years of education prefer walk and cycle rickshaw (Figure 4.27). People with less years of education are choosing cycle and MThW for making the shopping trips. As can be seen from Figure 4.28, irrespective of the family size the share of the walking trips is high compared to the other modes.

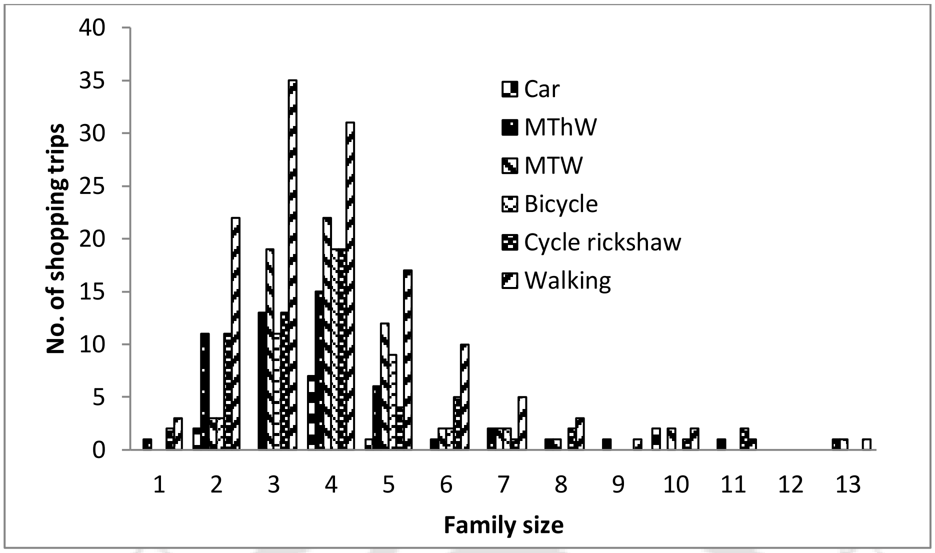


Figure 4.28: Distribution of shopping trips based on family size

Table 4.20: Statistics of the sampled individual's family size for various chosen mode

Modes	Average family size	Median	Mode	Standard Deviation
Car	4.75	4.00	4.00	2.60
MThW	4.11	4.00	4.00	2.70
MTW	4.34	4.00	4.00	1.91
Bicycle	4.04	4.00	4.00	1.13
Cycle rickshaw	4.07	4.00	4.00	2.15
Walking	4.08	4.00	3.00	2.09

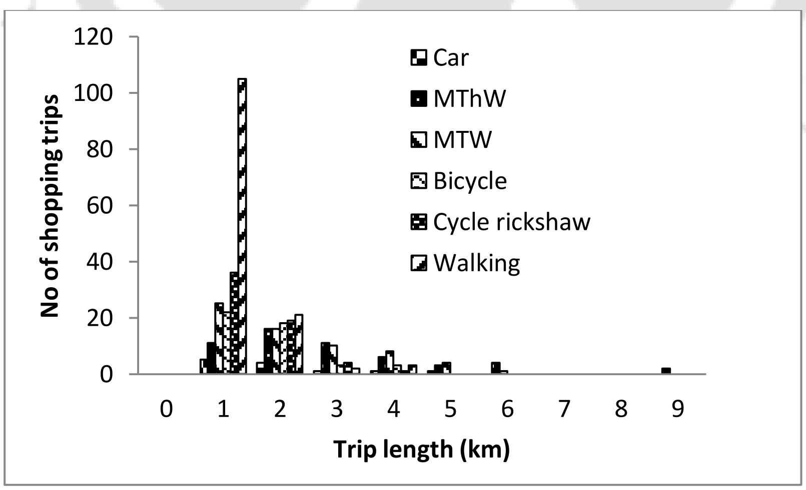


Figure 4.29: Distribution of shopping trips based on trip length

Table 4.21: Statistics of the sampled individual's trip length for various chosen modes

Modes	Average trip length(km)	Median	Mode	Standard Deviation
Car	1.57	1.55	--	1.22
MThW	2.52	1.99	0.92	1.89
MTW	1.75	1.14	0.76	1.39
Bicycle	1.12	1.04	1.15	0.76
Cycle Rickshaw	1.08	0.90	1.69	0.62
Walking	0.69	0.51	0.51	0.56

From Figure 4.29 it can be seen that most of the trips undertaken by walk mode, trip distance varies from 0-2 km. So it is clearly evident that people prefer walking while making the shopping trips when the trip distance is small. This also indicates that the presence of mixed land use reduces the average trip length.

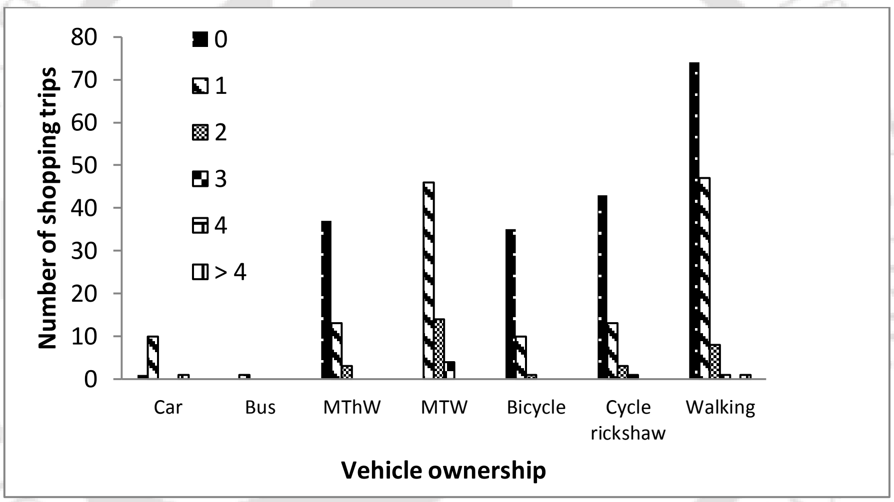


Figure 4.30: Distribution of shopping trips based on vehicle ownership (MTW + Car)

Table 4.22: Statistics of the sampled individuals' vehicle ownership for various chosen mode

Modes	Average private vehicle ownership	Median	Mode	Standard Deviation
Car	1.17	1.00	1.00	0.94
MThW	0.36	0.00	0.00	0.59
MTW	1.34	1.00	1.00	0.60
Bicycle	0.26	0.00	0.00	0.49
Cycle Rickshaw	0.37	0.00	0.00	0.66
Walking	0.55	0.00	0.00	0.81

From Figure 4.30 it can be said that those who do not own private motorized vehicles tend to use walk, cycle-rickshaw, MThW, and cycle as the modes of travel when making the shopping trips.

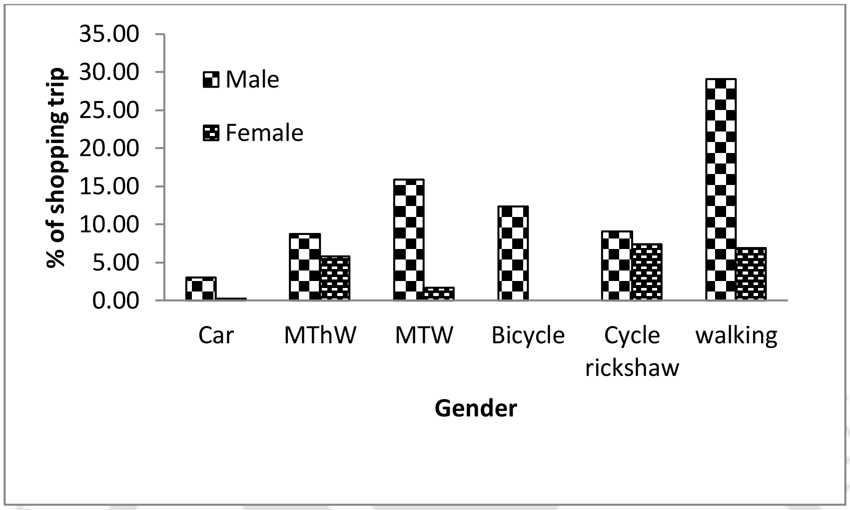


Figure 4.31: Distribution of shopping trips based on gender

Similar to work trips, female trip makers prefer MThW and cycle rickshaw in case of shopping trips (Figure 4.31).

4.5.3 Data related to perception and attitude

With an interest to understand the effect of perception of trip makers towards different modes, the questionnaire included a set of attitudinal questions. The questions were framed in such a way so that latent variables such as comfort, reliability, safety and flexibility can be estimated for each mode.

Six different modes, car, bus, MTW, MThW, bicycle, cycle-rickshaw were considered in the present analysis and their attributes were measured using Likert scale from 1 to 5 (extremely poor to excellent). Table 4.23 provides the mean rating by group for each variable. The data shown in the table indicate that car is the most comfortable travel mode compared to the other modes. In case of bus as travel mode, the level of comfort is the minimum. In case of reliability of transit and IPT, the reliability (measured using the statement 'ability to arrive on time') of bus as a travel mode is less than that of the MThW. Out of all the modes, bus and bicycle are regarded as the least reliable modes. With respect motorized modes, safety (from accident), safety (protection from weather) is least for MTW. Bus is rated to be least flexible mode.

Table 4.23: Mean perception rating for various modes

Perception attribute	Car	Bus	MThW	MTW	Bicycle	Cycle rickshaw
Comfortable in journey	4.11	2.99	3.23	3.73	3.00	3.57
Always Availability of comfortable seats	3.78	3.13	3.39	3.69	3.09	3.63
Very easy accessibility	3.69	3.08	3.50	3.90	3.08	3.31
Ability to reach destination in time	3.91	3.22	3.53	3.90	3.18	3.21
Can exactly calculate travel time prior to trip	3.90	3.16	3.37	3.91	3.10	3.21
safety from accident	3.83	3.55	3.43	3.24	2.95	3.15
Safety from theft	3.92	3.18	3.35	3.69	3.17	3.17
Safety from weather	3.89	3.55	3.49	2.20	2.20	3.01
Ability to make more trips	3.85	3.03	3.51	3.68	2.72	3.02
Can travel without changing vehicles	3.94	3.09	3.40	3.93	3.93	3.16

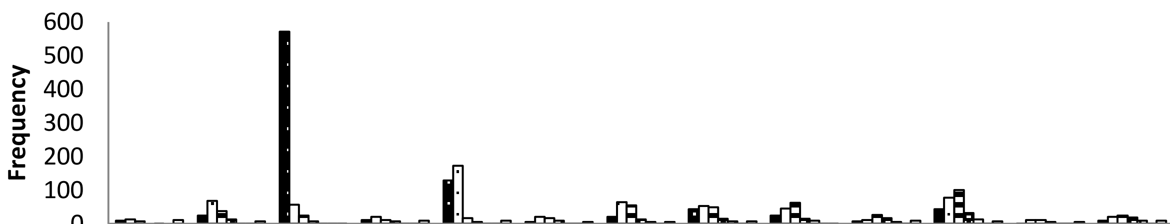
Agreement/disagreement statement data were collected to measure the effect of attitudes and perceptions on mode choice behavior. Table 4.24 shows the mean ratings of agreement disagreement statements.

Table 4.24: Mean ratings of agreement/disagreement statements

SN	**Agreement/disagreement statement**	Mean Rating
1	Personal vehicle (Car or MTW) is comfortable.	4.5
2	I am not comfortable when I travel with people I don’t know.	3.26
3	I always use the most convenient mode of transportation regardless of cost.	3.66
4	I always use the fastest route to destinations even if I have cheaper alternatives.	3.31
5	If fuel prices goes up further less likely to drive car to work.	3.01
6	If fuel prices goes further up less likely to drive MTW to work.	2.79
7	If I use public transport instead of car or two wheeler I have to cancel some of the activities.	3.68
8	It is hard to take public transport when travelling with children.	3.85
9	It is hard to take public transport with bags & luggage.	3.92
10	I know fully what buses I should take regardless of where I am going.	2.72
11	I need to have more flexibility to make many trips during working hours.	3.69
12	I don’t mind walking few minutes to destination.	3.96
13	People riding a bus help in reducing congestion.	2.99
14	Willing to pay more tax to improve bus service.	3.61
15	Bus is chosen when no other option is available.	3.70
16	Using bus service is cumbersome.	3.42
17	I don’t like transferring vehicles in the route.	3.92

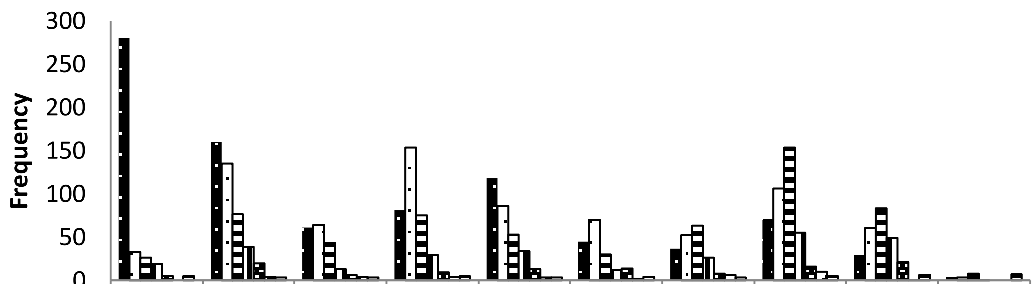
18	I will put up with crowds if I can reach destination early.	2.79
19	Walking on the road is comfortable.	2.92
20	Comfortable in walking without footpaths.	2.16
21	Trees provide ample shade in the roads near my house	3.39
22	There are interesting houses to look at in the roads near my house	3.44
23	Comfortable in walking in local shopping areas	2.96
24	Shared MThW is comfortable.	2.74
25	I would be willing to pay more if it would help environment	3.72
26	I would be willing to switch to different mode if it helps environment	2.88
27	Use of bus will improve the environment.	3.54
28	I am willing to change to bicycle mode if proper bicycle facilities/infrastructure is available.	2.95
29	I am willing to change to walk mode if proper pedestrian facilities are available and trip length is short.	3.76
30	It is safe to walk in our locality during day time.	3.70
31	It is safe to walk in our locality during night time.	3.55
32	Good footpaths are available in our locality.	2.05
33	Road facility to our locality is good.	3.46
34	Importance amenities, post office, bank etc nearby.	3.59
35	Good Para-transit (Motorized rickshaw/Mini-van) service is available to & from our area.	3.37
36	Frequent bus service is available to and from our area.	2.43

In order to understand the reason for not choosing bicycle and bus as travel modes, questions were asked about the reasons behind not choosing these modes by ranking suitable attributes. Figure 4.32 and 3.33, respectively summarize the perception ranking for not choosing bicycle and bus as travel modes.



	Inferior Good	Congested Roads	No Bicycle	Poor quality of Roads	Socially not acceptable	Difficulty in crossing intersection	Not Safe	Requires physical activity	Distance	Climate	Not Comfortable	No Bicycle facilities on the way to work	Time saving
Rank1	8	23	570	11	128	4	21	43	23	6	42	2	9
Rank2	12	67	56	21	172	20	63	52	45	11	77	10	21
Rank3	7	37	24	10	17	17	54	48	62	25	100	11	23
Rank4		12	6	6	5	9	13	15	15	16	31	5	19
Rank5	2	3	2	3	1	3	5	6	9	5	13	2	9
Rank6		2	1	1	2	1	3	1	3	3	2	1	1
>Rank6	11	7	2	8	9	4	4	6	2	9	6	5	9

Figure 4.32: Perception ranking for reasons not choosing bicycle as travel mode



	Not Available	Not Reliable (Not able to arrive on time)	Not flexible	Less frequency	Not comfortable	Bus stop is far away.	No direct bus	Takes more time	Seats are not available	Costly
Rank1	280	160	61	81	118	45	37	71	29	4
Rank2	33	135	64	154	86	70	52	106	60	3
Rank3	26	77	43	75	53	30	63	154	83	8
Rank4	19	39	13	29	34	12	26	55	49	1
Rank5	5	20	6	9	13	14	8	16	21	
Rank6		4	4	4	3	2	6	10	1	
>Rank6	5	3	3	5	3	4	3	5	6	7

Figure 4.33: Perception ranking for reason not choosing bus as travel mode

In case of bicycle mode choice, socially not acceptable and less comfortable are the two important reasons behind not choosing the cycle. In case of bus, unavailability from home, reliability and less comfort are the reasons given by the individuals for not choosing it.

4.6 Summary

This chapter presented the data that were collected as a part of the household survey conducted in the study area. Revealed preference questionnaire, with household details, personal details, vehicle ownership, perception and attitude related questionnaire along with SP survey completes the travel survey that was used in the subsequent chapters in estimating various econometric models.

From the exploratory data analysis it has been observed that the share of transit mode is higher (27%) for educational trips and percentage of NMT mode (65%) is higher for shopping trips. In case of work trips, individuals with lower income are using bicycle and walk modes. Individuals, with income less than Rs 40000.00 per month, prefer MTW and MThW compared to the other modes. People in the age group of 50-60 prefer car as travel mode. The mean lengths of walk, bicycle and cycle rickshaw trips are 0.9 km, 2.89

km and 1.56 km, respectively. In case of gender, female trip makers prefer MThW and cycle rickshaw whereas the male trip makers prefer motorcycle and bicycle.

In case of shopping trips, old people prefer walking as travel mode. In this case, the average trip length by walking is 0.69 km which is less than that of work trip. Considering the mean perception rating from the sample data, car is perceived to provide higher comfort and safety. Reliability (measured as the ability to reach destination in time) of transit mode is perceived to be the least. Flexibility and Reliability of MTW is higher but has poor perception with respect to safety. Further, in case of bicycle mode choice, socially not acceptable and less comfort are the two important reasons behind not choosing cycle. In case of bus, poor accessibility from home, reliability and comfort are the reasons given by individuals for not choosing bus.

Chapter 5

Analysis and modeling of mixed land use and its effects on travel parameters

Land use mix is one of the important measures of land use development pattern. It refers to the diversity of land uses within an area. When diverse land uses exist in a given area (generally census tract or municipal ward), it is expected that many trips originating from that area may have destinations in the same area. Land use mix is generally characterised using various indices like entropy, dissimilarity index, gini coefficient, herfindahl index, etc. The relationship between the land use mix and travel behaviour has been widely studied and evaluated by the researchers. In this study, effect of mixed land use on travel parameters has been studied in the context of smaller Indian cities. To understand the relationship between the land use and travel parameters, land use mix must be quantified using appropriate land use parameters. Various indices that quantify the mixed land use have been formulated and some commonly used land use indices have been modified suitably. Travel parameters related to work and shopping trips have been used to analyse the effects of land use mix. Detailed analysis of land use mix quantification using various existing parameters and the proposed new parameters is provided in the following sections.

5.1 Land use mix observed in the study area

Data collection methodologies, adopted in collecting the land use details of the study area, have been discussed in the chapter on data collection. Figure 5.1 shows the land use composition details observed in the Agartala city. From this figure, it can be observed that the study area includes significant quantity of vacant land, agriculture land, and water bodies. Once the vacant land and the land related to the water bodies and agriculture is neglected, residential land use is predominant and the remaining land uses are relatively small (Figure 5.2).

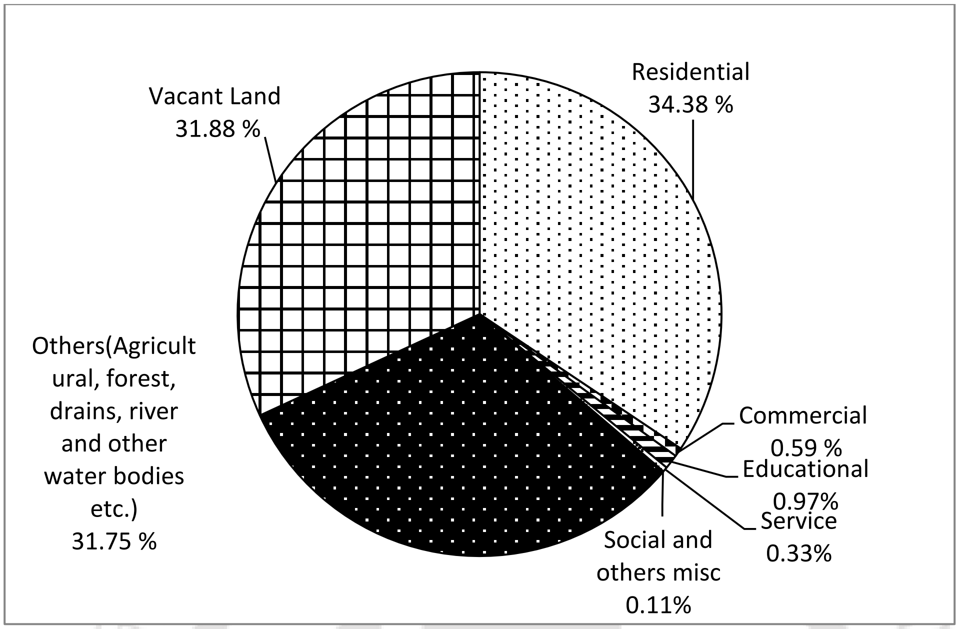


Figure 5.1: Land use distribution observed in the study area

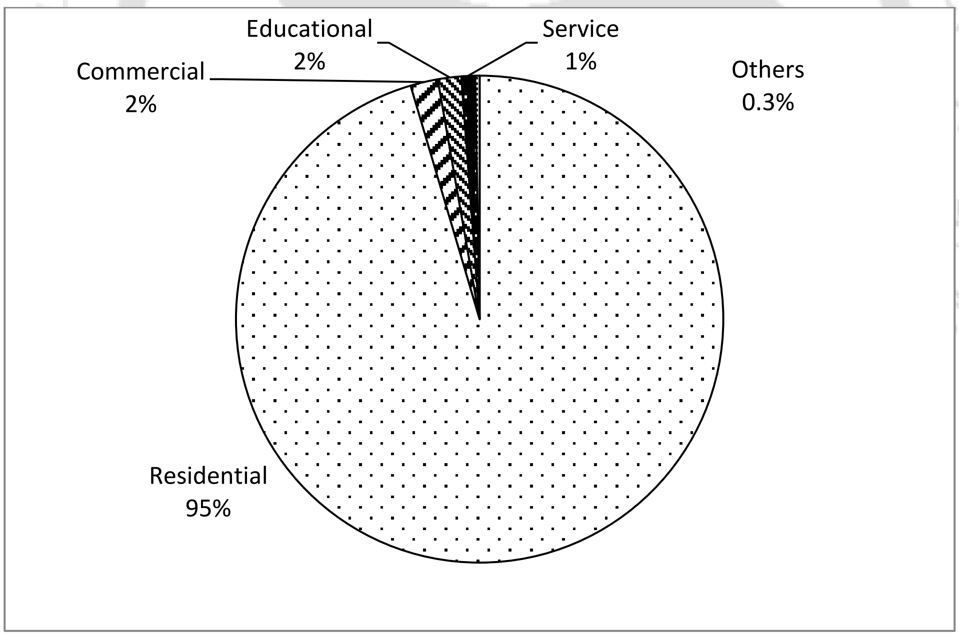


Figure 5.2: Land use distribution of the study area after excluding the vacant land and the land related to agriculture and water bodies

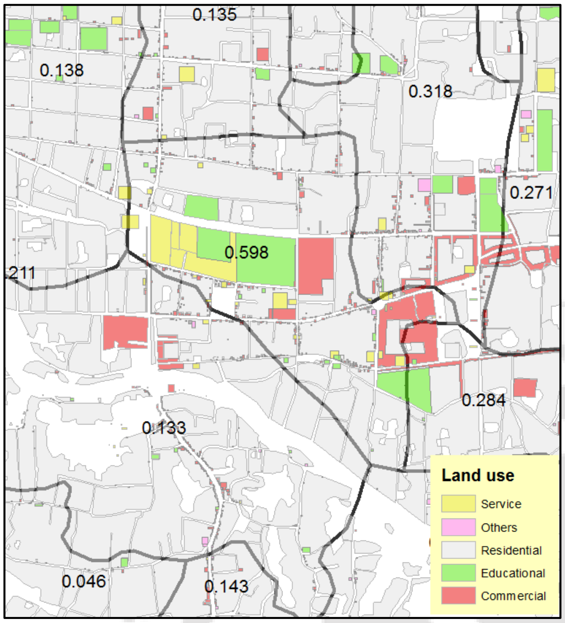


Figure 5.3: Entropy index measured for some of the census tracts of the study area

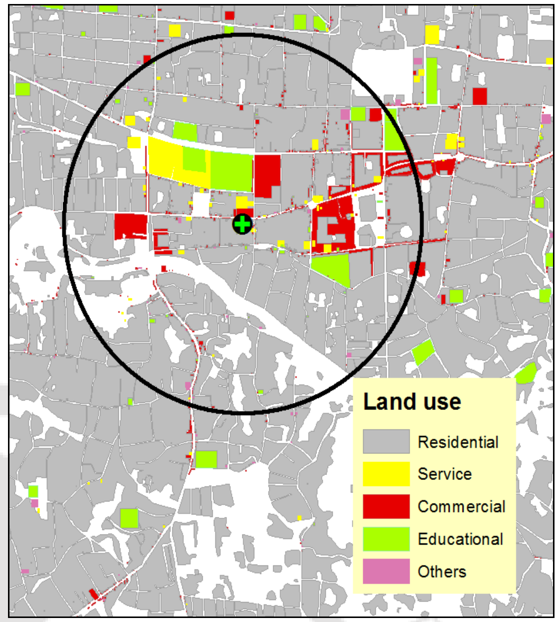


Figure 5.4: 1000 m buffer created around a sampled household for calculating the entropy index

5.2 Entropy Index

Entropy index is the most widely accepted and commonly used index for quantifying the land-use mix. Entropy quantifies the homogeneity of land use in a given area. Entropy is expressed as,

$$\text{Entropy} = \sum_j P_j \times \frac{\ln(P_j)}{\ln(J)} \tag{5.1}$$

where,

P_j = the proportion of the area of j^{th} land use category found in the tract being analyzed, and,

J = total number of land uses considered

Since the entropy is normalized using natural logarithm of the number of land uses, its value lies between 0 and 1, where 0 represents homogenous land use, and one indicates the tract of land is equally distributed across all the land use types.

5.2.1 Limitations of entropy index

The main drawback of the entropy index is that it could not represent the intensity of land use mixing properly. Entropy index takes the same value for two different scenarios having different land use patterns if the proportion of land use mix is same. Also, when the above discussed approach is used (in calculating the entropy for census tract or municipal ward as shown in Figure 5.3) in quantifying the mixed land use observed in smaller Indian cities, many households with varying travel behavior were characterized with similar entropy value. This is resulting due to the fact that people with varying socioeconomic backgrounds co-exist in the same census tract or municipal ward. To overcome this problem to some extent, entropy index is measured for each sampled household instead of the census tract. For this purpose, buffer of 1km radius (Figure 5.4) was created around each of the sampled household. Entropy Index values were computed for the buffer area created around the sampled households.

5.3 Dissimilarity Index (DI)

The dissimilarity index was used to compute the dissimilarity among the grid cells, constituting a tract (Cervero and Kockelman, 1997). According to Cervero and Kockelman (1997) dissimilarity index is calculated based on the points awarded to each actively developed hectare cell on the basis of the dissimilarity of its land use from those of eight adjacent hectare cells (see Figure 5.5a). The average of these point accumulations, across all the active hectares in a tract, is the dissimilarity index for that tract. It is calculated using the following equation;

$$\text{Dissimilarity Index} = \sum_k \frac{1}{K} \sum_i^8 \frac{X_{ik}}{8} \tag{5.2}$$

where,

K = number of actively developed grid-cells in a census tract or municipal ward,

X_{ik} =1 if land-use category of neighboring grid-cell differs from the central grid-cell, and

X_{ik} = 0, otherwise,

In calculating the DI, tract of land is divided into actively developed land parcels of uniform size known as cells. Thus, the DI presents more information about the type or intensity of mixing compared to the entropy index. The first step in calculating the DI (DI) is to find the points to be awarded to each cell based on the comparisons of the land use of the subject cell with that of the eight neighboring cells. Figure 5.5(a) represents a tract of land divided into cells with each cell representing a land use (R- Residential, S=Service, C=Commercial). The central cell is surrounded by three commercial and three service land uses which are different from its own land use. Hence, the central cell is awarded 6/8 points. Similarly, for all the developed cells (i.e., cells without vacant lands) DI was found out and averaged for the entire census tract. Similar to the entropy index, DI approaches the value of 1 as the land use mix increases.

C	R	C
C	R	S
S	S	R

(a)

C	R	C
C	R	C
C	C	R

(b)

Figure 5.5: Hypothetical land uses for awarding points to the central cell

5.3.1 Limitations of DI

As explained earlier, the points awarded to each actively developed cell play a major role in the calculation of DI for a given tract of land. Points were awarded based on the comparisons of the land uses of the eight neighboring cells with that of the subject/central cell. As shown in Fig 5.5, two different land use configurations around a residential land use, (a) having three commercial, three service and two residential land uses out of the eight neighboring cells; (b) having six commercial land uses and two residential land uses out of 8 neighboring cells; results in 6/8 points being awarded to the central cell.

The DI does not consider the type of land use of the adjacent cell, thus neglecting one of the important determinants of travel behavior, namely, land use interaction between the adjacent cells. Also, DI does not consider the mix of the land uses. Thus, the DI represents the dissimilarity of the adjoining cells but doesn’t incorporate the information about number of land use types around the central cell. To overcome this drawback, it is necessary to incorporate the information of number of land use types around the cell as well as the interaction between the adjacent cells. A new measure, termed as 'mix type index', is proposed in this study to consider the land use types around the central cell.

Another drawback of DI is the cell size, using which the census tract is divided. In many of the past studies, a cell size of 100m x 100m was used in the calculation of DI. Each cell of the census tract is allotted a land use type based on the dominant land use observed in that cell. With reference to Figure 5.7(a), the cell in the top right corner would be considered as residential land use as the residential area is dominant in that cell. In smaller Indian cities like Agartala, within the residential zones many commercial and educational establishments can be seen (Figure 5.7(a)). Hence, if bigger cell such as 100m x 100m is considered, there is a possibility that the information related to such land uses may be neglected. To represent these land uses, the cell size must be reduced so that small area features would also get due importance.

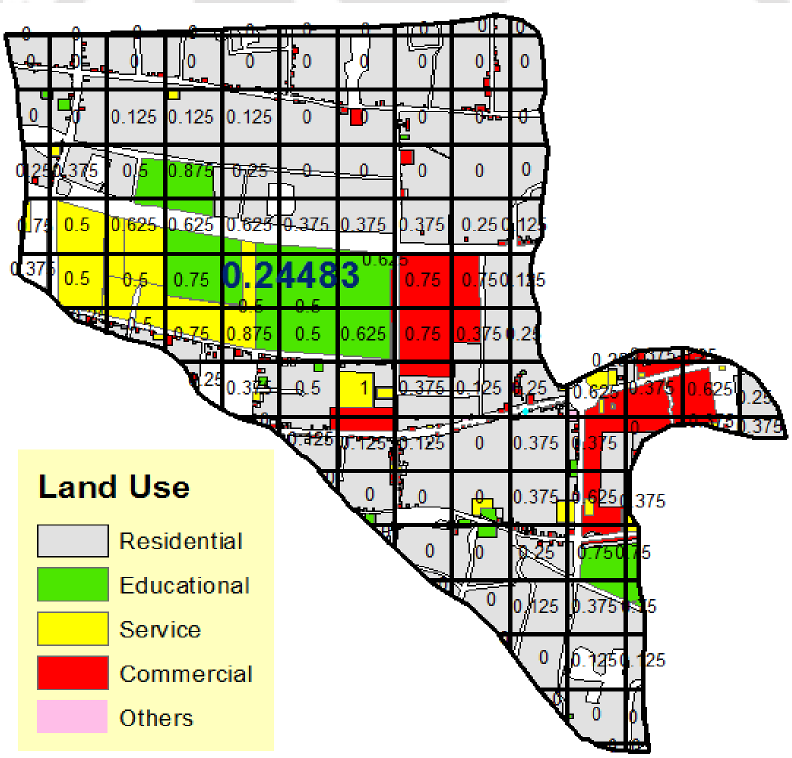


Figure 5.6: Points alloted to 100m x 100m cells for estimating DI at tract level (Ward 22, AMC) and estimated DI for the census tract

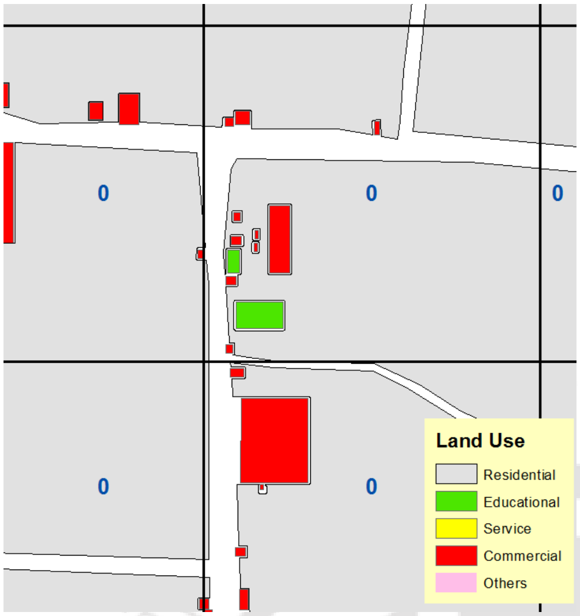
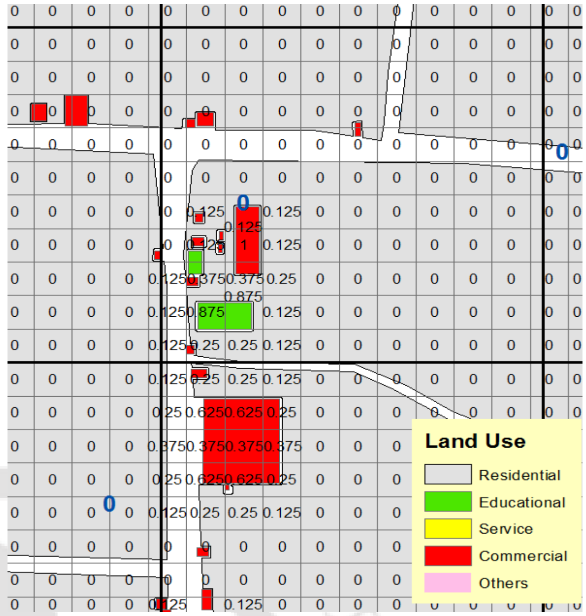


Figure 5.7 (a): Points alloted to 100m x 100m cell for estimating DI



5.7(b): Points alloted to 10m x 10m cell for estimating DI

In this study, 100m x 100 m, 50m x 50m, and 20m x 20m and 10m x 10m cell sizes were used and the frequency of cells, attributed with different land uses, has been computed. It has been observed that cell of 10m x 10m could capture the effect of minor land uses in a reasonable manner. When 10m x 10m cell size is considered, there was a two fold increase in the cells attributed with commercial land use. Figure 5.6 shows the DI value measured for a census tract using 100m x 100m cell (Ward no 22, Melarmath). If keenly observed from this figure (specifically the top right corner), it can be seen that many cells are awarded with zero points. Figure 5.7 (a), 5.7(b) shows point allotted to a part of this ward, one with 100 m cell size and other at 10 m cell size. Though the points attributed to a 100m x 100m cell is zero (since all the neighboring cells are of residential land use) when the same cell is further divided into 10m x 10m cells (7.b), some of the cells are attributed with non-zero points. Table 5.1 shows frequency of dominant land uses corresponding to different cell sizes. Further, instead of calculating the DI for a census tract, various tract sizes such as 60m x 60m, 100m x 100m, 500m x 500m and 1km x 1km have been tried. It was found that 500m x 500m and 1km x 1km tract sizes were relatively better in capturing the effect of land use on travel behavior. More specifically, 1km x 1km tract was found to be useful in capturing the effect of land use on non-motorized trip parameters. Hence, in further analysis 10m x 10m cell, and 1000000 square meters (1km x 1km) tract have been used.

Table 5.1: Frequency of dominant land uses corresponding to different cell sizes

Cell sizes	Frequency of dominant land uses in the study area				
	Commercial	Residential	Others	Educational	Service
100m x 100m	113	5180	22	115	37
20m x 20m	1880	80787	279	1991	721
10m x 10m	7425	243124	2593	6274	40039

5.4 Mix type index

As discussed earlier, DI has limitations in quantifying the land use mix of a census tract. To overcome this, a new measure called "Mix type Index" is proposed in this study. This measure allots points to each of the actively developed cells based on the mix of the land uses in the surrounding cells.

$$\text{Mixed type Index} = \frac{1}{K}\sum_k \frac{x_k}{(No.of\ distinct\ land\ uses\ in\ the\ study\ area)} \tag{5.3}$$

Where, x_k is the no. of distinct land uses observed in the surrounding cells of k including it; K is number of actively developed cells in a tract.

With reference to Figure 5.5, intensity of mixing is more in case of Figure 5.5(a) as there are three types of land uses present in the eight adjoining cells. On the other hand, in case of Fig 5.5(b), only two distinct land uses can be seen in the eight adjoining cells. Using the proposed index for case (a), the points allotted to the central cell is 3/5 (3 types of land uses present and five types of land uses considered in this study) and the points allotted in case (b) is 2/5. It can be said that the new index is able to incorporate the land use mix and also information about the differences in the land use around the central cell.

5.5 Area index

Hess et al. (2001) have explained the need to develop an index based on land use functional and spatial complementarity. Land use functional complementarity ensures the consideration of origins and destinations that are likely to be linked by travel. Land use spatial complementarity ensures that the land uses linked with travel are within adequate proximity. It was assumed that the mode choice would be different for different purposes inside the buffer zone of 1 km radius created around the trip origin (household in this case), to the outside of the buffer zone, thereby incorporating the land use spatial complementarity as well as functional complementarity. To adequately represent the spatial and functional complementarity, a new index termed as 'area index' has been proposed in the study.

Area index for work trip is defined as the ratio of the work areas in the buffer zone to the work areas in the whole study area, including that of the buffer zone. Commercial area, service area, and industrial area were considered as work areas. It was considered that the residential land uses are linked by travel to the work areas, thereby incorporating the land use functional complementarity. The ratio, when close to 1 indicates that most of the work places lie in the buffer zone, thus more non-motorized trips would be realized. The ratio when close to 0 indicates that most of the work places are outside the buffer zone and more motorized trips may be realized. Thus, area index explains the relationship between mode choice behavior and the amount of particular land use area available in the vicinity of the household. When area index was used for the analysis of shopping trips, the index value was computed based on the shopping space available in the buffer zone and the total shopping space in the study area. Area index values for the destinations of the trips have also been computed based on the hypothesis that the work and shopping space available in the vicinity of the destination influences the individual to make the non-home based trips.

5.6 Extraction of Land use data using ArcGIS

Land use parameters have been extracted by using ArcGIS software. Most of the procedure discussed below has been used by Ma and Chen (2013), except slight modifications. Figure 5.8 provides steps used in calculating DI and Mix type index. Detailed description of the same is given below;

- All the land use layers were added to ArcGis10.

- The grid cell was used as the unit of land use in computing the dissimilarity index. Therefore, a layer of grid cells was created over the entire study area. The steps used to create the grids include: Arc Toolbox/Data Management Tool/Feature Class/Create Fishnet. In the attribute table of fishnet layer, there is an id no. (FID) for every cell. A separate field, uni_id, is added to the attribute table in which same ids were inserted using field calculator option. This is done to avoid any problem in spatial joining.
- After creating the fishnet polygon layer, it was intersected with all the collected land use layers separately, one by one, using the intersect operator of the Arc Toolbox (Arc toolbox/analysis tools/overlay/intersect) of ArcGis10.
- On every intersected layer, the total land use area, falling on the each cell, was calculated using the summarize option with reference to uni_id field of the attributes of the intersected layers.
- The attribute table of fishnet is then exported to MS Excel, where, dominant land use was found out for every uni_id of fishnet, and a land use code was given for each of the land uses. Using the VLOOKUP function available in MS Excel, the first step of finding DI i.e. the points to be awarded to each of the cells were calculated. This table was then exported to ArcGis10 and joined to fishnet layer.
- The next step was to find out the number of actively developed cells in a given tract of land. To do this, another fishnet of bigger size was created over the study area (in this study, 1000m x 1000m have been used). This new fishnet layer was then intersected with the previous fishnet. The attribute table of the new fishnet layer was exported to MS Excel. In the MS Excel sheet, pivot table was used to calculate the number of developed land uses falling in each tract of land i.e. the K value. Also $\sum$ (Xik/8) was calculated for each cell of the tract. Finally, dissimilarity index was calculated using the formula for DI. The same concept was applied for finding the Mix type index, using slightly different commands of MS Excel.

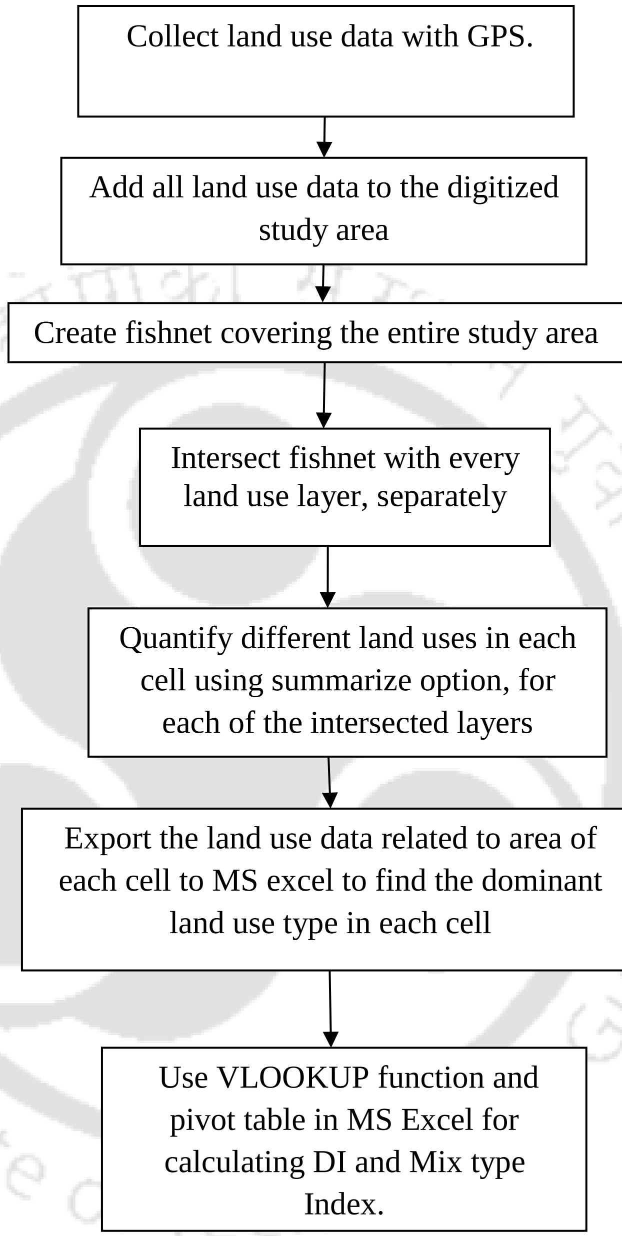


Figure 5.8: Steps for calculating DI and Mix type index

As mentioned earlier, entropy index was calculated for the municipal wards as well as for buffer zones of 1 km radius. Steps followed in calculating the Entropy index and Area index were given below:

- All the land use layers were added in ArcGis10.
- The layer of the household data which is a point shape files was added in ArcGis10.
- Buffer zones were created with a radius of 1km around the households.
- The new buffer layer was then intersected with every land use layer. Using the summarize option, all the land use areas falling in the buffer were calculated for every intersected layer.
- The attribute table of the household layer was then exported to MS Excel. In the MS Excel, using the corresponding formula, Entropy Index and Area Index were calculated.

5.7 Results and Analysis

The effect of socioeconomic and land use variables, on the travel parameters, is analyzed in this section. Only work and shopping trips have been considered for this analysis. Work trip indicates the travel from home to work, and the shopping trip indicates the travel from home to shop. Data used for the present analysis contain 686 work related trip information and 367 shopping related information. Return trips were not considered in the present analysis. As mentioned earlier, two types of analysis have been carried out, namely, studying the effect of land use parameter in explaining the variability of trip length (for work trips, mean trip length = 2.96 km and standard deviation = 2.40 km; for shopping trips mean trip length = 1.32 km and standard deviation =1.30 km), and to analyze the effect of land use on mode choice for work and shopping trips. In analyzing the former, linear regression models were prepared to understand the ability of land use parameters in explaining the variability of the trip length. In the second analysis, the effect of land use parameters on the utility of the various modes has been analyzed.

Table 5.2 Correlation matrix between different land use parameters

Land use parameters	DI(cell size of 100 m in census tract)	Entropy (census tract)	DI(10 m x 10 m cell in 1km x 1km tract)	Mix type index(10 m x 10 m cell in 1km x 1km tract)	Entropy (1000 m buffer)	Area Index at origin (1000m)
DI(cell size of 100 m in census tract)	1.00					
Entropy (census tract)	0.94	1.00				
DI(10 m x 10 m cell in 1km x 1km tract)	0.63	0.62	1.00			
Mix type index(10 m x 10 m cell in 1km x 1km tract)	0.62	0.62	1.00	1.00		
Entropy (1000 m buffer)	0.64	0.66	0.71	0.71	1.00	
Area Index at origin (1000m)	0.70	0.71	0.80	0.80	0.90	1.00

5.7.1 Effect of mixed land use on trip length

A base linear regression model was prepared considering the socioeconomic characteristics that are having strong influence on trip length for shopping and work purposes. In the subsequent models, only one land use parameter entered the model at a time as the land use parameters are correlated with each other (Table 5.2). Entropy measured using conventional approach, and using a buffer zone of 1000m radius, DI measured using conventional approach, using 100m x100m, and 10m x 10m cell sizes with 1000m x 1000m tract size, mix type Index measured using 10m x 10m cell with 1000m x 1000m tract, area index measured at origin and destination of the trip using 1000 m buffer, were considered in this analysis. Further a variety of land use development pattern measures such as the distance to transit, intersection density, centeredness, and road network length in a given area were also tested. Description of variables used in modeling is given in the Table 5.3.

Table 5.3: Description of variables used in modeling travel behavior.

SN	Variables	Description	
1	Age	Age in year	
2	Private Vehicle Ownership	Number of private vehicles in a household	
3	Education	Years of education.	
4	Gender	Dummy variable 1 for male respondent 0 for female respondent	
5	License	1 for respondent having license 0 for respondent not having license	
6	Family Size	Number of members in the family	
7	Income	**Gross monthly household income**	**Code**
		0-2000	1
		2001-5000	2
		5001-10000	3
		10001-15000	4
		15001-20000	5
		20001-30000	6
		30001-40000	7
		40001-50000	8
		50001-70000	9
		70001-90000	10
		90001-150000	11
		>150001	12
8	Bicycle Ownership	Dummy Variable 1 for households having bicycle 0 for households having no bicycle	
9	MTW Ownership	Dummy Variable 1 for households having Motorized Two wheeler 0 for households having no Motorized Two wheeler.	
10	Time	In vehicle travel time.	
11	Cost	Fare or fuel cost	
12	Area Index Entropy with buffer Entropy index(census tract) DI (census tract) DI(10m cell over 1km^2) Mix type index (10m cell over 1km^2)	Land use parameters obtained from GIS analysis.	
13	ASC_{Car} ASC_{Bus} ASC_{MThW} ASC_{MTW} $ASC_{Bicycle}$ $ASC_{Rickshaw}$ ASC_{Walk}	Alternate specific constants for different modes.	

In case of trip length for work purpose (Table 5.4), all the land use parameters were found to be significant in explaining the variability observed in trip length data. Entropy measured with 1000 m radius buffer, and DI measured using 10 m x 10 m cell with 1km x 1km tract, were found to be more significant than the conventional entropy index and dissimilarity measured for census tract. When the area index for both the origin and destination were considered, there was almost 395% increase in the model's ability in explaining the variability of trip length compared to the base model. It has been observed that with increasing land use mix at destination, the trip length reduces. This is due to the fact that with decrease in land use mix at the destination of the work trip, the destination would be out of the CBD area (where mix is high), as a result trip distance is higher. Further, when intersection density was considered there was significant improvement in the model. The coefficient of land use mix and intersection density at trip origin is negative which means an increase in intersection density and land use mix tends to reduce the trip length. From the model, it can be seen that older people generally have shorter trip length. Private vehicle ownership has significant effect in explaining the variability of trip length. With increase in private vehicles in a household, the trip length increases. As shown in Table 5.5, for shopping trips also land use parameters were found to have significant effect on trip length. All the land use mix parameters were found to be significant, when entered the model separately. In case of both shopping and work trips, older persons seem to go to closer destinations (shorter trip length). Similar to work trip, both the coefficient of land use mix and intersection density at the origin of the trip is negative, thereby reducing the trip length. Apart from age, private vehicle ownership has a significant effect in explaining the variability of trip length in the modeling.

5.7.2 Effect of land use on mode choice

To study the effect of mixed land use on mode choice, several binary logit models have been formulated. From the literature review, it has been observed that the mixed land use has significant impact on the probabilities of choosing non-motorized and public transit modes. In this study also binary logit models have been estimated to find the probabilities of choosing non-motorized and public transport modes. Models have been prepared for work trips as well as for shopping trips. Choice analysis of non-motorized and motorized, private motorized and public/intermediate public transport (IPT) has been carried out using

BIOGEME software (Bierlaire, 2003 and Bierlaire, 2008). MThW, which carries passengers on shared basis, has been considered as IPT. Finally, an MNL model has also been estimated to understand the significance of various land use parameters on the mode choice for work trips. MNL model has been estimated to understand the significance of land use parameters in explaining the observed component of the preference heterogeneity which otherwise is explained using the socioeconomic characteristics.

Result from the binary logit model estimated on motorized and non-motorized modes for work trips is shown in Table 5.6. Sign of the coefficients in the model are logically correct. Out of various socioeconomic variables tried in this study, years of education, gender, license status, and private vehicle ownership were found to be significant. As expected, having driving license, motorized vehicle ownership, and years of education reduces the utility of non-motorized modes. With respect to gender, male generally uses non-motorized modes. An increase in area index at origin and destination was also found to be enhancing the utility of non-motorized travel. Even when controlling for the socioeconomic characteristics, land use parameters have significantly improved the model. (Goodness of fit improved by 24.46%).

Table 5.4: Model for trip length per individual for work trips

Socioeconomic parameters	Base model	Base model with land use parameters									Final model includes Area index at origin and destination, and intersection density
		Entropy (census tract)	DI (cell size of 100m in census tract)	Area Index at origin (1000 m buffer)	Area Index (at origin and destination) (1000 m buffer)	Entropy (1000 m buffer)	DI(10 m x 10 m cell in 1km x 1km tract)	Mix type index(10 m x 10 m cell in 1km x 1km tract)	Intersection Density (No. of intersections in 1km x 1 km tract)		
Constant	4.144 (10.296)	4.348 (10.452)	4.354 (10.687)	4.447 (11.391)	4.847 (12.113)	4.94 (12.038)	4.381 (11.21)	4.425 (11.293)	4.959 (12.534)		5.325 (13.183)
Age	-0.031 (-3.516)	-0.03 (-3.548)	-0.029 (-3.396)	-0.022 (-2.618)	-0.021 (-2.559)	-0.025 (-2.962)	-0.02 (-2.338)	-0.02 (-2.322)	-0.017 (-2.037)		-0.016 (-1.891)
Private Vehicle Ownership	0.243 (1.924)	0.249 (1.97)	0.264 (2.094)	0.259 (2.125)	0.198 (1.629)	0.234 (1.91)	0.304 (2.48)	0.303 (2.467)	0.263 (2.191)		0.203 (1.702)
Land use parameter (Origin)		-0.765 (-1.878)	-1.054 (-2.746)	-2.22 (-6.656)	-2.051 (-6.166)	-2.183 (-6.189)	-2.33 (-6.351)	-2.317 (-6.289)	-0.032 (-7.908)		-0.803 (-1.947)
Land use parameter (Destination)					-1.014 (-3.773)						-1.046 (-3.972)
Intersection Density (Origin)											-0.025 (-4.942)
adj (R^2)	0.023	0.027	0.034	0.093	0.114	0.084	0.087	0.086	0.120		0.15
% increase in adj (R^2) from base model	0.000	17.391	47.826	304.348	395.652	265.217	278.261	273.913	421.739		552.174

Table 5.5: Model for trip length per individual for shopping trips

Socioeconomic parameters	Base model	Base model with land use parameters							Final model including Area index at origin and intersection density
		Entropy (census tract)	DI (cell size of 100m in census tract)	Area Index at origin (1000 m buffer)	Entropy (1000 m buffer)	DI(10 m x 10 m cell in 1km x 1km tract)	Mix type index(10 m x 10 m cell in 1km x 1km tract)	Intersection Density (No's of intersections in 1km x 1 km tract)	
Constant	2.377 (10.424)	2.576 (10.739)	2.571 (10.897)	2.698 (-12.043)	2.894 (11.42)	2.781 (12.176)	2.816 (12.223)	3.315 (13.851)	3.283 (13.762)
Age	-0.022 (-4.814)	-0.022 (-4.853)	-0.022 (-4.891)	-0.021 (-4.689)	-0.022 (-4.839)	-0.02 (-4.6)	-0.02 (-4.601)	-0.018 (-4.137)	-0.018 (-4.193)
Land use parameter (Origin)		-0.628 (-2.512)	-0.668 (-2.843)	-1.279 (-6.072)	-1.043 (-4.284)	-1.441 (-5.952)	-1.442 (-5.962)		-0.523 (-2.161)
Intersection Density (At origin)								-0.023 (-8.174)	-0.019 (-5.671)
adj (R^2)	0.057	0.07	0.075	0.14	0.099	0.137	0.137	0.2000	0.208
% increase in adj (R^2) from base model	0.000	22.807	31.579	145.614	73.684	140.351	140.351	250.877	264.912

Table 5.6: Non-motorized and motorized vehicle choice model for work trips

Socioeconomic parameters	Constant Only	Land use variables only(Area Index Origin and Destination, Intersection density	Socio-economic variables only	Base model with land use parameters								Final model Area index at origin and destination, intersection density
				Entropy (census tract)	DI (cell size of 100m in census tract)	Area Index at origin (1000 m buffer)	Area Index (at origin and destination) (1000 m buffer)	Entropy (1000 m buffer)	DI(10 m x 10 m cell in 1km x 1km tract)	Mix type index(10 m x 10 m cell in 1km x 1km tract)	Intersection Density (No. of intersections in 1km x 1 km tract)	
Constant (NMT)	-0.653 (-7.34)	-1.57 (-6.96)	0.952 (2.25)	0.788 (1.83)	0.819 (1.91)	0.731 (1.68)	0.409 (0.87)	0.31 (0.69)	0.664 (1.53)	0.618 (1.42)	0.450 (1.01)	0.179 (0.37)
Constant (Motorized vehicle)	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
Gender			0.403 (1.63)	0.391 (1.57)	0.388 (1.56)	0.479 (1.87)	0.537 (2.07)	0.428 (1.68)	0.391 (1.55)	0.398 (1.58)	0.500 (1.94)	0.569 (2.17)
License			-1.38 (-5.05)	-1.36 (-4.99)	-1.36 (-4.96)	-1.4 (-5.02)	-1.42 (-5.07)	-1.32 (-4.73)	-1.42 (-5.09)	-1.42 (-5.08)	-1.49 (-5.33)	-1.48 (-5.25)
Private Vehicle Ownership			-0.36(-2.18)	-0.369 (-2.24)	-0.387 (-2.34)	-0.424 (-2.49)	-0.402 (-2.35)	-0.423 (-2.49)	-0.453 (-2.65)	-0.451 (-2.64)	-0.373 (-2.21)	-0.387 (-2.27)
Education			-0.0983 (-3.35)	-0.105 (-3.53)	-0.104 (-3.52)	-0.131 (-4.21)	-0.126 (-4.04)	-0.122 (-3.95)	-0.117 (-3.85)	-0.117 (-3.85)	-0.152 (-4.72)	-0.15 (-4.57)
Land use parameter (Origin)		0.87 (2.51)		0.827 (2.1)	0.862 (2.32)	1.86 (5.22)	1.78 (4.96)	1.88 (4.96)	1.8 (4.68)	1.8 (4.68)	0.0251 (5.39)	1.01 (2.3)
Land use parameter (Destination)		0.771 (3.02)					0.528 (1.86)					0.512 (1.78)
Intersection Density (At origin)		0.0222 (1.54)										0.0169 (2.97)
adj (ρ^2)	0.07	0.099	0.188	0.191	0.193	0.223	0.225	0.219	0.215	0.215	0.227	0.234

Table 5.7: Non-motorized and motorized vehicle choice model for shopping trips

Socioeconomic parameters	constant	Land use variables only (Intersection density)	Socioeconomic variables only	Base model with land use parameters						
				Entropy (census tract)	DI (cell size of 100m in census tract)	Area Index at origin (1000 m buffer)	Entropy (1000 m buffer)	DI(10 m x 10 m cell in 1km x 1km tract)	Mix type index(10 m x 10 m cell in 1km x 1km tract)	Intersection Density (No of intersections in 1km x 1 km tract)
Constant (NMT)	0.531 (4.93)	-0.595 (-2.22)	-0.357 (-0.81)	-0.363 (-0.79)	-0.372 (-0.82)	-0.633 (-1.4)	-0.838 (-1.69)	-0.668 (-1.47)	-0.707 (-1.54)	-1.55 (-2.99)
Constant (Motorized vehicle)	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
Gender			0.399 (1.25)	0.399 (1.25)	0.398 (1.24)	0.459 (1.43)	0.454 (1.41)	0.462 (1.44)	0.464 (1.44)	0.632 (1.88)
License			-1.45 (-4.59)	-1.45 (-4.58)	-1.45 (-4.59)	-1.59 (-4.88)	-1.51 (-4.72)	-1.58 (-4.87)	-1.57 (-4.87)	-1.66 (-5.02)
Private Vehicle Ownership			-0.342 (-1.84)	-0.342 (-1.84)	-0.343 (-1.85)	-0.355 (-1.89)	-0.351 (-1.87)	-0.364 (-1.94)	-0.363 (-1.93)	-0.364 (-1.97)
Age			0.0288 (3.22)	0.0288 (3.22)	0.0288 (3.22)	0.0272 (3.04)	0.0283 (3.16)	0.0268 (3.00)	0.0269 (3.00)	-0.0241 (2.62)
Land use parameter (Origin)				0.0192 (0.04)	0.0595 (0.14)	1.22 (2.90)	0.965 (2.1)	1.24 (2.61)	1.25 (2.64)	
Intersection Density (Origin)		0.023 (4.55)								0.0278 (4.81)
adj (ρ^2)	0.045	0.083	0.163	0.159	0.159	0.176	0.168	0.173	0.173	0.206

Table 5.8: Motorized Private and motorized public/ IPT choice model for work trips

Socioeconomic parameters	Constant only	Land use variables only (Intersection density)	Socioeconomic variables only	Final model
Constant (Motorized Private)	0.147 (1.41)	-0.534 (-2.51)	-5.64 (-5.95)	-5.91 (-6.08)
Constant (Motorized Public/IPT)	Fixed	Fixed	Fixed	Fixed
Gender			0.951 (2.01)	1.02 (2.13)
License			2.56 (6.89)	2.53 (6.76)
Private Vehicle Ownership			1.10 (4.03)	1.14 (4.14)
Education			0.201 (3.54)	0.172 (2.92)
Intersection Density (Origin)		0.0164 (3.69)		0.0135 (1.89)
adj (ρ^2)	-0.000	0.024	0.471	0.474

Table 5.9: Motorized Private and motorized public/ IPT choice model for shopping trips

Socioeconomic parameters	Constant Only	Socioeconomic variables only
Constant (Motorized Private)	0.399 (2.29)	-5.05 (-2.72)
Constant (Motorized Public/IPT)	Fixed	Fixed
Education		0.223 (1.86)
License		2.33 (3.37)
Private Vehicle ownership		2.34 (4.03)
Age		-0.0425 (-1.80)
Gender		1.77 (2.05)
adj (ρ^2)	0.018	0.517

Table 5.10: Multinomial logit model for work trips with different mixed land use parameters

Variable description	Model Area index		Model Entropy with buffer		Entropy index (Census tract)		DI (Census tract)		DI (10 m cell over 1km^2)		Mix type index (10 m cell over 1 km^2)	
	Param.	t-value	Param.	t-value	Param.	t-value	Param.	t-value	Param.	t-value	Param.	t-value
ASC_{Car}	-0.134	-0.08	0.320	0.19	-0.358	-0.22	-0.494	-0.31	-0.208	-0.13	-0.139	-0.08
ASC_{Bus}	-4.420	-6.64	-4.440	-6.63	-4.360	-6.60	-4.360	-6.61	-4.430	-6.65	-4.430	-6.65
$ASC_{Auto\ Rickshaw}$	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00
$ASC_{Motor\ Cycle}$	-1.280	-1.52	-0.713	-0.80	-1.500	-1.75	-1.600	-1.91	-1.200	-1.42	-1.140	-1.34
$ASC_{Bicycle}$	-0.131	-0.10	-0.201	-0.15	-0.021	-0.02	0.030	0.02	0.199	0.15	0.188	0.15
$ASC_{Rickshaw}$	-1.880	-2.06	-1.920	-2.03	-1.670	-1.84	-1.700	-1.90	-1.540	-1.67	-1.550	-1.68
ASC_{Walk}	-0.536	-0.68	-0.485	-0.58	-0.253	-0.33	-0.268	-0.35	-0.010	-0.01	-0.028	-0.04
Land use Index $_{Public}$	1.390	1.81	2.420	2.75	1.320	0.88	2.810	0.82	2.000	2.52	2.060	2.56
Land use Index $_{NMT}$	2.720	3.24	3.270	3.39	3.470	2.34	8.580	2.55	2.300	2.71	2.370	2.76
Cost $_{Auto,Two-wheeler}$	-0.134	-3.84	-0.137	-3.85	-0.123	-3.74	-0.125	-3.79	-0.136	-3.90	-0.136	-3.89
Cost $_{Rickshaw}$	-0.040	-2.18	-0.042	-2.25	-0.046	-2.52	-0.044	-2.41	-0.047	-2.55	-0.047	-2.54
Bicycle Own $_{Car}$	-1.820	-1.71	-1.880	-1.71	-1.760	-1.68	-1.780	-1.70	-1.790	-1.69	-1.800	-1.69
Bicycle Own $_{Bicycle}$	4.840	4.05	4.740	3.99	4.780	4.05	4.730	4.00	4.850	4.07	4.860	4.07
Education $_{Rickshaw}$	0.115	1.97	0.124	2.13	0.121	2.10	0.123	2.14	0.131	2.28	0.131	2.28
Family Size $_{Car}$	-0.419	-2.14	-0.411	-2.06	-0.405	-2.10	-0.401	-2.07	-0.398	-2.02	-0.400	-2.03
Family Size $_{Bicycle}$	-0.557	-2.66	-0.537	-2.60	-0.544	-2.65	-0.534	-2.60	-0.548	-2.61	-0.549	-2.62
Gender $_{Walk}$	1.250	2.69	1.200	2.59	1.160	2.51	1.170	2.53	1.180	2.54	1.180	2.54
Income $_{Car}$	0.591	3.85	0.618	3.95	0.583	3.89	0.593	3.94	0.621	4.02	0.622	4.02
Income $_{Bus}$	0.157	1.68	0.155	1.65	0.159	1.71	0.158	1.70	0.155	1.66	0.155	1.59
License $_{Motor\ Cycle}$	3.360	5.81	3.460	5.78	3.300	5.83	3.330	5.87	3.510	5.96	3.510	5.96
Income $_{walk}$	-0.196	-1.94	-0.187	-1.85	-0.190	-1.88	-0.186	-1.85	-0.188	-1.87	-0.188	-1.86
Time $_{Auto,\ Two-wheeler}$	-0.024	-1.69	-0.024	-1.66	-0.025	-1.78	-0.025	-1.75	-0.024	-1.67	-0.024	-1.66
Time $_{Cycle}$	-0.069	-3.26	-0.070	-3.22	-0.075	-3.39	-0.075	-3.46	-0.077	-3.41	-0.077	-3.40
Time $_{Walk}$	-0.035	-2.00	-0.037	-2.11	-0.040	-2.24	-0.039	-2.24	-0.043	-2.38	-0.042	-2.37
MTW Ownership $_{Car}$	-1.750	-2.50	-1.690	-2.40	-1.690	-2.42	-1.720	-2.45	-1.800	-2.59	-1.800	-2.59
MTW Ownership $_{MTW}$	1.630	2.15	1.750	2.23	1.640	2.21	1.680	2.26	1.660	2.19	1.660	2.19
Initial log-likelihood	-679.526		-679.526		-679.526		-679.526		-679.526		-679.526	
Final log-likelihood	-315.303		-314.429		-317.873		-317.061		-316.907		-316.743	
Adjusted rho-square	0.499		0.500		0.495		0.497		0.497		0.497	

Results from the binary logit model estimated on non-motorized and motorized modes for shopping trips are shown in Table 5.7. There was a significant improvement in the model when land use parameters have entered the constants only model. When only socioeconomic parameters entered the model the improvement in the model was relatively better than the model with only land use variables. Similar to the work trip models, socioeconomic parameters such having driving license, motorized vehicle ownership and years of education reduces the observed preference heterogeneity of the non-motorized modes. Male in general uses NMT modes more often than females. Even when controlling for the socioeconomic characteristics, land use parameters have significantly improved the model (goodness of fit was improved by 33 %). Coefficient corresponding to area index for commercial area at origin was found to be significant, and implies a higher preference for non-motorized modes with increase in mixed land use. Intersection density, measured as number of intersections per 1000m x 1000m cell, was also found to be significant implying an increase with increase in intersection density.

Result from the binary logit model estimated on motorized private and public/IPT vehicles for work trips are shown in Table 5.8. It has been observed that the land use characteristics have minimum effect on the utility of private vehicles for work trips. Intersection density parameter has positive effect on the utility of private motorized vehicles which is counter intuitive in nature. This may be due to the presence of more number of upper middle class families, owning private vehicles, in mixed land use area. Table 5.9 describes the result from mode choice model on private motorized modes and public/IPT, for shopping trips. In this case, land use parameters were found to be insignificant, when entered in the constants only model. Percentage increase in the years of education has considerable effect on private mode choice. Further, driving license, age, gender, and vehicle ownership also affects the utility of private vehicles.

Results from the multinomial model estimated (using the RP data) on seven modes used for the work trips are shown in Table 5.10. Out of various land use mix parameters, Area Index, Entropy with 1km buffer, and DI with 10m x 10m cell were found to be significant and positive for public and non-motorized transport. Conventional DI and entropy index were not able to capture the effect of land use mix on the choice of public transport. However, conventional indices were significant in capturing the effect of mixed land use on NMT modes. In all further models formulated and estimated in this work using the RP data area index is used to capture the effect of mixed land use.

5.7.3 Elasticity analysis

Elasticities between the trip lengths and various land use indices, as well as between the probability of choosing non-motorized and public transport modes and the land use indices, were calculated and shown in Table 5.11. The Aggregate elasticities are calculated using the probability weighted sample enumeration (PWSE) technique, given below.

$$E_{X_{ikq}}^{\bar{P}_i} = (\sum_{q=1}^{Q} \bar{P}_{iq}\, E_{X_{ikq}}^{P_{iq}}) / \sum_{q=1}^{Q} \bar{P}_{iq} \quad (5.4)$$

where, $\bar{P}_i$ refers to the aggregate probability of choice alternative i,

$E_{X_{ikq}}^{P_{iq}} = \beta_{ik} X_{ikq}(1 - P_{iq})$, interpreted as the elasticity of the probability of alternate i for decision maker q with respect to marginal change in the kth attribute of the ith alternative (i.e. X_{ikq}) as observed by the decision maker q; $\bar{P}_{iq}$ is an estimated choice probability.

But in case of variables that appear in alternate specific forms in all alternatives, then it becomes a combination of direct response and multiple cross response. An identical change in variable such as area index will occur for the alternatives in which it appears as an alternative specific variable. Elasticity in this case is given by:

$$E_{X_{ikq}}^{\bar{P}_i} = \sum_{q=1}^{Q} \bar{P}_{iq}\, [\beta_{ik} X_{ikq}\left(1 - P_{iq}\right) - \sum_{j \neq i} \beta_{jk} X_{jkq}(P_{jq})] / \sum_{q=1}^{Q} \bar{P}_{iq} \quad (5.5)$$

Table 5.11: Elasticities of the travel parameters with respect to land use variables when single land use parameter entered in the model

Land use parameters	Trip length		Non-Motorized mode choice for work	Personal Motorized vehicle mode choice for work trip	Non-Motorized mode choice for shopping trips
	Work	Shopping			
Area Index(1000m buffer)	-0.232	-0.326	0.321	--	0.104
Entropy(1000m buffer)	-0.349	-0.241	0.495	--	0.146
Dissimilarity Index(10mx10m cell)	-0.252	-0.371	0.315	--	0.122
Mix-type index	-0.269	-0.397	0.336	--	0.131
Intersection Density	-0.478	-0.857	0.535	0.157	0.369

Table 5.12: Elasticities of mode choice with land use parameters, from MNL model

Modes	Elasticity (Area index)	Elasticity (Entropy 1000m buffer)	Elasticity (Entropy ward)	Elasticity (DI for census tract)	Elasticity(DI 10m x 10m grid)	Elasticity(Mix type index 10m x 10m grid)
Cycle	0.029	0.050	0.037	0.030	0.011	0.02
Cycle Rickshaw	0.265	0.290	0.190	0.168	0.02	0.119
Walk	0.206	0.242	0.150	0.140	0.004	0.105
Bus	0.0003	0.063	---	---	0.065	0.053
MThW	-0.06	-0.011	---	---	0.06	0.025

From the elasticities between the trip lengths and land use indices, it can be said that the trip length for shopping and work trips and the land use parameters such as entropy index measured using the buffer zones and mix type index are strongly associated. From the elasticities between the probability of choosing non-motorized modes for work trips and the land use indices, it can be said that the land use parameters have strong influence on the utilities of non-motorized modes. Area index, which was used to measure the land use area complementary, also has strong influence on travel parameters. An increase in area index by 1 % increases the probability of choosing NMT modes by 0.321 %. Land use quantified by entropy index (buffer) and intersection density has higher elasticity than area index. This may be due to basic difference in the way the land use mix parameters were quantified. Elasticity values estimated based on the MNL models are shown in Table 5.12. From the results shown in the table, it can be seen that the probability of choosing walk and cycle rickshaw increases with the increasing land use mix. In case of motorized modes, the elasticity values are not strongly related. The elasticity of NMT is higher than public transport modes. From the values shown in Table 5.12, it can be said that the modified land use indices have significant impact on the mode choice including the choice of non-motorized mode of transport.

Table 5.13: Improvement of R^2/ρ^2 (%) from existing conventional land use parameters

Land use parameters using the conventional approaches	**Models**	**Improvement in R^2/ρ^2**			
		Entropy index (1000m radius)	**DI (10m x10m)**	**Mix type index**	**Area index (1000m radius)**
DI (Census tract, 100x100m cell size)	Trip length(Work Trip)	147.06	155.88	152.94	173.53
Entropy Index (Census tract)	Trip Length(Work Trip)	211.11	222.22	218.52	244.44
DI (Census tract, 100x100m cell size)	Trip length(Shopping Trip)	32.00	82.67	82.67	86.67
Entropy Index (Census tract)	Trip Length(Shopping Trip)	41.43	95.71	95.71	100.00
DI (Census tract, 100x100m cell size)	NMT Mode choice (Work Trip)	13.47	11.40	11.40	15.54
Entropy Index (Census tract)	NMT Mode choice (Work Trip)	14.65	12.56	12.56	16.75
DI (Census tract, 100x100m cell size)	NMT Mode choice (Shopping Trip)	5.66	8.81	8.81	10.69
Entropy Index (Census tract)	NMT Mode choice (Shopping Trip)	5.66	8.81	8.81	10.69

Table 5.13 shows the improvements in the model performance in comparison to the model estimated with the land use parameters obtained using the conventional approaches. Inclusion of area index led to increased R^2 or ρ^2 for all the models, even when compared to the models with entropy calculated using the buffers. In case of MNL model with the area index or entropy calculated for buffer, there is a significant improvement in the adjusted rho square value compared to the models with the conventional land use mix indices.

When the performance of different indices is compared in terms of the model improvements, it can be said that area index quantifies the mixed land use effectively. Also, area Index is comparatively easy to calculate and interpret as it is only the ratio of areas associated to the nature of the trip. In case of shopping trips, the area index simply becomes the ratio of the area of commercial space in the buffer zone to the total extent of the commercial space in the study area. With the data on the extent of shopping space in the study area one can quantify the effect of mixed land use in a better way than generally used mixed land use variables.

5.8 Conclusions

In this study, the relationship between travel behavior and the land use variables has been modeled in the context of smaller Indian cities. Travel behavior was quantified in terms trip lengths, motorized/non-motorized mode choice, motorized private and public/IPT mode choice as well as the choice of individual modes. Land use mix has been measured using entropy, area index, DI, and mix type index. Approach used in calculating the entropy and DI was slightly different from that of conventional approach. Area index has been calculated for origin and destination of trip and this index takes different values based on the trip purpose. In quantifying the mix using modified DI and Mix type Index, cell size of 10m x 10m and grid size of 1km x 1km have been used. Entropy measured with 1000m buffer performed better than the entropy measured using the conventional method. The following are the important conclusions drawn out of the present work on land use and travel behavior interaction;

- When slightly modified approach was used for calculating the entropy and DI, both the land use parameters were able to quantify the land use mix and were consistently having significant effect on the travel parameters.
- The trip length of individuals, for both the shopping and work trips, was strongly correlated to the land use mix variables, even when controlling the socioeconomic characteristics. In case of work trips, entropy measured with 1000 m radius buffer, and DI measured using 10 m x 10 m cell for 1km x 1km tract were found to be more significant than the conventional entropy and dissimilarity indices measured for census tract. When the area Index for both the origin and destination were considered, there was almost 395% increase in the model's ability to explain the variability of trip length compared to the base model. The coefficient of land use mix and intersection density at trip origin is negative which means an increase in intersection density and land use mix tends to reduce the trip length. With respect to socioeconomic variables, older people generally have shorter trip length. Private vehicle ownership has significant effect in explaining the variability of trip length. With increase in private vehicles in a household, the trip length increases.
- All the land use variables were negatively correlated with the trip length for shopping trips. In case of both shopping and work trips, older persons seem to go to closer destinations (shorter trip length). Similar to work trip, both the coefficient of land use mix and intersection density at the origin of the trip is negative, thereby

reducing the trip length. Apart from age, private vehicle ownership has a significant effect in explaining the variability of trip length in the modelling.

- From the estimated binary logit models, it can be said that along with the socioeconomic parameters, land use parameters are also significant in explaining the mode choice. Area Index values, for both origin and destination, were found to be significant. Out of various socioeconomic variables tried in this study, years of education, gender, license status, and private vehicle ownership were found to be significant. As expected, having driving license, motorized vehicle ownership, and years of education reduces the utility of non-motorized modes. With respect to gender, male generally uses non-motorized modes. An increase in area index at origin and destination was also found to be enhancing the utility of non-motorized travel. Even when controlling for the socioeconomic characteristics, land use parameters have significantly improved the model. (Goodness of fit improved by 24.46%).
- A significant elasticity exists (0.535) between the intersection density and the utility of non-motorized modes, for work trips. This may be due to the availability of more cycle rickshaws in the areas where the intersection density is high. In case of shopping trips, land use mix measured by entropy and DI in conventional way was insignificant. This clearly shows the disadvantages of entropy index measured for census tract and DI measured using 100 m x 100 m cells for census tract. Land use mix measured using a slightly different approach, could sufficiently capture the variations in the mode choice.
- Similar to the work trip models, in case of shopping trips socioeconomic parameters such having driving license, motorized vehicle ownership and years of education reduces the observed preference heterogeneity of the non-motorized modes. Male in general uses NMT modes more often than females. Even when controlling for the socioeconomic characteristics, land use parameters have significantly improved the model (goodness of fit was improved by 33 %). Coefficient corresponding to area index for commercial area at origin was found to be significant, and implies a higher preference for non-motorized modes with increase in mixed land use. Intersection density, measured as number of intersections per 1000m x 1000m cell,

was also found to be significant implying an increase in non-motorised mode preference with increase in intersection density.

- In case of motorised private and public/IPT mode choice for work trips the utility of choosing private mode increases with the increase in the intersection density, which is counterintuitive. This may be due to the upper middle class people, owning vehicles, residing in mixed land use area. In this case the effect of mixed land use was found to be negligible. Same is the case with the shopping trips also.
- In case of work trips, land use quantified by entropy index (buffer) and intersection density has higher elasticity than area index. This may be due to basic difference in the way the land use mix parameters were quantified. An increase in area index by 1 % increases the probability of choosing NMT modes by 0.321 %.
- From the MNL models estimated on the mode choice for work trips, area index, entropy with 1km buffer. DI with 10m x 10m cell, and mix type index with 10m x 10m were found to be significant. The coefficients are positive for public and non-motorized transport which implies that the trip makers residing in the areas with mixed land use prefer public and non-motorized transport. Conventional DI and entropy indices were not able to capture the effect of land use mix for public transport; but they were significant in capturing the effect of mixed land use on NMT modes. It can be seen that the probability of choosing walk and cycle rickshaw increases with the increasing land use mix. An increase in area index by 1 % increases the probability of choosing cycle rickshaw by 0.265%. In case of motorized modes, the elasticity values with respect to choosing public transport are not strongly related. The elasticity of NMT is higher than public transport modes.

From these conclusions, it can be said that the land use mix variables, along with some of the other land use parameters, significantly influence the non-motorized transport mode choice for work as well as for shopping trips. Any change in the existing mixed land use pattern further shifts people towards the personal mode of travel.

Chapter 6

Estimation of choice models with RP, SP and combined SP-RP data

Choice models estimated using the RP data, SP data, and the combined data are discussed in this chapter. Modeling methodology used in developing these models has already been discussed in Chapter 3. Effects of mixed land use on travel behavior have been incorporated in the models estimated with the RP data. Mixed logit models have been estimated using the combined SP-RP data for motorized modes and these models (with mixed logit formulation) have been used to analyze the unobserved heterogeneity and correlations.

6.1 Mode choice models estimated with the RP Data

Models estimated with the RP data have been presented in this section. Several multinomial logit models were estimated for work trips made using seven different modes; Car, Bus, MThW, MTW, Bicycle, Cycle Rickshaw and Walk. As MThW mode is generally available to all the individuals and is taken as reference mode for model estimation. In case of modelling, determination of choice set for an individual is not straight forward and very difficult in case of small cities like Agartala where no network data and LOS related information was available. For example, the category of private motorized mode consists of two completely different modes, namely, MTW and car. Similarly, the category of public transport mode covers variety of modes such as cycle rickshaw, MThW, and various types of buses. The first two modes of this category offer greater flexibility/convenience but the travel attributes are different from that of the buses. Again the services offered by the MThW depend on the driver as there are no fixed regulations on its use. In this scenario, based on the vehicle ownership status and network characteristics, it is difficult to ascertain the choice set of any individual. The choice set has been identified by applying specific rules as given below;

1) Availability of MThW: MThW is available to all the individuals.

2) Availability of bus: Bus is available to all the individuals.

3) Availability of cycle rickshaw: Cycle rickshaw is available to all the passengers within the CBD area for trips shorter than 3km.

4) Availability of cycle: Based on cycle ownership

5) Availability of MTW: If the household has one MTW, it is available to the head of the family.

6) Availability of Car: Again car is available to the head of the family only.

Modal split observed in the RP data, for work trips, is shown in Figure 6.1. For modelling with RP data, as discussed earlier, only a subset of modes is available to the respondent. After removing the incomplete data, there were 561 data available for estimating the models. The following models have been estimated with the RP data.

i) Model with only modal attributes – Model 1
ii) Model with modal, socioeconomic and land use characteristics – Model 2
iii) Mixed logit model on the formulation of model 2–Model 3

The objective in doing so was to understand the significance of socioeconomic characteristics in explaining the observed component of the taste and preference heterogeneities. Model 1 was formulated only with the modal attributes. After various attempts, it has been observed that very few modal attributes, with logically correct signs, were significantly entering the model. This may be due to the absence of variability in the RP data as well as due to the correlations across the attributes. Model 2 has been formulated to understand the relevance of socioeconomic and the land use characteristics (explained in terms of 'area index') in explaining the preference heterogeneity. An attempt has also been made, while formulating model 2, to explain the taste heterogeneity using the observed socioeconomic characteristics of the individuals. None of the socioeconomic characteristics were found to be significant in explaining the taste heterogeneity. Model 3 is the same as model 2 but with additional error components and random parameters. This has been done to understand the presence of unobserved heterogeneities.

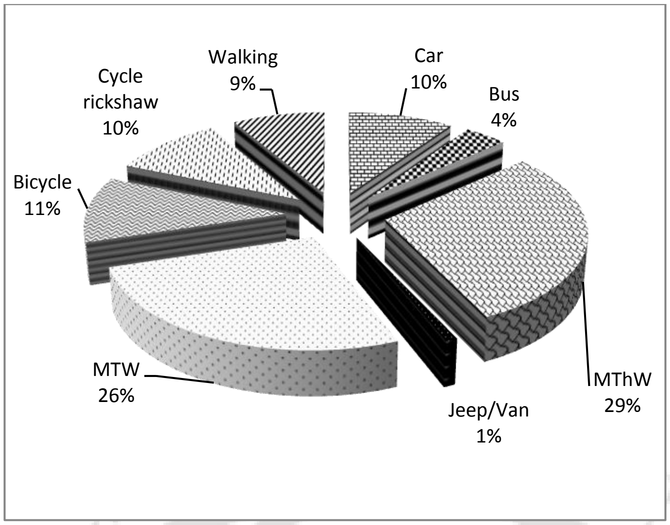


Figure 6.1: Modal share of the work trips in the RP data

Table 6.1 shows the results of MNL models estimated with the RP data. Detailed analysis of the results is presented in the following section.

6.2 Result from the models estimated with the RP data

MNL-1 was estimated by considering only mode related variables. Travel cost and travel time are the most important parameters representing mode’s characteristics. Individuals generally prefer a faster and cheaper mode of transport than slower and expensive modes. Travel time captured from network data has been used in the model rather than user specified travel time. In case of motorized modes, generic travel coefficient for MThW and MTW was considered. An increase in travel time of MThW, MTW, cycle and walk mode reduces the utility of the respective modes. Increasing travel cost, as expected, significantly reduces the utility of the modes (MThW, MTW and cycle rickshaw).

Table 6.1: Results from the models estimated with RP data

Variable description	RP-modal attribute only –MNL (1)		RP-modal with socioeconomic & land use data- MNL (2)		RP-modal with socioeconomic & land use data- Mixed logit (3)	
	coefficient	*t*-statistic	coefficient	*t*-statistic	coefficient	*t*-statistic
ASC_{Car}	1.000	2.88	-0.134	-0.08	0.392	0.15
ASC_{Bus}	-3.100	-8.48	-4.420	-6.64	-5.110	-5.87
ASC_{MThW}	-	-	-	-	-	-
ASC_{MTW}	2.200	10.45	-1.280	-1.52	-2.060	-1.44
$ASC_{Bicycle}$	2.110	5.38	-0.131	-0.10	-0.433	-0.32
$ASC_{Rickshaw}$	0.449	1.15	-1.880	-2.06	-2.240	-2.28
ASC_{Walk}	0.287	0.78	-0.536	-0.68	-0.861	-1.01
Area Index $_{Public}$			1.390	1.81	2.500	1.77
Area Index $_{NMT}$			2.720	3.24	3.890	2.66
Cost $_{MThW,MTW}$	-0.095	-3.22	-0.134	-3.84	-0.199	-2.97
Cost $_{Rickshaw}$	-0.048	-2.86	-0.040	-2.18	-0.048	-2.39
Bicycle Ownership $_{Car}$			-1.820	-1.71	-2.790	-1.62
Bicycle Ownership $_{Bicycle}$			4.840	4.05	4.910	4.03
Education $_{Rickshaw}$			0.115	1.97	0.130	2.12
Family Size $_{Car}$			-0.419	-2.14	-0.548	-1.61
Family Size $_{Bicycle}$			-0.557	-2.66	-0.566	-2.57
Gender $_{Walk}$			1.250	2.69	1.260	2.65
Income $_{Car}$			0.591	3.85	0.767	3.24
Income $_{Bus}$			0.157	1.68	0.148	1.35
License $_{MTW}$			3.360	5.81	5.000	3.10
Income $_{walk}$			-0.196	-1.94	-0.182	-1.74
Time $_{MThW, MTW}$	-0.019	-1.53	-0.024	-1.69	-0.023	-1.52
Time $_{Cycle}$	-0.066	-3.68	-0.069	-3.26	-0.075	-3.25
Time $_{Walk}$	-0.046	-2.32	-0.035	-2.00	-0.038	-2.08
MTW Ownership $_{Car}$			-1.750	-2.50	-2.310	-2.13
MTW Ownership $_{MTW}$			1.630	2.15	2.550	1.86
Sigma (Travel cost, MTW, MThW)					-0.072	-2.12
Error Component (Car, MTW)					2.220	*1.95*
Initial log-likelihood	*-679.526*		-679.526		-679.526	
Final log-likelihood	*-401.155*		-313.238		-312.286	
Adjusted rho-square	*0.393*		0.498		0.501	

Table 6.2: Elasticities of model parameters corresponding to the models estimated with the RP data

Variable description	RP-modal attribute only - MNL	RP-model with socioeconomic & land use data-MNL	RP-model with socioeconomic & land use data-Mixed logit
Area Index $_{Bus}$		0.0003	0.019
Area Index $_{MThW}$		-0.06	-0.06
Area Index $_{Cycle}$		0.029	0.027
Area Index $_{Cycle\ rickshaw}$		0.265	0.252
Area Index $_{Walk}$		0.206	0.288
Cost $_{MThW,}$	-0.275	-0.340	-0.462
Cost $_{MTW}$	-0.155	-0.133	-0.119
Cost $_{Rickshaw}$	-0.615	-0.446	-0.504
Time $_{MThW}$	-0.092	-0.104	-0.093
Time $_{MTW}$	-0.030	-0.021	-0.012
Time $_{Cycle}$	-0.314	-0.187	-0.176
Time $_{Walk}$	-0.438	-0.291	-0.308

Another multinomial logit model (MNL-2) was estimated considering the socioeconomic and land use variables along with the modal attributes. This model was prepared to understand the ability of socioeconomic characteristics and land use characteristics in explaining the observed component of the preference heterogeneity. Socioeconomic parameters such as income, age, gender, vehicle ownership, driving license, family size, and education were used to understand the preference heterogeneity in this model. It has been found that the utility of car reduces with the MTW and bicycle ownership. Household income has the positive effect on the utility of car. The utility of MTW increases with the possession of driving license and MTW ownership.

From Chapter 2 it has been found that mixed land use parameters significantly affect the mode choice behaviour, particularly the NMT modes. Mixed land use quantified by area index, which is explained in detail in Chapter 5, has been used in this model. Area index was introduced as a generic variable in the public and intermediate public transport (IPT) modes (bus and MThW) as well as for NMT modes (cycle, cycle rickshaw and walk). Coefficient of area index for both the public/IPT and NMT mode was positive and significant. This result calls for an important policy implication that can reduce the number of private vehicle trips. Utility of these modes is high in the localities with mixed land use. It may be due to the availability of various land uses that promotes public transport and NMT modes. This model, when compared to MNL-1 (with only mode related variables), offers better fit in terms of Rho-square value.

To know the effect of unobserved taste and preference heterogeneity mixed logit model (Model 3) has been estimated and tabulated in Table 6.1. In this model, coefficient for travel cost of MThW and MTW was considered as random following normal distribution. Correlation across the alternatives has been studied in Model 3. This model has confirmed the presence of both the unobserved taste heterogeneity (coefficient of cost for MTW and MThW) and strong error term correlation for two modes, car and MTW. Fitness of the model has improved by 27.48 % in comparison to the MNL model with only the modal attributes. As expected, compared to model-2 there is also a significant change in the coefficients of various attributes.

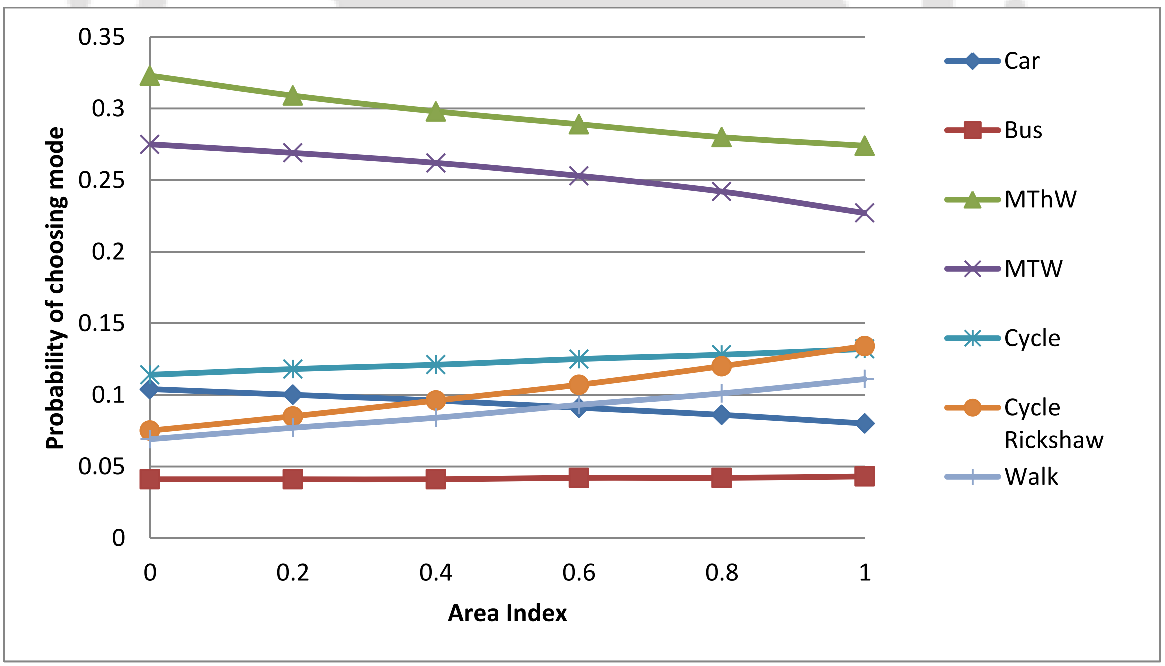


Figure 6.2: Variation in the probabilities corresponding to various modes

From Table 6.2, the elasticity of cost for MTW (-0.119) is lower than that of the cycle rickshaw (-0.504) and MThW (-0.462). Even the elasticity of travel time for MTW (-0.012) is much lower than the elasticity of travel time for cycle (-0.176), walk (-0.308), and MThW (-0.093). From this, it can be inferred that the users of MTW are less sensitive for cost and travel time and they may not switch to the other alternatives. When considering the effect of mixed land use on NMT modes, the elasticity of area index was found to be higher for cycle rickshaw (0.252) and walk mode (0.288) than cycle mode (0.027). Elasticity of area index in case of public transport, for bus is 0.019, and for MThW is -0.06. From this it can be inferred that walk and cycle rickshaw are more sensitive to the changes in the mixed land use. From Figure 6.2 (plot between area index and probability of choosing different modes), it can be seen that the probability of choosing MThW and MTW mode reduces with increase in area index. In case of MThW, the elasticity of choosing MThW reduces when area index is more than 0.4. The probability of choosing bus does not increase when area index increases. In case of NMT modes, it can be seen that with increase in area index, probability of choosing NMT modes increases, as reported earlier. This result suggests that with increase in area index, there will be increase in the share of non-motorized transport and decrease in the share of motorized transport. So, it can be inferred that with increase in land use mix, there will be increase in the probability of choosing NMT modes.

From the models estimated, it is evident that socioeconomic and land use parameters explains the preference heterogeneity and plays a significant role in the mode choice modelling. Coefficients of mixed logit model confirm the presence of both the unobserved taste heterogeneity and correlation across two modes, car and MTW. Since all the mode specific coefficients of LOS attributes could not be found out, models have been estimated with the SP data and the details are presented in the following section.

6.3 Models estimated with the SP data

It is well known fact that the attributes in RP data has little variability, i.e., the attribute value is not varied much. Moreover the attributes are highly correlated. Further, finalizing the choice set of RP data is very difficult. Compared to the RP data, SP data have certain advantages such as more variability in the attribute data, provision for new alternative, and new attributes. SP data have certain advantages over the RP data being having more variability; provision for analysing new proposed alternatives, new attributes and levels of attribute. As can be seen from the previous MNL models estimated with the RP data, some of the mode specific LOS variables are insignificant. As discussed earlier in Chapter 4, in the SP survey cost, time, comfort and frequency of transit were collected for four motorized modes namely car, bus, MThW and MTW. SP survey designs and the details of the collected data have been discussed in chapter 4. There were 2048 responses from the SP survey for work trips and are used in the model estimation. Figure 6.3 shows the modal split observed in the SP data.

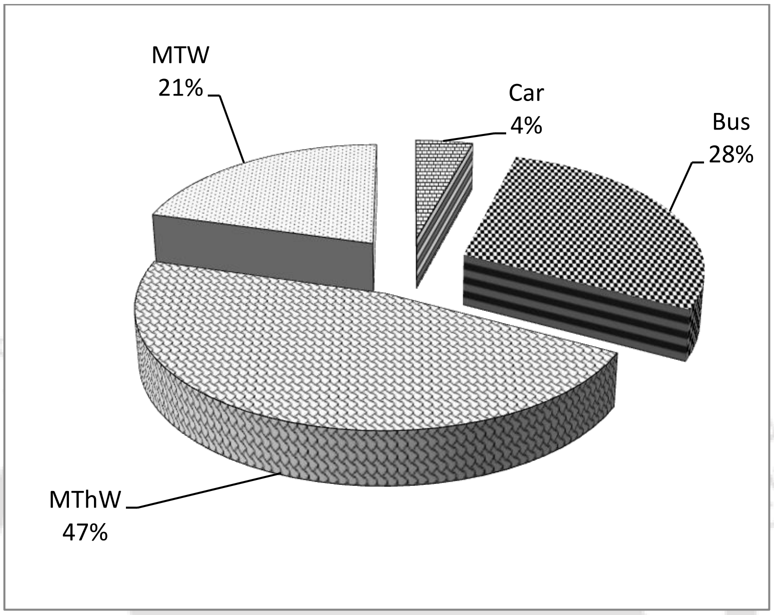


Figure 6.3: Modal split observed in the SP Data

6.3.1 Model estimation with SP data

The following models have been estimated with the SP data;

i) Base MNL model (**SP-1)**

ii) Mixed logit model with coefficient of travel time for bus as random parameter with normal distribution. **(SP-2)**

iii) Mixed logit model estimated with panel data with travel time coefficient as random parameter. **(SP-3)**

iv) Mixed logit model estimated with panel data with travel time coefficient as random parameter with triangular distribution. **(SP-4)**

v) Mixed logit model with panel data structure with normal distribution considering error components for car and MTW and for bus and MThW. **(SP-5)**

Table-6.3 shows the results from various models estimated with the SP data for four modes, namely, car, bus, MThW and MTW. All the estimated parameters are having expected sign. From the MNL model (SP Model -1) estimated with the SP data, income and gender were interacted with travel cost and comfort attributes, respectively. Comfort parameter was measured for two modes, bus and MThW.

Table 6.3: Results from the models estimated with the SP data

Variable description	SP- 1		SP-2		SP-3		SP-4		SP-5	
	Param.	t value	Param.	t value	Param.	t value	Param.	t value	Param.	t value
ASC $_{Car}$	2.590	6.17	2.610	5.91	3.28	4.76	2.798	5.56	3.480	6.23
ASC $_{Bus}$	-0.008	-0.04	0.184	0.77	0.17	0.54	-	-	0.250	0.95
ASC $_{MThW}$	-	-	-	-	-	-	-	-	-	-
ASC $_{MTW}$	1.860	9.47	1.880	8.51	2.00	6.73	1.939	8.80	2.620	7.20
Frequency $_{Bus}$	-0.028	-6.61	-0.031	-5.97	-0.04	-5.73	-0.029	-7.17	-0.033	-5.76
Time $_{Car}$	-0.038	-2.06	-0.045	-2.30	-0.12	-1.95	-0.071	-2.13	-0.062	-2.73
Time $_{Bus}$	-0.050	-8.60	-0.097	-5.47	-0.13	-4.95	-0.080	-8.06	-0.112	-5.51
Time $_{MThW}$	-0.036	-7.14	-0.054	-5.22	-0.07	-4.63	-0.056	-7.30	-0.064	-5.57
Time $_{MTW}$	-0.041	-5.92	-0.048	-5.73	-0.04	-4.12	-0.062	-5.46	-0.055	-5.33
SIGMA $_{Time\ Bus}$			0.071	4.28	-0.10	-4.37	0.080	8.06	0.084	4.31
SIGMA $_{Time\ Car}$					-0.14	-1.84	0.071	2.13		
SIGMA $_{Time\ MThW}$			0.029	1.69	0.06	2.83	0.056	7.30	-0.040	-2.64
SIGMA $_{Time\ MTW}$							0.062	5.46		
Cost car/inc	-0.351	-4.20	-0.388	-4.31	-0.41	-3.38	-0.381	-4.15	-0.392	-3.50
Cost bus/inc	-0.542	-6.92	-0.658	-6.10	-0.90	-5.87	-0.592	-6.78	-0.719	-6.12
Cost MThW/inc	-0.335	-5.83	-0.423	-5.55	-0.59	-5.40	-0.389	-5.98	-0.453	-5.41
Cost MTW/inc	-0.561	-6.73	-0.784	-7.15	-1.12	-6.63	-0.657	-7.19	-0.981	-5.91
Cost * gen Car	-0.098	-5.76	0.052	-4.12	-0.07	-3.60	-0.053	-4.09	-0.066	-4.04
Cost * gen Bus	-0.074	-3.30	-0.078	-2.90	-0.09	-2.48	-0.079	-3.20	-0.094	-3.03
Cost*gen MThW	-0.102	-6.10	-0.116	-5.30	-0.15	-4.79	-0.114	-5.94	-0.135	-5.01
Cost*gen MTW	-0.029	-1.56	-0.032	-1.46	-0.05	-1.78	-0.032	-1.52	-0.051	-1.58
Comfort * gen bus	0.561	2.66	0.719	3.32	0.85	3.11	0.627	3.25	0.861	3.48
Comfort * gen MThW	0.312	1.80	0.402	1.99	0.69	2.55	0.404	2.19	0.549	2.34
Comfort * income bus	0.149	5.40	0.161	4.93	0.20	4.88	0.161	5.60	0.176	4.63
Comfort * inc MThW	0.124	4.68	0.141	4.64	0.16	4.00	0.128	4.61	0.157	4.28
Error component (Car,Bus,MTW)					-1.43	-7.16				
Error component Car, MTW									1.640	3.61
Error Bus, MThW									-0.707	-0.77
Init log-likelihood	-1793.736		-1793.736		-1793.736		-1793.736		-1793.736	
Final log-likelihood	-1224.746		-1216.631		-1201.056		-1203.051		-1178.683	
Adjusted rho-square	0.306		0.309		0.317		0.310		0.330	

Income and gender were found to be significant in explaining the taste heterogeneity towards the comfort of bus and MThW. Coefficients of the interactions are positive and the parameters are statistically significant. Coefficient associated to the frequency of bus, measured as time interval between the services, is negative and significantly affects the utility of bus. With less frequency, the waiting time of the people increases, thereby decreasing the utility of the bus.

In SP Model-2, a mixed logit model was estimated to check the presence of unobserved taste heterogeneity towards travel time. The coefficient of the travel time is considered as random variable following normal distribution. This model confirms the presence of unobserved taste heterogeneity towards the travel time of bus and MThW. For the other two modes, the parameters of the random coefficient were found to be insignificant. The standard deviations of the random parameters for bus and MThW were found to be statistically significant.

Since the SP data contains repeated observations from the same individual, a mixed logit specification (SP Model-3) for repeated choices i.e., integral of the product of logit probabilities, one for each observation, has been incorporated. Panel formulation captures the correlation across the observations from the same individual. With the panel specification, there has been significant improvement in the fitness of the model along with the improvement in parameter estimates.

As evident from many of the past studies, the use of unbounded distributions like normal distribution may lead to both the positive and negative travel time coefficients. In order to avoid this constrained triangular distribution was used to model the random parameter of the travel time and the revised model (SP Model-4) has been estimated. As opposed to SP model-3, in this case random parameters of travel time, corresponding to all the four modes, were found to be significant and confirm the presence of unobserved taste heterogeneity.

Table 6.4: Elasticities of model parameters corresponding to the models estimated with the SP data

Variable description	SP Model (1)	SP Model (2)	SP Model (3)	SP Model (4)	SP Model (5)
Frequency measured in Time interval) $_{Bus}$	-0.284	-0.288	-0.313	-0.279	-0.288
Time $_{Car}$	-0.200	-0.098	-0.442	-0.310	-0.231
Time $_{Bus}$	-0.361	-0.571	-0.655	-0.514	-0.605
Time $_{MThW}$	-0.021	-0.005	-0.067	-0.033	-0.021
Time $_{MTW}$	-0.185	-0.212	-0.176	-0.262	-0.209
Costcar/inc	-0.552	-0.263	-0.542	-0.564	-0.483
Costbus/inc	-0.473	-0.523	-0.575	-0.484	-0.514
Cost$_{MThW}$/inc	-0.259	-0.277	-0.305	-0.276	-0.271
Cost$_{MTW}$/inc	-0.359	-0.509	-0.620	-0.400	-0.517
Cost * gen $_{(Car)}$	-0.615	0.163	-0.466	-0.402	-0.421
Cost * gen$_{(Bus)}$	-0.206	-0.194	-0.182	-0.207	-0.215
Cost*gen$_{(MThW)}$	-0.244	-0.229	-0.244	-0.25	-0.250
Cost*gen$_{(Mc)}$	-0.082	-0.090	-0.129	-0.089	-0.124
Comfort * income $_{(bus)}$	0.212	0.205	0.246	0.216	0.216
Comfort * gen $_{(bus)}$	0.116	0.137	0.151	0.122	0.156
Comfort * gen $_{(MThW)}$	0.039	0.042	0.063	0.048	0.055
Comfort * inc $_{(MThW)}$	0.108	0.101	0.106	0.102	0.111

In order to study the correlation across the alternatives, a mixed logit model with common error components (SP Model-5), has been estimated. The model confirmed the presence of correlation across two modes, car and MTW along with the taste heterogeneity confirmed by statistically significant standard deviation of travel time coefficient for bus and MThW. Fitness of the model has also improved when compared to the previous models.

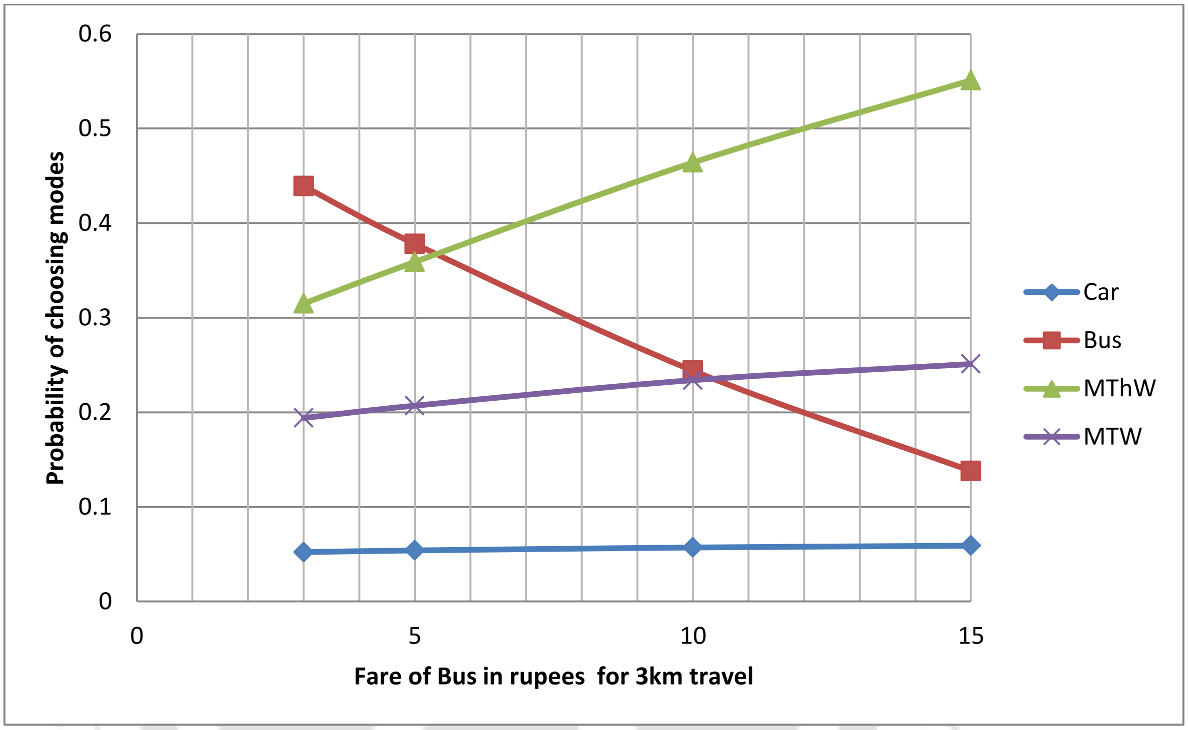


Figure 6.4 Probability of choosing travel mode with respect to increasing bus fare (with SP-5 model).

The coefficient for comfort (interacted with income and gender) is higher for bus compared to MThW, and so are the elasticities. This could lead to further policy decisions of improving the comfort level of transit so that people can be motivated to use the transit mode. Further, with increase in the frequency of trips of transit, the utility of bus mode increases. From the taste heterogeneity towards the travel cost explained using the income (coefficient of Cost/Income), it can be seen that the users of car and MThW are less sensitive to the cost compared to the users of bus and MTW.

From the elasticity values presented in Table 6.4, it can be seen that high elasticity exists between the cost and the choice probabilities. From the elasticities based on SP Model-5 it can be seen that high elasticity exists between travel cost (interacted with income) and the probabilities of choosing bus (-0.514), MTW (-0.517) and car (-0.483). The users of these modes may shift to other modes of transport if the travel cost is increased. In case of car users, as most of the people are from middle income group, an increase in price will force the users to the other available modes. This provides a very good policy implementation scenario in which demand for mode could be controlled by regulating the fares. The same is reflected in Figure 6.4 (based on SP-5 model), showing the probability of choosing modes with respect to the bus fare, for 3 km trips. When the fare of bus is increased from Rs 3 to Rs 10, the reduction in probability of choosing bus is about 0.2. When the fare of bus is Rs 5, the probability of choosing both bus and MThW modes are same, and with increase in further fare, the probability of choosing bus reduces and MThW increases.

From these models, it can be concluded that the utility of transit can be increased which will lead people getting motivated to shift to public mode of transport from the other modes by regulating fare and improving comfort and reliability (combination of frequency and travel time).

6.4 **Models estimated with the combined SP-RP data**

For travel related data, the RP data has issues related to correlation of attributes. The cost/fare and travel time attributes in real market situation may be correlated and finding both cost and travel time parameter for each mode is very difficult to determine. Further, if looked carefully the possibility of correlation between travel time and travel cost parameter is higher as fares are determined based on the trip length. Also, travel times are similar across alternatives during peak hour period due to congestion. As such it is very difficult to properly estimate time and cost parameter from RP data due to these limitations. As SP data is hypothetical, it was decided to combine SP and RP data to get better parameter estimates. In combining the data, cost parameter has been taken as common attribute for both the data sets and travel time parameter was estimated only from SP data to avoid any problem due to correlation. The estimated model parameters are confounded with the scale factors for each data set. As a result data enrichment includes pooling of two choice data sources under the restriction that common parameters are equal while controlling the scale factors.

6.4.1 SP-RP modeling

Two different models have been estimated using the combined SP-RP data. The following are the models estimated;

1) Multinomial logit model with interaction between LOS variables and socioeconomic characteristics
2) Mixed logit model with panel formulation, error components and random parameters.

While RP data was combined with SP data there was a significant improvement in the model when compared to the models estimated with only SP data. From Table 6.5, it can be seen that the value of λ^{SP} is greater than one and statistically significant i.e., error in the model estimated with only the RP data has more variance. In the estimated models, the signs of the coefficients are logical. From the estimated elasticity of combined SP-RP model it can be seen that the elasticity of cost parameter for bus is higher than the other modes.

Further, two mixed logit models have been estimated to check the effect of state dependence i.e., the effect of revealed choice on stated choice of the individual. The first model (shown in Table 6.6) does not consider the state dependency and is similar to the mixed SP-RP model shown in Table 6.5 except the interactions between the LOS and socioeconomic variables. To understand the effect of state dependence, the data file was prepared in such a way that dummy $\delta_{state\text{-}effect}$ parameter is one when RP choice and the SP choice are same for an individual. This parameter is very significant as it provides information whether the individuals using a particular mode are willing to shift to other modes. In this model, state dependency has been checked separately for all the four modes, i.e., the dummy parameters are mode specific. The standard deviation characterizing the unobserved taste heterogeneity for travel time, for bus, is significant in the model without state dependency effect (Table 6.6). It can also be observed from the result shown in Table 6.6 that the state dependence is significant in case of both MTW and MThW users.

Table 6.5: Results from the models estimated with combined SP-RP data

Variable description	MNLSP-RP		MMNL-1 SP-RP		MNLSP-RP Elasticity	MMNL-1 SP-RP Elasticity
	Param.	t-statistics	Param.	t-statistics		
ASC $_{Car}$	2.680	6.60	2.650	6.62		
ASC $_{Bus}$	-1.970	-8.90	-1.960	-8.81		
ASC $_{MThW}$	--	--	--	--		
ASC $_{MTW}$	2.44	8.08	2.560	8.14		
ASC $_{Car\ SP}$	0.976	2.05	1.290	2.24		
ASC $_{Bus\ SP}$	-0.007	-0.09	0.104	0.93		
ASC $_{MThW\ SP}$	--	--	--	--		
ASC $_{MTW\ SP}$	0.778	1.81	1.040	2.09		
Frequency $_{Bus}$	-0.012	-1.76	-0.013	-2.01	-0.115	-0.117
Time $_{Car\ SP}$	-0.017	-1.32	-0.024	-1.61	-0.016	-0.024
Time $_{Bus\ SP}$	-0.021	-1.78	-0.043	-2.02	-0.158	-0.262
Time $_{MThW\ SP}$	-0.015	-1.78	-0.022	-2.05	-0.072	-0.087
Time $_{MTW\ SP}$	-0.017	-1.75	-0.021	-2.00	0.068	-0.071
Cost(car)/inc	-0.109	-2.93	-0.108	-3.31	-0.042	-0.044
Cost(bus)/inc	-0.219	-1.80	-0.277	-2.08	-0.197	-0.215
Cost(MThW/inc	-0.144	-1.77	-0.181	-2.03	-0.105	-0.107
Cost(MTW)/inc	-0.285	-1.83	-0.375	-2.32	-0.239	-0.307
Comfort * income bus	0.063	1.73	0.068	1.93	0.086	0.083
Cost * gen $_{Car}$	-0.021	-1.56	-0.029	-1.69	-0.195	-0.238
Cost * gen$_{Bus}$	-0.031	-1.52	-0.034	-1.65	-0.087	-0.085
Cost*gen$_{MThW}$	-0.041	-1.69	-0.050	-1.92	-0.093	-0.094
Comfort * gen bus	0.231	1.58	0.337	1.83	0.046	0.061
Comfort * gen $_{MThW}$	0.126	1.28	0.206	1.57	0.014	0.018
Comfort * inc $_{MThW}$	0.052	1.71	0.059	1.91	0.039	0.036
SIGMA $_{Time\ Bus}$			0.034	1.96		
Error components (Car ,MTW)			-0.684	-1.78		
Error components (Bus ,MThW)			-0.166	-0.23		
SP to RP scale factor	2.400	1.81	2.440	2.10		
Init log-likelihood	-2162.184		-2162.184			
Final log-likelihood	-1391.708		-1348.477			
Adjusted rho-square	0.345		0.364			

MTW being the cheaper and flexible mode, people are less willing to change it. When the state dependency is incorporated in the model unobserved taste heterogeneity is coming out to be insignificant. This model also shows a reduction in travel time and travel cost sensitivity compared to the MMNL model without state dependency. The fit measures in the model indicate improvement due to inclusion of state dependence factor, the adjusted

ρ^2 value increases from 0.375 to 0.461.

Table 6.6: Result from the models estimated for state dependency effect

Variable description	MMNL-2 without interactions		MMNL-3 with state dependency	
	Param	t value	Param	t value
Mode Constants				
ASC_{Car}	2.650	4.98	2.560	4.77
ASC_{Bus}	-2.440	-8.56	-2.500	-8.66
ASC_{MThW}	-	-	-	-
ASC_{MTW}	2.270	7.82	2.210	7.94
$ASC_{Car\ SP}$	0.609	2.44	0.750	2.13
$ASC_{Bus\ SP}$	0.160	1.44	0.099	1.26
$ASC_{MThW\ SP}$	-	-	-	-
$ASC_{MTW\ SP}$	0.594	2.54	0.247	1.85
$\text{Comfort}_{\text{Bus}}$	0.579	2.56	0.370	2.05
Comfort_{MThW}	0.424	2.52	0.287	2.03
Cost_{Car}	-0.032	-2.72	-0.026	-2.20
Cost_{Bus}	-0.088	-2.59	-0.055	-2.07
$\text{Cost}_{\text{MThW}}$	-0.071	-2.56	-0.050	-2.06
Cost_{MTW}	-0.074	-2.63	-0.051	-2.08
$\text{Frequency}_{\text{Bus}}$	-0.011	-2.44	-0.008	-2.00
$\text{Time}_{\text{Car SP}}$	-0.019	-1.71	-0.013	-1.34
$\text{Time}_{\text{Bus SP}}$	-0.041	-2.49	-0.013	-2.01
$\text{Time}_{\text{MThW SP}}$	-0.021	-2.54	-0.012	-2.04
$\text{Time}_{\text{MTW SP}}$	-0.014	-2.28	-0.011	-1.88
$\text{SIGMA}_{\text{Time Bus}}$	0.034	2.46		
Error comp (Generic)	0.414	2.53	0.280	2.06
SP to RP scale	3.630	2.58	4.090	2.08
$\text{State Dependence}_{\text{MThW}}$			0.696	2.08
$\text{State Dependence}_{\text{MTW}}$			1.190	2.08
Init log-likelihood	-1957.991		-1957.991	
Final log-likelihood	-1204.359		-1034.056	
Adjusted rho-square	0.375		0.461	

Table 6.7: Value of travel time (VOT) in Rs/hour for different modes

Models	**VOT Car**	**VOT Bus**	**VOT MThW**	**VOT MTW**
MNL (RP)			10.746	10.746
MMNL (RP)			6.934	6.935
SP Model-1	52.080	28.260	32.940	26.940
SP Model-3	90.120	41.760	44.460	34.740
SP Model-4	144.960	45.480	38.400	14.160
SP Model-5	76.560	47.940	43.680	20.580
SP-RP MMNL-1	108.620	48.620	38.740	20.480

Table 6.8: Hourly wage rate calculated from VOT of different modes.

Mode	Average monthly salary(Rs)	Hourly wage rate(Rs)	VOT from model SP-RP MMNL-1	Ratio of VOT to the Hourly wage rate
Car	70000.00	336.54	108.62	32.27
Bus	20000.00	96.15	48.62	50.56
MThW	15000.00	72.12	38.74	53.72
MTW	30000.00	114.23	20.48	17.93

Results on value of travel time (VOT) estimated based on various models are shown in Table 6.7. From the MNL model estimated using the RP data, the value of travel time saving is Rs. 10.746/hr for MThW. From the SP MNL-1 the value of travel times are Rs. 52.08/hr for car, Rs. 28.26/hr for bus, Rs. 32.94/hr for MThW and Rs. 26.94/hr for MTW. In case of SP-RP MMNL-1, value of travel time for car is Rs. 108.62/hr, for bus is Rs 48.62/hr and for MThW is Rs 38.74/hr and for MTW is Rs 20.48/hr. But in case of mixed logit model with random term following triangular distribution the value of travel time has increased compared to other SP models. In case of combined SP–RP model it seems that the value of travel time is more reasonable than that of the other models. Though there is no previous study in the study area to compare to value of travel time but it can be approximately compared with the monthly wage of individual. In case of work trips, i.e. going to work and return trip (not including work trip made during working hours) value of travel time will be less than per hour wage rate. The Value of travel time calculated in all the cases are less than the average hourly wage rate. Details of the value of travel time corresponding to trip makers various chosen modes in terms of hourly wage rate are provided in Table 6.8. The value of travel time estimated for modes other than car were less. This might be because, in case of modes other than car people are giving more importance to the cost parameter compared to the other attributes. Further, a low value of travel time for other modes may be due to the reason that the individual has less income and therefore prioritise his/her expenditure to consume goods/services other than travel.

6.5 Summary and conclusions

In this study, important determinants of travel behaviour have been analysed in the context of smaller Indian cities using the case study of Agartala. Discrete choice models have been prepared to model the mode choice behavior for work trips. Utilities of various modes have been formulated using the socioeconomic, land use and modal related attributes. The following are the important conclusions drawn from the models estimated in this chapter.

1) From the models estimated using the RP data, it can be said that the socioeconomic variables have significant effect on the utility of different modes. Availability of vehicles or vehicle ownership is directly related to the increase in the utility of the respective mode. The utility of MTW increases when the individuals have driving license. Utility of NMT (Bicycle, Cycle Rickshaw and Walking) and public transit/IPT modes (Bus and MThW) significantly increases in case of the individuals living in the locality with mixed land use. The elasticity for area index was found to be higher for cycle rickshaw (0.252) and walk mode (0.288) than cycle mode (0.027).
2) From the models estimated with the SP data, it can be observed that the coefficient of comfort (interacted with income and gender) is higher for bus than MThW. This implies that the policy makers have to play with the comfort of transit parameter so that people can be motivated to use transit mode. From the elasticity values estimated from the SP models, it can be seen that a high elasticity value exists for cost parameter in case of bus, indicating higher sensitivity of the bus users towards the travel cost. Same is the case with the users of MTW i.e., if any other competitive mode (in terms of cost) is available there is a scope for modal shift.
3) When the models are estimated with combined SP-RP data there was a significant improvement in the model performance when compared to the models estimated with only SP data. It can also be observed that state dependence has significant effect on both MTW and MThW users.
4) Compared to the value of travel times obtained from the models estimated with either RP or SP data, the value of travel times obtained from the MMNL model, estimated using combined SP-RP data seems to be realistic based on VOT values with respect to average incomes of the individuals using various modes.

Chapter 7

Effect of latent variables on travel behavior

Mode choice has been mainly studied and modeled as a rational process based upon the random utility theory. Traveler's attitude and perception are some of the important determinants of travel behavior recognized by the researchers for a considerable time. Though there has been considerable interest in analyzing and modeling the effect of these variables, not many works were carried out in this area. This may be due to two primary reasons (Kuppam et al. 1999); detailed data regarding the trip maker's attitudes and preferences are not collected in traditional household survey; and difficulties involved in forecasting these variables. Due to these reasons attitudinal data are not considered useful from practical stand point.

In the area of mode choice modeling, various methods have been used by the researchers to explicitly capture the effect of psychological factors. One method is to include the psychological factors or the indicator variables (such as the responses to survey questions related to the individuals' attitudes and perceptions) directly in the utility function. Ben-Akiva et al. (2002) rejected the use of indicator variables directly in the choice model as they are highly dependent on the phrasing of survey questions besides the difficulties in forecasting. Another method is to carry out factor analysis on the indicators and use the fitted variables in the modal utility (Yanez et al. 2010). General framework of this methodology is based on the assumption that each individual's utility for an alternative depends on a number of observable explanatory variables along with the unobservable variables (Latent variables). Sequential and simultaneous estimations methods were used to estimate the parameters involved in this modelling framework. Sequential method is a two-stage approach where the structural equation model (SEM) is estimated first and then entering the resulting latent variables in the choice models. However, Ben-Akiva et al. (2002) have shown that sequential approach results in inconsistent and inefficient parameter estimates. The latent variables contain measurement errors, and to get consistent estimates the choice probability must be integrated over the distribution of factors obtained from the factor analysis model. Alternately, this integrated model can be estimated simultaneously resulting in efficient estimates, but this method is computationally much more demanding. For specification, apart from utility equation, this type of choice model requires a) measurement equation that links the latent variables to its observable indicator variables measured on a Likert scale and b) structural equation that links the explanatory variables to the latent variables.

Latent variables are the factors related to the individual's behavior and perceptions, and difficult to quantify due to lack of proper measurement scale as different persons may perceive them differently. Moreover, identification of the latent variables requires supplementing a survey, either revealed or stated, with questions that capture user's perception about some aspects of the alternatives. Likert scale is commonly used to measure the attitude and other social factors. Respondent might totally or partially agree, or disagree with a given statement that represents the attributes being assessed. Attitude was measured on a 5 point Likert scale.

Formulation and estimation of a hybrid choice model, formulated to understand the effect of psychological factors on mode choice behavior, is presented in the following sections. Maximum likelihood estimation method was used for model simulation where likelihood function is defined over the joint probabilities of observing the choice and the indicators of latent components. This estimation is done using PYTHON BIOGEME which allows for estimation of latent variable model with simultaneous estimation (Bierlaire and Fetiarison (2009)). In the following section, qualitative aspects of modes have been considered as latent and the estimation of corresponding model is presented. In the ensuing section, travel time is considered to be latent, to model the effects of erroneous network skimmed travel time data.

7.1 Mode choice with latent variables

Data used in this work are taken from the Agartala Household survey which was discussed in detail in the Chapter on data collection. Responses to the qualitative questions provided important insights about the individuals' opinions on travel mode. The respondent had to rate his/her level of agreement on a five point Likert scale ranging from a total disagreement (response of 1) to a total agreement (response of 5). Indicator variables were grouped suitably based on prior assumptions to construct four latent variables, comfort, safety, flexibility, and reliability. Three explanatory variables education, age, gender were used to model the latent variables. Figure 7.1 shows the grouping of indicator variables to construct the latent variables.

In this analysis, two different models were estimated (a) Multinomial logit model (MNL) as base model; and (b) an integrated choice and latent variable model. In the model outlined in Figure 7.2, it is hypothesized that there is an underlying latent attitude among individuals towards different travel modes which vary across the respondents. The structure of the hybrid choice model is provided in the following section. Description of variables used in the models is presented in Table 7.2.

Table 7.1: Observed mean perception rating for different modes

Perception attribute	Car	Bus	MThW	MTW	Cycle	Cycle rickshaw
Comfortable in journey	4.11	2.99	3.23	3.73	3.00	3.57
Always availability of comfortable seats	3.78	3.13	3.39	3.69	3.09	3.63
Very easy accessibility	3.69	3.08	3.50	3.90	3.08	3.31
Ability to reach destination in time	3.91	3.22	3.53	3.90	3.18	3.21
Can exactly calculate travel time prior to trip	3.90	3.16	3.37	3.91	3.10	3.21
Safety from accident	3.83	3.55	3.43	3.24	2.95	3.15
Safety from theft	3.92	3.18	3.35	3.69	3.17	3.17
Safety from weather	3.89	3.55	3.49	2.20	2.20	3.01
Ability to make more trips	3.85	3.03	3.51	3.68	2.72	3.02
Can travel without changing vehicles	3.94	3.09	3.40	3.93	3.93	3.16

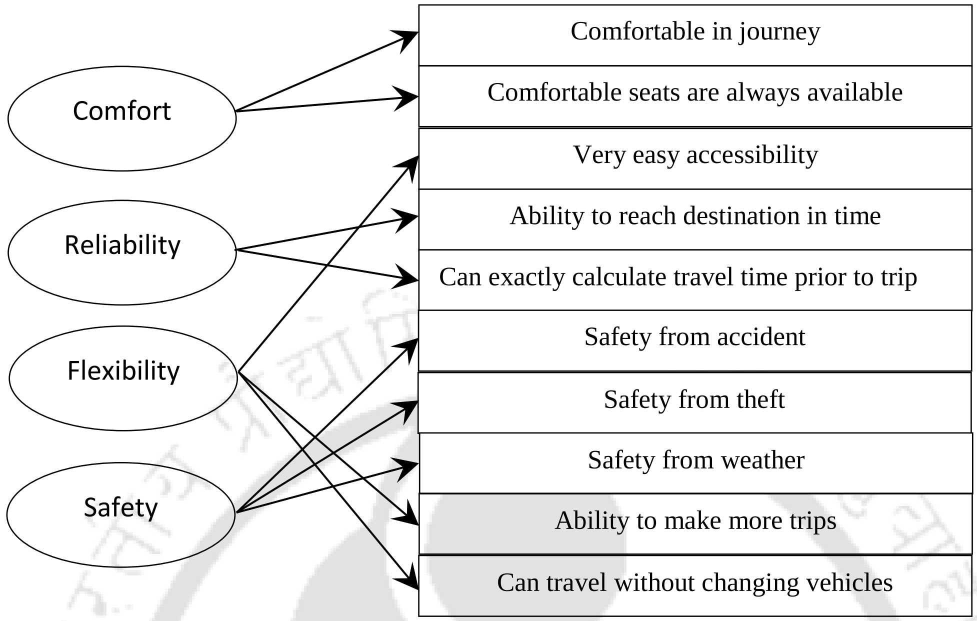


Figure 7.1: Construction of latent variable as a function of indicator variables

Table 7.2: Description of variables used in modeling latent variable choice models

SN	Variables	Description	
1	Age	Age in years	
2	Private Vehicle Ownership	Number of private vehicles in a household	
3	Education	Years of education.	
4	Gender	Dummy variable 1 for male respondent 0 for female respondent	
5	License	1 for respondent having license 0 for respondent not having license	
6	Family Size	Number of members in the family	
7	Income	Gross monthly household income	Code
		0-2000	1
		2001-5000	2
		5001-10000	3
		10001-15000	4
		15001-20000	5
		20001-30000	6
		30001-40000	7
		40001-50000	8

		50001-70000	9
		70001-90000	10
		90001-150000	11
		>150001	12
8	Bicycle Ownership	Dummy Variable 1 for households having bicycle 0 for households having no bicycle	
9	MTW Ownership	Dummy Variable 1 for households having Motorized Two wheeler 0 for households having no Motorized Two wheeler.	
10	Time	In vehicle travel time.	
11	Cost	Fare or fuel cost	
12	Area Index	Land use parameters obtained from GIS analysis.	
13	LV$comfort_{car}$	Comfort of car measured by indicator variable	
14	LV$flex_{\mathrm{MTW}}$	Flexibility of MTW measured by indicator variable	
15	LVflex $_{\mathrm{MThW}}$	Flexibility of MThW measured by indicator variable	
16	LV Education	Education in number of years used in structural equation	
17	LV Age	Age in number of years used in structural equation	
18	LV Gender	Gender as dummy variable used in structural equation	
19	ASC_{Car} ASC_{Bus} ASC_{MThW} ASC_{MTW} $ASC_{Bicycle}$ $ASC_{Rickshaw}$ ASC_{Walk}	Alternate specific constants for different modes.	

7.1.1 Specification for structural equation model

Various specifications of structural equation model have been tried and the final model used in the present analysis is shown below. Reliability and safety could not be modeled using any of the socioeconomic characteristics. In the final model comfort related to car, flexibility associated to MTW and MThW are found to be significant in explaining the mode choice behavior. Socioeconomic variables that are significantly entering the SEM were also found to be significant in modeling the systematic component of the preference heterogeneity (from the previously estimated models). In the final SEM model, age, gender, and education variables are used in modeling the comfort of car and flexibility of the MTW and MThW.

$$LV\textit{comfort}_{car} = \lambda_{1car} * \text{education} + \lambda_{2car} * \text{age} + \lambda_{3car} * \text{gender} + \zeta_{car} \quad (7.1)$$

$$LV\textit{flex}_{MTW} = \lambda_{1\,MTW} * \text{education} + \lambda_{2\,MTW} * \text{age} + \lambda_{3\,MTW} * \text{gender} + \zeta_{MTW} \quad (7.2)$$

$$LV\text{flex}_{MThW} = \lambda_{1\,MThW} * \text{education} + \lambda_{2\,MThW} * \text{age} + \lambda_{3\,MThW} * \text{gender} + \zeta_{MThW} \quad (7.3)$$

where, λ is the associated parameter to be estimated and ζ is the error term normally distributed with zero mean and standard deviation σ.

7.1.2 Specification for the measurement equations

Measurement equations explain the relationship between the latent variables and the indicator variables. Following are the possible significant relationships between various indicator variables and the latent variables associated to Car, MTW, and MThW. All the relationships are found to be logically convincing.

$$\text{Comfortable journey}_{car(Indicator)} = \Upsilon_{2Car} * LV\textit{comfort}_{car} + \delta_{1car} \quad (7.4)$$

$$\text{Comfortable seat}_{car(Indicator)} = \Upsilon_{2Car} * LV\textit{comfort}_{car} + \delta_{2car} \quad (7.5)$$

$$\text{Accessibility}_{MTW\,(Indicator)} = \Upsilon_{1\,MTW} * LV\textit{flex}_{MTW} + \delta_{1\,MTW} \quad (7.6)$$

$$\text{Ability to make more trips}_{MTW\,(Indicator)} = \Upsilon_{2\,MTW} * LV\textit{flex}_{MTW} + \delta_{2\,MTW} \quad (7.7)$$

$$\text{Travel without changing vehicle}_{MTW\,(Indicator)} = \Upsilon_{3\,MTW} * LV\textit{flex}_{MTW} + \delta_{3MTW} \quad (7.8)$$

$$\text{Accessibility}_{MThW\,(Indicator)} = \Upsilon_{1\,MThW} * LV\textit{flex}_{MThW} + \delta_{1\,MThW} \quad (7.9)$$

$$\text{Ability to make more trips}_{MThW\,(Indicator)} = \Upsilon_{2\,MThW} * LV\textit{flex}_{MThW} + \delta_{2\,MThW} \quad (7.10)$$

$$\text{Travel with changing vehicle}_{MThW\,(Indicator)} = \Upsilon_{3\,MThW} * LV\textit{flex}_{MThW} + \delta_{3\,MThW} \quad (7.11)$$

where, Υ is the associated parameter to be estimated and δ is the error term normally distributed with zero mean and standard deviation σ.

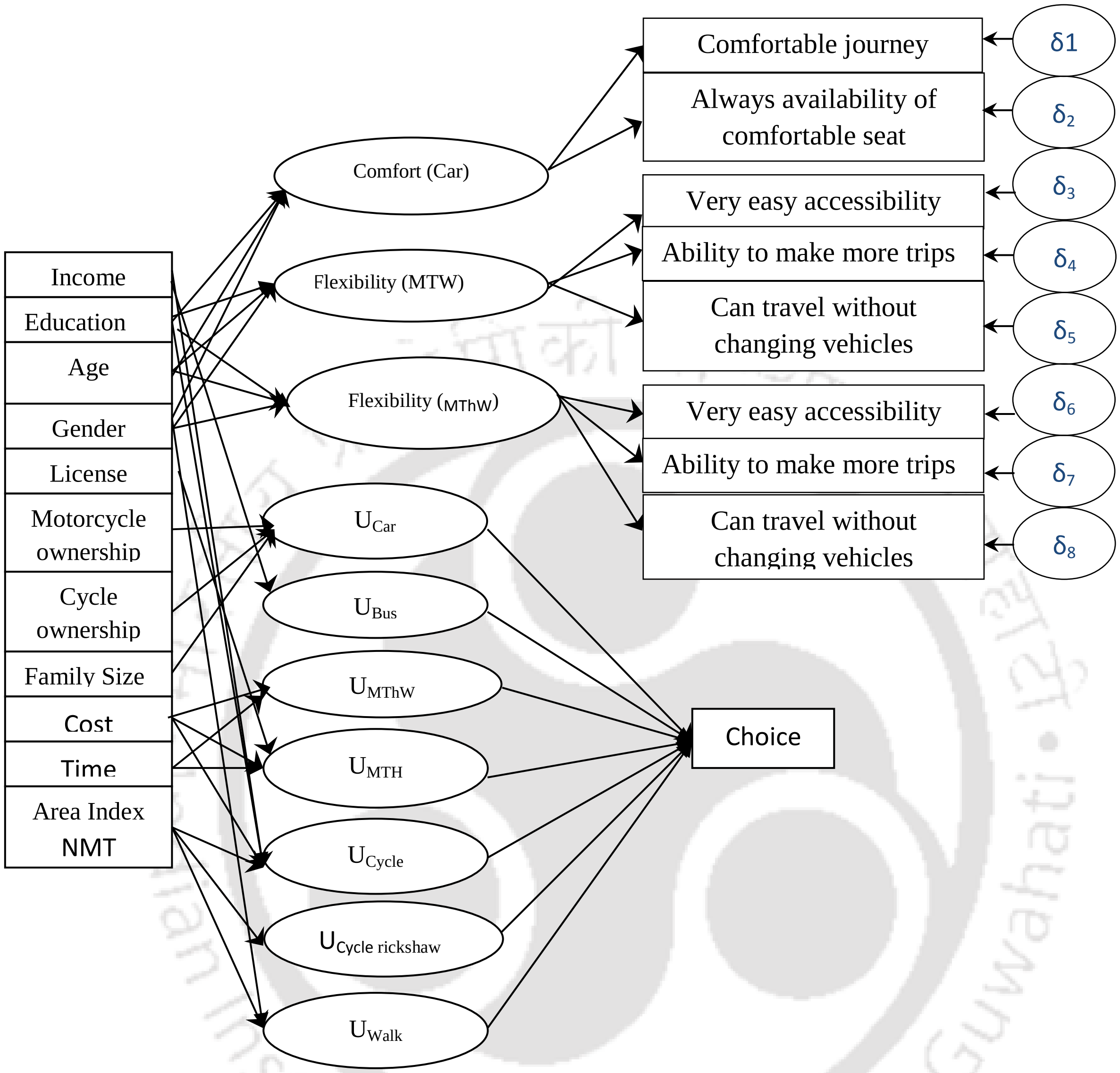


Figure 7.2: Schematic diagram of simultaneous latent variable and mode choice model

7.2 Travel time as Latent variable

Travel time is one of the most important modal attributes used to capture the observed component of modal utility. As discussed earlier, travel time data are estimated from the road network database. Mcfadden (2000) reported that accurate measurement of travel time and cost component at individual level is susceptible to bias and measurements taken from individuals or users generally show systematic biases which affects disaggregate models. Walker et al. (2010) suggested latent variable approach to correct the measurement error which generally gets introduced in the travel demand modelling. They have treated true travel time as latent variable and measured travel time as an indicator variable to latent or true travel time. In this study a similar methodology has been applied to correct the measurement error, if any, by estimating a hybrid choice model.

To determine the travel time more efficiently and accurately, travel time was measured on different roads in the city and then loaded on the network (detailed explanation is provided in chapter 3). As explained previously, in the small cities bus stops and MThW stops are not defined properly and therefore calculation of travel time for a particular trip contains error. Further, travel time corresponding to access and egress components of the trip is not properly provided by the individuals. Keeping in view of the above sources of error involved in travel time measurement, mode choice model was estimated keeping the true travel time as latent variable which is a function of measured travel time (indicator variable) from the network data. Out of various socioeconomic characteristics only 'years of education' was found to be significant in modeling the latent variable 'true travel time'. The framework for error in measurement for travel time correction is given in Figure 7.3.

7.3 Results and Discussion

Table 7.3 shows the result obtained from a basic MNL model and two hybrid choice models with the specification described in Figure 7.2 and Figure 7.3. For comparative analysis, same utility functions of the MNL model were used in the hybrid choice model. The parameters were estimated with the extended version of software package BIOGEME (2009).

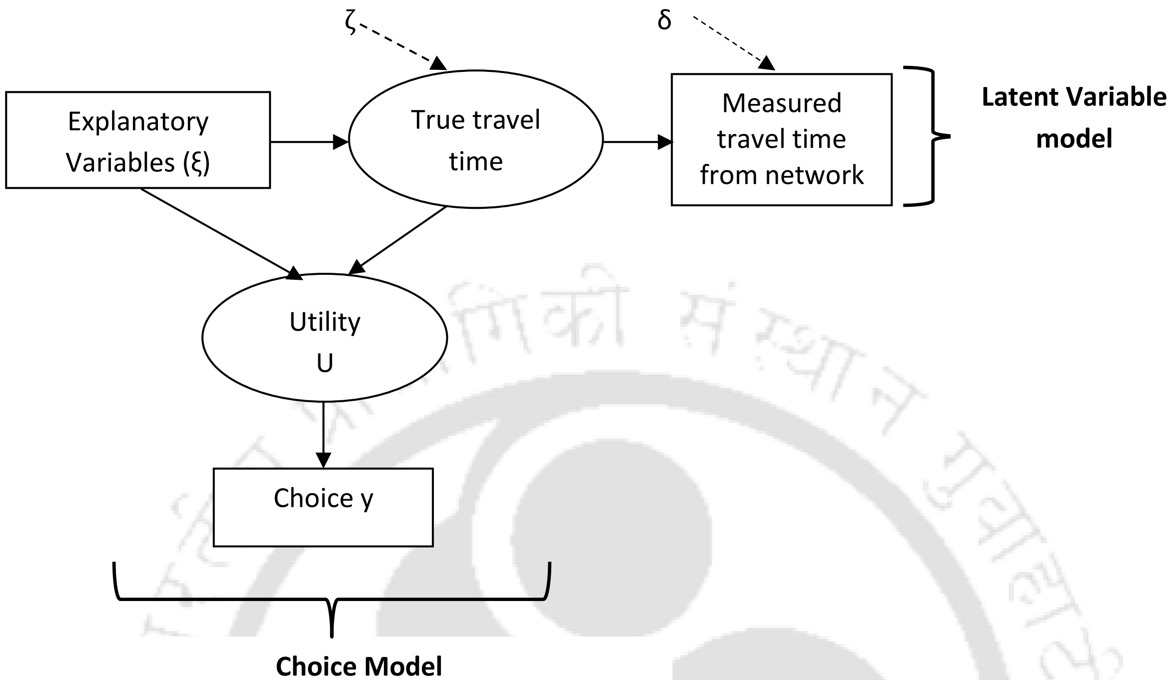


Figure 7.3: Framework for measurement error correction for travel time

Finding out right specification for the hybrid choice model is computationally intensive process (La paix et al. 2011). After testing various specifications best model is presented in Table 7.3. The first column in the Table 7.3 represents base model, second column shows model for measurement correction for travel time and third column shows the effect of psycological factors in mode choice behavior.

It can be seen from the results shown in Table 7.3, education, age and gender are found to be statistically significant in explaining the latent variable 'comfort car'. Education, age, and gender are found to be significant in case of flexibility of MTW and MThW. In the above cases, the socioeconomic characteristics have positive effect on the latent variables which seems to be logically correct.

Table 7.3: Result from the MNL and hybrid choice models

Variable description	MNL Model		Model with latent travel time		Model with perception latent variables	
	Param	t-statistic	Param	t-statistic	Param	t-statistic
ASC $_{Car}$	-0.616	-0.39	-0.762	-0.48	-6.03	-1.36
ASC $_{Bus}$	-4.420	-6.70	-4.360	-6.67	-1.55	-1.28
ASC $_{MThW}$	0.000					
ASC $_{MTH}$	-1.750	-2.17	-1.320	-1.60	-5.76	-1.73
ASC $_{Bicycle}$	-0.197	-0.16	-0.072	-0.05	3.37	1.87
ASC $_{Rickshaw}$	-0.597	-1.18	-0.324	-0.62	1.08	0.80
ASC $_{Walk}$	-0.196	-0.27	0.749	0.91	2.63	2.00
Age $_{Bus}$						
Area Index * Income $_{NMT}$	0.375	4.01	0.353	3.66	0.389	3.63
Cost $_{MThW, MTW}$	-0.128	-3.87	-0.121	-3.57	-0.142	-3.81
Cost $_{Rickshaw}$	-0.040	-2.16	-0.048	-2.52	-0.040	-2.07
Bicycle Ownership $_{Car}$	-1.700	1.77	-1.710	-1.60	-3.20	-2.21
Bicycle Ownership $_{Bicycle}$	4.830	4.07	4.800	4.01	4.72	3.92
Education $_{Rickshaw}$					0.129	2.04
Family Size $_{Car}$	-0.398	-2.04	-0.373	-1.93	-0.661	-2.83
Family Size Bicycle	-0.567	-2.74	-0.550	-2.63	-0.591	-2.75
Gender Walk	1.290	2.81	1.260	2.60	1.50	3.14
Income Car	0.588	3.95	0.596	3.95		
Income Bus	0.166	1.71	0.161	1.68	0.212	2.12
License $_{MTW}$	3.290	5.92	3.350	5.80	3.04	4.75
Income walk	-0.299	-2.90	-0.303	-2.74	-0.232	-2.15
Time $_{MThW, MTW}$	-0.025	-1.72	-0.056	-1.65	-0.029	-1.91
Time Cycle	-0.066	-3.24	-0.113	-2.93	-0.0738	-3.23
Time Walk	-0.037	-2.09	-0.112	-3.61	-0.034	-1.98
MTW Ownership Car	-1.760	-2.52	-1.780	-2.50	-2.04	-2.21
MTH Ownership $_{MTW}$	1.710	2.32	1.720	2.36	1.93	2.31
Comfort $_{Car}$					1.82	2.91
Flexibility $_{MThW}$					0.518	2.80
Flexibility $_{MTW}$					1.05	2.35
			Latent variable model			
LV Education(Car)					0.260	9.62
LV Age(Car)					0.075	9.45
LV Gender(Car)					0.980	5.31
Comfortable journey (Car)					0.536	11.77
Comfortable seats(Car)					0.497	11.71
LV Education ($_{MTW}$)					0.291	12.67
LV Age ($_{MTW}$)					1.07	6.77
LV Gender ($_{MTW}$)					0.051	9.76

Easy accessibility ($_{MTW}$)					0.524	15.42
Ability to make more trips ($_{MTW}$)					0.558	15.44
Travelling without changing vehicles ($_{MTW}$)					0.553	15.44
LV Education ($_{MThW}$)					0.228	11.08
LV Age ($_{MThW}$)					1.01	6.24
LV Gender ($_{MThW}$)					0.064	10.53
Easy accessibility ($_{MThW}$)					0.518	14.67
Ability to make more trips ($_{MThW}$)					0.501	14.67
Travelling without changing vehicles ($_{MThW}$)					0.516	14.67
VOT $_{MThW,\ MTW}$	Rs11.72/hr		Rs 27.77/hr		Rs 12.25/hr	
Init log-likelihood	-679.526		-15655.969		-37278.722	
Final log-likelihood	*-315.526*		-9406.292		-6309.949	
Adjusted rho-square	0.502		0.397		0.830	

When compared the results of hybrid choice model with the MNL base model, there are significant changes observed in the values of the estimated parameters. This is due to the absence of latent variable in the base MNL model that led to the underestimation of the significance of most of the socioeconomic variables in explaining the preference heterogeneity. The parameters obtained from hybrid choice models are statistically more significant than parameters obtained from the base MNL model. Adjusted Rho square is higher for hybrid choice model (0.83) than the MNL model (0.502).

Hybrid choice model with latent travel time variable was estimated with an intention to correct the measurement error in travel time and to get a better estimate of travel time co-efficient. Travel time is taken as generic in case of MThW and MTW. When compared with the MNL base model, there are significant changes observed in the value of estimated parameters corresponding to travel time. There was a significant improvement in the coefficient of travel time for MThW and MTW due to which the value of travel time has increased from Rs 11.72/hour to Rs 27.77/hour that can be considered nearer to the actual VOT for this two modes. These values of travel time, obtained from the hybrid choice model, are relatively higher compared to the result obtained from the mixed logit models estimated with the RP data and SP-RP data. From these results, it can be said that inclusion of latent variables in the choice model lead to better performance of models by improving the model fit and also improving the coefficient estimates.

Chapter 8

Summary and Conclusions

Mode choice analysis and modeling of the trip makers living in a small sized Indian city, Agartala, capital of a north-east Indian state, has been carried out in this study. In this process, this study has tried to address the issues related to the LOS variables and the choice set availability. Using discrete choice modeling framework, this study has addressed three important aspects of choice modeling. First is to analyze and understand the effect of mixed land use on mode choice in the context of smaller Indian cities. Second aspect is to determine the important determinants of mode choice through an appropriately specified choice model. Third aspect is to understand the effects of psychological factors and the error in travel time data on choice modeling using hybrid choice modeling approach. Important conclusions drawn out of the present study are also categorized into three parts and presented below.

8.1 Land use effects on travel behavior

In this study, impact of land use variables on travel behavior has been analyzed in the context of smaller Indian cities. Travel behavior was quantified in terms of trip lengths, motorized/non-motorized mode choice, motorized private and public/IPT mode choice as well as the choice of individual modes. Land use mix has been measured using entropy, area index, DI, and mix type index. Approach used in calculating the entropy and DI is slightly different from that of the conventional approach. Area index has been calculated for origin and destination of trip and this index takes different values based on the trip purpose. In quantifying the mix using modified DI and Mix type Index, cell size of 10m x 10m and grid size of 1km x 1km have been used. Entropy measured with 1000m buffer performed better than the entropy measured using the conventional method.

The following are the important conclusions drawn from the study on interaction between land use and travel behavior;

1. When slightly modified approach was used for calculating the entropy and dissimilarity indices, both the land use parameters were able to quantify the land use mix and were consistently having significant effect on the travel parameters.
2. Trip lengths of the individuals, for both the shopping and work trips, were strongly correlated to the land use mix variables, even when controlling the socioeconomic characteristics. In case of work trips, Entropy measured with 1000 m radius buffer, and DI measured using 10 m x 10 m cell for 1km x 1km tract were found to be more significant than the conventional entropy and dissimilarity indices measured for census tract. When the area index for both the origin and destination were considered there was almost 395% increase in the model's ability to explain the variability of trip length compared to the base model.
3. All the land use variables were negatively correlated with the trip length for shopping trips.
4. Elasticity analysis shows that, an increase in land use mix (area index) by 1 % will reduce trip length by 0.23% and 0.326 % for work and shopping trips respectively.
5. From the estimated binary logit models, it can be said that both socioeconomic parameters and land use parameters are significant in explaining the mode choice.
6. In case of non-motorized and motorized mode choice for work trips, land use variables were found to be significantly influencing the goodness of fit of the model. Area Index values, for both origin and destination, were found to be significant.
7. A significant elasticity exists (0.535) between the intersection density and the utility of non-motorized modes, for work trips. This may be due to the availability of more cycle rickshaws in the areas where the intersection density is high. In case of shopping trips, land use mix measured by entropy and DI in conventional way were insignificant. This clearly shows the disadvantages of entropy index measured for census tract and DI

measured using 100 m x 100 m cells for census tract. Land use mix measured using a slightly different approach, could sufficiently capture the variations in the mode choice.

8. In case of motorized private and public/IPT mode choice for work trips, the utility of private mode increases with the increase in the intersection density which is counterintuitive. This may be due to the upper middle class people, owning vehicles, and residing in mixed land use area. In this case effect of mixed land use was found to be negligible. Same is the case with the shopping trips also.
9. From the MNL model estimated on the mode choice for work trips, Area Index, Entropy with 1km buffer and DI with 10m x 10m cell, Mix type index with 10m x 10m cell were found to be significant. The coefficients are positive for public and non-motorized transport which implies that the trip makers residing in the areas with mixed land use prefer public and non-motorized transport. Conventional DI and entropy indices were not able to capture the effect of land use mix for public transport; but they were significant in capturing the effect of mixed land use on NMT modes. It can be seen that the probability of choosing walk and cycle rickshaw increases with the increasing land use mix. An increase in area index by 1 % increases the probability of choosing cycle rickshaw by 0.265%. In case of motorized modes, the elasticity values with respect to choosing public transport are not strongly related. The elasticity of NMT is higher than public transport modes.

Overall, it can be concluded that the land use mix has significant influence on the trips made by non-motorized modes and the public transport mode. From the elasticity's between the area index and the trip lengths associated to the non-motorized and public transport modes, it can be said that any reduction in the land use mix may lead to significant changes in the trip lengths.

8.2 Mode choice models

In this study, important determinants of mode choice behaviour have been determined. Discrete choice models have been prepared to model the mode choice behavior for work trips. Utilities of various modes have been formulated using the socioeconomic, land use and modal related attributes. The following are the important conclusions drawn from the models estimated for mode choice behavior;

1. From the models estimated using the RP data, it can be said that the socioeconomic variables have significant effect on the utility of different modes. Availability of vehicles or vehicle ownership is directly related to the increase in the utility of the respective mode. The utility of MTW increases when the individuals have driving license. Utility of Non-Motorized Transport (Bicycle, Cycle Rickshaw and Walking) and public transit/IPT modes (Bus and MThW) significantly increases in case of the individuals living in the locality with mixed land use. The elasticity for Area index was found to be higher for cycle rickshaw (0.252) and walk mode (0.288) than cycle mode (0.027).
2. From the models estimated with the SP data, it can be observed that the coefficient of comfort (interacted with income and gender) is higher for bus than MThW. This implies that the policy makers have to play with the comfort of transit so that people can be motivated to use transit mode. From the elasticity values estimated from the SP models, it can be seen that a high elasticity value exists for cost parameter in case of bus, indicating higher sensitivity of the bus users towards the travel cost. Same is the case with the users of MTW i.e., if any other competitive mode (in terms of cost) is available there is a scope for modal shift.
3. When the models are estimated with combined SP-RP data, there was a significant improvement in the model performance when compared to the models estimated with only SP data. It can also be observed that state dependence has significant effect on both MTW and MThW users.
4. Compared to the value of travel times obtained from the models estimated with either RP or SP data, the value of travel times obtained from the MMNL model, estimated using combined SP-RP data seems to be realistic based on VOT values with respect to average incomes of the individuals using various modes.

Based on the elasticities calculated from SP-5 model, it can be said that mode shift is influenced by fare of bus, MTW. Further, it has been observed that there is a significance of bus fare on the choice probability corresponding to various modes.

Results from the mode choice analysis also show that modal share can be controlled by regulating fares. From the elasticity values are reported in the thesis from analysis of SP survey, the change in mode share can be controlled by regulating the fare of the mode. Combination of both a) increase in mixed land use b) Regulation of fare/ price can bring about desired modal share.

8.3 Hybrid choice models

To understand the effect of attitudinal factors on mode choice behavior, hybrid choice models were estimated. Comfort, reliability, flexibility and safety are the latent variables which were constructed by grouping relevant indicator variables. Further, in order to accommodate the measurement error of the travel time data, this variable has been considered as latent and a hybrid choice model has been estimated. The following are the important conclusions drawn from the results of hybrid choice models;

1. The parameters obtained from hybrid choice models are statistically more significant than parameters obtained from the base MNL model. When latent variables like comfort of car, flexibility of MTW, and flexibility of MThW were included in the model, the fitness of the model improved considerably, adjusted rho square increased to 0.83 than the base MNL model (0.502).
2. In hybrid choice model for travel time measurement error correction, there was a significant improvement in the coefficient of travel time for MThW and MTW due to which the value of travel time has increased from Rs 11.72/hour (from the MNL model estimated using the RP data) to Rs 27.77/hour.

From these results, it can be said that inclusion of latent variables in the choice model lead to better parameter estimates and improvement in the model fit.

8.4 Research Contribution

Contribution of this thesis towards understanding the mode choice behavior of trip makers of smaller Indian cities is briefly given below.

1. Mixed land use indices have been modified and the modified indices were found to be significant in quantifying the land use mix compared to the conventional land use parameters.
2. In this work, a detailed land use database has been prepared for Agartala city. This database would be useful for further research on the land use effect on travel behavior.
3. Important determinants of travel behavior found through this study can be used for formulating appropriate transport policies in the context of smaller Indian cities.

8.5 Further scope of study

Based on the findings of the present study, it can be said that mixed land use, attitudinal variable play major role in mode choice. To establish these findings, further studies can be carried out on the following aspects :

1. In characterization of land use mix, DI can further be modified to account for land use mix based on the location of the household. Land use mix in multilevel buildings is neglected in this study hence further studies can be taken up in this area.
2. Modeling of residential sorting effect may be carried out to understand the actual effect of land use mix on mode choice.
3. This study provides overall understanding of mode choice behavior for smaller Indian cities. There are many aspects such as modeling the choice set availability, determination of LOS data of non-chosen modes which can be taken up in further studies.
4. Modeling of travel time as latent variable may be carried out using the travel time data obtained from the individuals as indicator variable.
5. Modeling the effect of land use mix for different trip purposes.

References

Adamowicz, W., Swait, J., Boxall, P., Louviere, J., and Williams, M. (1997). “Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation.” Journal of environmental economics and management, 32(1), 65-84.

Ahmed, M. A., and Datta, R. N. (2006). “Utility of paratransit modes in cities of Assam, India.”Transportation Research Record: Journal of the Transportation Research Board, 1971(1), 107-115.

Algers, S., Bergstroem, P., Dahlberg, M., and Dillen, J.L. (1998). “Mixed logit estimation of the value of travel time.” Working Paper, Department of Economics, Uppsala University, Uppsala, Sweden.

Ben-Akiva, M. Walker, J., Bernardino, A., Gopinath, D., Morikawa, T. and Polydoropoulou, A. (2002) “Integration of Choice and Latent Variable Models,” In Perpetual Motion: Travel Behavior Research Opportunities and Application Challenges, Edited by Mahamassani H., Elsevier Science, 431-470.

Ben-Akiva, M., and Morikawa, T. (1990). "Estimation of switching models from revealed preferences and stated intentions." Transportation Research Part A: General, 24(6), 485-495.

Best, H., and Lanzendorf. M. (2005). “Division of labour and gender differences in metropolitan car use: an empirical study in Cologne, Germany.” Journal of Transport Geography, 13(2), 109-121.

Bhat, C. R. (1998). “Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling.” Transportation Research Part A: Policy and Practice, 32(7), 495-507.

Bhat, C. R., and Castelar, S. (2002). “A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay area.” Transportation Research Part B: Methodological, 36(7), 593-616.

Bhat, C.R. (2000), “Incorporating Observed and Unobserved Heterogeneity in Urban Work Mode Choice Modeling,” Transportation Science, 34(2), 228-238.

Bhat, C.R., and Koppleman, F.S. (2006). *A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models*. <http://www.ce.utexas.edu/prof/bhat/COURSES / LM_Draft_060131 Final-060630.pdf>.

Bhat, C.R., and Gossen, R. (2004). “A mixed multinomial logit model analysis of weekend recreational episode type choice.” Transportation Research Part B, 38 (9), 767–787.

Bierlaire, M. (2003). “BIOGEME: A free package for estimation of discrete choice models.” Proceedings of 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Bierlaire, M.(2008). *Estimation of Discrete choice models with BIOGEME 1.8*, biogeme.epfl.ch

Bierlaire, M., and Fetiarison, M. (2009). "Estimation of discrete choice models: extending biogeme." Proceedings of the 9th Swiss Transport Research Conference (STRC), Ascona, Switzerland.

Bierlaire, M., Hurtubia, R., and Flotterod, G. (2010). "Analysis of implicit choice set generation using a constrained multinomial logit model." Transportation Research Record: Journal of the Transportation Research Board, 2175, 92-97.

Bolduc, D., Ben-Akiva, M., Walker, J., and Michaud,A. (2005). "Hybrid Choice Models with Logit Kernel: Applicability to Large Scale Models," in (M.E.H. Lee-Gosselin and S.T. Doherty, eds.) Integrated Land-Use and Transportation Models: Behavioral Foundations, Emerald Group Publishing, 275-302.

Bradley, M. A., and Daly, A. J. (1991). "Estimation of Logit Choice Models Using Mixed Stated Preference and Revealed Preference Information." Paper presented to the 6th International Conference on Travel Behavior, Quebec, May 22-24.

Bradley, M., and Daly, A. (1994). "Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data." Transportation, 21(2), 167-184.

Brownstone, D., Bunch, D. S., Golob, T. F., and Ren, W. (1996). "A transactions choice model for forecasting demand for alternative-fuel vehicles." Research in Transportation Economics, 4, 87-129.

Cao, X., Mokhtarian, P. L., and Handy, S. L. (2009). "Examining the impacts of residential self-selection on travel behaviour: a focus on empirical findings." Transport Reviews, 29(3), 359-395.

Cervero, R. (1989a). "America's Suburban Centres: The Land Use Transportation Link." Unwin Hyman, London.

Cervero, R. (1989b). "Land-Use Mixing and Suburban Mobility." UC Berkeley: University of California Transportation Center. Retrieved from: <http://escholarship.org/uc/item/4nf7k1v9>. Accessed on 10 September 2013

Cervero, R. (1991). "Land Uses and Travel at Suburban Activity Centers." UC Berkeley: University of California Transportation Center. Retrieved from :< http://escholarship.org/uc/item/0d08h1bz >. Accessed on 10 September 2013

Cervero, R. (1995). "Mixed Land-Uses and Commuting: Evidence from the American Housing Survey." Transportation Research A, Policy and Practice, 30(5), 361-377.

Cervero, R., and Duncan, M. (2003). "Walking, bicycling and urban landscapes: evidence from the San Francisco Bay Area." Am J Public Health; 93(9), 1478–83.

Cervero, R., and Kockelman, K. (1997). “Travel Demand and the 3ds: Density, Diversity, and Design.” Transportation Research Part D: Transport and Environment. 2(3), 199-219.

Cervero, R., and Radisch, C. (1996). “Travel choices in pedestrian versus automobile oriented neighborhoods.” Transport Policy, 3(3), 127-141.

Chapman, J., and Frank, L. (2007). “Integrating Travel Behavior and Urban Form Data to Address Transportation and Air Quality Problems in Atlanta.” Report retrieved from < http://hdl.handle.net/ 123456789/8137>

Chatterjee, K. (2011). “Modelling the dynamics of bus use in a changing travel environment using panel data.” Transportation, 38(3), 487-509.

Cherchi, E., and Ortuzar, J. D. D. (2003). “Alternative specific variables in non-linear utilities: influence of correlation, homoscedasticity and taste variations.” In 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Chintakayala, P. K., and Maitra, B. (2010). “Modeling generalized cost of travel and its application for improvement of taxies in Kolkata.” Journal of Urban Planning and Development, 136(1), 42-49.

City Development Plan-Agartala. Report prepared by LEA Associates South Asia Pvt.Ltd in 2006. Retrieved from :< http://agartalacity.tripura.gov.in/PDF/Utilities/CityDevP.pdf>

Daly, A., Hess, S., and Train, K. (2012). "Assuring finite moments for willingness to pay in random coefficient models." Transportation, 39(1), 19-31.

Das, S. S., Maitra, B., and Boltze, M. (2009). “Valuing Travel Attributes of Rural Feeder Service to Bus Stop: Comparison of Different Logit Model Specifications.” Journal of Transportation Engineering, 135(6), 330-337.

Dissanayake, D., and Morikawa, T. (2010). “Investigating household vehicle ownership, mode choice and trip sharing decisions using a combined revealed preference/stated preference Nested Logit model: case study in Bangkok Metropolitan Region” Journal of Transport Geography, 18(3), 402-410.

Dissanayake, D., Kurauchi, S., Morikawa, T., and Ohashi, S. (2012). “Inter-regional and inter-temporal analysis of travel behaviour for Asian metropolitan cities: Case studies of Bangkok, Kuala Lumpur, Manila, and Nagoya.” Transport Policy, 19(1), 36-46.

Enam, A., and Choudhury, C. F. (2011). “Methodological issues in developing mode choice models for dhaka, bangladesh.” Transportation Research Record: Journal of the Transportation Research Board, 2239, 84-92.

Ewing, R., and Cervero, R. (2010). “Travel and the Built Environment: A Meta-Analysis.” J. Am. Plan. Assoc. 76(3), 265–294.

Frank, L., and Pivo, G. (1994). "The Impacts of Mixed Use and Density On the Utilization of Three Modes of Travel: The Single Occupant Vehicle, Transit, and Walking." Transportation Research Record: Journal of the Transportation Research Board, 1466, 44-52.

Giuliano, G., and Dargay, J. (2006). "Car ownership, travel and land use: a comparison of the US and Great Britain." Transportation Research Part A, 40(2), 106-124.

Hanson, S., and Hanson, P. (1980). "Gender and urban activity patterns in Uppsala, Sweden." Geographical Review,10(3), 291-299.

Hanson, S., and Johnston, I. (1985). "Gender differences in work-trip length: explanations and implications." Urban Geography, 6(3), 193-219.

Hensher, D. A., and Bradley, M. (1993). "Using stated response choice data to enrich revealed preference discrete choice models." Marketing Letters, 4(2), 139-151.

Hensher, D. A., Rose, J. M., and Greene, W. H. (2008). "Combining RP and SP data: biases in using the nested logit 'trick'–contrasts with flexible mixed logit incorporating panel and scale effects." Journal of Transport Geography, 16(2), 126-133.

Hensher, D., Rose, J.M., and Greene, W.H. (2005). *Applied Choice Analysis A primer*, Cambridge University Press. New York, USA.

Hess, P.M., Moudon, A.V., and Logsdon, M.G. (2001). "Measuring Land Use Patterns for Transportation Research" Transportation Research Record: Journal of the Transportation Research Board,1780, 17-24.

Hess, S., Bierlaire, M., and Polak J. W. (2005). "Estimation of value of travel-time savings using mixed logit models" Transportation Research Part A, 39(2-3), 221-236

Johansson, M.V, Heldt, T., and Johansson, P. (2006). "The effects of attitude and personality traits on mode choice." Transportation Research Part A, 40(6), 507-525

Kitamura, R., and Kostyniuk, L.P. (1986). "Maturing motorization and household travel: The case of nuclear-family households." Transportation Research A, 20(3), 245-60.

Kuppam, A.R., Pendyala, R.M., and Rahman, S. (1999). "Analysis of the Role of Traveler Attitudes and Perceptions in Explaining Mode Choice Behavior." Transportation Research Record 1676, Journal of the Transportation Research Board, pp. 68-76.

Kitrinou, E., Polydoropoulou, A., and Bolduc, D. (2009). "Development of Integrated Choice and Latent Variable (ICLV) models for the residential relocation decision in island areas." International Choice Modelling Conference, Institute for Transport Studies, University of Leeds.

Kostyniuk, L.P., and Cleveland, D.E. (1978). “Gender-role identification in the methodology of transportation planning.” Research and special programs’s administration, US department of transportation,Proceedings of The Conference on Women’s Travel Issues, edited by Sandra Rosenbloom. 17-20,Washington DC,USA.

Limtanakool, N., Dijst, M., and Schwanen, T. (2006). “The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips.” Journal of transport geography, 14(5), 327-341

Litman, T. (2012). “Land use impacts on transport: How land use factors affect travel behavior.” Report from Victoria Transport Policy Institute, retrieved from < http://www.vtpi.org/landtravel.pdf>.

Louviere, J., Hensher, D., and Swait, D. (2000). *Stated Choice Methods, Analysis and Application*, Cambridge University Press, U.K.

Ma, Y., Chen, X (2013). “Geographical and statistical analysis on the relationship between land-use mixture and home based trip making and more: case of Richmond,Virginia.” Journal of Urban and Regional Analysis, 1, 5-44.

Maat, K., Wee, V.B., and Stead, D. (2005). “Land use and travel behaviour: expected effects from the perspective of utility theory and activity-based theories.” Environment and Planning B: Planning and Design, 32(1), 33-46.

Manski, C. (1977). “The structure of random utility models.” Theory and Decision, 8(3), 229-254.

Martinez, F., Aguila, F., and R. Hurtubia (2009). “The constrained multinomial logit model: A semi-compensatory choice model.” Transportation Research Part B: Methodological, 43(3), 365–377.

McFadden, D. (2000). “Disaggregate Behavioral Travel Demands RUM Side: A 30-Year Retrospective.” Resource Paper, Department of Economics. University of California, Berkeley.

Messenger, T., and Ewing, R. (1996). “Transit-oriented development in the sun belt.” Transportation Research Record: Journal of the Transportation Research Board, 1552(1), 145-153.

Morikawa, T. (1989). “Incorporating stated preference data in travel demand analysis” Doctoral dissertation, Massachusetts Institute of Technology, Retrieved from < http://dspace.mit. edu/handle/1721.1/14326>.

MoUD, Govt. of India report on “Study on Traffic and Transportation Policies and strategies in urban areas in India,” May 2008.

Newbold, K. B., Scott, D. M., Spinney, J. E. L., Kanaroglou, P., and Paez, A. (2005). “Travel behavior within Canada's older population: a cohort analysis.” Journal of Transport Geography, 13(4), 340-351.

Ortuzar, J.D., and Willumsen, L.G. (2011). *Modelling Transport*, 4th Edition, Wiley, West Sussex, England.

Paix, L., Bierliare, M., Elisabetta, C., and Monzon, A. (2011). “How urban environment affects travel behaviour? Integrated Choice and Latent Variable Model for Travel Schedules.” 2nd International Choice Modelling Conference. Leeds, GB.

Phanikumar, C. V., and Maitra, B. (2006). “Valuing urban bus attributes: an experience in Kolkata.” Journal of Public Transportation, 9(2), 69-87.

Phanikumar, C. V., and Maitra, B. (2007). “Willingness-to-pay and preference heterogeneity for rural bus attributes.” Journal of transportation engineering, 133(1), 62-69.

Pinjari, A. R., and Bhat, C.R. (2006). “Nonlinearity of response to level-of-service variables in travel mode choice models.” Transportation Research Record: Journal of the Transportation Research Board, 1977, 67-74.

Pinjari, A.R., Pendyala, R.M., Bhat, C.R., and Waddell, P.A. (2007). “Modeling Residential Sorting effects to Understand the Impact of the Built Environment on Commute Mode Choice.” Transportation, 34 (5), 557–573.

Rajamani, J., Bhat, C.R., and Handy, S.L. (2003). “Assessing Impact of Urban Form Measures on Non-work Trip Mode Choice After Controlling for Demographic and Level-Of-Service Effects.” Transportation Research Record: Journal of the Transportation Research Board, 1831, 158–165.

Raveau, S., Alvarez-Daziano, R., Yanez, M.F., Bolduc, D., and Ortuzar, J.D. (2010). “Sequential and Simultaneous Estimation of Hybrid Discrete Choice Models Some New Findings.” Transportation Research Record: Journal of the Transportation Research Board, 2156, 131-139

Revelt, D., and Train, K. (1998). “Mixed Logit with Repeated Choices: Household’s Choices of Appliance Efficiency Level,” The Review of Economics and Statistics, MIT Press, 80(4), 647-657.

Report on household data collection plan, Retrieved from < http://www.public.asu.edu/~rpendyal/ HouseholdTravelSurveyDataCollectionPlan.pdf>

Sikka,N. (2012). “Understanding travelers' route choice behavior under uncertainty.” Doctoral thesis, University of Iowa. Retrieved from <http://ir.uiowa.edu/etd/2988/>

Simma, A., and Axhausen, K. W. (2004). “Interactions between travel behaviour, accessibility and personal characteristics: The case of the Upper Austria Region.” European Journal of Transport and Infrastructure Research, 3(2),179-198.

Srinivasan, K., Pradhan, G., Maheswara, N. (2007). “Commute Mode Choice in a Developing Country: Role of Subjective Factors and Variations in Responsiveness Across Captive, Semicaptive, and Choice Segments.” Transportation Research Record: Journal of the Transportation Research Board, 2038, 53-61.

Srinivasan, S., and Rogers, P. (2005). “Travel behavior of low-income residents: studying two contrasting locations in the city of Chennai, India.” Journal of Transport Geography, 13(3), 265-274.

Steimetz, S. S., and Brownstone, D. (2005). “Estimating commuters’“value of time” with noisy data: a multiple imputation approach.” Transportation Research Part B: Methodological, 39(10), 865-889.

Sun, X., Chester, G.W., and Kasturi, T. (1998). “Household Travel, Household Characteristics, and Land Use: An Empirical Study from the 1994 Portland Activity-Based Travel Survey.” Transportation Research Record: Journal of the Transportation Research Board, 1617, 10-17.

Swait, J. (2001). “Choice set generation with the generalized extreme value family of discrete choice models.” Transportation Research Part B:Methodological, 35(7), 643-666.

Swait, J., Louviere, J. J., and Williams, M. (1994). “A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice.” Transportation, 21(2), 135-152.

Tracy, A., Peng, S., Sadek, A., and Wang, Q. (2011). “Assessing the Impact of the Built Environment on Travel behavior: A case study of Buffalo, New York.” Transportation, 38(4), 663-678.

Train, K. (2009). *Discrete choice methods with simulation*. Cambridge university press, New York, USA.

Tsai, C.H., Mulley, C., and Geoffrey, C. (2012). “The Spatial Interaction between Public Transport Demand Land use characteristics in the Sydney Greater Metropolitan Area.” Road & Transport Research: A Journal of Australian and New Zealand Research and Practice, 21(4), 62-73.

Vance, C., and Hedel, R. (2007). “Impact of urban form on automobile travel: Disentangling causation from correlation.” Paper presented at the 86th annual meeting of the Transportation Research Board, Washington, DC.

Vonkova, H. (2011). “The Use of Subjective Survey Data: Anchorine Vignettes and Stated Preference Methods.” Dissertation report, Tilburg University.

Walker, J.L., Li., J., Srinivasan, S., and Bolduc, D. (2010). “Travel Demand Models in the Developing World: Correcting for Measurement Errors,” Transportation Letters: The International Journal of Transportation Research, 2(4), 231-243.

Washington, S., Ravulaparthy, S., Rose, J. M., Hensher, D., and Pendyala, R. (2014). “Bayesian imputation of non-chosen attribute values in revealed preference surveys.” Journal of Advanced Transportation, 48(1), 48-65.

Yanez, M.F., Raveau, S., and Ortuzar, J.D. (2010). “Inclusion of latent variables in Mixed Logit models: Modelling and forecasting.” Transportation Research Part A: Policy and Practice, 44(9), 744–753.

Zegras, P. C. (2007). "The built environment and motor vehicle ownership and use: Evidence from Santiago de Chile." Paper presented at the 86th annual meeting of the Transportation Research Board, Washington, DC.

List of Publications from the present work

International Conference

Sarkar, P. P., and Mallikarjuna, C. (2013). "Effect of Land Use on Travel Behaviour: A Case Study of Agartala City." Procedia-Social and Behavioral Sciences, 104, 533-542.

Bordoloi, R., Mote, A., Sarkar, P. P., and Mallikarjuna, C. (2013). "Quantification of Land Use diversity in the context of mixed land use." Procedia-Social and Behavioral Sciences, 104, 563-572.

National Conference

Sarkar, P. P., and Mallikarjuna, C. (2011). "Travel behavior analysis using a simple O-D survey:A case study of Agartala City." IUT conference, Delhi, India.

International Journals, Communicated

Sarkar, P. P., and Mallikarjuna, C. (2014) "Quantification and Analysis of Land use effects on travel behavior in smaller Indian cities: A case study of Agartala." Journal of Urban Planning and Development, ASCE.

International Journals – In preparation

Sarkar, P. P., and Mallikarjuna, C. (2014) "Determinants of parameters affecting travel behavior for smaller Indian cities: A case study of Agartala." In preparation

Sarkar, P. P., and Mallikarjuna, C. (2014) "Effect of latent variables in mode choice analysis in the context of a small sized Indian city." In preparation

HOUSEHOLD DETAILS.

1.	Household ID	
2.	Zone	
3.	Head of the Family (Name)	
4.	Full Postal Address	Latitude Longitude
5.	Contact details	Landline: Mobile: E-mail :
6.	Type of Housing	☐ Detached or single house ☐ Duplex ☐ Triplex or 4-plex ☐ Row house ☐ Apartment ☐ Hut made of bamboo ☐ Hut made of Mud

Data related to Present & Five Year Back:

		Present	Five years back
7.	Number of members in the family		
8.	Number of Active adults		
9.	Number of senior adults (Retired)		
10.	Number of employed adults		
11.	Number of unemployed adults		
12.	Number of self employed		
13.	Numbers of licensed drivers in the family		
14.	Number of Members with disability		
15.	Number of Full time student		
16.	Status of the household	☐ Owned ☐ Rented ☐ Provided by employer	☐ Owned ☐ Rented ☐ Provided by employer

17.	Aggregated monthly household income without any deduction (in rupees) i.e. grosses income of family or all the person staying together.	Increase in monthly household income without any deduction in last five years (in rupees)
	☐ 0 – 2000 ☐ 2001 – 5000 ☐ 5001 – 10000 ☐ 10001 – 15000 ☐ 15001 – 20000 ☐ 20001 – 30000 ☐ 30001 – 40000 ☐ 40001 – 50000 ☐ 50001 – 70000 ☐ 70001 – 90000 ☐ 90001 – 150000 ☐ > 150001	☐ No increase in income ☐ Decrease in income ☐ Increase in income (Tick the appropriate amount increase/decrease of income) ☐ 0 – 500 ☐ 501 – 1000 ☐ 1001 – 2000 ☐ 2001 – 3000 ☐ 3001– 5000 ☐ 5001 – 7000 ☐ 7001 – 10000 ☐ 10001 - 15000 ☐ 15001 – 20000 ☐ 20001 – 30000 ☐ > 30000

Details of the vehicles owned by household

Vehicle Number	**1**	**2**	**3**	**4**	**5**
Type of Vehicle (Please enclose the appropriate)	Passenger Car/ Van Good Van Truck Motor Cycle/Scooter Bicycle Other Vehicle Type	Passenger Car/ Van Good Van Truck Motor Cycle/Scooter Bicycle Other Vehicle Type	Passenger Car/ Van Good Van Truck Motor Cycle/Scooter Bicycle Other Vehicle Type	Passenger Car/ Van Good Van Truck Motor Cycle/Scooter Bicycle Other Vehicle Type	Passenger Car/ Van Good Van Truck Motor Cycle/Scooter Bicycle Other Vehicle Type
Make of Vehicle (Eg: Hyundai/Suzuki etc)					
Model of vehicle (Eg:i10/Swift etc)					
Year of Manufacture					
Month &Year of Purchase/Ownership(or vehicle registration number)					
Fuel type	Petrol Diesel LPG/CNG Electric	Petrol Diesel LPG/CNG Electric	Petrol Diesel LPG/CNG Electric	Petrol Diesel LPG/CNG Electric	Petrol Diesel LPG/CNG Electric
Who pays for the cost of Running Vehicle?					
Private or Business Use					
Monthly use in (KM)					
Owner's name					
Odometer reading(Total Km reading)					
Daily use in KM					

AGARTALA HOUSEHOLD TRAVEL SURVEY 2012

Personal dataset.

Present and Five years back

Sl.No .	Data Entity	Data (Present)	Data (Five Year Back Data)
1.	Household No		X
2.	Person Name & Contact nos	Name	Contact No
3.	Gender		X
4.	Age		X
5.	Are you married?		
6.	Age of child < than 14 years.(**Write Age**)		
7	Do you have a license (Tick the Appropriate)?	☐ Four wheeler ☐ Two wheeler ☐ None	☐ Four wheeler ☐ Two wheeler ☐ None
8.	Are you disabled?		
9.	Employer Type	☐ Government ☐ Private ☐ Self	☐ Government ☐ Private ☐ Self
10.	Employment Status	☐ Full time ☐ Part time ☐ Self employed ☐ Retired ☐ Home maker ☐ Unemployed ☐ Others ---------------	☐ Full time ☐ Part time ☐ Self employed ☐ Retired ☐ Home maker ☐ Unemployed ☐ Others --------------
11.	Are you a student? (Tick)	☐ Yes ☐ No	☐ Yes ☐ No
12.	**Highest Educational Qualification**		
13.	**Occupation**		
14.	Mode of travel you use very frequently (Rank it)	☐ Car ☐ Bus ☐ Auto ☐Jeep/Cruiser ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk	☐ Car ☐ Bus ☐ Auto ☐Jeep/Cruiser ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk
15.	Mode of travel you don't use	☐ Car ☐ Bus ☐ Auto ☐Jeep/Cruiser ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk	☐ Car ☐ Bus ☐ Auto ☐Jeep/Cruiser ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk
16.	Most preferred and least preferred travel mode	Most preferred / least preferred	Most preferred / least preferred

17.	If you are not using bicycle then please rank the reasons for not using bicycle. Rank all the parameters according to preference, rank 1 as most suitable.	☐ Inferior Good ☐ Congested Roads ☐ No Bicycle ☐ Poor quality of Roads ☐ Socially not acceptable ☐ Difficulty in crossing intersection ☐ Not Safe ☐ Requires physical activity ☐ Distance ☐ Climate ☐ Not Comfortable ☐ No Bicycle facilities on the way to work ☐ Time saving Others please specify
18.	If you are using bicycle then please rank the reasons for using bicycle Rank all the parameters according to preference, rank 1 as most suitable.	☐ Fitness/ Health concerns ☐ Pleasure/ Enjoyment ☐ Environmental concerns related to Automobile use ☐ Convenience/ Speed ☐ Avoid driving in congested conditions ☐ Limited Auto parking ☐ Socially Safe ☐ Flexible ☐ Regular Habit of cycling Others please specify
20.	Please rank the reason if you are not choosing a bus as mode of transport? Rank all the parameters according to preference, rank 1 as most suitable.	☐ Not Available ☐ Not Reliable (Not able to arrive on time) ☐ Not flexible ☐ Less frequency ☐ Not comfortable ☐ Bus stop is far away. ☐ No direct bus ☐ Takes more time ☐ Seats are not available ☐ Costly
21.	What may be done so that you may be motivated to use bus? Rank all the parameters according to preference, rank 1 as most suitable.	☐ More frequency ☐ Better buses ☐ More reliable service ☐ Bus stop in our area ☐ Direct Bus ☐ Buses with fewer stops ☐ Less fare ☐ Faster Buses

22.	Please rank the reasons for not choosing walking? Rank all the parameters according to preference, rank 1 as most suitable.	☐ Distance ☐ Too slow : Takes too long time ☐ Weather ☐ Dislike walking ☐ Difficult to carry things ☐ Inconvenience ☐ Fear of Crime ☐ No time ☐ Darkness ☐ No sidewalks Others please specify	
23.	Please rank the reasons for choosing walking. Rank all the parameters according to preference, rank 1 as most suitable.	☐ Exercise/ Health ☐ For Pleasure/ I like to walk ☐ Walk the Dog ☐ Get outdoors/ Fresh air ☐ Walk to store ☐ Walk to work Others please specify	
30.	What is the one way distance between home and work place?		
33.	How much do you personally pay for transportation?	Fare (Rs --------/month) Fuel (Rs --------/month)	Fare (Rs --------/month) Fuel (Rs --------/month)
37.	A few new buses were introduced in last year under JNNURM scheme. Please rate the above statement from 0-5 scale. 1 – Totally disagree 2 – Disagree 3 – Agree somewhat 4 – Agree 5 – Fully Agree	☐ Ride in the new bus is comfortable. ☐ Many people started using bus services instead of their personal vehicle. ☐ The new bus service is reliable (Arrives at proper time). ☐ The fare is reasonable. ☐ New bus service is providing better service.	
39	Your approx income (Monthly) without any deductions.(in rupees)	Increase/decrease in monthly income in last five years.(in rupees)	
	☐ 0 – 2000 ☐ 2001-5000 ☐ 5001 – 10000 ☐ 10001-15000 ☐ 15001 – 20000 ☐ 20001-30000 ☐ 30001 – 40000 ☐ 40001-50000 ☐ 50001 – 70000 ☐ 70001-90000 ☐ 90001 – 150000 ☐ > 150001	☐ No increase in income ☐ Decrease in income ☐Increase in income (Tick the appropriate amount increase/decrease of income) ☐ 0 – 500 ☐ 501 – 1000 ☐ 1001 – 2000 ☐ 2001 – 3000 ☐ 3001-5000 ☐ 5001 – 7000 ☐ 7001 – 10000 ☐ 10001 - 15000 ☐ 15001 – 20000 ☐ 20001 – 30000 ☐ > 30000	

AGARTALA HOUSEHOLD TRAVEL SURVEY 2012

Rate the following perception. Strongly disagree= 1 Disagree= 2 Agree somewhat= 3 Agree= 4 Strongly Agree = 5

Sl.no.	Statements	Ratings				
1.	Personal vehicle (Car or Two wheeler) is comfortable.	1	2	3	4	5
2.	I am not comfortable when I travel with people I don't know.					
3.	I always use the most convenient mode of transportation regardless of cost.					
4.	I always use the fastest route to destinations even if I have cheaper alternatives.					
5.	If fuel prices goes up further less likely to drive car to work.					
6.	If fuel prices goes further up less likely to drive motor cycle to work.					
7.	If I use public transport instead of car or two wheeler I have to cancel some of the activities.					
8.	It is hard to take public transport when travelling with children.					
9.	It is hard to take public transport with bags & luggage.					
10.	I know fully what buses I should take regardless of where I am going.					
11.	I need to have more flexibility to make many trips during working hours.					
12.	I don't mind walking few minutes to destination.					
13.	People riding a bus help in reducing congestion.					
14.	Willing to pay more tax to improve bus service.					
15.	Bus is chosen when no other option is available.					
16.	Using bus service is cumbersome.					
17.	I don't like transferring vehicles in the route.					
18.	I will put up with crowds if I can reach destination early.					

19.	Walking on the road is comfortable.					
20.	Comfortable in walking without footpaths.					
21.	Trees provide ample shade in the roads near my house					
22.	There are interesting houses to look at in the roads near my house					
23.	Comfortable in walking in local shopping areas.					
24.	Shared auto is comfortable.					
25.	I would be willing to pay more if it would help environment.					
26.	I would be willing to switch to different mode if it helps environment.					
27.	Use of bus will improve the environment.					
28.	I am willing to change to bicycle mode if proper bicycle facilities/infrastructure is available.					
29.	I am willing to change to walk mode if proper pedestrian facilities are available and trip length is short.					
	Safety in Neighborhood.					
30.	It is safe to walk in our locality during day time.					
31.	It is safe to walk in our locality during night time.					
32.	Good footpaths are available in our locality.					
	Accessibility of Neighborhood.					
33.	Road facility to our locality is good.					
34.	Importance amenities, post office, bank etc nearby.					
35.	Good para-transit (Auto/Mini-van) service is available to & from our area.					
36.	Frequent bus service is available to and from our area.					

Please rate the travel modes as per the parameters given below:

Ratings.	
1.	Extremely Poor.
2.	Poor.
3.	Average.
4.	Good.
5.	Excellent.

Sl.no	Parameter	Car	Bus	Auto	Two – wheeler	Jeep/Cruiser/Magic	Non-Motorized (Three wheeler)	Bi-Cycle
1.	Comfortable in journey.							
2.	Always availability of comfortable seat							
3.	Very easy accessibility							
4.	Ability to reach the destination always in time							
5.	Can exactly calculate travel time prior to trip							
6.	Safety w.r.t accident							
7.	Safety from theft							
8.	Protection from weather							
9.	Able to make more trips in short time							
10.	Can travel without changing vehicles/stops.							

SP SURVEY FORM NO – 0,1 (3KM)

AGARTALA HOUSEHOLD TRAVEL SURVEY 2012

Serial no.	CAR		BUS				AUTO RICKSHAW			MOTOR CYCLE		Write the chosen mode
	Cost	Time(min)	Cost	Time	Comfort	Frequency	Cost	Time	Comfort	Cost	Time	
1	₹20	00:18	₹5	00:18		Every 5 min	₹10	00:06		₹7	00:07	
2	₹ 9	00:3:30	₹15	00:09		Every 30 min	₹3	00:09		₹4	00:04:30	
3	₹25	00:04:30	₹3	00:04:30		Every 15 min	₹15	00:18		₹9	00:3:30	
4	₹30	00:07	₹10	00:06		Every 45 min	₹5	00:04:30		₹5	00:18	

Time	No of Travel Activity out of home
Morning	
Noon & Afternoon	
Evening	
Night	

For next tour go to next page

Work Trip ⇨

Mode	start-time	End-time	start-time(R)	End-time(R)
Car				
Bus				
Auto				
Motor-cycle				
Cycle				
Rickshaw				
Jeep				
Walk				

Date of travel: / /2012 TRAVEL DIARY Tour -

Name of the Interviewee: Address: Contact No: Gender: M / F

If you have not made any trips tick here. ☐

Trip No	Origin (Beginning of Trip)	Destination (End of Trip)	Route selected in the trip. Mention the name of the route & mark the same in the map provided.	Breaks/Stops/Transfer (Write the place or station & the time consumed)	Destination (Address with land Mark)	Starting time of trip	End time of trip	How did you travel? (Mode) 1. Bus 2. School/College Bus 3. Car 4. Auto 5. Jeep/Max 6. Van (Tata Magic etc) 7. Scooter / Motor Cycle 8. Rickshaw 9. Bicycle 10. Walking	Distance (Km) & time in (mins)	How many people were travelling with you?	Were you driver or passenger in the trip?	Walk ingress time & distance for public transport & para-transit	Walk egress time & distance for public transport & para-transit	Fuel / Fare cost	Parking Cost	**Purpose** 1. Work 2. Education 3. Business 4. Shopping (including major & minor groceries) 5. Social (Visit friend & relatives) 6.Medical (Doctors appointment, visit to drug store etc) 7. Lunch/ Dining/Coffee. 8. Recreation (Pleasure driving, visit to parks &play grounds) 9. Pick up or drop some one. (Family member, friend or children) 10.Religoius Events 11. Cultural Events 12. Vegetable Purchasing If Others please write in the space below.	Activity Duration Time (Mins)	Reason for choosing the mode.(Please rank the attributes)	If you travel by bus/Para-transit how you did go to bus stop. (walked/drove/dropped off).	Waiting time if travelled by bus or para-transit.	Other Modes available for this trip).There may be more than two answers.
5.				Place Time in Transfer					Time Distance			Time Distance	Time Distance					Low Cost Comfort Less Travel Time Flexibility (*Easy to change travel plan immediately*) More safe No other alternative Reliability(Ability to arrive on time)			☐ Car ☐ Bus ☐ Auto ☐Jeep ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk
6.				Place Time in Transfer					Time Distance			Time Distance	Time Distance					Low Cost Comfort Less Travel Time Flexibility (*Easy to change travel plan immediately*) More safe No other alternative Reliability(Ability to arrive on time)			☐ Car ☐ Bus ☐ Auto ☐Jeep ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk
7.				Place Time in Transfer					Time Distance			Time Distance	Time Distance					Low Cost Comfort Less Travel Time Flexibility (*Easy to change travel plan immediately*) More safe No other alternative Reliability(Ability to arrive on time)			☐ Car ☐ Bus ☐ Auto ☐Jeep ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk
8.				Place Time in Transfer					Time Distance			Time Distance	Time Distance					Low Cost Comfort Less Travel Time Flexibility (*Easy to change travel plan immediately*) More safe No other alternative Reliability(Ability to arrive on time)			☐ Car ☐ Bus ☐ Auto ☐Jeep ☐ Motor Cycle ☐Bicycle ☐Rickshaw ☐Walk