University of Alberta

Investigation of Commuting Mode Choice with respect to TDM Policies

by

Hamid uz Zaman

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

in

Transportation Engineering

Department of Civil and Environmental Engineering

Spring 2010
Edmonton, Alberta

Examining Committee

Dr. Khandker M. Nurul Habib, Civil and Environmental Engineering (Supervisor)

Dr. Mohamed Al-Hussein, Civil and Environmental Engineering

Dr. Zhi-Jun Qiu, Civil and Environmental Engineering

Dr. M David Checkel, Mechanical Engineering

DEDICATION

To my Parents

ABSTRACT

Travel Demand Management (TDM) is now considered one of the most important aspects of transportation planning and operation. The prime objective of TDM is to develop a sustainable transportation system utilizing the existing infrastructure. It is now a well known fact that excessive use of single occupancy vehicle causes numerous problems like traffic congestion, environmental pollution etc. Thus, from TDM perspective, it is of great importance to analyze travel behaviour in order to influence people to reduce car use and choose more sustainable modes such as – carpool, public transit, park & ride, walk, bike etc. This study attempts an in-depth analysis of commuting mode choice behaviour using workplace commuter survey data from the City of Edmonton. Unlike traditional mode choice models, this study uses both instrumental and latent variables to better understand the choice process and analyzes their sensitivities with respect to TDM policies.

ACKNOWLEDGEMENT

First of all, I am grateful to Almighty Allah, the most Gracious, most Merciful; nothing comes to reality without His wish.

I would like to express my profound gratitude to my supervisor Dr. K.M. Nurul Habib. This thesis would not be possible without his expert guidance and encouragement. Working with him has been an amazing experience for me.

I would like to thank the Transportation Department of the City of Edmonton for sharing the workplace commuter survey data, which has been used in this study.

I am also thankful to my family, friends and colleagues, especially Md. Tazul Islam, for their continuous support and cooperation throughout this study period.

Hamid Zaman

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

Traffic congestion and associated problems are now a major concern for transport planners, politicians and the public (Saleh and Sammer, 2009). These transport related problems require immediate attention, particularly as many past policies have failed to deal with them adequately. The traditional approach of 'predict and provide' for dealing with traffic congestion is no longer viable, because building more infrastructure for increasing travel demand is not a sustainable solution. It is now widely accepted that unrestrained demand for travel by car cannot be sustained. Moreover, the transportation sector is considered one of the major contributors through human intervention for environment pollution. Transport accounts for 14% of global greenhouse gas emission, with a vast majority of these emissions produced by the road transport sector (Hensher 2008). The major transportation related problems such as congestion, inadequate mobility, various economic, social and environmental costs are associated with high levels of automobile travel (Litman 2009). Therefore, measures taken to address the problems have shifted from 'predict and provide' to 'predict and manage', which is, in other words, travel demand management.

Successfully administered travel demand management (TDM) programs manage congestion and reduce the number of vehicles on road, while maintaining full accessibility for individuals (Ungemah et al 2009). In order to accomplish that, one of the major objectives of TDM is to reduce 'drive alone' or single-occupancy vehicle travel and influence the demand to be shifted to alternate modes of travel (Winters 2000). It is a fact that in many developed countries the other modes of transport are not true competitors of drive alone (Day, 2008). However, most of the transport agencies across the world are now considering building a transport system that has attractive alternate modes such as transit, carpooling, vanpooling, cycle and ride, walking, and encouraging people to use these modes more often. Numerous research works are available and still ongoing in transportation field focusing on the attractiveness and effectiveness of various TDM measures. Some of these works attempt to reveal the factors behind people's preference towards car use (Gardner et al 2007, Lucas 2009), while some focus on how various measures, programs or factors can effectively lead towards sustainable transportation system (Garling et al 2007, Ko et al 2009, Buliung et al 2009, Dill & Voros 2007).

Traditionally, car use and mode choice behavior has been studied under the random utility framework where an individual's mode choice decision is considered to be made on the basis of attributes of the alternatives and the individual (Domarchi et al, 2008). Some examples of such alternative attributes are travel time, travel cost (also referred to as utilitarian or level of service attributes) etc., while age, gender and income are some examples of individuals' attributes (also referred to as generic variables). Although the traditional approach of using these attributes to define mode choice has been quite successful over the years, recent publications suggest that understanding of latent behavioral factors is also important to better capture one's mode choice decision. Johansson (2006) and Domarchi et al (2008) have addressed these issues in their recent studies. Beside encouraging people to use alternate modes instead of single occupancy vehicles, programs like voluntary car use reduction, flexible office hours and compressed workweek are also implemented by many agencies as travel demand management tools. Research works are also being carried out seeking to determine the effectiveness of these programs (Sundo & Fuji 2005, Zhou & Winters 2008, Litman 2009).

Therefore, it is understood from the study of many recent publications that people's mode choice behavior is a complex phenomenon which depends on many instrumental and latent factors. Discrete choice analysis is an effective tool to analyze such travel behavior and it is being extensively used over the past decades. Literature on discrete choice analysis techniques and their applications in transportation suggest that a number of discrete choice models are in practice today. The most common of these are logit, GEV, probit, mixed logit etc. (Trains 2002). Each of these models has its own strength and limitations, but multinomial logit and nested logit are the two most widely used discrete choice models for travel behavior analysis (Hensher 1998). However, these models are developed based on likelihood estimation of stated and/or revealed preference data and are capable of representing mode choice probabilities at a disaggregate level. So, we understand that availability of a sufficient amount of mode choice data is required to develop a model.

As mentioned earlier, many transportation agencies are adopting measures that are beneficial in terms of sustainability. The City of Edmonton also addresses sustainability issues through its new Transportation Master Plan - "The way we move" (2009). Transportation mode shift, sustainability, integrated transportation and land use policy are some of the strategic goals set out in the latest master plan. The transportation department of the City of Edmonton conducted a workplace commuter survey in 2007, which was performed under the process of developing TDM program for the City staff. The survey consists of both revealed and stated preference data and also incorporates questions on some ongoing TDM policies. Moreover, the final cleaned dataset contains 7 days observation for about 3000 individuals living in various locations in the city. This gives us an excellent opportunity to make use of a large scale panel dataset and develop a mode choice model using both utilitarian and latent variables. The effectiveness of TDM measures such as compressed workweek and flexible office hours can also be analyzed using the dataset. Therefore, this study attempts to make use of Edmonton's workplace commuter survey data for in-depth analysis of commuter mode choice behavior and to investigate the effectiveness of some TDM measures.

1.2 Objectives of the Study

The broad objective of this study is to investigate the factors affecting the mode choice behavior for commuter trips. The specific objectives are as follows:

- Analyze workplace commuter survey data to understand the current modal share in the Edmonton.
- Develop a mode choice model using panel data to investigate heterogeneity across the individuals as well as temporal distribution.
- Calculate elasticity for level of service attributes to reveal the level of their effects on mode choice probabilities.
- Develop a discrete choice model with latent variables in order to capture the behavioral component in choice decision and compare the results with observed distribution.
- Analyze the effectiveness of some TDM programs such as compressed workweek and flexible office hours.

1.3 Methodology

First, necessary cleaning of noise in the survey data is performed. Then the dataset is analyzed to understand the mode choice scenario on the basis of revealed data. Then the utilitarian variables such as in-vehicle travel time, walk time & wait time (in case of transit), number of transfers in transit etc. are collected for all individuals using Google Map feature. Once the data preparation is complete, the suitable discrete choice model is selected on the basis of literature reviews. Then the estimation of mode is performed using the maximum likelihood estimation technique in BIOGEME.

Different models are developed using different specifications and combination of variables. The models incorporate both instrumental and latent variables to represent the utility of each alternate. After the estimation of the models, a few models are chosen based on the statistical significance of the parameters. The chosen models are validated using validation data, which is previously separated from the whole sample for validation. Then, elasticities of several important variables are calculated to understand their effects on mode choice from the TDM point of view. Chapter 3 describes the modeling frameworks used in this study and Figure 3.1 in chapter 3 graphically represents a methodological overview.

1.4 Thesis Outline

This thesis contains six chapters. The current chapter, Introduction, presents a brief outline of the background, objective and the methodology of the study.

Chapter 2 presents a literature review of recent publications on travel behavior analysis, various TDM measures and discrete choice modeling. The chapter also covers the basic concepts of choice models, elasticity measures and panel data analysis.

Chapter 3 is the Methodology chapter, and it discusses the conceptual framework of the study and also describes probability equations of multinomial and nested logit models. It also describes the formulation of the models used in this study for mode choice probability analysis.

Chapter 4 presents a brief description of workplace commuter survey questionnaire and the dataset used in this study. It also describes the methods used for collecting of level of service attribute data for all alternatives and for defining choice sets.

Chapter 5 presents the analyses and model results. The final mode choice models are described and compared in terms of parameter significance and goodness-of-fit measures. It also presents the elasticity for the level of service attributes that are important in terms of TDM policy implications. This chapter also presents the models where latent variables derived from stated choice questions have been incorporated.

And finally, chapter 6, Conclusions and Recommendation presents an overview of the study and its results. Policy implication, limitations of this study and future recommendations are also presented in the chapter.

CHAPTER 2

LITERATURE REVIEW

Travel Demand Management (TDM) has received increasing attention during this decade among the agencies, practitioners and researchers working in the field of transportation planning. Numerous research works have taken place and are still on-going on the evaluation of various TDM measures that can help in developing a sustainable transportation system. As TDM products and services include encouragement to use alternatives to the single-occupant vehicles (interchangeably used with 'drive alone' in this paper), it is clear that influencing travel behavior is an important factor in order to achieve the goals of TDM (Winters 2000). Investigation of factors affecting the mode choice behavior for home based work trips (commuter trips) is of particular importance, as it constitutes the highest percentage (20%) of all types of trips (Edmonton Household Travel Survey 2005). Since this study attempts to investigate the influence of some factors and TDM measures on commuter mode choice behavior, review of literatures on travel behavior analysis, travel demand modeling is essential. This chapter briefly presents some relevant literatures and recent works on the subject matters.

2.1 Travel Demand Management

Travel Demand Management or Transportation Demand Management (both TDM) is a general term, which includes strategies and programs that encourage more efficient use of transport resources (Litman 2003). It deals with developing and evaluating various strategies and policies to reduce traffic congestion, environmental pollutions and energy consumptions, as well as to produce benefits like improved traffic safety, consumer cost savings etc.

There is a wide range of different TDM strategies using various approaches to influence travel decisions. Some of them improve the existing transport options, some provide incentives to change mode, time or destination, while some improve land use accessibility and some involve transportation policy reforms (VTPI 2009). Some of the important strategies that this study is concerned with are – flextime, bike/transit integration, HOV priority, parking pricing, encouraging non-motorized modes etc. It is believed that all of these measures can effectively influence people's behavior to shift from single occupancy vehicle to other alternatives.

2.2 Mode Choice and Travel Behavior Analysis

Plenty of literature can be found where the various aspects of travel behavior have been addressed. Much recent literature can also be found focusing on the attractiveness and effectiveness of environment friendly alternative modes such as carpooling, cycling, walking etc. This section presents those, which in particular focused on people's mode choice behavior and the factors that can influence people to switch to TDM desirable alternatives.

Traditionally, car use and mode choice behavior has been studied under the random utility framework, where an individual's mode choice decision is considered to be made based on some attributes of the alternatives and the individual. Although this approach has been successfully applied in the development of useful models, recent publications suggest that capturing complex human behavior through analyzing latent and/or affective variables can be more effective in understanding the actual mode choice behavior of an individual (Domarchi et. al, 2008). Because, it is believed that in real life, individual heterogeneity, such as different preference level towards safety, comfort, flexibility etc., also affect the choice of mode (Johansson et al 2006). In a study on effect of attitude and personality traits on mode choice, Johansson et al (2006) addresses how unobservable or latent variables can affect mode choice decision. Domarchi et al (2008) also performed a study on socio-psychological factors of individuals, and showed that people's mode choice behavior can be better captured by analyzing attitudinal variables. This means that mode choice decisions are influenced not only by the instrumental variables but also by some latent variables governed by attitude, habit and behavior. However, collecting such variables in large scale samples is often difficult because of lengthy questionnaires.

In recent years, while some researchers are focusing on the behavioral aspect of mode choice scenario, works are also being performed on evaluating the effectiveness of various measures that reduce car use and promote alternative, more sustainable modes. Gardner and Abraham (2007) conducted a study applying the grounded theory analysis in order to reveal the reasons behind using cars for work trips. The study was performed on regular private car commuters in England, and was designed to explore driving decisions from the driver's perspective. The study also focused on relationship between utilitarian and affective motives, and driver's motivations to travel demand management policy making. The authors identified five core motives that influence commuters to use car for their work trips: journey time concerns, journey based affect, effort minimization, personal space concern and monetary costs. The analysis has also revealed that many users have misconceptions regarding journey times and underestimation of car-related monetary cost, which in turn affect their mode choice and lead them using car so often. However, the study did not use a large sample and the grounded theory is sometimes criticized for its qualitative data analysis procedure (Thomas and James 2006).

Gardner's (2007) study was on exploring the reason behind using car for commuter trips. On the other hand, Habib et. al. (2009) did an investigation on people's behavior and perception towards transit, where a multinomial logit model combined with latent variable mode was developed to capture unobserved factors influencing people's choice of transit. The study used the 2007 transit customer satisfaction survey data from Calgary, and mainly concentrated on analyzing the attitude of transit users towards different attributes of transit service. The results suggested that transit users in Calgary value reliability and convenience over ride comfort. This finding complies with the survey data of workplace commuter survey 2007, where 35.2% respondents mentioned that they would use public transit more often if it was more reliable. However, the study did not use any data from users of other alternates and hence the comparative analyses of different modes are missing.

While the above studies performed in-depth analysis to reveal the reasons behind choosing a particular mode, many studies also analyzed the impact of various factors and TDM measures on commuter mode choice. A discrete choice analysis on impact of road pricing and parking charges on commuter mode choice was performed by Washbrook et. al. (2006) using stated choice data from a Greater Vancouver suburb. The objective of the study was to analyze how road pricing and parking cost can influence people's commuting mode choice to be shifted from single occupancy vehicle (or drive alone) towards more sustainable alternates such as carpool and public transit. The survey respondents were provided with scenarios specific time and cost variables for each mode, and were asked to select the most feasible mode for commuting. The results of the study suggest that improving travel time for alternate travel modes above a base level of service has only a small effect on mode choice and does little to reduce the demand for SOV travel. On the other hand, increasing the cost of SOV travel by introducing new charges or increased cost of travel has a substantial effect of demand for driving alone. This is an important finding in terms of policy implication. Waerden et al (2009) carried out similar study for shopping trips, where consumers' response to introduction of paid parking at regional shopping mall is analyzed. The results suggest that changes in expenditure can affect shopping location, duration, frequency as well as mode choice. But the study focused on shopping trips, and the result only shows the short-term effects of paid parking; long term effects may be different though. Another study by Weinberger et al (2009) shows how parking facility can affect car ownership and mode share of commuter trips.

However, Garling et. al. (2007) has shown that although coercive TDM measures, such as increasing cost for or prohibiting car use, can be effective to reduce car use but these are difficult to implement because of public opposition and political infeasibility. Therefore, such measures should be applied simultaneously with non-coercive measures, for example encouraging reduction of car use or incentive for using alternate modes. Ko and Cho (2009) carried out a study to evaluate voluntary program to reduce car use. The study attempted to assess the effectiveness of Weekly No-Driving Day (WND) program in Seoul, South Korea. Seoul Metropolitan Government (SMG) started the WND program in July 2003 to reduce vehicle trips by encouraging citizens to leave their cars home at least once a week (Monday to Friday). Volunteer Participants select days on which to refrain from car use. For the study, a field survey was carried out to investigate the compliance rate and it was found that daily car use was reduced by 1.3%. A questionnaire survey was also conducted to determine which drivers are more likely to participate in program. The results indicated that car-dependents drivers, who are frequent and regular user of cars, are less likely to participate in the program. This result complies with the findings by Johansson (2006) and Domarchi (2008). Lucas (2009) also performed a study on car use dependency, which aimed to gain deeper insight into the causes of car use. The study took place in the United Kingdom and mainly focused on implication and impact of car use reduction programs and identifying factors behind car dependence.

Besides encouraging car use reduction, the transport agencies are also putting emphasis on increasing the attractiveness of alternative modes such as carpooling, cycle and ride, cycling, skating etc. Research works are also being performed how incentives can be offered for the use of alternate modes, which may encourage the users to select from a wide range of attractive alternatives alongside drive alone. For example, carpool represents one of many possible alternatives to single-occupancy vehicle use for work trips. Buliung et al (2009) investigated the factors behind successful carpool formation where recent attempts to encourage carpool formation in Canada are examined. It was found that web-based applications that facilitate connections between potential carpoolers can be effective for carpool formation. Several studies have also been performed on assessing demand and policy implication of bicycle as an alternate sustainable mode. Transportation Research Board (TRB) published Transportation Research

Record No. 2031 in 2007, which focuses only on different aspects of bicycles and motorcycles. For example, Dill and Voros (2007) worked on bicycle demand in the United States and used a survey data from Portland, Oregon to explore the factors affecting bicycle demand.

We understand that mode choice is influenced by various instrumental and latent variables. As mentioned earlier, beside the level of service attributes of the alternatives, mode choice can also vary depending upon many user specific factors, such as home to work distance, age, gender, departure time, working hour etc. Recent advancements in transportation research are considering joint modeling approach for trip timing and mode choice decisions, in order to better replicate the actual choice process (Habib et. al. 2009). Although this new approach has opened a whole new arena for research, but this is beyond the scope of this study. Rather, the focus of this study stays within effects of mode and individual specific instrumental and latent variables on individual mode choice behavior. Such variables are level of service attributes (travel time and cost) of the alternate modes and individual specific variables (age, gender, travel distance), along with TDM measures such as flexible office hour and compressed work week and also some latent variables expressing people’s perception towards different aspect of available modes.

Sundo and Fulii (2005) carried out a study on effect of compressed work week on commuters’ daily activities. The study was performed using survey data on some selected employees of the University of Philippines after the Philippine government implemented an experimental two month compressed working week scheme in 2002. During the compressed working week, the individuals worked four days a week for 10 hours each day, instead of 8 hours in five days. Naturally it resulted in a change in the daily activity patterns of people, and also in their travel pattern in terms of departure time and travel times. After comparative analysis of commuting travel behaviour patterns and activity durations before and during the compressed working week, it was found that the commuting travel times were significantly reduced during the compressed working week, which indicates that compressed work week scheme has the potential to reduce travel times for each commuter. However, this travel time reduction was probably because of the shift in departure time for the people having compressed work week, and since the number of people under the scheme was very small, carrying out similar investigation with larger sample size is necessary. Zhou and Winters (2008) analyzed the trend and determinants of compressed workweek (CWW) by using Washington State commute trip reduction (CTR) data. The result of the MNL model suggests that people are more likely to participate in CWW when the employers are supportive and the CWW program has been implemented for longer period. Home to work distance was also identified as a key factor for successful CWW implementation.

In the online TDM encyclopedia of Victoria Transport Policy Institute (VTPI), Litman (2009) illustrates various benefits of flexible office hour, which allows an employee the flexibility of varying his/her office hour. This measure can be effective in the sense that it reduces traffic congestion, and supports ride sharing and public transit use because the trip maker can match his/her office hours with the transit schedule. Many organizations across Canada (City of Edmonton, City of Winnipeg, The Royal Bank, Toronto Star) are implementing this program. However, to the best of our knowledge, not many studies have addressed comprehensive analysis on the level of influence that these measures can have on individual mode choice.

The preceding sections have presented some previous works on travel behavior research. In all of these works, the researchers have used different analytical and modeling techniques to achieve better representation of facts. Likewise, in this study, we are going to follow some specific modeling techniques in order to reveal the effect of variables and their elasticities on mode choice. The following sections of this chapter focus on the theoretical background of the modeling and analysis techniques used in this study.

2.3 Travel Demand Modeling and Discrete Choice Analysis

Demand forecasting is an essential element in transportation system analysis, and it is mainly concerned with the behavior of users of transportation services and facilities (Ben-Akiva & Lerman, 1985). This is obvious that it is not possible to forecast transportation demands efficiently without proper understanding of complex behavioral aspects behind the decisions taken by its users. A mathematical function expressing an individual's choice (such as mode choice, route choice etc.) as a function of some variables is hence required to predict such decisions, and discrete choice analysis is used for the purpose. Discrete-choice models are widely used in the field of transportation planning, to represent the choice of one among a set of mutually exclusive alternatives (Koppelman & Sethi 2000). Walker (2005) identified the key advantages of discrete choice in transportation modeling practice as its policy relevance, integrated model systems and market segmentation. These models are based on random utility maximization theory which was conceived by Marschak and Block as a probabilistic representation of individual choice (Ibanez 2006). Many types of discrete choice models are presently in practice and some recognized models are logit, probit, generalized extreme value, ordered probability etc. Each of these models has its appropriateness and applications in different choice situations and extensive research on this field is still ongoing with the advent of advanced computational ability.

2.4 Discrete Choice Models

As mentioned above, several types of discrete choice models exist in practice, and the most prominent ones are logit, generalized extreme value (GEV), probit and mixed logit (Train 2002). Each of these models uses different assumptions and they are derived under different specifications of the density of unobserved factors. These models are briefly discussed in the following section.

2.4.1 Logit Models

The most widely used discrete-choice model to represent choice probabilities is the family of logit (Train 2002). Its popularity is because of the fact that the logit formula for choice probabilities takes a closed mathematical form and can be interpreted very easily. The logit formula was originally derived by Luce (1959) from the assumption of Independent and Irrelevant Alternatives (IIA) property, which was later proven to be consistent with utility maximization theory. The relation of the logit formula to the extreme value distribution of the unobserved utility leads to the formulation of logit model. Among the logit models, the Multinomial Logit is considered as the workhorse for the empirical analysis of travel behavior in respect of discrete choices (Hensher 1998).

The well-known Multinomial Logit Model, developed by McFadden (1974), is being extensively used in discrete choice modeling over the past three decades. As described by Koppelman and Sethi in the Handbook of Transport Modelling (Hensher & Button 2000), earlier choice models were based on the assumption that the error terms were multivariate normal or independently and identically type I extreme value (gumbel) distributed. The multivariate normal assumption leads to the multinomial probit model, while the independent and identical gumbel assumption leads to the multinomial logit. Although probit model allows complete flexibility in the variance-covariance structure of the error terms, its use requires numerical integration of a multidimensional normal distribution, which is computationally cumbersome. On the other hand, MNL has a closed form of probability equation and is easily interpretable, and hence it has been so popular. In his Nobel Prize lecture McFadden described the history of development of MNL (McFadden 2000).

As mentioned above, multinomial logit models are widely used in transportation planning for discrete choice analysis. Using a stated preference data from Edmonton, Hunt (2001) developed a multinomial logit model to carry out an analysis on sensitivities of various elements of transportation and urban forms.

2.4.2 Generalized Extreme Value (GEV) Models

Although multinomial logit model is the most widely used choice model because of its simple mathematical structure and ease of estimation, its IIA assumption imposes the restriction that the distribution of the random term is independent and identical over alternatives (Wen & Koppelman 2000), which is likely to be violated in many choice contexts. This led the researchers to attempt relaxing the IIA assumptions, and thus other models have been developed. As the name implies, generalized extreme-value models are based on generalization of extreme-value distribution. It allows correlation in unobserved factors over alternatives and collapses to logit model when this correlation is zero. The correlation can be flexible depending on the type of GEV model. For example, a comparatively simple GEV model places the alternatives into several groups called nests, with unobserved factors having the same correlation for all alternatives within a nest, and no correlation for alternatives in different nests.

Like logit models, GEV models also have closed form equations and hence easy to estimate. The most common example of GEV model is Nested Logit model.

Ortuzar (2001) describes the historical development of Nested Logit, where it appears that NL has been fully developed and described in mid-seventies, and since then it has been one of the most preferred models for discrete-choice analysis. The NL model is characterized by nesting (grouping) subsets of alternatives that are similar to each other with respect to excluded characteristics than they are with the other alternatives (Koppelman & Bhat 2006). Alternatives in a common nest exhibit higher degree of similarities than alternatives in different nests. The nesting can be expressed by tree structures which offer relaxation in the IIA assumption.

Various applications of nested logit models in travel demand analysis is found today. Earlier application can be seen in business travel analysis on Ontario-Quebec corridor (Forinash & Koppelman 1993). A set of nested logit structures were estimated in the study, to examine the differential sensitivity of various combinations to changes in service quality of rail. Several nested logit models with different nesting structure were developed and their forecasting accuracy was compared with that of multinomial logit model. The paper demonstrated a statistically significant rejection of multinomial logit model in favor of three alternative nested logit models. In Alberta, both City of Edmonton and City of Calgary use nested logit models for mode choice analysis (City of Calgary, 2001

Regional Transportation Model). The model was developed in Edmonton using Edmonton dataset and it was calibrated for Calgary. Separate time-of day and mode choice models were estimated using ALOGIT software. Edmonton regional transportation model also uses nested logit for mode choice (Blaschuk et al. 2007). The model has a three-level nesting structure: 2 nests in top level are mechanical and metabolic. Each of these has 2 branches; transit & auto for mechanical, and walk & cycle for metabolic. The auto nest contains car and park & ride at the bottom level. However, the model does not consider the minor modes such as carpool and ride, cycle and ride within choice set. Nested logit model has also been used to develop auto-ownership model for Edmonton (Hunt et al., 2005).

2.4.3 Probit Models

Probit models are based on the assumption that the unobserved factors are normally distributed (Train 2002). In this type of model, any pattern of correlation and heteroskedasticity can be accommodated. When applied to sequences of choices over time, the unobserved factors are assumed to be jointly normal over time as well as over alternatives. Probit models are very advantageous because of its capacity in handling correlations over alternatives and time (Trains 2002). But, its normal distribution assumption may not be appropriate in many choice situations. Nevertheless, econometric estimation of multinomial probit models have been largely focused during the past decades. Bolduc (1999) illustrated the practical technique to estimate multinomial probit

(MNP) models in transportation. He applied maximum simulated likelihood technique for developing a mode choice model using commuter trip data from central business district of Santiago, and found that this technique can be very effective for the purpose. However, probit models do not have easy closed form equations as logit models, and hence require simulation which is computationally cumbersome.

2.4.4 Mixed Logit Models

Mixed logit is a highly flexible model that can approximate any random utility model (McFadden et al 2000). It is advantageous over standard logit because it can allow random taste variation, unrestricted substitution patterns and correlation in unobserved factors over time. The mixed logit probability can be derived from utility-maximization based on random coefficients. Mathematically, the utility of person *n* from alternative *j* is expressed as –

$$U_{nj} = \beta_n^{'} x_{nj} + \varepsilon_{nj}$$

And the mixed logit probability is given by –

$$P_{ni} = \int \left(\frac{e^{\beta^{'} x_{ni}}}{\sum_j e^{\beta^{'} x_{ni}}} \right) f(\beta) d\beta$$

Where, alternate *i* is chosen from a set of *j* alternates.

The error component of mixed logit can be divided into two parts – one having vector of random terms with zero mean, and the other having IID (Independently and Irrelevantly Distributed) extreme value. Hensher and Greene (2003) illustrated the specification and estimation techniques of mixed logit using three stated choice datasets from New Zealand and Australia. The result of the study revealed that mixed logit has the capacity to better capture the behavioral variability in the choice making process. However, to produce a good result, it requires large amount of high quality data which is not always available.

The discrete choice models mentioned above are based on random utility maximization (RUM) theory, where the modal preference is captured by specifying utility functions for each alternatives (Habib et al. 2010). As mentioned earlier, most of the models use revealed data on utilitarian variables to represent the mode choice. But, recent studies have shown that incorporating latent variables in the analysis can improve the understanding of mode choice. RUM based discrete choice modeling framework with latent variables has very recently been developed, but is not yet in practice widely. This type of model is called hybrid choice model (HCM) and it has the ability to incorporate stated choice variables, latent variables representing attitude or behavior, heterogeneity across decision-makers (Ben-Akiva et al. 2002). Habib (2010) has recently developed a hybrid choice model using revealed and stated choice data from Edmonton, which shows that it can better capture the latent behavioral factors influencing people's mode choice.

2.5 Elasticity Measures

Any demand study requires elasticity measures of the responsiveness of demand to changes in policy relevant variables (Dunne 1984). By definition elasticity is the ratio of percent change in probability to the percent change in attribute. Thus, elasticity measure is a useful tool to demonstrate how a particular variable affect the choice probabilities for the alternatives, and hence is very important for policy evaluation. As for discrete choice models in this study, the responsiveness of choice probability of a particular alternative with respect to changes in variables such as auto parking cost, transit travel time is determined. This section briefly discusses the elasticity measuring techniques in multinomial and nested logit models.

For MNL, the direct and cross elasticity equations are the same for all alternatives, whereas, elasticity expressions for NL changes for alternatives depending on the nesting structure (Koppelman & Bhat 2006). Direct elasticity is expressed as change in choice probability of an alternative for unit changes in the value of explanatory variables associated with that particular alternative. The equations for calculating direct elasticity of variables for MNL and NL are shown below (Koppelman & Bhat, 2006):

MNL: For changes in alternative *j*: $(1-P_j)\beta_{LOS}LOS_j$

NL: For changes in non-nested Alternative *j*: $(1-P_j)\beta_{LOS}LOS_j$

NL: For changes in nested alternative *k*:

$$\left[(1-P_k)+\left(\frac{1-\theta_N}{\theta_N}\right)(1-P_{k|N})\right]\times\beta_{LOS}LOS_k$$

Here: *LOS* stands for Level of Service attributes, *P* stands for probability of choosing an alternate.

On the other hand, cross elasticity is expressed as change in choice probability of an alternative with respect to changes in explanatory variables associated with other alternatives. Equations for calculating cross elasticity of variables for NL models are shown below (Koppelman & Bhat, 2006). It is to be noted that, in the case of nested logit models, cross elasticity of a particular variable varies among the alternatives, depending on whether the alternatives belong to same or different nests. This is because of the difference in scale parameters between different nests, and the elasticity equation incorporates this by adding φ in the equation. Therefore, cross elasticity within the same nest comes to be larger than that for different nests. The mathematical expressions of elasticity for MNL and NL are given below:

MNL: For changes in Alternative *j*: $-P_j \beta_{LOS} LOS_j$

NL: Effects on non-nested alternatives:

For changes in non-nested Alternative *j*: $-P_j \beta_{LOS} LOS_j$

For changes in nested Alternative *k*: $-P_k \beta_{LOS} LOS_k$

NL: Effects on nested alternatives:

For changes in non-nested Alternative *j*: $-P_j \beta_{LOS} LOS_j$

For changes in nested Alternative *k*: $-\left[P_k + (\frac{1-\theta_N}{\theta_N}) P_{k|N} \right] \times \beta_{LOS} LOS_k$

The models developed in this study contain dummy variables such as flexible office hour, compressed work week, which are important parameters for mode choice. Elasticity of these variables can be important for policy implication in

TDM perspective. Rajamani et al. (2003) performed a study on assessing the impact of urban form measures for nonwork trip mode choice, where elasticities of dummy variables have been calculated using the following method.

To compute the aggregate level elasticity of a dummy variable (such as flexible office hour), the entire sample is first divided into two subgroups based on the value (1 or 0) of that dummy variable. Then for each subgroups, the value of the dummy is changed (zero to one for one subgroup, and one to zero for the other), and the shifts in aggregate level mode share is calculated respectively. Then the ratio of the difference of the shifts between the two subgroups, to the aggregate mode shares in the entire sample, gives the aggregate level elasticity of the dummy variable.

2.6 Panel Data Analysis

A longitudinal, or panel data refers to a set of observations that comprises of multiple observations of each individuals over time (Hsiao 2003). Panel data is advantageous over cross sectional data because of its capacity to capture the complexity of human behavior, and hence panel data models have become increasingly popular among applied researchers.

In transportation planning, panel data is of particular importance in order to analyze travel behavior and travel demand forecasting. During the late 80s and early 90s, use of panel data in travel behavior analysis have been given specific emphases in several international conferences; such as in 1988 travel behavior conference at Oxford, the 1989 World Conference on Transport Research in

Yokohama, and in 1990 Third International Conference on Survey Methods (Kitamura 1990).

The advantage of panel data analysis in transportation planning includes capturing unobserved contributing factors or latent variables in travel decision facilitating more precise measurement of behavioral changes, reducing sampling errors resulting reduced sample size requirements compared with repeated cross-sectional surveys, identifying temporal variation in travel behavior, and so on. However, not many works have been performed in transportation studies using panel data due to fact that the survey design and data collection process can be cumbersome (Kitamura 1990).

In a study on commuter mode choice, Dargay (2005) estimated dynamic panel data model using eleven years of data from the British Household Panel Survey (BHPS). In the study, the changes in the behavior of the individuals as a result of changes in external factors and in their socio-economic and demographic characteristics were studied. The results showed that the heterogeneity specified in the random effect model was significant, indicating difference between individuals behavior not accounted for in the explanatory variables included in the survey. But, considering the fact that car was the predominant mode (64% to 70%), the study developed a binary probit mode choice model with two modes only – car and non-car. It is to be noted that, binary probit models are applicable only in the situations having two alternates. If there are more than two alternates in the choice set, binary models are not applicable.

Zureiqat (2009) performed a panel data study on the response of public transport users to fare changes. In the study, a discrete ticket choice model and a discrete-continuous model mode choice and time-of-day choice were developed using 2.5 years data from London's public transport system. The ticket choice model was at disaggregate level and it was estimated as a multinomial logit model using Biogeme. Elasticities of transit fare were also determined for policy analysis. The study showed that effective transport demand study requires capturing the complexity of transit fare structures and long duration panel data can be very effective for forecasting purpose.

2.7 Summary

From the discussions presented in the chapter, it is found that numerous research works have taken place to reveal the actual process behind people's mode choice behavior. Various factors influencing the mode choice trend in TDM perspective have also been looked after by many researchers. But to the best of our knowledge, not many works have made use of large scale panel data containing both instrumental and affective variables to investigate their effects on commuter mode choice. This research therefore attempts to develop a mode choice model to investigate the sensitivities of observed and unobserved factors influencing people's commuting mode choice.

CHAPTER 3

METHODOLOGY

It is important to develop a sound methodology of work in order to systematically approach the mode formulation. Different parts of this thesis work are based on the theoretical background presented in the previous chapter. The literature was used for conceptual development of the model structures and policy analysis performed in this study. This chapter mainly presents the techniques used for data preparation and the mathematical frameworks for model development. Figure 3.1 shows the flowchart of work performed in this study.

3.1 Model Structures

From the discussion presented in Section 2.4, we know that different types of models are available for discrete choice modeling. Among those, multinomial logit and nested logit models are most widely used in transportation modeling. Probit and mixed logit models are computationally burdensome, as they do not have closed form probability equations. In this respect, multinomial and nested logit models are easier to estimate and interpret. Therefore, in this thesis, we have used these models to analyze the mode choice. However, as mentioned before, the dataset used in this study has 7 days mode choice observations for each individual, which gives us an opportunity to perform panel data analysis. Although mixed logit models are particularly effective for analyzing heterogeneity, this study attempts to develop a simpler nested logit model with panel data. Thus, the utility of each alternate is expressed as a function of mode specific and generic variables, with an additional random coefficient to described the variation of choice across time. This section mainly presents an overview of formulation of probability equations for multinomial and nested logit models.

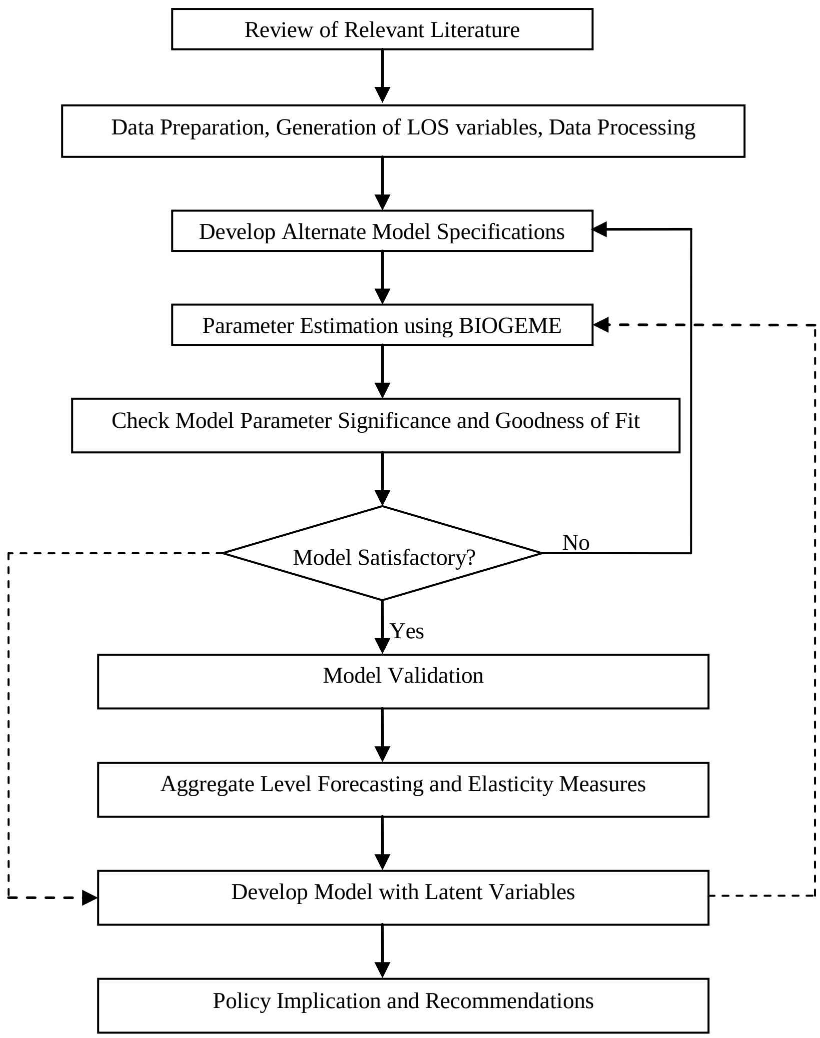


Figure 3.1: Methodological Flowchart

3.1.1 Multinomial Logit

The MNL model is derived thorough the application of utility maximization concepts to a set of alternatives *i*, from which individual *t* chooses the one having maximum utility. It is assumed that the total utility has two components, namely systematic utility V_{it}, and a random component ε_{it}. So,

$$U_{it} = V_{it} + \varepsilon_{it}$$

The systematic deterministic component of the utility function is commonly specified as linear in parameters and includes variables that represent the attributes of the alternatives (eg. Travel time, cost), the decision context (eg. time of travel), and the characteristics of the decision-maker (eg. Gender, age). Assuming the error term in the above equation as type I extreme value gumbel distributed and the alternatives as independent and irrelevant alternatives, leads to the famous MNL model:

$$P_{it} = \frac{e^{V_{it}}}{\sum_{i=1}^{j} e^{V_{it}}}$$

where P_{it} is the probability that alternative *i* is chosen by individual *t*, and *j* is the total number of alternatives within the choice set. The closed form of the MNL model is easy to use and interpret. So it is widely used in choice contexts in various fields including travel related choice situations such as mode choice, destination, car ownership, residential locations etc.

3.1.2 Nested Logit Model

As this study uses a nested logit model to represent the mode choice probabilities, understanding of the basic theory behind the NL model is important. This section, therefore, presents the formulation of an NL model used in this study.. Here, it is considered that there are eight modes in the choice set (presented in Figure 3.1), and they are grouped in three nests; namely – Auto, Transit and Non-motorized. The modes are at the lower level of the nest, where the auto nest consists of drive alone and carpool, the transit nest includes walk, auto & bike accessed transit modes, and the non-motorized nest contains cycle and walk/jog/skate modes. The utility functions of one alternative from each nest can be written as –

$$U_{DA} = V_{Auto} + V_{DA} + \varepsilon_{Auto} + \varepsilon_{DA}$$

$$U_{PT} = V_T + V_{PT} + \varepsilon_T + \varepsilon_{PT}$$

$$U_W = V_{NM} + V_W + \varepsilon_{NM} + \varepsilon_W$$

Where,

U = total utility; V = systematic utility; ε = random utility having type I extreme value distribution.

Auto = Auto nest; T = Transit nest; NM = Non-motorized nest

DA = Drive alone; PT = Public Transit (walk access); W = Walk/jog/skate

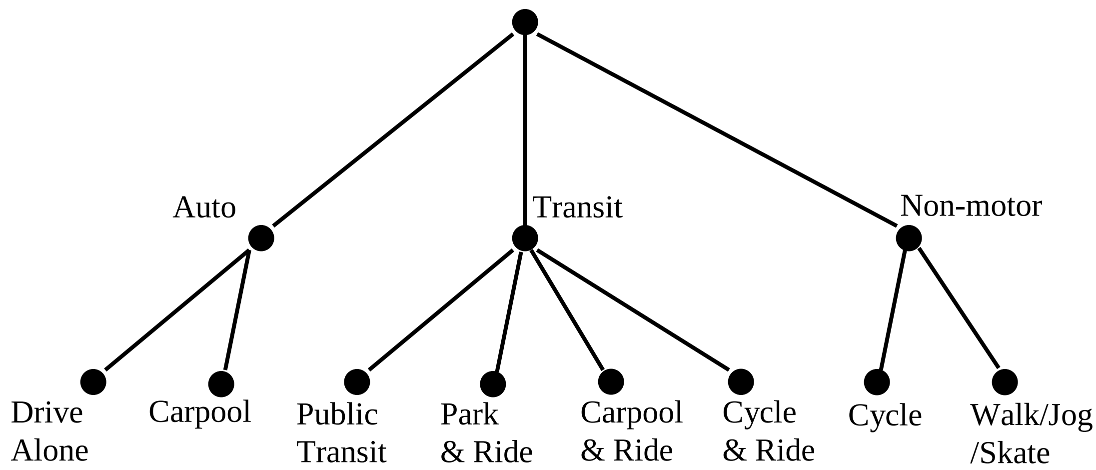


Figure 3.2: Two level Nesting Structure

The systematic portion of utility comprises of the alternate specific and/or user specific variables *x*, multiplied by coefficient *β*. The systematic utility function is entered into estimation software in order to estimate the values of *β*. The utility equations for nested logit used in this study are presented later in this chapter.

However, once the coefficient values are estimated and systematic utility values (*V*) are known, the choice probabilities for the lower level of the nested alternatives (drive alone, public transit, walk/jog/skate) or the conditional probabilities (probability of choosing an alternative, given that the nest is chosen), can be calculated by using the following equations:

$$\Pr(DA \mid Auto) = \frac{\exp\left(\frac{V_{DA}}{\theta_{Auto}}\right)}{\sum \exp\left(\frac{V_i}{\theta_{Auto}}\right)}$$

$$\Pr(PT \mid T) = \frac{\exp\left(\frac{V_{PT}}{\theta_T}\right)}{\sum \exp\left(\frac{V_j}{\theta_T}\right)}$$

$$\Pr(W \mid NM) = \frac{\exp\left(\frac{V_W}{\theta_{NM}}\right)}{\sum \exp\left(\frac{V_k}{\theta_{NM}}\right)}$$

Here *i*, *j* and *k* represent all alternates within auto, transit and non-motorized nests respectively. And θ is commonly referred to as the logsum parameter, and is the inverse of scale parameter μ for corresponding nests. The value of θ ranges between 0 and 1, and it can be mathematically expressed as follows:

$$\theta = \frac{1}{\mu}$$

The choice probability for the upper level of the nest (Auto, transit and non-motorized) can be written as –

$$\Pr(Auto) = \frac{\exp(V_{Auto} + \theta_{Auto} \times IV_{Auto})}{\exp(V_{Auto} + \theta_{Auto} \times IV_{Auto}) + \exp(V_T + \theta_T \times IV_T) + \exp(V_{NM} + \theta_{NM} \times IV_{NM})}$$

$$\Pr(T) = \frac{\exp(V_T + \theta_T \times IV_T)}{\exp(V_{Auto} + \theta_{Auto} \times IV_{Auto}) + \exp(V_T + \theta_T \times IV_T) + \exp(V_{NM} + \theta_{NM} \times IV_{NM})}$$

$$\Pr(NM) = \frac{\exp(V_{NM} + \theta_{NM} \times IV_{NM})}{\exp(V_{Auto} + \theta_{Auto} \times IV_{Auto}) + \exp(V_T + \theta_T \times IV_T) + \exp(V_{NM} + \theta_{NM} \times IV_{NM})}$$

where, IV is called the inclusive value parameter and it represents the expected value of the maximum utility of lower level alternates. For example IV_T is mathematically expressed as –

$$IV_T = \log\left[\sum \exp\left(\frac{V_j}{\theta_T}\right)\right]$$

Finally, the probabilities of choosing the nested alternatives can be determined by multiplying the conditional probabilities with the nest probability:

$$\Pr(DA) = \Pr(DA \mid Auto) \text{ x } \Pr(Auto)$$

$$\Pr(PT) = \Pr(PT \mid T) \text{ x } \Pr(T)$$

$$\Pr(W) = \Pr(W \mid NM) \text{ x } \Pr(NM)$$

As mentioned earlier, NL can handle potential violation of IIA property by grouping similar alternatives in one nest. This is a big advantage of NL over MNL. Moreover, the value of the logsum parameter θ is useful to interpret the correctness of the nested logit model.

3.2 Model Specifications

Modeling mode choice requires defining separate utility equations for each alternate available in the system. Eight alternate are found from the workplace commuter survey data, which are as follows:

1. Drive Alone
2. Carpool
3. Public Transit (Bus / LRT)
4. Park & Ride
5. Carpool & Ride

6. Cycle & Ride

7. Cycle

8. Walk, Jog or Inline Skate

Details on finalizing the above set of alternates are discussed in the data preparation chapter. However, the utility functions of each of these alternates have been developed using the variables available in the survey data. Some of these variables are level of service variables which are alternate specific, such as – travel time, cost etc, while some are user specific or generic variables, for example – age, gender etc. Different variables are also identified in terms of their type – such as continuous, ordered, dummy etc. Using these variables, utility functions for each alternate have been written, a generic form of which is presented below:

$V_i = ASC_i + \beta_{IVTT} * (\mathrm{IVTT})_i + \beta_{P\_Cost} * (\mathrm{P\_Cost})_i + \beta_{D\_Cost} * (\mathrm{D\_Cost})_i + \beta_{H\_W\_Dist\_i} * (\mathrm{Hwd}) + \beta_{Age\_i} * (\mathrm{Age}) + \beta_{Gen\_i} * (\mathrm{Gen}) + \beta_{Emp\_i} * (\mathrm{Emp}) + \beta_{Flex\_i} * (\mathrm{Flex}) + \beta_{Comp\_i} * (\mathrm{Comp}) + \sigma_i$

Once the base model is estimated using utility equations similar to that shown above, the latent variables have then been incorporated in the base model. The latent variables have been drawn from the responses of stated preference questions in the workplace commuter survey. For example, the respondents were asked to choose from a set of options against the question "under what circumstances would you consider carpooling more often?" Options were such as – "if I could find compatible carpool partner" or "if I didn't have to use car for work purpose". The selected options have been assigned 1 while the others are 0, and then they have been added in the utility functions as dummy variables. Figure 3.2 shows the types of variables used in the mode choice utility equations.

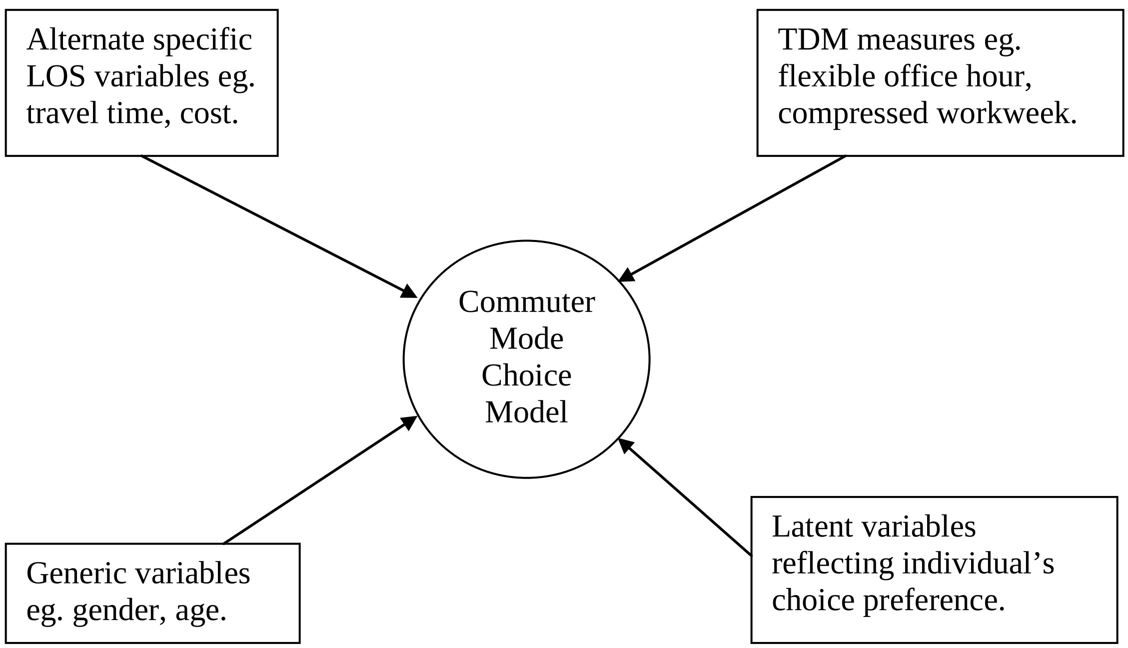


Figure 3.3: Variable Types in Model Input Data

3.3 Model Estimation

The mode parameter estimation has been performed using estimation software BIOGEME version 1.8. BIOGEME (Bierlaire 2009) is a powerful tool for maximum likelihood estimation of generalized extreme value models and is extensively used by many researchers worldwide. For estimating a model in BIOGEME, it requires the data set to be prepared in a specified format. So, after necessary data processing, the estimation dataset has been prepared for estimation. A model specification file is also required where details of model specifications, such as utility functions, variables definitions, nesting structure, dummy variable, random parameter, number of draws etc. are given.

CHAPTER 4

DATA FOR EMPIRICAL INVESTIGATION

This study develops a commuter mode choice model and policy evaluation framework from a Transportation Demand Management perspective. Any study in travel behavior analysis requires an authentic sample with sufficient observations having revealed preference or stated preference mode choice data. In revealed preference data, the respondents' actual choice preference in real situations is collected, whereas in stated preference the respondents are presented with hypothetical choice situations and the stated choice preference data are collected (Train 2002). This study uses the dataset from a survey conducted by the Transportation Department of the City of Edmonton. This chapter discusses about the data source and preparation for model input.

4.1 The Workplace Commuter Survey

The Transportation Department of the City of Edmonton conducted a workplace commuter survey in 2007. The survey was performed under the process of developing a Transportation Demand Management (TDM) program for City of Edmonton staff and all participants were targeted among the City of Edmonton employees. The objective of the TDM program was to promote alternate modes of transportation that may include walking, cycling, public transit and carpooling among the city staff, with a vision to reducing the number of single-occupant vehicles commuting to the workplace, and thus encouraging travel behavior that reduces traffic congestion, energy consumption and vehicle emissions. The purpose of the survey was to obtain better understanding of City staff attitude and behaviors towards alternate modes of travel and to assist in the development of potential commuter options.

The survey questionnaire contained a total of 40 questions in three sections. The questionnaire is briefly discussed here, while the details questions are presented in Appendix A. The first section focused on user characteristics such as age, gender, work department, home & work postal codes, employment status, work start and leave hours, home to work distance, and availability of flexible office hours and compressed work week.

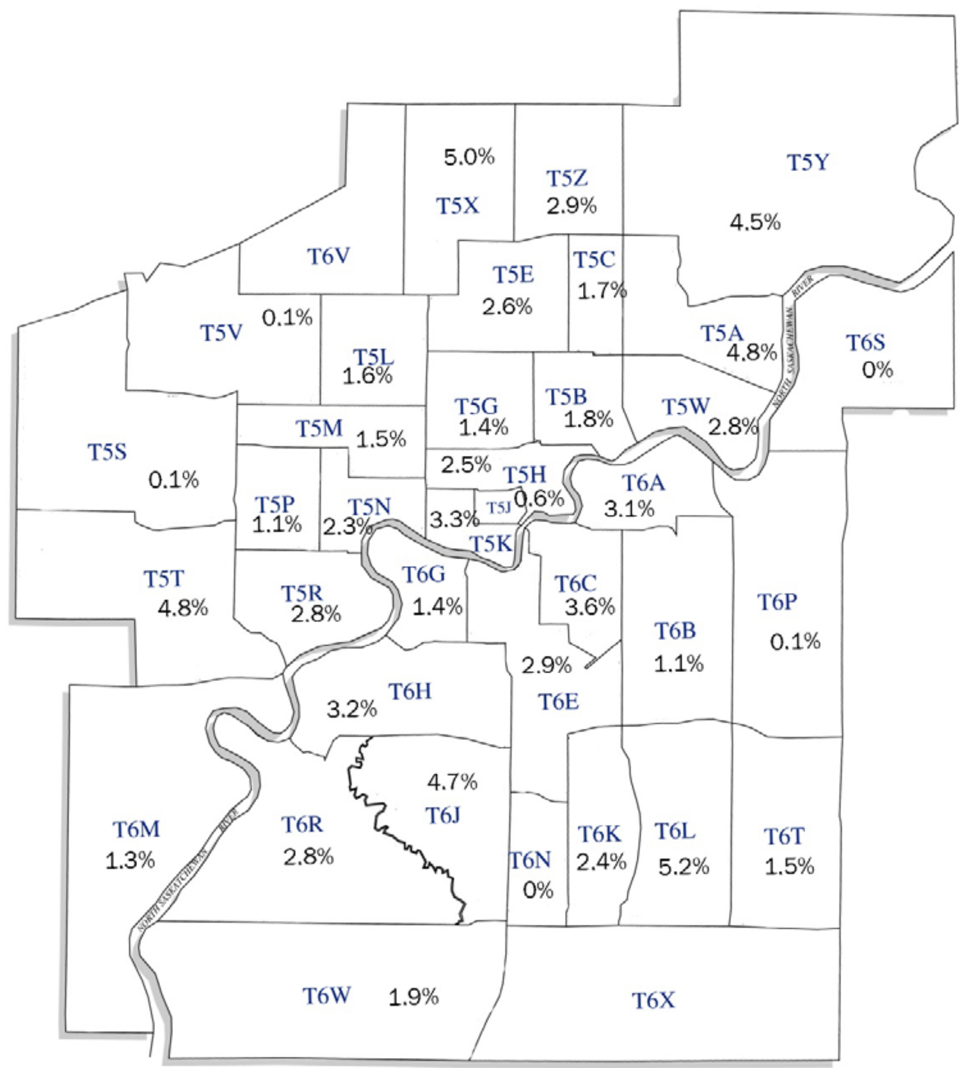
The next section gathered revealed information on means of commuting to work. Here, weeklong mode choice information has been collected. Each respondent had to select the mode actually used to commute on each day of the previous week. Thus, a longitudinal or panel dataset has been collected, where 7 mode choice observations for each individual are available. The respondents were given the following options to choose from: Did not Work, Drive Alone, Carpool, Public Transit, Park & Ride, Carpool & Ride, Cycle & Ride, Cycle, Walk Jog or Inline skate, Work from Home & Others. No trips were made in the cases where the individuals chose either "Did not work" or "Work from Home", and very few respondents (less than 1% in average) reported that they used the "other" modes for the trips. Therefore, the total number of transport mode alternates available in the system is eight, as listed above.

Questions on frequency of the use of five alternates (public transit, carpool, cycle, walk/jog/skate, telework) during the previous 12 months were also included in this section, along with a stated preference question on each of the alternates, asking under which circumstances the respondent would consider using each mode more often. For each of the 5 alternates, the respondents were given a set of preference scenarios, which would foster more frequent use of the modes.

The last section of the survey questionnaire contained queries on commuting details for selected alternates – drive alone, carpool, public transit, cycle, and walk/jog/skate. Each respondent had to answer the questions related to his/her most used alternate only. In this section, revealed information on travel time, cost, number of transfers, mode of payment, availability of parking space etc. was asked for each alternate.

Although a total of 3723 respondents completed and returned the survey, fewer respondents (2932) covered all three sections, while each section was completed by 3469, 3184 and 3055 respondents respectively. From the initial analysis of the data, it has been found that the distribution of the respondents is reasonable in terms of home & work locations, gender, age, employment status, etc.

City of Edmonton is the third largest employer in Edmonton with over 9,000 employees (FL 2010). As can be seen from Figure 4.1, the home locations of the survey respondents are distributed evenly across the city. The ratio of female to male among the survey respondents are 48.5% to 51.5%, which is pretty close to that from 2006 Edmonton census (50.3% to 49.7%). The distribution of age groups of the respondents is also similar to that found in Edmonton household travel survey (2005), with the highest frequency of age group being 45 to 54 years (36.6%).



Note: The above map shows the zones within the city boundary only. Apart from those shown in the figure, the following postal code zones have significant observations:

T7X – Spruce Grove: 1.7%

T8A – Sherwood Park: 4.1%

T8N – St. Albert: 6.0%

Figure 4.1: Distribution of Survey Participants' Home Locations across Edmonton

The preliminary data analysis suggests that drive alone is by far the predominant mode (60.4%) used by the survey respondents. The aggregate level mode share percentages are shown in Figure 4.2, where it can be seen that after drive alone, public transit (127.7%) and carpool (9.3%) are used by a fair percentage of individuals, but the rest of the modes are less than 5%. From Figure 4.3, it is understood that aggregate mode share percentages do not significantly vary along days of week, and on each weekday about 60% of respondents used drive alone to commute to work. But, in disaggregate or individual level, it is found from the data that most individuals (69%) used the same mode on all work days. The rest (31%) used different modes on different workdays. This is where the importance of panel data comes in. Why is an individual having the same characteristics and access to alternates with the same attributes using different modes on different days? It is anticipated that there are some latent factors or variables controlling human behavior which a simple cross-sectional model cannot capture. Therefore, this study makes use of the panel data so that this variability can be understood.

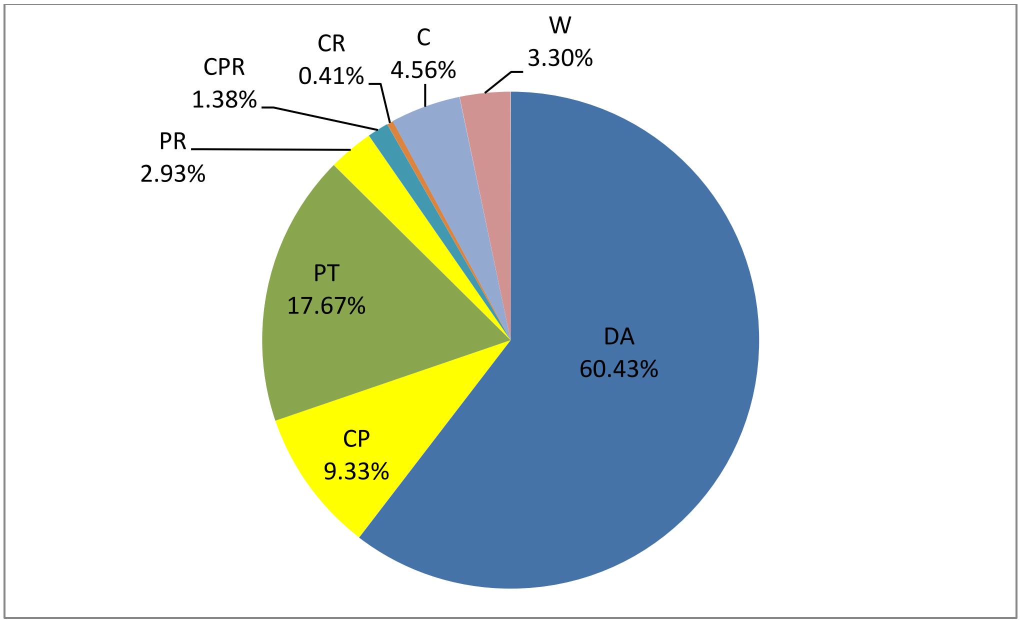


Figure 4.2: Observed Mode Share Percentages for the Whole Sample

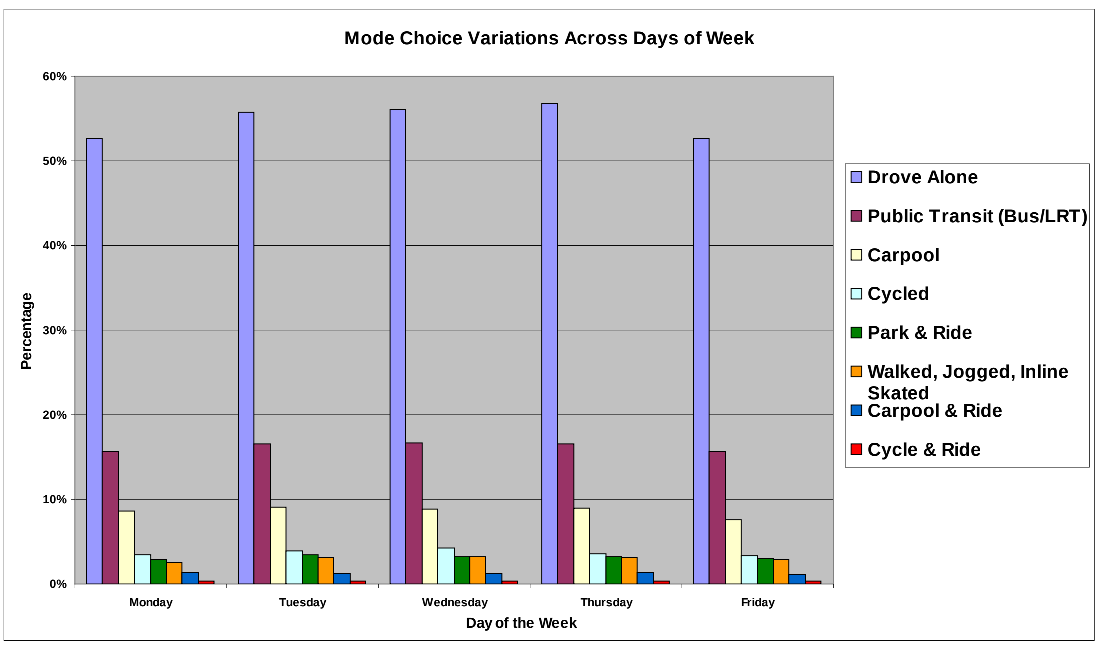


Figure 4.3: Mode Share Variations across Days of Week

4.2 Level of Service Attributes

It is mentioned in the preceding article that the survey has collected details of commuting mode actually chosen by each respondent. This implies that the Level of Service (LOS) attributes of all modes for all individuals are not available from the survey data. For example, if an individual had chosen drive alone mode for his/her commute, he/she only provided the travel time and cost information for drive alone. But, there may have been other alternates within his/her choice set which the respondent did not choose. As a result, for each individual, LOS variables of only one mode are available from the survey, even if his/her choice set contained two or more alternates. To develop a mode choice model LOS variables (Travel time, travel cost etc.) for each mode within the choice set is necessary, and these data has been collected using the web-based Google Maps feature from Google Inc. (Google Maps 2009). This feature allows the user to obtain travel time information for a specific trip, if the origin and destination postal codes are given as input. Travel time for driving, public transit and walking is available for most North American cities. It should be mentioned that, transit travel time information is also available at the ETS (Edmonton Transit System) website, where a Google maps link is provided. However, this study has made use of Google maps because of its user friendly interface and faster speed. Figure 4.1 shows the postal code zones in Edmonton, while Figure 4.4 shows the interface of Google Maps, in obtaining public transit travel time information for a trip from T6H to T5B.

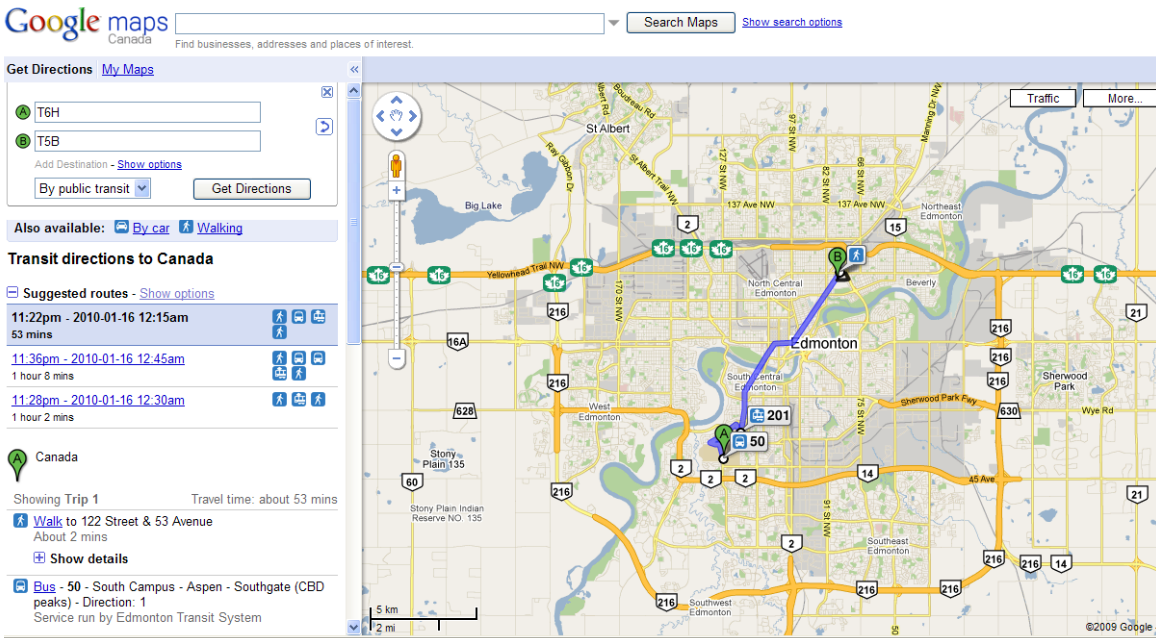


Figure 4.4: Edmonton Transit Information on Google Maps

4.3 Obtaining LOS Data

This section discusses the process followed to obtain the LOS variables for each mode for each respondent, along with the assumptions made while gathering LOS data for the alternates.

- General: As mentioned before, the respondents' home and work postal code zones have been used while obtaining travel time data for the work trip. Since, the survey collected the first three digits of the postal codes, the center of home and work postal code zones have been taken as the origin and the destination respectively. There were a few exceptions, as in the case of postal code T5Y. The north part of the zone is yet to be developed and most of the localities are in the southern part. In this case, considering the center of the zone is not reasonable, and hence the approximate center of major development within the zone has been assumed to be the origin of trips.

From the survey data, some individuals were found to living and working in the same postal code zones. Intrazonal travel time for auto and especially public transit was necessary for such observations, and nearest neighbor method was used for the purpose (Travel Forecasting Guidelines, 1992). The method involves averaging half of travel time taken for trips to the adjacent zones. For example, to calculate the intrazonal travel time of zone T5L, time taken from this zone to its adjacent zones T5E, T5G, T5M, T5V, T6V were first collected. Then each travel time was halved and the average value was taken as the intrazonal travel time for T5L.

- Drive Alone: Drive alone level of service variables used in this study are travel time, parking cost and driving cost. Since most of the respondents (about 60%) used drive alone for commuting to work, the survey collected travel time and parking cost information for these respondents. However, significant variation in travel time reporting was observed. For example, respondents having same home and work locations reported wide range of home to work distance and/or travel time. To eliminate this error, the survey collected distance and travel time information was cross-checked using Google Maps, and necessary corrections have been made. For the respondents who did not provide drive alone variables (because they used other modes to commute), their travel time and parking cost were considered same as reported by other respondents having same home and work locations. Home to work distance in km has been used to calculate the driving cost. Here per km driving cost has been assumed to be \$0.165, based on standard Cobalt LT operating cost. (Canadian Automobile Association, 2008).

- Carpool: Carpool variables are similar to drive alone, only with an exception that travel time increases by a small margin, which has been assumed to be 5 minutes in most cases. Analysis of survey data revealed that 89% of respondents who chose carpool reported 2 persons in the carpool. Therefore, a 2 person carpool has been assumed for all individuals and parking & driving costs have been taken as half of those for drive alone (assuming cost is equally shared by 2 carpool partners).
- Public Transit: Google Maps generates several route choice options for a particular trip in public transit, having varying travel time. In this study, the travel time for options with lesser number of transfers has been chosen, unless travel time varied significantly between the options. It is understood that travel time in public transit can vary significantly depending on the time of day; because of varying transit frequency during the day. Reasonably, Google Maps allows the user to enter the time of travel, and so the actual time of travel, as reported by the respondents could be used in obtaining travel time data. For each individual, public transit In-Vehicle Travel Time (IVTT), wait time at starting point (which is often assumed to be half of transit frequency), wait time for transfers, walk time (from origin and to destination), (summation of total wait and walk time is referred to as Out-of-Vehicle Travel Time or

OVTT), and number of transfers (0, 1, and 2 or more) data have been obtained. Transit fare has been considered on ETS current fare rates – $2.5 flat for 90 minutes of travel. Many respondents have their home locations outside the city boundary, such as St. Albert or Sherwood Park. To obtain transit time and fare data for them, respective agency websites (such as Strathcona County website for City of Sherwood Park) have been visited and required information has been collected.

- Park & Ride: Increasing emphasis has recently been put by the City of Edmonton on park & ride options to the city center, especially complying with the expansion of LRT system. However, at present there are few park & ride locations available, such as Clareview, Belvedere and Commonwealth Stadium park & ride facilities. So, for park & ride the LOS variables have been obtained considering home to above-mentioned parking facilities as drive alone and then taking transit up to the work location. Cost of parking has been assumed to be zero.
- Carpool & Ride: Carpool & ride used the same data as park & ride variables, except for added carpool time and half the driving cost.
- Cycle & Ride: Edmonton Transit System has bicycle parking facilities in several locations in the city. The locations, as found in the City of Edmonton website are:
  - Belvedere LRT Station
  - Belvedere Transit Centre
  - Century Park Transit Centre

- Clareview LRT Station
- Coliseum LRT Station
- Grandin LRT Station
- Millgate Transit Centre
- Mill Woods Transit Centre
- South Campus LRT Station
- Southgate Transit Centre
- Stadium LRT Station
- Westmount Transit Centre

The trip makers are assumed to cycle the nearest facility from home on cycle and then take public transit to the work location. Cycling distance from home and public transit time to destination have been taken from Google, and the cycling time from home to transit station has been calculated assuming average cycling speed of 15 km/hr.

- Cycle: Cycle time is obtained using home to work distance, assuming the same (15 km/hr) average cycling speed. Cost is assumed to be zero.
- Walk / Jog / Inline Skate: Walking has been assumed to be dominant within this alternate, as the survey did not have any queries regarding this. However, walking speed has been taken as 4.5 km/hr to compute time for each individual. Cost has been taken as zero.

4.4 Defining the Choice Set

Defining choice set is an important part of data preparation. It is evident that not all the modes are available to all the people, or in other words, a particular individual may not have all alternates within his/her choice set. For example, people living far from their workplaces would not have “walking” in their choice set. Therefore, choice set has been defined for each individual on the basis of respondent characteristics, reported preference and some assumptions. Some of stated preference answers have been considered while defining the choice sets. These are given below:

- The survey did not have any household car ownership related question. So availability of car to a particular person is not known, and thus all respondents have been assumed to have access to a vehicle to drive alone with. However, drive alone is considered not within the choice set for the respondents who mentioned (in one of the stated preference questions) that they did not drive.
- Carpool is considered available to most of the respondents, except those who mentioned that they would never consider carpooling to work.
- For evaluation whether cycling would be within the choice set of an individual, his/her home to work distance is considered. Analysis of survey data shows that, maximum distance cycled by most of the cyclists (more than 80%) is 15 km, and for individuals having lesser home to work distance cycling was included within the choice set.

- Similar approach is taken for defining walking. Maximum distance covered by walk/jog/skate was found to be 7.5 km, and based on this the availability of this mode is defined.
- For park & ride and carpool & ride availability, parking facilities within and outside the city have been considered. For example, there are two free parking facilities in Sherwood Park and Strathcona County, which allows the residents of that zone to have park & ride within their choice set.

4.5 Data Arrangement

After all the required information was gathered, the dataset was sorted and arranged so that it can be read by the model estimation software BIOGEME. As mentioned earlier in this chapter, some data were found missing in each section of the survey. Such erroneous observations can cause error in the estimation process and hence have been deleted. The final datasheet has been arranged in such a manner that each row contains one mode choice observation and all the explanatory variables, with subsequent rows containing data from the same individuals. Data rows containing chosen mode either “did not work” or “others” or “telework” have been omitted from the final datasheet. Respondents choosing “others” as their home or work postal code zones were excluded from the datasheet, as required LOS information cannot be gathered if the postal codes are not known.

After needful noise cleaning and arrangement, the dataset comprised of 2932 individuals with 13522 observations. It should be mentioned that the dataset does not contain an equal number of observations for all individuals. So, the final dataset is an unbalanced panel dataset. However, the entire dataset has been randomly divided into two sets, generating random number against each observation. 80% of the dataset (10817 observations) has been used for model estimation, and 20% (2705 observations) has been used for validation.

CHAPTER 5

ANALYSIS AND MODEL RESULTS

As discussed in chapter 1, the prime objective of this study is to develop mode choice models using panel data analysis, and to observe the sensitivity of modal shares with respect to changes in explanatory variables, especially those that are important for TDM strategies. Before developing the model specifications, some simple analyses on mode choice against various factors were performed. For mode choice modeling, several multinomial and nested logit models with different nesting structure have been estimated and examined. The model that yielded the best result in terms of forecasting aggregate level modal shares has been taken as the final model, for which the elasticities of some important variables have been computed. This chapter presents the data analysis results, model specifications and results with validation, and the elasticity measures of some important variables.

5.1 Data Analysis

As seen from Figure 4.2, the mode share (averaged over all observations) is heavily auto dependent, with a percentage of over 60%. The data also suggest that the aggregate level mode share percentages do not vary significantly over the week (Figure 4.3). However, variations in mode choice at an individual level are expected based on availability of flexible office hours and a compressed work week. 53.3% of the individuals have reported that they have flexible office hours, while 58.72% people have compressed work week options.

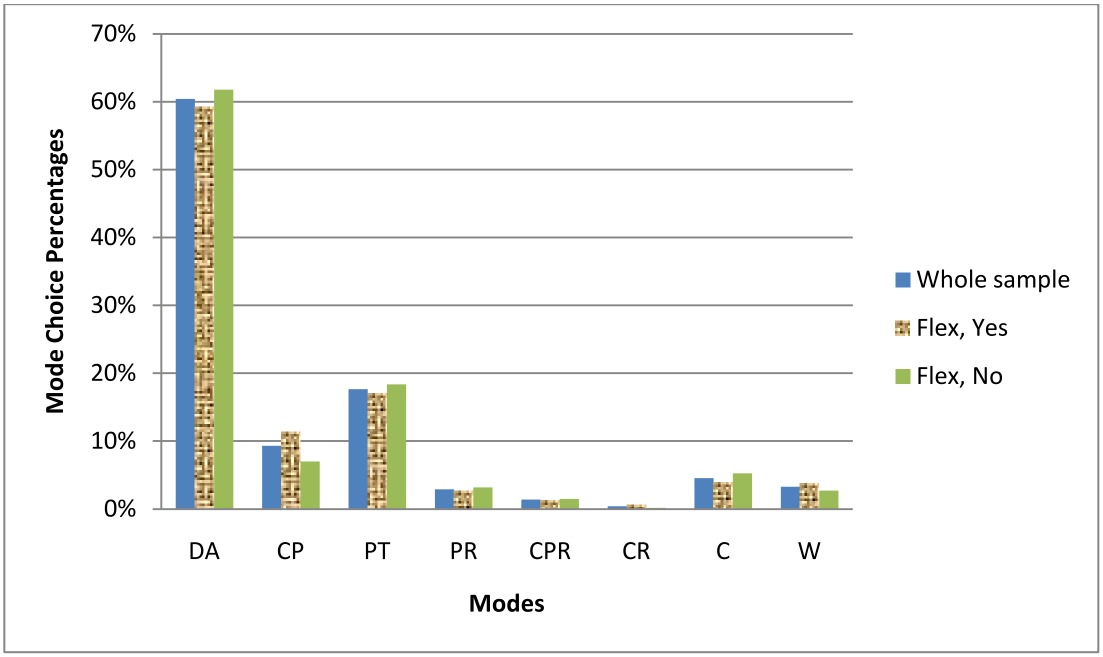


Figure 5.1: Variation in Mode Choice for Flexible Office Hour

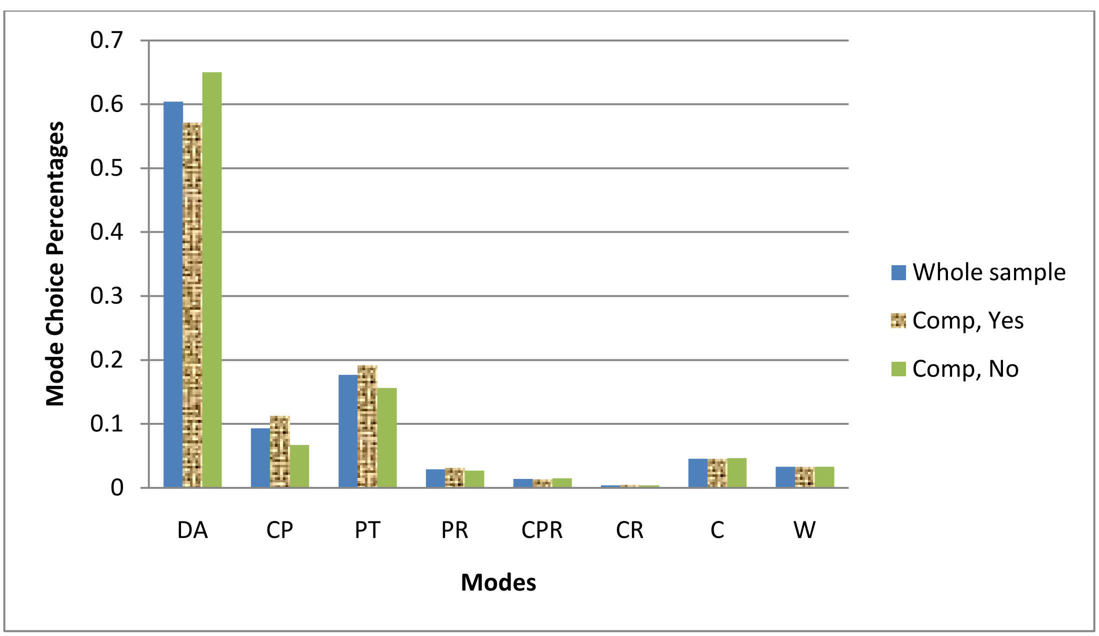


Figure 5.2: Variation in Mode Choice for Compressed Work Week

From the above figures it can be seen that these two measures have some effect (although little) on drive alone, carpool and public transit percentages. On the other hand, flexible office hours and a compressed work week have a fair amount of effect on work arrival time (Figure 5.3 and 5.4), which in turn is related with mode choice.

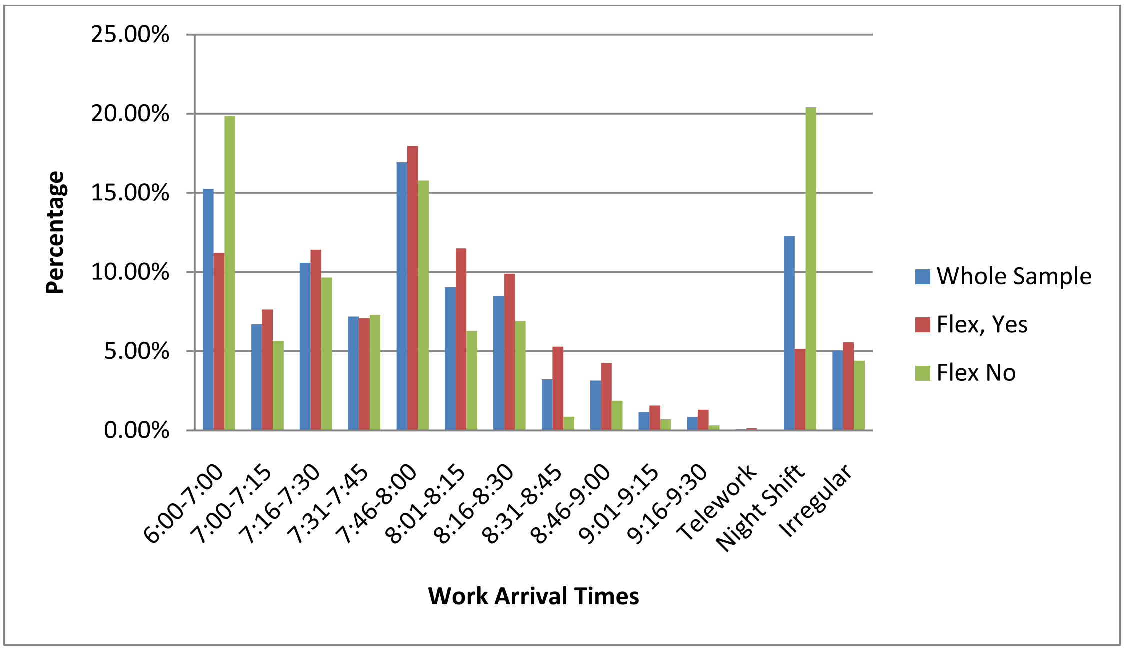


Figure 5.3: Variation in Work Arrival Time for Flexible Office Hour

Figure 5.5 shows mode share distribution with home to work distance. As one can imagine, for shorter distances the walk percentage is significant. Public transit is used mostly for a distance between 6 to 10 km and people prefer using drive alone for longer distances.

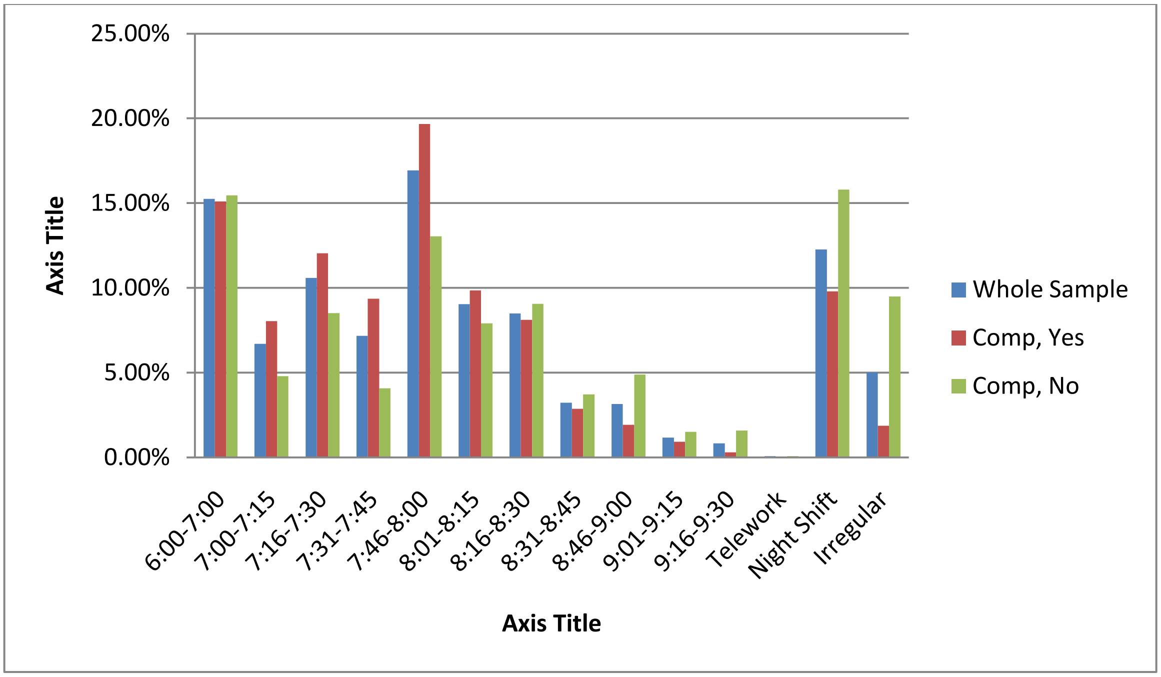


Figure 5.4: Variation in Work Arrival Time for Compressed Work Week

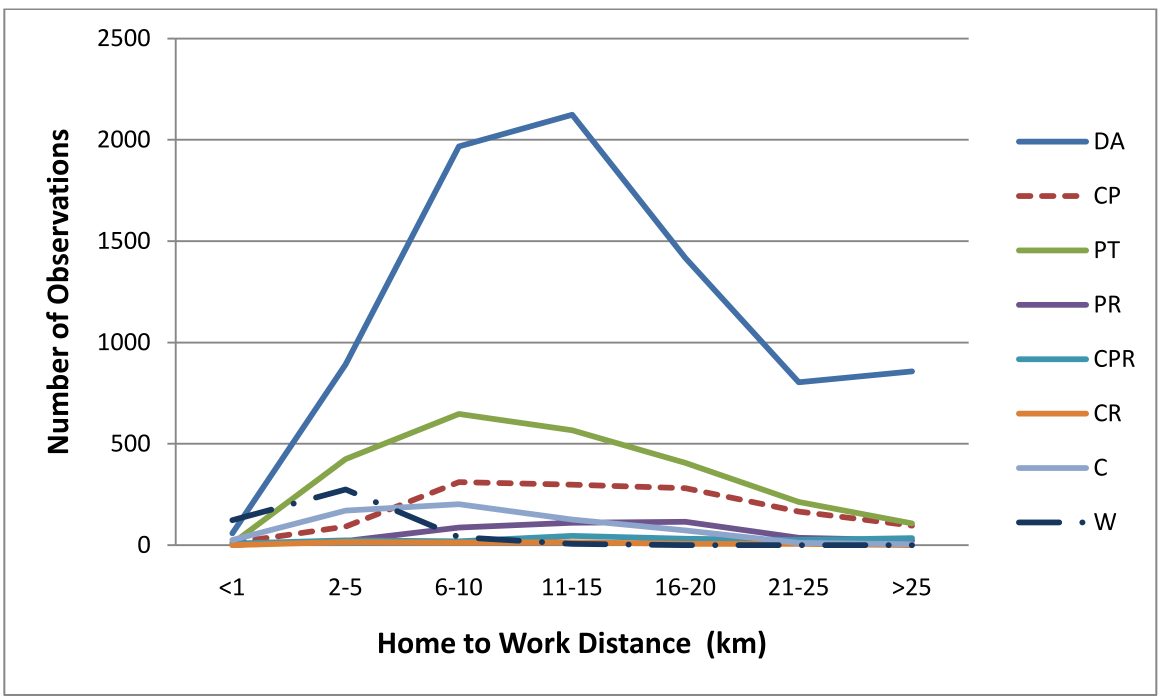


Figure 5.5: Mode Share Distribution with Home to Work Distance (km)

5.2 Model Specifications and Results

From the discussions in the literature review, it is clear that logit models are appropriate for discrete choice mode choice analysis. This study therefore develops logit models to describe the choice phenomena in the study area. Using the final arranged datasets, several logit models have been developed with different specifications. Initially multinomial logit models were attempted, which did not yield satisfactory results in terms of statistical significance and signs of the parameters. This is pretty much expected because it is evident that some of the alternates (such as drive alone and carpool) have common characteristics and thus may violate the IIA assumption. Attempts were thereby taken to develop nested logit models.

Different combinations of nesting structure are possible for even the simple nested logit models. It is understood that alternates having similar characteristics can cause potential violation of IID (Independently and Irrelevantly Distributed), and hence they are grouped together in form of a nest. So, NL model specifications have been developed with different nesting structures and specifications with different combinations of explanatory variables, and the parameters have been model estimated using BIOGEME. After several trials, models yielding better results have been finalized and validated, and they have been compared.

5.3 Model Specifications

This section describes the specifications of three NL models that appear to have given more satisfactory results than the others. The models are:

Model A: 3 nest single level NL using survey respondent reported travel time data.

Model B: 3 nest single level NL using Google reported travel time data.

Model C: 4 nest single level NL using Google reported travel time data.

Nesting structures of Model A & B are presented in figure 5.1. It can be seen that nest A contains the auto only modes – drive alone and carpool. Nest B includes the alternates where transit (bus and/or LRT) is used, either for the entire trip (public transit) or for part of the journey (park & ride, carpool & ride, cycle & ride). The non-motorized modes are grouped in Nest C. The scale parameters at the lower level are set to unity and the upper level scale parameters (φ) are estimated. For the estimation of the nest parameters, Nest A scale parameter has been fixed to 1, allowing the other nest parameters to be estimated. The reason for fixing Nest A is that it contains the alternates which are available to most of the respondents.

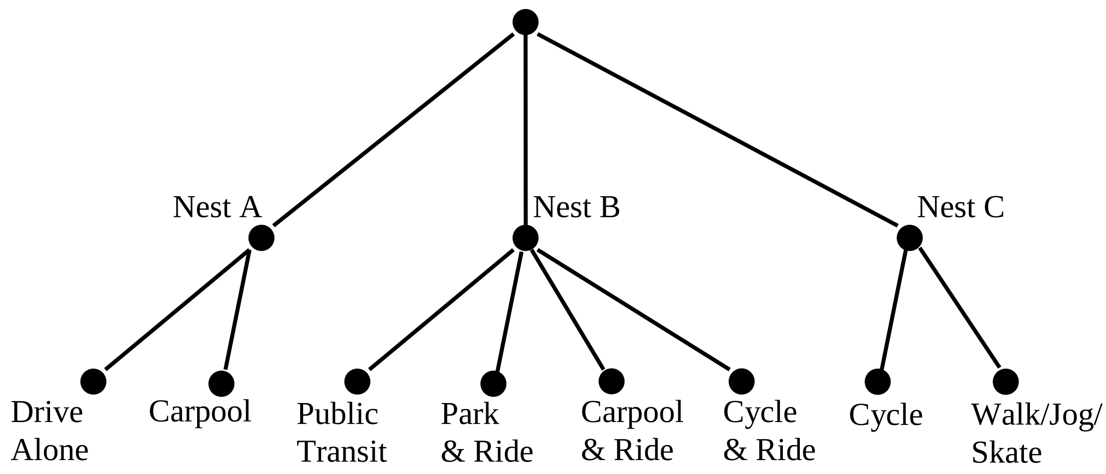


Figure 5.6: Nesting Structure of Model A & B

Model A uses travel time data as reported by the survey respondents. It has already been mentioned in the previous chapter that some of these data were found to be inconsistent with home to work distance (km) due to human error in reporting time. Two models with similar specification (A & B) have been developed to observe the difference in model results. The utility function of Model A consists of the major time and cost variables, and user characteristics such as age, gender, home to work distance etc. It also contains dummy variables like employment status, flexible office hour, compressed work week etc.

Model B has similar specifications but uses the travel time data gathered from Google Maps, where time data are consistent with home to work distance. However, this has created co-linearity problem because for drive alone and carpool modes, travel time and cost are directly proportional with distance.

Therefore, home to work distance variable has been omitted from the model in order to avoid the co-linearity problem.

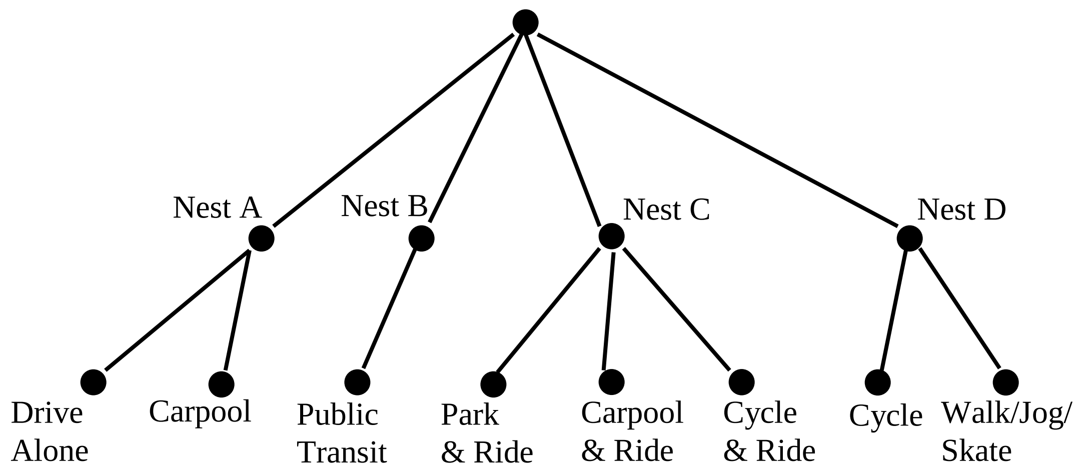


Figure 5.7: Nesting Structure of Model C

Model C has a different nesting structure, as it consists of 4 nests (Figure 5.2). Public transit (bus and/or LRT) mode has been separated from the park & ride options and kept as a single nest. This means, walk access transit mode has been separated from the auto access and bike access transit modes. The other nests are the same as before. Within Model C, two trials have been given - one with fixing the Nest A scale parameter to 1, and the other with fixing Nest B scale parameter to 1. The reason behind fixing the public transit nest parameter to 1 is evident, as the nest contains only one alternate.

Since this is a panel data analysis, an important aspect of model specification is the random parameters. A random coefficient (σ) has been added with the utility functions of each mode, so that the random effect of human behavior and latent variables can be captured through the models. The random coefficient for drive alone has been set to 1, allowing the other sigmas to be estimated. This random parameter will allow for capturing the heterogeneity across the individuals by developing a distribution of mode share probabilities.

The utility equations used in Model B are given below:

1. $V_{DA} = \beta_{IVTT} * (\mathrm{IVTT})_{\mathrm{DA}} + \beta_{P\_Cost} * (\mathrm{P\_Cost})_{\mathrm{DA}} + \beta_{D\_Cost} * (\mathrm{D\_Cost})_{\mathrm{DA}} + \sigma_{DA}$

2. $V_{CP} = ASC_{CP} + \beta_{IVTT} * (\mathrm{IVTT})_{\mathrm{CP}} + \beta_{P\_Cost} * (\mathrm{P\_Cost})_{\mathrm{CP}} + \beta_{D\_Cost} * (\mathrm{D\_Cost})_{\mathrm{CP}} + \beta_{Age\_CP} * (\mathrm{Age}) + \beta_{Gen\_CP} * (\mathrm{Gen}) + \beta_{Emp\_CP} * (\mathrm{Emp}) + \beta_{Flex\_CP} * (\mathrm{Flex}) + \beta_{Comp\_CP} * (\mathrm{Comp}) + \sigma_{CP}$

3. $V_{PT} = ASC_{PT} + \beta_{IVTT} * (\mathrm{IVTT})_{\mathrm{PT}} + \beta_{OVTT} * (\mathrm{OVTT})_{\mathrm{PT}} + \beta_{D\_Cost} * (\mathrm{Fare})_{\mathrm{PT}} + \beta_{Age\_PT} * (\mathrm{Age}) + \beta_{Gen\_PT} * (\mathrm{Gen}) + \beta_{Emp\_PT} * (\mathrm{Emp}) + \beta_{Flex\_PT} * (\mathrm{Flex}) + \beta_{Comp\_PT} * (\mathrm{Comp}) + \sigma_{PT}$

4. $V_{PR} = ASC_{PR} + \beta_{IVTT} * (\mathrm{IVTT})_{\mathrm{PR}} + \beta_{OVTT} * (\mathrm{OVTT})_{\mathrm{PR}} + \beta_{D\_Cost} * (\mathrm{Fare})_{\mathrm{PR}} + \beta_{Age\_PR} * (\mathrm{Age}) + \beta_{Gen\_PR} * (\mathrm{Gen}) + \beta_{Emp\_PR} * (\mathrm{Emp}) + \beta_{Flex\_PR} * (\mathrm{Flex}) + \beta_{Comp\_PR} * (\mathrm{Comp}) + \sigma_{PR}$

5. $V_{CPR} = ASC_{CPR} + \beta_{IVTT} * (\mathrm{IVTT})_{\mathrm{CPR}} + \beta_{OVTT} * (\mathrm{OVTT})_{\mathrm{CPR}} + \beta_{D\_Cost} * (\mathrm{Fare})_{\mathrm{CPR}} + \beta_{Age\_CPR} * (\mathrm{Age}) + \beta_{Gen\_CPR} * (\mathrm{Gen}) + \beta_{Emp\_CPR} * (\mathrm{Emp}) + \beta_{Flex\_CPR} * (\mathrm{Flex}) + \beta_{Comp\_CPR} * (\mathrm{Comp}) + \sigma_{CPR}$

6. $V_{CR} = ASC_{CR} + \beta_{IVTT} * (\mathrm{IVTT})_{\mathrm{CR}} + \beta_{OVTT} * (\mathrm{OVTT})_{\mathrm{CR}} + \beta_{D\_Cost} * (\mathrm{Fare})_{\mathrm{CR}} + \beta_{Age\_CR} * (\mathrm{Age}) + \beta_{Gen\_CR} * (\mathrm{Gen}) + \beta_{Emp\_CR} * (\mathrm{Emp}) + \beta_{Flex\_CR} * (\mathrm{Flex}) + \beta_{Comp\_CR} * (\mathrm{Comp}) + \sigma_{CR}$

7. $V_C = ASC_C + \beta_{IVTT} * (\text{Time})_C + \beta_{Age\_C} * (\text{Age}) + \beta_{Gen\_C} * (\text{Gen}) + \beta_{Emp\_C} * (\text{Emp}) + \beta_{Flex\_C} * (\text{Flex}) + \beta_{Comp\_C} * (\text{Comp}) + \sigma_C$

8. $V_W = ASW_W + \beta_{IVTT} * (\text{Time})_W + \beta_{Age\_W} * (\text{Age}) + \beta_{Gen\_W} * (\text{Gen}) + \beta_{Emp\_W} * (\text{Emp}) + \beta_{Flex\_W} * (\text{Flex}) + \beta_{Womp\_W} * (\text{Comp}) + \sigma_W$

5.4 Model Results

The estimated parameters for each of the models are presented in Table 5.1. It is found for all three models that the generic parameters for the level of service variables are statistically significant (one tail t-value for 95% confidence 1.64 is considered as critical) and have appropriate signs. Model A has the home to work distance as a variable, and the alternate specific parameters corresponding to this are significant for all modes. Parameters for the dummy variables are reasonable in all three models. However, it can be seen that some random parameters have come to be insignificant in Model A, while the other two models yield better result.

Table 5.1: Model Estimation Results

Name	Description	Model A		Model B		Model C	
		Value	T-Stat	Value	T-Stat	Value	T-Stat
ASC_{DA}	Alternate Specific Constants	0	Fixed	0	Fixed	0	Fixed
ASC_{CP}		-2.52	-4.91	-3.05	-19.64	-13.7	-16.93
ASC_{PT}		-1.24	-2.26	3.53	3.23	-4.04	-6.83
ASC_{PR}		-2.35	-3.23	-6.21	-9.21	-18.1	-10.12
ASC_{CPR}		-3.29	-2.31	-29.3	-10.63	-18.0	-7.89
ASC_{CR}		-10.7	-1.33	-23.7	-8.34	-17.9	-8.70
ASC_C		-3.56	-3.9	-6.26	-11.38	-10.0	-11.61
ASC_W		-	-	-10.38	-10.97	-14.2	-12.16

β_{IVTT}	Generic Parameters	-0.0202	-6.61	-0.0605	-10.29	-0.0439	-8.38
β_{OVTT}		-0.0443	-9.8	-0.155	-6.42	-0.0643	-2.99
β_{P_Cost}		-0.036	-1.98	-0.237	-2.84	-0.486	-4.88
β_{D_Cost}		-0.242	-9.01	-0.59	-6.59	-0.262	-3.46
β_{Dist_DA}	Home to Work Distance (km)	0	Fixed	-	-	-	-
β_{DIST_CP}		-0.025	-4.91	-	-	-	-
β_{DIST_PT}		-0.0359	-5.7	-	-	-	-
β_{DIST_PR}		0.0241	2.38	-	-	-	-
β_{DIST_C}		0.0895	4.17	-	-	-	-
β_{DIST_W}		-0.462	-6.72	-	-	-	-
β_{Gen_DA}	Gender Dummy (Female = 1, Male = 0)	0	Fixed	0	Fixed	0	Fixed
β_{Gen_CP}		0.169	2.46	0.796	2.49	0.576	1.93
β_{Gen_PT}		0.541	8.59	1.27	4.01	1.84	6.08
β_{Gen_PR}		1.41	8.19	7.34	8.19	7.94	8.53
β_{Gen_CPR}		0.397	1.91	6.46	6.31	-	-
β_{Gen_CR}		-	-	-4.7	-3.62	-5.45	-4.10
β_{Gen_C}		-1.61	-11.23	-5.3	-11.51	-5.50	-11.89
β_{Gen_W}		0.34	1.92	-4.65	-8.10	-4.25	-7.78
β_{Age1_PR}	Age between 18-24 yrs = 1	-2.33	-2.97	-	-	-	-
β_{Age2_PR}	Age between 25-34 yrs = 1	-1.26	-1.77	-	-	-	-
β_{Age5_C}	Age between 55-64 yrs = 1	-2.13	-2.22	-	-	-	-
β_{Age3_W}	Age between 35-44 yrs = 1	-2.58	-3.03	-	-	-	-
β_{Age4_W}	Age between 45-54 yrs = 1	-1.82	-2.18	-	-	-	-
β_{EMP_DA}	Employment Status (FT = 1, PT = 0)	0	Fixed	0	Fixed	0	Fixed
β_{EMP_PT}		-0.502	-5.31	-2.04	-4.88	-1.34	-3.66
β_{EMP_CPR}		-0.702	-2.37	-	-	-3.27	-3.23
β_{EMP_CR}		-	-	-	-	-5.30	-4.02
β_{EMP_C}		-0.47	-2.54	-	-	-	-
β_{EMP_W}		0.449	1.82	1.29	2.17	1.71	2.45
β_{Flex_DA}	Flexible Office Hour (Yes =1)	0	Fixed	0	Fixed	0	Fixed
β_{Flex_CP}		0.465	6.65	1.18	3.66	1.34	4.45
β_{Flex_PT}		-	-	-1.27	-3.67	-0.803	-2.83
β_{Flex_PR}		-	-	-3.42	-4.62	-2.96	-3.93
β_{Flex_CPR}		-	-	3.35	3.14	-	-
β_{Flex_CR}		1.59	3.32	5.44	3.27	3.14	2.67
β_{Flex_C}		-0.58	-4.92	-1.78	-3.97	-1.41	-3.40

β_{Flex_W}		0.646	3.56	0.866	1.52	2.20	4.18
β_{Comp_DA}	Compressed Work Week (Yes = 1)	0	Fixed	0	Fixed	0	Fixed
β_{Comp_CP}		0.657	8.47	2.10	6.08	2.05	6.33
β_{Comp_PT}		0.516	7.39	-	-	-0.462	-1.54
β_{Comp_PR}		-	-	1.87	2.59	2.26	3.19
β_{Comp_CPR}		-	-	-2.28	-1.90	1.73	2.14
β_{Comp_CR}		-	-	-	-	-6.36	-5.37
β_{Comp_W}		0.644	3.49	1.11	1.92	-	-
σ_{DA}	Random Parameters	1.0	Fixed	1.0	Fixed	1.0	Fixed
σ_{CP}		-0.129	-4.04	-10.1	-22.02	-9.72	-19.80
σ_{PT}		-	-	-6.62	-22.91	-6.90	-23.13
σ_{PR}		0.165	2.33	-11.8	-12.30	-10.6	-13.63
σ_{CPR}		0.341	3.42	-19.9	-12.29	-13.7	-9.84
σ_{CR}		-	-	11.2	10.14	13.1	10.21
σ_{C}		-	-	8.49	19.99	8.73	18.09
σ_{W}		-	-	15.0	16.31	13.0	15.98
φ_A	Scale Parameter for Nest A	1.0	Fixed	1.0	Fixed	1.0	-
φ_B	Scale Parameter for Nest B	0.702	14.89	0.307	11.69	1.0	Fixed
φ_C	Scale Parameter for Nest C	0.293	11.64	0.221	9.92	0.173	11.17
φ_D	Scale Parameter for Nest D	-	-	-	-	0.266	9.29
Number of Individuals		2702		2696		2696	
Number of Observations		13459		10767		10767	
Number of Estimated Parameters		91		48		49	
Null log-likelihood		-21429.464		-17145.739		-17145.739	
Initial log-likelihood		-22318.422		-13530.169		-13530.169	
Final log-likelihood at convergence		-13567.549		-5207.540		-5201.890	
ρ-square value		0.367		0.696		0.697	
Adjusted ρ-square value		0.363		0.693		0.694	

Notes:

DA = Drive Alone

CP = Carpool

PR = Park & Ride

CPR = Carpool & Ride

CR = Cycle & Ride

C = Cycle

W = Walk, Jog, Inline Skate

IVTT = In-Vehicle Travel Time
OVTT = Out-of-Vehicle Travel (walk time + wait time)
P_Cost = Cost of Parking
D_Cost = Cost of Driving / Transit Fare

The model goodness-of-fit parameter (Adjusted ρ-square value) for model A is satisfactory but the values for models B & C are better. Final log likelihood values at convergence are also reasonable in all cases. The nest parameters for Model A and B are satisfactory, carrying values less than 1.0. However, it is found that the nest parameter in model C has an error. In model C, the nest B parameter has been fixed to 1, and the other nest values are expected to be within a range between 0 and 1. However, the value of nest A has become 1.0 with insignificant t-stat. A similar result was obtained for Nest B when an alternate model was tried, fixing the nest A parameter value to 1. This implies that the nesting structure used in Model C is not appropriate, and so model A and B have been chosen for further validation.

5.5 Model Validation

Validation of models A and B has been performed using an Excel spreadsheet. The estimated parameters have been plugged into the utility equations to compute the utility values for each mode. Then the nested logit equations (described in chapter 3) have been used to compute the marginal probability of each alternate. Since this is a panel dataset, a series of random numbers has been generated to develop a series of mode share probabilities for each individual. Then the average values are taken and summed to get the aggregate level mode share percentages.

The aggregate level mode share percentages are compared with the observed mode share percentages to check the compatibility of the model. This approach is also useful for forecasting purposes.

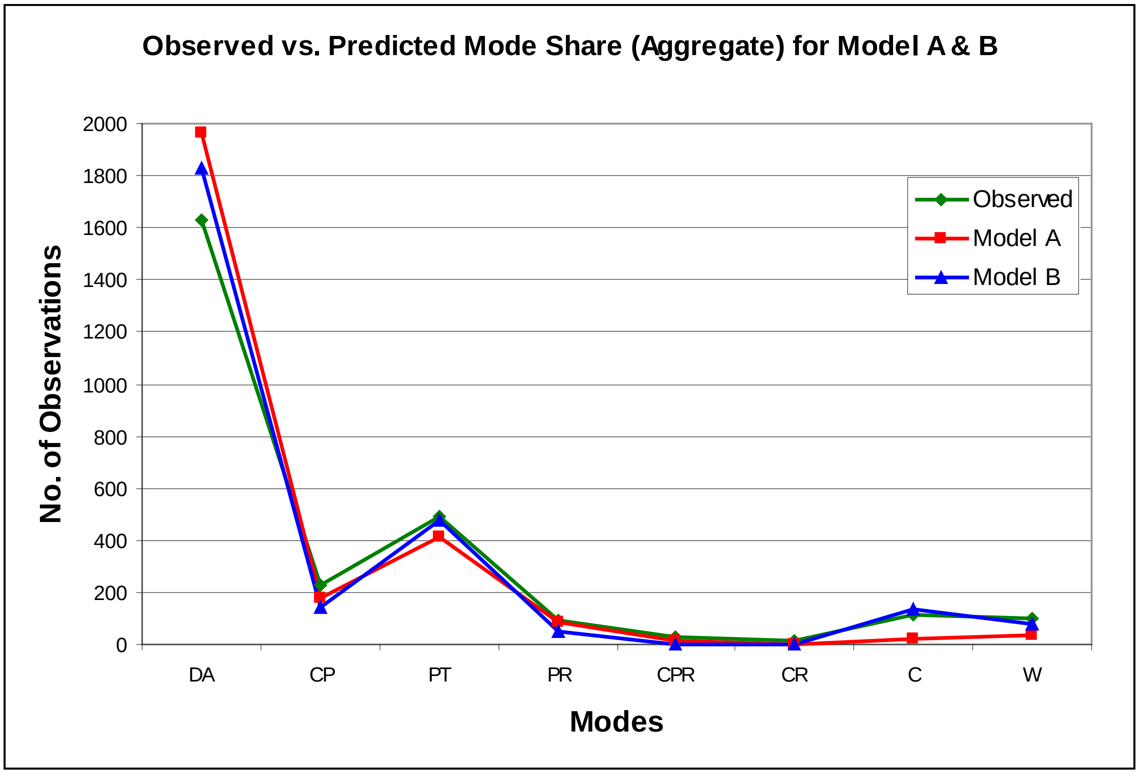


Figure 5.8: Observed vs. Predicted Mode Shares (Aggregate level)

From Figure 5.8, it can be seen that both models A and B yield satisfactory prediction of aggregate level mode share percentages, with a slight over-prediction of drive alone probability. This may have happened because the validation has been performed using separate data with much smaller number of observations. However, from the figure it is evident that Model B yields better results than Model A, and hence model B has been chosen as the final model. Model B is further used for elasticity calculations of variables.

5.6 Elasticity Measures

This section presents the elasticity values of some important variables. The elasticity equations described in Chapter 2 have been used to determine the elasticity values for variables estimated in Model B.

First, the direct elasticity of the level of service variables have been calculated. As can be seen in Table 5.2, elasticity of time and cost for drive alone are not very high. This is expected in the case of a predominant mode such as drive alone. All other alternates have reasonably high elasticity to In-vehicle travel time (IVTT) (all less than -1). The out-of-vehicle travel time (OVTT) elasticity of the transit oriented alternates have come to be reasonably high which refers that these modes are sensitive to travel time. The elasticity values of drive alone parking cost did not come as high as expected. This may have been because the survey indicates that 35.2% of respondents reported parking to be free, while another 32.5% respondent reported that their parking cost is paid for by their employer. So, the sensitivity of mode share for changes in parking price is not very high. For direct elasticity of transit fare, all the transit oriented alternates are found to be sensitive.

Table 5.2: Direct Elasticity of LOS variables

Modes\LOS Variables	IVTT	OVTT	Parking Cost	Driving Cost / Fare
Drive Alone	-0.339	-	-0.215	-0.383
Carpool	-1.314	-	-0.334	-0.567
Public Transit	-1.827	-2.407	-	-1.845
Park & Ride	-2.754	-2.628	-	-6.597
Carpool & Ride	-3.793	-2.676	-	-6.157
Cycle & Ride	-1.746	-2.667	-	-5.295
Cycle	-3.218	-	-	-

However, submodels were developed to check the significance of the parking cost on auto modes by adding “free parking” as dummy variable. The parameter came as statistically significant, which implies that if there was no free parking or the employers did not pay the parking charges, cost of parking could have an elastic effect on auto modes.

In this study, one particular interest related to elasticity measures was to make use of panel analysis to develop a distribution of elasticity across individuals. This analysis has been performed for the two major modes of transport used by the people. Figure 5.9 shows the distribution of direct elasticity of drive alone time and cost variables. To obtain this, a series of elasticity of each LOS variable was generated for each individual. Then the elasticity values were averaged for each individual and the distribution across the sample was plotted. It is seen that all three elasticities generate very similar distribution patterns, which reach a peak at a value around 0.2 (location parameter) and gradually declines after values less than -0.5.

For the level of service variables of Public Transit, as it can be seen in Figure 5.10, no particular trend is observed for public transit in-vehicle travel time. It is spread over an elasticity range from 0 to -3. But, out-of-vehicle travel time and transit fare shows distinct peaks close to values -0.3 and -0.6 respectively. This implies that transit OVTT and fare affect the use of public transit with higher consistencies.

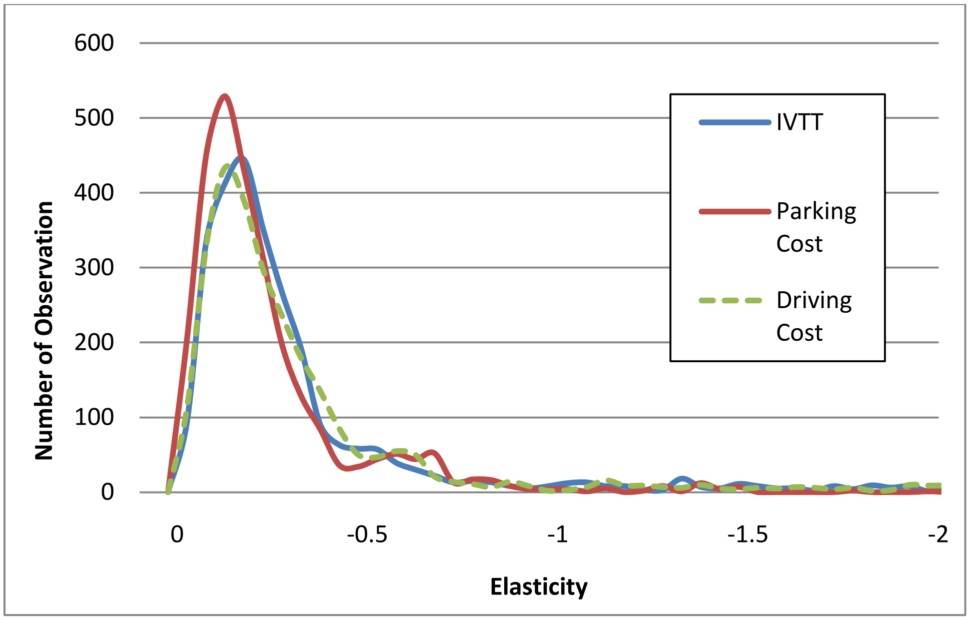


Figure 5.9: Distribution of Elasticity for Drive Alone LOS Variables

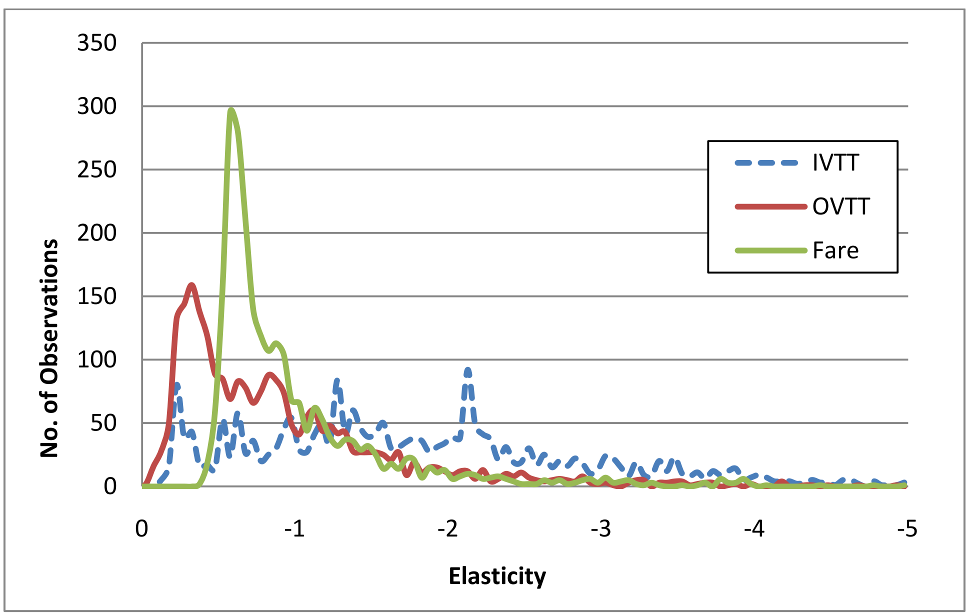


Figure 5.10: Distribution of Elasticity for Public Transit LOS Variables

Table 5.3 presents the cross elasticity values for drive alone and public transit variables. It is observed that the cross elasticity values are not very high, which suggests that the mode shifting process in the study area is not governed by these. However, for public transit variables, alternates belonging to the same nest are more sensitive to changes as indicated by the high cross-elasticity values.

Table 5.3: Cross Elasticity of Drive Alone & Public Transit LOS

Change in	Effects on	IVTT	OVTT	Parking Cost	Driving Cost / Fare
Drive Alone	All other alts.	0.748	-	0.494	0.818
Public Transit	Alts. in other Nests	0.078	0.088	-	0.087
Public Transit	Alts. in same nest	2.916	3.817	-	3.208

Table 5.4: Elasticity of Dummy Variables

Modes \ LOS Variables	Flexible Office Hour	Compressed Work Week
Drive Alone	-0.159	0.098
Carpool	-1.053	-1.168
Public Transit	0.374	0.121
Park & Ride	2.298	-1.130
Carpool & Ride	-1.276	0.474
Cycle & Ride	-1.996	0.017
Cycle	1.299	0.089
Walk/Jog/Skate	-0.697	-0.458

Finally, elasticity of dummy variables such as flexible office hour and compressed work week, have been determined. These elasticities can be particularly important in TDM perspective. The elasticity values in Table 5.4 suggest that the transit oriented alternates are moderately sensitive to flexible office hour option, while compressed work week option does not have a significant effect on mode choice.

5.7 Model Results with Latent Variables

Taking Model B as the base model, the latent variables have then been incorporated in the model. In each model, the three highest reported variables have been added to the model and checked for significance. Since it is seen from Figure 4.3 that mode choice does not significantly vary over time, separate models for workdays have been developed. It is found that all the models yield similar results. Table 5.5 presents the estimation results for models using Wednesday and Friday data. It can be seen that most of the latent variables are statistically significant. However, the parameter signs are not consistent. This implies that the straight forward options for stated preference questions do not always reveal the facts about behavioral or attitudinal aspect of mode choice phenomena. Rather the scaling of preference against stated preference questions is more appropriate to explore the reasoning behind one’s mode choice behavior, which has recently been proposed by some researchers (Domarchi et al., 2008).

Table 5.5: Estimation Results for Model with Latent Variables

Name	Description	Model D Wednesday		Model E Friday	
		Value	T-Stat	Value	T-Stat
ASC_{DA}	Alternate Specific Constants	0	Fixed	0	Fixed
ASC_{CP}		-2.14	-6.43	-1.98	-6.14
ASC_{PT}		-0.485	-1.63	-	-
ASC_{PR}		-1.57	-2.89	-2.79	-3.32
ASC_{CPR}		-1.45	-2.34	-1.55	-2.22
ASC_{CR}		-3.67	-2.80	-4.11	-2.72
ASC_{C}		-4.70	-6.59	-4.12	-5.48
ASC_{W}		-	-	-	-
β_{IVTT}	IVTT for PT, PR, CPR, CR	-0.0057	-1.31	-	-
β_{OVTT}	OVTT for PT, PR, CPR, CR	-0.0348	-4.10	-0.0374	-4.19
B_{D_Cost1}	Driving Cost for DA, CP	-0.0549	-1.83	-0.0471	-1.52
β_{D_Cost2}	Cost (Transit Fare + Driving cost) of PT, PR, CPR, CR	-0.176	-3.69	-0..218	-4.13
β_{DIST_C}	Home to Work Distance (km) for Cycling	-	-	0.453	6.94
β_{DIST_W}	Home to Work Distance (km)	-0.476	-3.29	-0.482	-3.14
β_{Gen_DA}	Gender Dummy (Female = 1, Male = 0)	0	Fixed	0	Fixed
β_{Gen_CP}		-	-	-	-
β_{Gen_PT}		0.690	5.88	0.727	5.87
β_{Gen_PR}		1.28	4.41	1.46	4.19
β_{Gen_C}		-1.22	-3.47	-1.46	-3.52
β_{EMP_DA}	Employment Status (FT = 1, PT = 0)	0	Fixed	0	Fixed
β_{EMP_PT}		-0.452	-2.24	-0.485	-2.31
β_{EMP_PR}		-0.803	-2.04	-	-
β_{EMP_CPR}		-1.19	-2.32	-1.28	-2.17
β_{EMP_C}		-1.39	-2.79	-1.50	-2.89
β_{Flex_DA}	Flexible Office Hour (Yes =1)	0	Fixed	0	Fixed
β_{Flex_CP}		0.418	2.77	0.418	2.60
β_{Flex_W}		-	-	0.795	2.11
β_{Comp_DA}	Compressed Work Week (Yes = 1)	0	Fixed	0	Fixed
β_{Comp_CP}		0.606	3.75	0.689	3.96

β_{Comp_PT}		0.502	4.04	0.568	4.31
β_{Comp_PR}		0.422	1.57	-	-
Coefficients of Latent Variables used as dummy					
β_{01_CP}	Carpool	-1.32	-6.40	-1.20	-5.62
β_{03_CP}		-0.161	-0.79	-	-
β_{05_CP}		-0.282	-1.17	-	-
β_{03_PT}	Public Transit	0.110	0.88	-	-
β_{01_PT}		0.587	4.75	0.598	4.58
β_{04_PT}		-0.588	-3.92	-0.618	-3.85
β_{05_C}	Cycling	0.468	1.38	-	-
β_{013_C}		-1.88	-2.44	-2.00	-2.02
β_{08_C}		-0.833	-1.89	-1.26	-2.29
β_{09_W}	Walk/Jog/Inline Skate	-1.80	-2.76	-1.86	-2.56
β_{04_W}		-1.35	-2.08	-1.52	-1.97
B_{11_W}		-2.23	-2.88	-2.05	-2.76
φ_A	Scale Parameter for Nest A	1.0	Fixed	1.0	Fixed
φ_B	Scale Parameter for Nest B	0.998	6.07	0.860	6.16
φ_C	Scale Parameter for Nest C	0.294	7.13	0.293	6.76
Number of Individuals		2444		2280	
Number of Observations		2444		2280	
Number of Estimated Parameters		55		55	
Null log-likelihood		-3780.82		-3534.20	
Initial log-likelihood		-3780.82		-3534.20	
Final log-likelihood at convergence		-2497.63		-2278.75	
Likelihood ratio test		2566.38		2510.91	
ρ-square value		0.339		0.355	
Adjusted ρ-square value		0.325		0.340	

Notes: Stated Preference questions and options for latent variables:

Under what circumstances would you consider CARPOOLING more often?

- I could find compatible carpool partner β_{01_CP}
- I could still drive alone if I needed to β_{03_CP}
- I had a guaranteed ride home in case of emergency or unscheduled overtime β_{05_CP}

Under what circumstances would you consider using PUBLIC TRANSIT more often?

- Transit service was faster, more frequent or more reliable β_{03_PT}
- It cost less β_{01_PT}
- Fewer transfer β_{04_PT}

Under what circumstances would you consider CYCLING to work more often?

- The cycling routes between my home and work were safer or more convenient β_{05_C}
- I lived closer to work β_{013_C}
- I felt safer cycling in traffic β_{08_C}

Under what circumstances would you consider WALKING, JOGGING or INLINE SKATING to work more often?

- I lived closer to work β_{09_W}
- The routes between my home and work were safer or more convenient β_{04_W}
- I didn't have to use my personal vehicle for work purposes β_{11_W}

The above table shows the estimation results of nested logit models with latent variables. It is found that the model captured the effect of attitude and/or behavior on the mode choice process, which gives an insight of mode choice preference. This finding can be very important in terms of implementing TDM policies. Yet, a hybrid choice model has also been developed using the same dataset, in order to explore a better understanding of attitudinal variables. Habib et al. (2010) showed that hybrid choice model using the stated preference questions in the survey as latent captivity parameters can significantly improve the model parameters.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary of Results

This study analyzes a revealed and stated preference dataset for mode choice with an attempt to reveal the effect of various factors on mode choice. The specific objectives of the study were to analyze mode choice data in terms of instrumental and latent variables and develop a model to capture the actual choice behavior.

The results of the study suggest that mode choice is a complex behavioral process and cannot be fully described through the use of level of service variables only. The process depends on alternate specific level of service and user specific attributes, as well as latent variables reflecting an individual's choice preference. The model results in this study bolster this hypothesis, as most of the above-mentioned variables have come to be statistically significant. The models also incorporate some TDM programs as dummy variables as an attempt to investigate their effectiveness. It appears from the results that flexible office hours and compressed workweek programs can effectively affect the mode choice of commuter trips.

This study also analyzed the elasticity of LOS variables, which are very important in terms of policy evaluation. It is found that drive alone is not sensitive to travel time and cost. For most of the model specifications, the drive alone time and cost parameters have come to be positive, which is not intuitive. This implies that car use is not significantly dependent on level of service attributes, rather it is governed by behavioral and attitudinal factors. It is shown that transit oriented modes are highly sensitive to in-vehicle travel time, out-of-vehicle travel time and fare. This gives useful insight of the choice process and this finding should be considered for effective improvement in transit modal shares.

One of the important findings of this study is about correctness of revealed preference data. It has been found that the reported travel time data or home to work distance data can sometimes be anomalous. In order to check such anomaly, two identical models have been developed with same specification: one using travel time data from the survey, the other using travel time data from Google. It is found that Google data has consistency across the sample and thus results in a better model.

Although this study has made use of panel data through incorporating a random coefficient in the utility function, results suggest that these random coefficients do not have significant impact on the mode choice probabilities. The reason may be that the dataset has only 7 days of observation which is not sufficient to capture the mode choice variation over time.

6.2 Policy Implications

- The effectiveness of TDM programs such as flexible office hours and a compressed workweek have been analyzed in this study. The findings suggest that flexible office hour have significant impact on carpool, park & ride, carpool & ride, cycle & ride, and cycling. On the other hand, a compressed workweek has significant effect on carpool and park & ride modes. Therefore, it is understood that in order to influence people's mode choice towards these alternative sustainable modes, the employers can implement these two TDM programs.
- Apparently, it is anticipated that increase in parking cost or operating cost for drive alone can reduce the use of car. But from the analysis performed in this study, the elasticity of these variables has come to be less than 1.0, which implies that increased cost of car alone is not a significant factor to reduce car use, rather these should be implemented jointly with other programs that impose reduced car use as well as encourage increased use of alternate modes.
- For successful implementation of TDM measures and influencing people's commuter mode choice towards more sustainable alternates, the effect of latent or behavioral factors should be considered. For example, in response to the stated preference question "under what circumstances would you consider carpooling more often?", the parameter for option "if I could find compatible carpool partner" has come to be statistically significant with a negative sign. This implies that getting a compatible carpool partner would increase the utility of carpool for the individual. Therefore, to increase carpool modal

share, carpool information sharing programs can be developed and implemented by the employers.

6.3 Limitations of the Study

The limitations of this study can be summarized as following:

- The study has been limited to developing multinomial and nested logit models only. More complex cross-nested logit or mixed logit models could be developed to better describe the choice behavior.
- Limitations associated with dataset are that the survey did not collect information regarding individuals income, car ownership etc. However, these variables have important effects on mode choice behavior.
- Many recent research works suggest that joint models combining departure time and mode choice are more accurate than traditional mode choice models. Although the dataset has the departure time information, joint models have not been attempted due to complexity.

6.4 Recommendations for Future Works

It is found from the results of this study that models containing both instrumental and latent variables can help us understanding the mode choice behavior of individuals. Although the models developed here could capture many important factors that influence this behavior, further research may be performed to develop more complex models using advanced computational technology which may help better understand the mode choice process. Attempting alternate approach for latent variable modeling can also be useful to understand the behavioral factors affecting mode choice.

References

Ben-Akiva, M., Lerman, S. *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, 1985.

Ben-Akiva, M. Hybrid Choice Models: Progress and Challenges. Marketing Letters 13:3, Kluwer Academic Publishers, pp 163-175, 2002.

Bierlaire, M. Estimation of Discrete Choice Models with BIOGEME 1.8, 2009. http://transp-or.epfl.ch/page63023.html. Accessed: June 2009.

Blaschuk, C.R., Brownlee, A.T., Hunt, J.D. Estimating Congestion Costs Using a Transportation Demand Model of Edmonton, Canada. 11th TRB National Transportation Planning Applications Conference, 2007.

Bolduc, D. A Practical Technique to Estimate Multinomial Probit Models in Transportation. Transportation Research Part B, vol. 33, pp 63-79. TRB 1999.

Buliung, R.N., Kalina, S., Catherine, H., Lanyon, R. Driving Factors behind Successful Carpool Formation and Use. Transportation Research Record Vol. 2118, pp 31-38. TRB 2009.

Canadian Automobile Association. Driving Costs, 2008 Edition (updated June 2008). http://www.caa.ca/documents/DrivingCostsBrochure-july08-eng-WEB.pdf. Accessed in July 2009.

Cherchi, E. Modelling individual preferences, State of the art, recent advances and future directions. Methodological development in activity-travel demand analysis workshop, 12th international conference on travel behavior research, 2009.

City of Calgary Regional Transport Model. Vol. 1: Personal Travel Model Description. Transportation Department, City of Calgary, 2006.

City of Edmonton Postal Code Zoning Map, Canada Post website. http://www.canadapost.ca/cpc2/addrm/hh/maps/FSA/AB04.pdf. Accessed: Jan 2010.

City of Edmonton website. Edmonton Transit System (ETS), Route Schedules. http://www.edmonton.ca/transportation/edmonton-transit-system-ets.aspx. Accessed: January 2010.

City of Edmonton, Transportation website. Bicycle Parking at ETS Facilities. http://www.edmonton.ca/transportation/ets/bicycle-parking-at-ets-facilities.aspx. Accessed: January 2010.

Dargay, J.M. Choice of Commuting Mode: A Panel Data Analysis. 8th NECTAR Conference, 2005.

Day, N. Modeling Mode Choice and Trip Timing for Work Trip. M.Sc. thesis. Department of Civil Engineering, University of Toronto. 2008.

Dill, J., Voros, K. Factors Affecting Bicycling Demand: Initial Survey Findings from the Portland, Oregon, Region. Transportation Research Record No. 2031, pp 9-17. TRB 2007.

Domarchi, C., Tudela, A., Gonzalez, A. Effect of attitude, habit and affective appraisal on mode choice: an application to university workers. Transportation, vol. 35, no. 5, 2008.

Dunne, J.P. Elasticity Measures and Disaggregate Choice Models. Journal of Transport Economics and Policy, vol. 18, No. 2, pp 189-197, 1984.

Forinash, C.V., Koppelman, F.S. Application and Interpretation of Nested Logit Models of Intercity Mode Choice. Transportation Research Record, No. 1413, pp 98, TRB 1993.

FoundLocally- FL (2010), http://foundlocally.com/Edmonton/Hr/Jobs-Employers.htm Accessed in March 2010.

Gardner, B., Abraham, C. What Drives Car Use? A Grounded Theory Analysis of Commuters' Reason for Driving. Transportation Research Part F 10, pp 187-200. 2007.

Garling, T., Schuitema, G. Travel Demand Management targeting reduced private car use: effectiveness, public acceptability and policy feasibility. Journal of social issues, vol. 63, no. 01, pp 139-153. 2007.

Google Maps, Google Inc. http://maps.google.ca/maps?hl=en&tab=wl. Accessed: June through December, 2009.

Habib, K.M.N., Day, N., Miller, E.J. An investigation of commuting trip timing and mode choice in the Greater Toronto Area: Application of a joint discrete-continuous model. Transportation Research Part A, vol. 43, pp 639-653. 2009.

Habib, K.M.N., Kattan, L., Islam, M.T. Why do the People Use Transit: A Critical Investigation on Calgary Transit Users Using Latent Discrete Choice Modeling. 88th Annual Meeting of Transportation Research Board, Washington DC. 2009.

Habib, K.M.N., Zaman, H. Effects of Incorporating Latent and Attitudinal Information in Mode Choice Model. Accepted for presentation at the 45th CTRF conference, 2010.

Hensher, D.A. Climate Change, Enhanced Greenhouse Gas Emissions and Passenger Transport – What can we do to make a difference? Transportation Research Part D, volume 13, pp 95-111. TRB 2008.

Hensher, D., Greene, W. The Mixed Logit Model: The State of Practice. Transportation, volume 30, number 2, pp 133-176. Springer Netherlands, 2003.

Hsiao, C. *Analysis of Panel Data*. 2nd Edition. Cambridge University Press, 2003.

Hunt, J.D. A Stated Preference Analysis of Sensitivities to Elements of Transportation and Urban Form. Presented at TRB conference, 2001.

Hunt, J.D., Brownlee, A.T. Developing an Auto Ownership Model for Edmonton. Presented at Annual Conference of Transportation Association of Canada, 2005.

Ibanez, J.N. Consistency of Nested Logit Models with Utility Maximization. European Transport Conference. 2006.

Johansson, M.V., Heldt, T., Johansson, P. The Effects of Attitudes and Personality Traits on Mode Choice. Transportation Research Part A 40 (2006), pp 507-525.

Kitamura, R. Panel Analysis in Transportation Planning: An Overview. Transportation Research Part A, Vol. 24A, No. 6, pp 401-415, 1990.

Ko, J., Cho, Y., Choi, J., Kim, T. Evaluation of Travel Demand Management Strategies Using Importance-Performance Analysis. Transportation Research Record No. 2118, pp 67-74. TRB 2009.

Ko, J., Cho, Y. Voluntary Program to Reduce Car Use. Transportation Research Record No. 2118, pp 1-7. TRB 2009.

Koppelman, F.S., Bhat, C. *A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models*. U.S. Department of Transportation, Federal Transit Administration, 2006.

Koppelman, F.S., Sethi, V. Closed Form Discrete Choice Models. *Handbook of Transport Modelling*. 1st Edition. Elsevier, Oxford, 2000.

Litman, T. Transportation Elasticities: How Prices and Other Factors Affect Travel Behavior. Victoria Transport Policy Institute, 2009. www.vtpi.org. Accessed: January 2010.

Litman, T. The Online TDM Encyclopedia: mobility management information gateway. Transport Policy vol. 10, pp 245-249, 2003.

Lucas, K. Actual and Perceived Car Dependence: Likely Implications of Enforced Reductions in Car use for Livelihood, Lifestyles, and Well-Being. Transportation Research Record No. 2118, pp 8-15. TRB 2009.

McFadden, D. Conditional Logit Analysis of Qualitative Choice Behavior. Frontiers of Econometrics, ed. by P. Zarembka, New York: Academic Press, pp 105-142, 1974.

McFadden, D. Economic Choice. Nobel Prize Lecture, 2000. Economic Sciences 2000.

McFadden, D., Train, K. Mixed MNL Models of Discrete Response. Journal of Applied Econometrics vol. 15, pp 447-470, 2000.

Ortuzar, J.D. On the Development of the Nested Logit Model. Transportation Research Part B, vol. 35, issue 2, pp 213-216, 2001.

Rajamani, J., Bhat, C.R., Handy, S., Knaap, G., Song, Y. Assessing Impact of Urban Form Measures on Nonwork Trip Mode Choice After Controlling for Demographic and Level-of-Service Effects. Transportation Research Record, Vol. 1831, pp 158-165, 2003.

Saleh, W., Sammer, G. *Travel Demand Management and Road User Pricing*. Ashgate Publishing Limited, 2009.

Strathcona County website. Transit Schedules & Routes, Service to Edmonton. http://www.strathcona.ab.ca/departments/Transit/Schedules_Routes/service-to-edmonton.aspx. Accessed: September 2009.

Summary Report on Weekday Travel by Residents on the Edmonton Region. Household Travel Survey, Alberta Infrastructure and Transportation, 2005. City of Edmonton website. Accessed: December 2009.

Sundo, M.B., Fulii, S. The effects of a compressed work week on commuters' daily activity pattern. Transportation Research Part A, vol. 39, pp 835-848. 2005.

The Way We Move, Transportation Master Plan 2009. City of Edmonton website. http://www.edmonton.ca/city_government/documents/TransportationMasterPlan.pdf. Accessed February 2010.

Thomas, G. and James, D. (2006). Re-inventing grounded theory: some questions about theory, ground and discovery, British Educational Research Journal, 32, 6, 767–795.

Train, K.E. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2002.

Travel Forecasting Guidelines, Department of Transportation, State of California, 1992.

Turnbull, K.F., Evans, J.E., Levinson, H.S. Traveler Response to Transportation System Changes. Chapter 3: Park-and-Ride/Pool. TCRP Report 95, Transportation Research Board. 2004.

Ungemah, D., Casey, D. Transportation Demand Management Benchmark: Results from 2008 TDM Program Survey. Transportation Research Record No. 2118, pp 55-66. TRB 2009.

Victoria Transport Policy Institute. Alternative Work Schedules: Flextime, Compressed Work Week, Staggered Shifts. TDM Encyclopedia, 2008. www.vtpi.org. Accessed: January 2010.

Waerden P., Borgers, A., Timmermans, H. Consumer Response to Introduction of Paid Parking at a Regional Shopping Center. Transportation Research Record No. 2118, pp 16-23, TRB 2009.

Walker, J.L. Discrete Choice Models in Transportation: Proof of Concept. 84th Annual Meeting of Transportation Research Board, workshop 111. Washington DC, 2005.

Washbrook, K., Haider, W., Jaccard, M. (2006). Estimating commuter mode choice: A discrete choice analysis of the impact of road pricing and parking charges. Transportation, Vol. 33, pp 621-639.

Weinberger, R., Seaman, M., Johnson, C. Residential Off-Street Parking Impacts on Car ownership, Vehicle Miles Traveled, and Related Carbon Emissions. Transportation Research Record No. 2118, pp 24-30, TRB 2009.

Winters, P.L Transportation demand management. 79th annual meeting of the transportation research board, Washington D.C. 2000.

Zureiqat, H.M. Fare Policy Analysis for Public Transport: A Discrete-Continuous Modeling Approach using Panel Data. M.Sc. thesis, MIT, 2008.

Zhou, L., Winters, P. Empirical Analysis of Compressed Workweek Choices. Transportation Research Record No. 2046, pp 61-67. TRB 2008.

Appendices

Appendix A
Workplace Commuter Survey Questionnaire

Section 1: General Information

1. Which City of Edmonton Department do you work for?
2. What are the first three digits in your home postal code?
3. What are the first three digits in your work location postal codes?
4. Gender
5. Age category
6. Employment status
7. What time do you typically arrive at work?
8. What time do you typically leave work?
9. Do you have the flexibility to vary your arrival and departure hours?
10. Do you work a compressed work week? (In a compressed work week, employees work fewer but longer days)
11. Approximately how many kilometers is it from your home to work (one way)?

Section 2: Means of Commuting

12. How did you commute to work during your last full week of work?
13. In the last 12 months, how often have you used Public Transit to commute to work?
14. Under what circumstances would you consider using Public Transit more often?
    - It cost less
    - Increased cost of driving (fuel, parking etc.)
    - Transit service was faster, more frequent or more reliable
    - Fewer transfers
    - I could buy tickets and passes at my workplace
    - I could get transit information and advice at my workplace
    - The bus stop / LRT station near home/work was closer or easier to walk to
    - The bus stop near home/work had better shelter, seating or lighting

- My bike could be taken onto the bus/LRT at all times and on all routes
- There was public transit service where I live
- I had a guaranteed ride home in case of emergency or unscheduled overtime
- I didn't have to use my personal vehicle for work purposes
- I worked more regular hours
- I already take public transit as much as possible
- I would not consider taking public transit to work
- Other

15. In the last 12 months, how often have you Carpooled to commute to work, either as a driver or a passenger?
16. Under what circumstances would you consider Carpooling more often?
    - I could find compatible carpool partners
    - I could carpool temporarily or occasionally
    - I could still drive alone when I needed to
    - Carpool parking was available that was more convenient or cheaper than regular parking
    - I had a guaranteed ride home in case of emergency or unscheduled overtime
    - I didn't have to use my personal vehicle for work purposes
    - I worked more regular hours
    - I already carpool as much as possible
    - I would not consider carpooling to work
    - Other
17. In the last 12 months, how often have you Cycled to work?
18. Under what circumstances would you consider Cycling to work more often?
    - Shower, changing and locker facilities were provided
    - Shower, changing and locker facilities were improved
    - Bicycle parking was more secure and sheltered
    - Bicycle parking was provided or more spaces were available
    - The cycling routes between my home and work were safer or more convenient
    - I had help finding safe, enjoyable cycling routes
    - I could find someone else or a group to cycle with
    - I felt safer cycling in traffic
    - The dress code was more relaxed
    - I had a guaranteed ride home in case of emergency or unscheduled overtime
    - I didn't have to use my personal vehicle for work purposes
    - I worked more flexible hours

- I lived closer to work
- I already cycle as much as possible
- I would not consider cycling to work
- Other

19. In the last 12 months, how often have you Walked, Jogged or Inline Skated to work?
20. Under what circumstances would you consider walking, jogging or inline skating to work more often?
    - Shower, changing and locker facilities were provided
    - Shower, changing and locker facilities were improved
    - The dress code was more relaxed
    - The routes between my home and work were safer or more convenient
    - I had help finding safe, enjoyable walking, jogging or inline skating routes
    - I could find someone or a group to walk, jog or inline skate with
    - I felt safer on pathways
    - I had a guaranteed ride home in case of emergency or unscheduled overtime
    - I lived closer to work
    - I worked more flexible work hours
    - I didn't have to use my personal vehicle for work purposes
    - I already walk, jog or inline skate to work as much as possible
    - I would not consider walking, jogging or inline skating to work
    - Other
21. In the last 12 months, how often have you Teleworked?
22. Under what circumstances would you consider Teleworking more often?
    - My employer permitted me to telework
    - My job was more suitable for teleworking
    - I felt personally well-suited to telework (personality, working style, etc.)
    - I had better office space or equipment at home
    - I had a faster connection to the office network from home
    - I had full computer/email access from home
    - My workplace had drop-in office space for teleworkers
    - I felt comfortable working away from my colleagues
    - There were personal financial incentives
    - I already telework full-time
    - I already telework as much as possible
    - I would never consider teleworking

- Other

Section 3: Commuting Details

23. What is your typical one-way travel time to work when Driving Alone?
24. On a daily basis what is the cost of your parking space?
25. If there is a cost to your parking space, is it paid for by your employer?
26. Why do you Drive Alone to work?
    - Vehicle needed for work
    - Vehicle needed to transport family, shopping or errands
    - Vehicle needed for health or disability reasons
    - Long or irregular work hours
    - Free or inexpensive parking
    - Fast travel time
    - Safety or security
    - Convenience, comfort or enjoyment
    - I can't find anyone to Carpool with
    - I live too far away from work to Cycle, Walk, Jog or Inline Skate
    - No Public Transit where I live
    - Weather
    - Other options are not available or feasible
    - Other
27. What is your typical one-way travel time to work when you participate in a Carpool?
28. How many people are typically in your carpool?
29. On a daily basis what is the cost of your Carpool's parking space?
30. If there is a cost to your parking space, is it paid for by your employer? If so, how much?
31. Why do you commute in a Carpool?
    - Cost savings
    - Safety or security
    - Convenience, comfort or enjoyment
    - Can use travel time productively
    - I was able to find compatible CARPOOL partners
    - Health or disability reasons
    - Social interaction

- Environmental benefits
- Reduce vehicle congestion
- I live too far from work to Cycle, Walk, Jog or Inline Skate
- Don't like to drive or don't drive
- Don't have access to a vehicle
- Other options not available or feasible
- Other

32. Which Public Transit services do you typically use to get to work?
33. What is your typical one-way travel time to work when you commute by Public Transit?
34. How many times do you to transfer when commuting by Public Transit (one-way)?
35. How do you typically pay your public transit fare?
36. Are your public transit fares paid for by your employer? If so, how much?
37. Why do you commute by Public Transit?
    - Cost savings
    - Fast travel time
    - Safety or security
    - Convenience, comfort or enjoyment
    - Can use travel time productively
    - Health or disability reasons
    - Companionship
    - Reduce vehicle congestion
    - Environmental benefits
    - Poor weather conditions
    - Don't like to drive or don't drive
    - Driving is too stressful
    - Don't have access to a vehicle
    - Other
38. What is your typical one-way travel time to work when you Cycle, Walk, Jog or Inline Skate to work?
39. During which months do you typically cycle, walk, jog or inline skate to work?
40. Why do you cycle, walk, jog or inline skate to work?
    - Cost savings
    - I live close to work

- Safety or security
- Convenience, comfort or enjoyment
- Exercise or fitness
- Take advantage of good weather
- Social interaction
- Reduce vehicle congestion
- Environmental benefits
- Workplace incentives
- Don't like to drive or don't drive
- Don't have access to a vehicle
- Other options not available or feasible
- Other

Appendix B
Final Datasheet

Table B1: Sample Final Datasheet (General and Mode Choice Information)

	General Information								
Q. No. >	1	4	5	6	7	8	9	10	11
Sl	Department	Sex	Age	Employment Status	Time of Arriving at Work	Time of Leaving Work	Flexible Office Hour	Compressed Work Week	Approx. Home to Work Distance
40168	0	1	3	0	3	1	1	1	7
40172	0	2	3	0	4	8	1	1	7
40173	0	2	3	0	3	4	1	1	4
40185	0	2	2	0	5	5	1	1	6
40262	0	1	4	0	5	4	1	1	6
40345	0	2	2	0	3	2	1	1	4
40348	0	2	4	0	6	6	1	1	3
40352	0	1	2	0	5	5	1	1	3
40357	0	1	3	0	5	5	1	1	2
40366	0	1	5	0	2	1	1	1	7
40370	0	1	3	0	6	10	1	1	4
40385	0	2	3	0	5	5	1	1	1
40420	0	1	3	1	11	13	1	1	1
40617	0	2	4	1	5	5	1	1	3
42092	7	1	3	1	7	5	1	1	3
42122	7	2	4	1	7	5	1	1	4
42187	3	2	4	2	5	6	2	2	4
42250	7	1	4	3	13	13	2	2	3
42253	7	2	6	1	1	3	2	1	2

	Means of Commuting														
Q. No. >	12							13							
	Mode used for Commuting During Last Week								Under Which Circumstances wou						
Sl	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Frequency of Public Transit Use in last 12 months	It cost less	Increased cost of driving (fuel, parking etc.)	Transit service was faster, more frequent or more reliable	Fewer transfers	I could buy tickets and passes at my workplace	I could get transit information and advice at my workplace	The bus stop / LRT station near home/work was closer or easier to walk to
40168	3	3	3	3	3	0	0	1	0	0	0	0	0	0	0
40172	1	1	1	1	1	0	0	3	0	0	0	0	0	0	0
40173	3	1	3	2	1	0	0	1	1	0	0	0	0	1	0
40185	3	3	3	3	7	0	0	1	0	0	0	0	0	0	0
40262	2	2	2	1	0	0	0	4	0	0	1	0	0	0	0
40345	1	1	1	1	0	0	0	4	1	0	0	0	0	0	0
40348	0	0	0	0	0	0	0	3	1	0	1	0	0	0	0
40352	7	7	7	7	7	0	0	1	1	0	1	0	0	0	0
40357	2	2	2	2	2	0	0	2	1	0	0	0	0	0	0
40366	1	1	1	1	1	0	0	4	0	0	0	0	0	0	0
40370	1	1	1	1	1	0	0	3	0	0	1	0	0	0	0
40385	8	8	8	8	8	0	0	3	0	0	1	0	1	0	1
40420	7	7	7	7	7	0	0	4	1	0	0	0	0	0	0
40617	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0
42092	1	1	1	1	1	0	0	4	0	1	0	0	0	0	0
42122	1	1	1	2	1	0	0	4	1	1	0	0	0	0	0
42187	1	1	1	1	1	0	0	3	0	1	1	0	0	0	0
42250	1	1	0	0	1	1	1	4	0	1	0	0	0	0	0
42253	1	1	1	1	1	0	0	3	0	0	0	0	0	0	0

Table B2: Sample Final Datasheet contd. (Stated Preference Latent Variables)

Transit										Carpool																
14										15	16										17					
ances would the respondent consider using Transit more often											Under Which Circumstances would the respondent consider Carpooling more often											Under Wh				
The bus stop / LRT station near home/work was closer or easier to walk to	The bus stop near home/work had better shelter, seating or lighting	My bike could be taken onto the bus/LRT at all times and on all routes	There was public transit service where I live	I had a guaranteed ride home in case of emergency or unscheduled overtime	I didn't have to use my personal vehicle for work purposes	I worked more regular hours	I already take public transit as much as possible	I would not consider taking public transit to work	Other	Frequency of Carpool in last 12 months	I could find compatible carpool partners	I could carpool temporarily or occasionally	I could still drive alone when I needed to	Carpool parking was available that was more convenient or cheaper than regular parking	I had a guaranteed ride home in case of emergency or unscheduled overtime	I didn't have to use my personal vehicle for work purposes	I worked more regular hours	I already carpool as much as possible	I would not consider carpooling to work	Other	Frequency of Cycle in last 12 months	Shower, changing and locker facilities were provided	Shower, changing and locker facilities were improved	Bicycle parking was more secure and sheltered	Bicycle parking was provided or more spaces were available	The cycling routes between my home and work were safer or more convenient
0	0	0	0	0	0	0	1	0	0	4	0	0	0	1	0	0	0	0	0	0	4	0	0	0	0	0
0	0	0	0	0	0	1	0	0	1	4	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	3	1	0	0	1	0	0	1	0	0	0	4	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0	4	1	1	1	0	0	0	0	0	0	0	2	1	0	0	0	0
0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	4	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	3	1	1	1	1	1	0	0	0	0	0	4	0	0	0	0	1
0	0	0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	4	1	0	1	0	0
0	0	0	0	0	0	0	0	0	1	4	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0	3	1	0	0	0	0
0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	4	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	3	0	1	1	0	0	0	1	0	0	0	4	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	4	0	1	1	0	0	0	0	0	1	0	4	1	0	0	0	0
0	0	0	1	1	0	0	0	0	0	4	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	4	0	0	0	0	0
0	0	0	0	0	1	0	0	0	1	3	0	0	0	0	0	0	0	0	0	1	4	0	0	0	0	0
0	1	0	0	1	0	0	0	0	0	2	1	0	1	1	1	0	0	0	0	0	4	0	1	0	1	0
0	0	0	0	0	0	0	0	0	0	4	1	1	1	1	0	0	0	0	0	0	4	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0
0	0	0	0	1	1	0	0	0	1	4	0	0	0	0	0	0	0	0	1	0	4	0	0	0	0	0

Table B3: Sample Final Datasheet contd. (Mode Specific Revealed and Stated Variables)

Drive Alone																	Carpool										
23	24	25	26														27	28	29	30							
			Why Do You DRIVE ALONE To Work? (Select all that apply)																		Why Do You Commute In						
What Is Your Typical One-Way Travel Time To Work When DRIVING ALONE?	On A Daily Basis What Is The Cost Of Your Parking Space?	If There Is A Cost To Your Parking, Is It Paid For By Your Employer?	Vehicle needed for work	Vehicle needed to transport family, shopping or errands	Vehicle needed for health or disability reasons	Long or irregular work hours	Free or inexpensive parking	Fast travel time	Safety or security	Convenience, comfort or enjoyment	I can't find anyone to Carpool with	I live to far away from work to Cycle, Walk, Jog or Inline Skate	No Public Transit where I live	Weather	Other options are not available or feasible	Other	What Is Your Typical One-Way Travel Time To Work When You Participate In A CARPOOL?	How Many People Are Typically In Your CARPOOL?	On A Daily Basis What Is The Cost Of Your CARPOOL'S Parking Space?	If There Is A Cost To Your CARPOOL'S Parking Space, Is It Paid For By Your Employer And If So How Much?	Cost savings	Safety or security	Convenience, comfort or enjoyment	Can use travel time productively	I was able to find compatible CARPOOL partners	Health or disability reasons	Social interaction
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
8	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
6	5	4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	6	1	1	4	1	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
8	3	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	8	1	3	2	1	0	0	0	0	0	
5	2	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1	3	2	0	0	1	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	1	0	1	0	1	0	0	0	
9	2	3	1	0	0	1	0	0	0	0	0	0	0	0	1	0	9	1	2	3	0	0	1	0	0	0	
7	3	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1	3	2	0	0	1	0	0	0	
3	5	3	1	1	0	0	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
5	2	3	1	0	0	0	1	1	0	1	0	1	0	1	0	0	5	1	2	3	0	0	1	0	1	0	
7	2	2	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
5	1	1	0	0	0	1	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
2	2	4	0	0	0	1	1	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	

Table B4: Sample Final Datasheet contd. (Level of Service Variables)

Drive Alone				Carpool				Public Transit						Park & Ride								Carpool & Ride						
Avg. One Way travel time (min) when DRIVE ALONE	One way Cost of Parking (Whole day parking cost/2)	One way Cost of Driving	Is Drive Alone in the choice set? 1 if yes, 0 if no.	Avg. One Way travel time (min) in CARPOOL	Cost of Parking	Cost of Driving	Is Carpool in the choice set? 1 if yes, 0 if no.	IVTT	Walk Time	Wait Time (for transfer)	No. of Transfers	Transit Fare	Availability	Home to Transit IVTT	Transit IVTT	Walk Time	Wait Time	Auto Cost (Parking cost assumed to 0)	Transit Cost	No. of Transfers in Transit	Availability	Home to Transit IVTT	Transit IVTT	Walk Time	Wait Time	Auto Cost (Parking cost assumed to 0)	Tran Cost
32	2.93	4.95	1	37	1.46	2.48	1	65	8	0	0	5	1	3	50	3	5	0.30	5.00	0	1	8	50	3	5	0.15	
45	1.25	4.95	1	50	0.63	2.48	1	65	8	0	0	5	1	3	50	3	5	0.30	5.00	0	1	8	50	3	5	0.15	
28	8.75	2.06	1	33	4.38	1.03	1	34	8	5	0	5	1	8	27	3	7	1.32	5.00	0	1	13	27	3	7	0.66	
5	2.93	3.71	1	10	1.46	1.86	1	4	9	0	0	2.5	1	0	7	5	3	0.33	2.50	0	1	5	7	5	3	0.17	
45	3.75	3.71	1	50	1.88	1.86	1	0	0	0	0	0	0	17	29	3	4	1.83	5.00	0	1	22	29	3	4	0.92	
23	1.25	2.06	1	28	0.63	1.03	1	34	8	5	0	5	1	8	27	3	7	1.32	5.00	0	1	13	27	3	7	0.66	
14	2.93	1.24	1	19	1.46	0.62	1	22	5	0	0	2.5	1	0	0	0	0	0.00	2.50	0	0	5	0	0	0	0.00	
10	2.93	1.24	1	15	1.46	0.62	1	17	3	0	0	2.5	1	0	0	0	0	0.00	2.50	0	0	5	0	0	0	0.00	
10	2.93	0.58	1	15	1.46	0.29	1	9	15	7	1	2.5	1	2	7	5	3	0.36	2.50	0	1	7	7	5	3	0.18	
55	1.25	4.95	1	60	0.63	2.48	1	65	8	0	0	5	1	3	50	3	5	0.30	5.00	0	1	8	50	3	5	0.15	
35	3.75	2.06	1	40	1.88	1.03	1	35	8	7	1	7.5	1	26	13	1	3	2.90	5.00	0	1	31	13	1	3	1.45	
5	2.93	0.08	1	10	1.46	0.04	1	4	9	0	0	2.5	1	0	7	5	3	0.33	2.50	0	1	5	7	5	3	0.17	
5	2.93	0.08	1	10	1.46	0.04	1	4	9	0	0	2.5	1	0	7	5	3	0.33	2.50	0	1	5	7	5	3	0.17	
0	2.93	1.24	1	5	1.46	0.62	1	22	9	0	0	2.5	1	7	7	5	3	0.73	2.50	0	1	12	7	5	3	0.36	
13	8.75	1.24	1	18	4.38	0.62	1	24	11	4	1	2.5	1	0	0	0	0	0.51	2.50	0	0	5	0	0	0	0.26	
23	1.25	2.06	1	28	0.63	1.03	1	34	8	5	0	5	1	8	27	3	7	1.32	5.00	0	1	13	27	3	7	0.66	
35	1.25	2.06	1	40	0.63	1.03	1	0	0	0	0	0	0	0	0	0	0	0.00	2.50	0	0	5	0	0	0	0.00	
23	2.93	1.24	1	28	1.46	0.62	1	28	13	9	1	2.5	1	0	0	0	0	0.00	2.50	0	0	5	0	0	0	0.00	
8	1.25	0.58	1	13	0.63	0.29	1	38	14	8	2	2.5	1	0	0	0	0	0.00	2.50	0	0	5	0	0	0	0.00	

Appendix C
Sample Model Specification File for BIOGEME

Sample 1: Nested Logit Model for Panel Data

```
[ModelDescription]
"Edmonton Workplace Commuter Survey - Nested Logit model"


[Choice]
Choice


[Beta]
// Name Value  LowerBound UpperBound  status (0=variable, 1=fixed)

ASC1  0      -100  100   1
ASC2  0      -100  100   0
ASC3  0      -100  100   0
ASC4  0      -100  100   0
ASC5  0      -100  100   0
ASC6  0      -100  100   0
ASC7  0      -100  100   0
ASC8  0      -100  100   0

BETA101      0      -100  100   0
BETA102      0      -100  100   0
BETA103      0      -100  100   0
BETA104      0      -100  100   0

BETA701      0      -100  100   1
BETA702      0      -100  100   0
BETA703      0      -100  100   0
BETA704      0      -100  100   0
BETA705      0      -100  100   0
BETA706      0      -100  100   0
BETA707      0      -100  100   0
BETA708      0      -100  100   0

BETA801      0      -100  100   1
BETA802      0      -100  100   0
BETA803      0      -100  100   0
BETA804      0      -100  100   0
BETA805      0      -100  100   0
BETA806      0      -100  100   0
BETA807      0      -100  100   0
BETA808      0      -100  100   0

BETA811      0      -100  100   1
BETA812      0      -100  100   0
BETA813      0      -100  100   0
BETA814      0      -100  100   0
```

```
BETA815     0     -100  100   0
BETA816     0     -100  100   0
BETA817     0     -100  100   0
BETA818     0     -100  100   0

BETA821     0     -100  100   1
BETA822     0     -100  100   0
BETA823     0     -100  100   0
BETA824     0     -100  100   0
BETA825     0     -100  100   0
BETA826     0     -100  100   0
BETA827     0     -100  100   0
BETA828     0     -100  100   0

ZERO        0     -100  100   1
SIGMA1      1     -100  100   1
SIGMA2      0     -100  100   0
SIGMA3      0     -100  100   0
SIGMA4      0     -100  100   0
SIGMA5      0     -100  100   0
SIGMA6      0     -100  100   0
SIGMA7      0     -100  100   0
SIGMA8      0     -100  100   0


[Utilities]
// Id Name  Avail  linear-in-parameter expression (beta1*x1 +
beta2*x2 + ... )

1     DA    DA_Av ASC1 * one + BETA101 * DA_IVTT +    BETA102 *
DA_OVTT +   BETA103 * DA_D_Cst +    BETA104 * DA_P_Cst + BETA701 *
dum11 +     BETA801 * dum12 +       BETA811 * dum13 +
      BETA821 * dum14 +       ZERO [ SIGMA1 ] * one

2     CP    CP_Av ASC2 * one + BETA101 * CP_IVTT +    BETA102 *
CP_OVTT +   BETA103 * CP_D_Cst +    BETA104 * CP_P_Cst + BETA702 *
dum11 +     BETA802 * dum12 +       BETA812 * dum13 +
      BETA822 * dum14 +       ZERO [ SIGMA2 ] * one

3     PT    PT_Av ASC3 * one + BETA101 * PT_IVTT +    BETA102 *
PT_OVTT +   BETA103 * PT_D_Cst +    BETA104 * PT_P_Cst + BETA703 *
dum11 +     BETA803 * dum12 +       BETA813 * dum13 +
      BETA823 * dum14 +       ZERO [ SIGMA3 ] * one

4     PR    PR_Av ASC4 * one + BETA101 * PR_IVTT +    BETA102 *
PR_OVTT +   BETA103 * PR_D_Cst +    BETA104 * PR_P_Cst + BETA704 *
dum11 +     BETA804 * dum12 +       BETA814 * dum13 +
      BETA824 * dum14 +       ZERO [ SIGMA4 ] * one

5     CPR   CPR_Av      ASC5 * one + BETA101 * CPR_IVTT +
      BETA102 * CPR_OVTT +    BETA103 * CPR_D_Cst +   BETA104 *
CPR_P_Cst + BETA705 * dum11 +       BETA805 * dum12 +
      BETA815 * dum13 +       BETA825 * dum14 +       ZERO
[ SIGMA5 ] * one

6     CR    CR_Av ASC6 * one + BETA101 * CR_IVTT +    BETA102 *
CR_OVTT +   BETA103 * CR_D_Cst +    BETA104 * CR_P_Cst + BETA706 *
```

```
dum11 +     BETA806 * dum12 +       BETA816 * dum13 +
      BETA826 * dum14 +       ZERO [ SIGMA6 ] * one

7     C     C_Av  ASC7 * one + BETA101 * C_IVTT +     BETA102 *
C_OVTT +    BETA103 * C_D_Cst +     BETA104 * C_P_Cst + BETA707 *
dum11 +     BETA807 * dum12 +       BETA817 * dum13 +
      BETA827 * dum14 +       ZERO [ SIGMA7 ] * one

8     W     W_Av  ASC8 * one + BETA101 * W_IVTT +     BETA102 *
W_OVTT +    BETA103 * W_D_Cst +     BETA104 * W_P_Cst + BETA708 *
dum11 +     BETA808 * dum12 +       BETA818 * dum13 +
      BETA828 * dum14 +       ZERO [ SIGMA8 ] * one


[Draws]
500

[PanelData]
Id
ZERO_SIGMA1
ZERO_SIGMA2
ZERO_SIGMA3
ZERO_SIGMA4
ZERO_SIGMA5
ZERO_SIGMA6
ZERO_SIGMA7
ZERO_SIGMA8


[Expressions]
// Define here arithmetic expressions for name that are not
directly
// available from the data
one = 1

dum11 = ( Gender_F >= 1 )
dum12 =     (     Emp_FT      >=    1     )
dum13 =     (     Flex_Yes    >=    1     )
dum14 =     (     Comp_Yes    >=    1     )


[NLNests]
NESTA 1     -1    1     1     1 2
NESTB 1     -1    1     0     3 4 5 6
NESTC 1     -1    1     0     7 8


[Model]
// Currently, only $MNL (multinomial logit), $NL (nestelogit),
$CNL
// (cross-nested logit) and $NGEV (Network GEV model) are valid
keywords
//
$NL
```

Sample 2: Nested Logit Model with Latent Variables

```
// File trial12C for Wednesday.mod

[ModelDescription]
"Edmonton Workplace Commuter Survey - Nested Logit model with
Latent Variables"

[Choice]
Wed

[Beta]
// Name Value  LowerBound UpperBound  status (0=variable, 1=fixed)

ASC1  0      -100  100   1
ASC2  0      -100  100   0
ASC3  0      -100  100   0
ASC4  0      -100  100   0
ASC5  0      -100  100   0
ASC6  0      -100  100   0
ASC7  0      -100  100   0
ASC8  0      -100  100   0

BETA102     0     -100  100   0
BETA104     0     -100  100   0
BETA105     0     -100  100   0
BETA106     0     -100  100   0
BETA107     0     -100  100   0
BETA108     0     -100  100   0

BETA701     0     -100  100   1
BETA702     0     -100  100   0
BETA703     0     -100  100   0
BETA704     0     -100  100   0
BETA705     0     -100  100   0
BETA706     0     -100  100   0
BETA707     0     -100  100   0
BETA708     0     -100  100   0

BETA801     0     -100  100   1
BETA802     0     -100  100   0
BETA803     0     -100  100   0
BETA804     0     -100  100   0
BETA805     0     -100  100   0
BETA806     0     -100  100   0
BETA807     0     -100  100   0
BETA808     0     -100  100   0

BETA811     0     -100  100   1
BETA812     0     -100  100   0
BETA813     0     -100  100   0
BETA814     0     -100  100   0
BETA815     0     -100  100   0
BETA816     0     -100  100   0
BETA817     0     -100  100   0
BETA818     0     -100  100   0
```

```
BETA821     0     -100  100   1
BETA822     0     -100  100   0
BETA823     0     -100  100   0
BETA824     0     -100  100   0
BETA825     0     -100  100   0
BETA826     0     -100  100   0
BETA827     0     -100  100   0
BETA828     0     -100  100   0

BETA831     0     -100  100   0
BETA832     0     -100  100   0
BETA833     0     -100  100   0
BETA834     0     -100  100   0
BETA835     0     -100  100   0
BETA836     0     -100  100   0
BETA837     0     -100  100   0
BETA838     0     -100  100   0
BETA839     0     -100  100   0
BETA840     0     -100  100   0
BETA841     0     -100  100   0
BETA842     0     -100  100   0


[Utilities]
// Id Name  Avail  linear-in-parameter expression (beta1*x1 +
beta2*x2 + ... )

1     DA    DA_Av ASC1 * one + BETA105 * DA_Cost + BETA701 * dum11
+ BETA801 * dum12 + BETA811 * dum13 + BETA821 * dum14

2     CP    CP_Av ASC2 * one + BETA105 * CP_Cost + BETA702 * dum11
+ BETA802 * dum12 + BETA812 * dum13 + BETA822 * dum14 + BETA831 *
dum15 + BETA832 * dum16 + BETA833 * dum17

3     PT    PT_Av ASC3 * one + BETA102 * PT_IVTT + BETA104 *
PT_OVTT + BETA106 * PT_D_Cst + BETA703 * dum11 + BETA803 * dum12 +
BETA813 * dum13 + BETA823 * dum14 + BETA834 * dum18 + BETA835 *
dum19 + BETA836 * dum20

4     PR    PR_Av ASC4 * one + BETA102 * PR_IVTT + BETA104 *
PR_OVTT + BETA106 * PR_D_Cst + BETA704 * dum11 + BETA804 * dum12 +
BETA814 * dum13 + BETA824 * dum14

5     CPR   CPR_Av      ASC5 * one + BETA102 * CPR_IVTT + BETA104
* CPR_OVTT + BETA106 * CPR_D_Cst + BETA705 * dum11 +  BETA805 *
dum12 + BETA815 * dum13 + BETA825 * dum14

6     CR    CR_Av ASC6 * one + BETA102 * CR_IVTT + BETA104 *
CR_OVTT + BETA106 * CR_D_Cst + BETA706 * dum11 + BETA806 * dum12 +
BETA816 * dum13 + BETA826 * dum14

7     C     C_Av  ASC7 * one + BETA107 * H_W_Dist + BETA707 *
dum11 + BETA807 * dum12 + BETA817 * dum13 + BETA827 * dum14 +
BETA837 * dum21 + BETA838 * dum22 + BETA839 * dum23
```

```
8     W     W_Av  ASC8 * one + BETA108 * H_W_Dist + BETA708 *
dum11 + BETA808 * dum12 + BETA818 * dum13 + BETA828 * dum14 +
BETA840 * dum24 + BETA841 * dum25 + BETA842 * dum26


[Expressions]
// Define here arithmetic expressions for name that are not
directly
// available from the data
one = 1

DA_Cost = ( DA_D_Cst + DA_P_Cst )
CP_Cost = ( CP_D_Cst + CP_P_Cst )
PT_D_Cst = PT_Cost
PT_OVTT = ( PT_Walk + PT_Wait )
PR_IVTT = ( PR_AIVTT + PR_TIVTT )
PR_OVTT = ( PR_Walk + PR_Wait )
PR_D_Cst = ( PR_A_Cst + PR_T_Cst )
CPR_IVTT = ( CPR_AIVTT + CPR_TIVTT )
CPR_OVTT = ( CPR_Walk + CPR_Wait )
CPR_D_Cst = ( CPR_A_Cst + CPR_T_Cst )
CR_IVTT = ( CR_C_Tim + CR_TIVTT )
CR_D_Cst = CR_Cost


dum11 = ( Gender >= 1 )
dum12 = ( Emp >= 1 )
dum13 = ( Flex >= 1 )
dum14 = ( Comp >= 1 )
dum15 = ( Q16O1 >= 1 )
dum16 = ( Q16O3 >= 1 )
dum17 = ( Q16O5 >= 1 )
dum18 = ( Q14O3 >= 1 )
dum19 = ( Q14O1 >= 1 )
dum20 = ( Q14O4 >= 1 )
dum21 = ( Q18O5 >= 1 )
dum22 = ( Q18O13 >= 1 )
dum23 = ( Q18O8 >= 1 )
dum24 = ( Q20O9 >= 1 )
dum25 = ( Q20O4 >= 1 )
dum26 = ( Q20O11 >= 1 )


[NLNests]
NESTA 1     -1    1     1     1 2
NESTB 1     -1    1     0     3 4 5 6
NESTC 1     -1    1     0     7 8


[Model]
// Currently, only $MNL (multinomial logit), $NL (nestedlogit),
$CNL
// (cross-nested logit) and $NGEV (Network GEV model) are valid
keywords
//
$NL
```

Appendix D
Sample Model Estimation Output File from BIOGEME

BIOGEME Version 1.8 [Sat Mar 7 14:36:56 CEST 2009]

Michel Bierlaire, EPFL

This file has automatically been generated.

01/06/10 17:31:41

Edmonton Workplace Commuter Survey - Mixed Nested Logit model

Model:	Mixed Nested Logit for panel data
Number of draws:	30
Number of estimated parameters:	48
Number of observations:	10767
Number of individuals:	2696
Null log-likelihood:	-17145.739
Init log-likelihood:	-13530.169
Final log-likelihood:	-5207.540
Likelihood ratio test:	23876.398
Rho-square:	0.696
Adjusted rho-square:	0.693
Final gradient norm:	+6.212e-002
Diagnostic:	Convergence reached...
Iterations:	262
Run time:	02h 44:29
Variance-covariance:	from finite difference hessian
Sample file:	10817datafor modeling.dat

Utility parameters

Name	Value	Std err	t-test	p-value		Robust Std err	Robust t-test	p-valu
ASC1	0.00	fixed						
ASC2	-13.9	0.709	-19.64	0.00		1.09	-12.82	0.00
ASC3	-2.18	0.676	-3.23	0.00		1.39	-1.57	0.12
ASC4	-18.3	1.98	-9.21	0.00		3.40	-5.38	0.00
ASC5	-29.6	2.78	-10.63	0.00		4.40	-6.71	0.00
ASC6	-23.7	2.84	-8.34	0.00		5.02	-4.71	0.00
ASC7	-9.48	0.833	-11.38	0.00		1.69	-5.60	0.00
ASC8	-13.8	1.26	-10.97	0.00		1.97	-6.98	0.00
BETA101	-0.0605	0.00588	-10.29	0.00		0.00962	-6.29	0.00
BETA102	-0.155	0.0241	-6.42	0.00		0.0565	-2.73	0.01
BETA103	-0.590	0.0895	-6.59	0.00		0.157	-3.76	0.00
BETA104	-0.237	0.0832	-2.84	0.00		0.149	-1.59	0.11
BETA701	0.00	fixed						
BETA702	0.796	0.320	2.49	0.01		0.546	1.46	0.14
BETA703	1.27	0.316	4.01	0.00		0.557	2.27	0.02
BETA704	7.34	0.897	8.19	0.00		1.51	4.86	0.00
BETA705	6.46	1.02	6.31	0.00		2.54	2.54	0.01
BETA706	-4.70	1.30	-3.62	0.00		2.25	-2.09	0.04
BETA707	-5.30	0.461	-11.51	0.00		0.858	-6.18	0.00
BETA708	-4.65	0.574	-8.10	0.00		0.875	-5.31	0.00
BETA801	0.00	fixed						
BETA802	0.296	0.415	0.71	0.47	*	0.650	0.46	0.65
BETA803	-2.04	0.419	-4.88	0.00		0.787	-2.60	0.01
BETA804	-0.496	0.810	-0.61	0.54	*	1.34	-0.37	0.71
BETA805	0.377	0.982	0.38	0.70	*	2.42	0.16	0.88
BETA806	-1.57	1.91	-0.82	0.41	*	6.06	-0.26	0.80
BETA807	0.0180	0.570	0.03	0.97	*	1.15	0.02	0.99

Sample BIOGME Output File, contd..

BETA818	0.866	0.571	1.52	0.13	*	0.896	0.97	0.33	*
BETA821	0.00	fixed							
BETA822	2.10	0.346	6.08	0.00		0.573	3.67	0.00	
BETA823	0.294	0.373	0.79	0.43	*	0.821	0.36	0.72	*
BETA824	1.87	0.721	2.59	0.01		1.35	1.38	0.17	*
BETA825	-2.28	1.20	-1.90	0.06	*	2.31	-0.99	0.32	*
BETA826	-1.74	1.38	-1.26	0.21	*	2.83	-0.61	0.54	*
BETA827	-0.109	0.421	-0.26	0.80	*	0.720	-0.15	0.88	*
BETA828	1.11	0.576	1.92	0.05	*	0.973	1.14	0.26	*
SIGMA1	1.00	fixed							
SIGMA2	-10.1	0.458	-22.02	0.00		0.804	-12.56	0.00	
SIGMA3	-6.62	0.289	-22.91	0.00		0.502	-13.19	0.00	
SIGMA4	-11.8	0.958	-12.30	0.00		1.76	-6.70	0.00	
SIGMA5	-19.9	1.62	-12.29	0.00		2.52	-7.88	0.00	
SIGMA6	11.2	1.10	10.14	0.00		1.73	6.47	0.00	
SIGMA7	8.49	0.425	19.99	0.00		0.712	11.93	0.00	
SIGMA8	15.0	0.920	16.31	0.00		1.32	11.41	0.00	
ZERO	0.00	fixed							

Model parameters

Name	Value	Std err	t-test 0	p-value	t-test 1	p-value	Robust Std err	Robust t-test 0	p-value	Robust t-test 1	p-value
NESTA	1.00	fixed									
NESTB	0.307	0.0263	11.69	0.00	-26.34	0.00	0.0491	6.26	0.00	-14.10	0.00
NESTC	0.221	0.0223	9.92	0.00	-34.96	0.00	0.0375	5.89	0.00	-20.76	0.00

Utility functions

7	C	C_Av	ASC7 * one + BETA101 * C_IVTT + BETA102 * C_OVTT + BETA103 * C_D_Cst + BETA104 * C_P_Cst + BETA707 * dum11 + BETA807 * dum12 + BETA817 * dum13 + BETA + ZERO [SIGMA7] * one
2	CP	CP_Av	ASC2 * one + BETA101 * CP_IVTT + BETA102 * CP_OVTT + BETA103 * CP_D_Cst + BETA104 * CP_P_Cst + BETA702 * dum11 + BETA802 * dum12 + BETA812 * dum13 + dum14 + ZERO [SIGMA2] * one
5	CPR	CPR_Av	ASC5 * one + BETA101 * CPR_IVTT + BETA102 * CPR_OVTT + BETA103 * CPR_D_Cst + BETA104 * CPR_P_Cst + BETA705 * dum11 + BETA805 * dum12 + BETA815 * dum1 * dum14 + ZERO [SIGMA5] * one