AN ABSTRACT OF A THESIS

DISCRETE – CONTINUOUS MODEL OF HOUSEHOLD VEHICLE OWNERSHIP AND TRIP GENERATION

Kiranmai Chirumamilla

Master of Science in Civil Engineering

The primary objective of the study was to model household vehicle ownership and trip generation. A literature review was conducted and shortcomings in the current model structure of trip generation and vehicle ownership were identified. An appropriate decision structure that addresses the primary shortcoming was formulated, namely that it captured the interdependency between vehicle ownership and trip generation decisions of a household. A descriptive analysis identified the variables that influence these two decisions relating to mobility and travel. The variables were used in developing two model structures that capture the interdependency between the number of vehicles a household owns and the daily number of trips it makes by using the data collected in the travel behavior survey undertaken in the Nashville metropolitan region in 1998.

Based on the study results, the proposed models were found to work well and to be able to predict the vehicle ownership and travel choices of households in the dataset, indicating that they had behavioral merit. Socioeconomic variables such as household income, household size, and number of employed household members were found to strongly influence household vehicle ownership and trip-making decisions.

DISCRETE – CONTINUOUS MODEL OF HOUSEHOLD VEHICLE OWNERSHIP AND TRIP GENERATION

A Thesis

Presented to

The Graduate School

Tennessee Technological University

by

Kiranmai Chirumamilla

In Partial Fulfillment

Of the Requirements for the Degree

MASTER OF SCIENCE

Civil Engineering

December 2008

CERTIFICATE OF APPROVAL OF THESIS

DISCRETE – CONTINUOUS MODEL OF HOUSEHOLD VEHICLE OWNERSHIP AND TRIP GENERATION

by

Kiranmai Chirumamilla

Graduate Advisory Committee:

D Badoe

Dr. Daniel A. Badoe, Chairperson

12/3/08

Date

Dr. Steven Click

2008-12-03

Date

Roy C Loutzenheiser

Dr. Roy C. Loutzenheiser

12-3-08

Date

Approved for the Faculty:

Francis Otuonye

Associate Vice President for Research

and Graduate Studies

12/8/08

Date

ACKNOWLEDGEMENTS

I would like to express my profound thanks to the people who have helped me make this project successful. Foremost, I wish to thank the honorable chairperson of my graduate committee, Dr. Daniel A. Badoe, for his wisdom, efforts, resources, excellent guidance, and immense patience throughout the duration of this thesis. It was an honor to work with him. The invaluable suggestions offered by him and the diligence with which he assisted in my research helped me a long way in realizing this thesis. I would also like to thank my other graduate committee members Dr. Steven Click and Dr. Roy C. Loutzenheiser for their reviewing and valuable comments on my thesis work.

I wish to express my heartfelt appreciation for the staff of the Department of Civil and Environmental Engineering for helping me with my degree. I would like to thank the Department of Civil and Environmental Engineering and the Centre for Energy Systems Research for providing financial support for my graduate study. I am thankful to the Nashville Metropolitan Organization for providing the database used in the research.

I am thankful to my parents, family members, friends, and colleagues for their continuous affection and support.

TABLE OF CONTENTS

Page

LIST OF TABLES v

LIST OF FIGURES vi

CHAPTER 1 1

INTRODUCTION 1

LITERATURE REVIEW 5

2.1 Introduction 5

2.2 Trip Generation Modeling 5

2.3 Vehicle Ownership Modeling 6

2.3.1 Aggregate Cross-Sectional Studies 7

2.3.2 Disaggregate Cross-Sectional Studies 8

2.3.3 Time-Series Studies 14

2.3.4 Panel Studies 16

CHAPTER 3 19

DATA AND DESCRIPTIVE ANALYSIS 19

3.1 Introduction 19

3.2 Survey 19

3.3 Descriptive Analysis 20

CHAPTER 4 40

METHODOLOGY 40

4.1 Modeling Household Trip Generation and Household Vehicle Ownership Level 40

4.1.1 Joint Discrete/Continuous Model 43

4.1.2 Joint Discrete-Continuous Model without Stratification 47

4.2 Model Assessment 49

4.2.1 Goodness of Fit Measure for Vehicle Ownership (Discrete Choice) Model 49

4.2.2 Goodness of Fit Measure for Household Trip Generation Model 50

4.2.3 Signs and Statistical Significance of the Estimated Coefficients 51

CHAPTER 5 52

RESULTS AND DISCUSSION OF RESULTS 52
5.1 Variables..52
5.2 Model Estimation Output – Model with Selectivity Bias Correction Term..56
5.3 Model Estimation Output – Discrete-Continuous Model without Stratification 64
5.3.1 Vehicle Ownership Model..64
5.3.2 Trip Generation Model 67
CHAPTER 6..69
CONCLUSIONS AND FUTURE RESEARCH 69
6.1 Summary and Conclusions..69
6.2 Future Research..70
REFERENCES 72
VITA..74

LIST OF TABLES

Page

Table 3-1: Trip rate by household Vehicle ownership....................20

Table 3-2: Trip rate by Household Size....................22

Table 3-3: Household trip-rate by Life Cycle:....................23

Table 3-4b:Trip rate of a household by number of vehicles owned and household income....................28

Table 3-5a: Frequency distribution of households by vehicle ownership and county of residence....................30

Table 3-5b: Household trip rate by vehicle ownership and county of residence....................31

Table 3-6a: Distribution of households by vehicle ownership and household size....................33

Table 3-6b: Household trip-rate by vehicle ownership and household size....................34

Table 3-7a: Vehicle Ownership by life cycle....................35

Table 3-7b: Household trip rate by vehicle ownership and life cycle....................36

Table 3-8: Household trip rate by household income....................38

Table 5-1: Definition of variables specified in the Vehicle Ownership Models....................55

Table 5-2: Estimated coefficients of the variables in the model.....................57

Table 5-3: Estimates of coefficients of trip generation variables for households owning zero vehicles....................60

Table 5-4: Estimates of coefficients of trip generation variables for households owning one vehicle....................61

Table 5-5: Estimates of coefficients of trip generation variables for households owning two vehicles....................61

Table 5-5: Estimates of coefficients of trip generation variables for households owning two vehicles....................63

Table 5-6: Vehicle Ownership Model with Embedded Trip Generation Equation....................65

Table 5-7: Trip Generation Model with Expected Vehicle Ownership as Explanatory Variable....................68

LIST OF FIGURES

Page

Figure 3-1: Trip rate against Household Vehicle Ownership 21

Figure 3-2: Household trip rate versus Household Size 22

Figure 3-3: Household trip-rate versus life cycle 24

Figure 3-4: Frequency distribution of vehicle ownership conditional on household income. 29

Figure 3-5: Vehicle ownership by county of residents 31

Figure 3-6: Frequency distribution of households by vehicle ownership conditional on household size 34

Figure 3-7: Frequency distribution of households by vehicle ownership conditional on life cycle of household 36

Figure 3-8 Household trip-rate by household income. 38

CHAPTER 1

INTRODUCTION

An important step in the planning process for transportation systems of metropolitan regions is the travel-forecasting step. In the majority of metropolitan regions, the approach for generating these travel forecasts is embodied in what is referred to as the four-step urban transportation modeling system (UTMS). The first of this four-step procedure is trip generation, which is concerned with predicting the number of trips generated by the land use activities in each of the traffic zones into which an urban region is partitioned. The second step is trip distribution, which is concerned with predicting the spatial distribution of travel within the region. The third step is mode split, which is concerned with predicting the number of trips that would be made by each of the competing modal alternatives. The fourth and the final step is traffic assignment, which is concerned with predicting the vehicular flows on the urban region's transportation network.

Trip generation is modeled as a function of several explanatory variables notable amongst which is the number of vehicles owned by the members of a household. The current practice assumes that the structure of the decision process governing how many vehicles a household owns and the number of trips members of a household make daily is a recursive structure, in other words, a dependency that goes in one direction. More explicitly, the theory assumes that the household vehicle ownership decision, which is viewed as a medium term decision is, made first, while the number of trips made daily is treated as a short run decision, made conditional on the number of vehicles a household owns (Meyer and Miller, 2001; Ben-Akiva and Lerman, 1985).

The source of travel and travel-related data required by the above approach for modeling vehicle ownership and trip generation in metropolitan regions is the Household Travel Behavior Survey. An important requirement is that the size of sample be large enough to support modeling at a disaggregate level. Given the high cost of conducting household travel surveys (they range in cost from $150 - $400 per household), many metropolitan regions are not able to afford the collection of necessary data and therefore rely on census data. Census data, for reasons of preserving anonymity, are provided at an aggregate level. The modeling methods are therefore based on the use of aggregate data; several metropolitan regions, including Nashville, have developed a cross-classification model of auto ownership using the applicable aggregate census data. Output from the vehicle ownership model is then employed in the trip generation step for predicting the trips at traffic zone-level directly. A merit with the use of census data is that it is collected every ten years by the federal government hence planners are assured of getting new data at specified time-periods for “new” model development without having to be concerned with finding the financial and other resources to collect data periodically. However, use of data at the aggregate level results in the burying of variability in the data and, therefore, appropriate sensitivities of residents of a region to policy variables cannot be correctly ascertained.

An important thing to note from the above discussion is that in modeling trip generation, the vehicle ownership level of each household is treated as known, indicating that the determination of vehicle ownership is exogenous to the travel-forecasting phase. Trip generation, by this, does not affect household vehicle ownership level. Careful thought though about the behavior of individuals constituting households would suggest that difficulties encountered in households as they seek to undertake trips relating to day-to-day non-work activities in addition to work activities would prompt changes in a household's vehicle ownership level until an equilibrium point is reached in terms of activity participation away from home requiring vehicular travel and the vehicles available to support this activity participation level. This would suggest interdependency in these two decisions.

Other concerns with current method include, first, the assumption that they make that vehicle ownership level of a household in the future can be predicted with certainty hence only a single value of vehicle ownership is obtained for each household. In reality, the best the analyst can do is to predict the probability of a household having a specified number of vehicles given the relevant characteristics of the household.

Second, modeling methods based on aggregate data such as employed in Nashville and several other regions, by virtue of their construct, do not lend themselves to investigating the impact of policy variables on auto ownership.

It is therefore hypothesized in this thesis that the decision of the daily number of trips a household undertakes and the number of vehicles a household owns at a given point in time (as with models based on cross-sectional data, a strong assumption is made here that each household at the time of survey was at the equilibrium point) are interdependent decisions and hence should be modeled as such.

The thesis document is organized as follows: Chapter 2 presents the review of the relevant literature on vehicle ownership modeling and trip generation. Chapter 3 presents the descriptive analysis of the travel database. Chapter 4 presents the methodology. Chapter 5 presents and discusses the model estimation results. Finally, Chapter 6 presents the conclusions of the research and recommendations for modeling practice.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

There have been several studies on the modeling of trip generation and auto-ownership, respectively. This chapter begins with a review of trip generation modeling followed by a review of vehicle ownership modeling.

2.2 Trip Generation Modeling

Trip generation is in general modeled using either cross-classification analysis or multiple linear regression analysis (Meyer and Miller, 2000; Ortuzar and Willumsen, 2001; McNally, 2000). Both these methods are very well documented in the literature. The explanatory variables specified in both linear and cross-classification models are socio-economic variables, with a few applications including spatial variables as well. Studies have sought to include some measure of accessibility in trip generation models in the quest to have trip generation be sensitive to transport system characteristics, but this has largely been unsuccessful because the specified accessibility variables have consistently emerged from model-estimation with a counter-intuitive sign (Ortuzar and Willumsen, 2001). Vehicle ownership has consistently been one of the most important variables influencing the number of trips produced by a household (Ryan and Han, 1999). All studies of trip generation encountered in the literature treat household/zonal vehicle ownership as a known input for trip generation. The implication is that vehicle ownership is determined exogenously to the first step of the travel-forecasting phase.

Several studies have been conducted on the modeling of vehicle ownership for input into some stage of the transportation planning process. Some of them are discussed below.

2.3 Vehicle Ownership Modeling

A review on vehicle availability modeling by Cambridge Systematics, Inc,. (1997) described the state of practice and state of art vehicle availability models developed by transportation planners and researchers at metropolitan planning organizations and state DOT. They provided a detailed description of the advancement of model structures from linear regression to cross-classification to logit and probit models making use of the aggregate and disaggregate data. The data used in these studies are the census data obtained from the Census Transportation Planning Package (CTPP) or the Public Use Micro data Sample or Travel Survey data. Basic vehicle availability models, developed completely on aggregate census data, used the common variables of household size, income, and location. Limitation of these models is that they do not contain any variables explaining the impact of other transportation services on vehicle availability. More advanced practice models based on travel survey data include additional variables such as additional household characteristics and locational descriptors, transit and highway accessibility, and pedestrian environment variables. The state-of-the-art models include extended features which combine vehicle availability and vehicle type choice and dynamic models that focus on variability in vehicle acquisition. These models are complex but provide better information on household activity and travel behavior.

2.3.1 Aggregate Cross-Sectional Studies

The Southeastern Wisconsin Regional Planning Commission (Cambridge Systematics, Inc. (1997)) developed a model (Milwaukee model) for estimating the average automobile availability by zone. The model was developed using a combination of census data and MPO land use data at the zonal level. Probabilities of auto availability level and the average household sizes were estimated by linear regression procedure. The model is simple and requires limited number of data sources but it does not provide variables required for input to the trip generation process.

The Puget Sound Regional Council (Cambridge Systematics, Inc. (1997)) developed and implemented a vehicle availability model (Seattle model) based on a combination of PUMS data and CTPP data. Estimation of households by vehicle availability level was developed in two stages. Initially PUMS data were used to develop cross-tabulations of households with the dimensions of household size, income group, number of workers, and vehicles available and then applied to each zone for which a cross-tab of households by the first three variables is available from the CTPP data to provide an initial estimate. Estimates of the model parameters were obtained through linear regression analysis at the second stage. Limitations of the model are lack of pedestrian environment variables and aggregate household data.

Thakuriah and Liao (2005) analyzed vehicle ownership expenditure of American households. They used the Consumer Expenditure Survey data collected in 1980 by the U.S. Census Bureau for the Bureau of Labor Statistics, which contains data on socio-demographic, economic, financial, and subsidy characteristics as well as the income and expenditure data of the households for the purpose of analysis. The model used in this paper is a two-stage sample selection model proposed by Heckman (1979). The first stage was to model the household decision of vehicle ownership and the second stage was to model the expenditure intensity decision for the sub-sample of households given the vehicle ownership decision. Estimates for the different regions were obtained and their results compared. It was found that the location and region effects on vehicle ownership and expenditure were not constant but varied by type of location and type of region. The coefficient of determination and standard error of the residuals were reasonable. Of concern with this study are first the quarterly expenditure data, which was converted into annual data. These data were at the traffic zone level rather than at the individual level. Errors due to aggregation therefore affected the results. Second, the interpretation of coefficients was not straightforward due to the involvement of dummy variables.

2.3.2 Disaggregate Cross-Sectional Studies

Whelan (2007) developed an econometric model of household car ownership in Great Britain. More specifically, the study developed discrete choice models of how many vehicles household owns as a function of market saturation, license holding, household income and structure, household employment, company car provision, and purchase and cost of use. The models were developed to generate forecasts across Britain to the year 2031. Socio-demographic characteristics of British residents were considered important to predicting vehicle ownership. National travel survey data obtained from Transport Statistics Great Britain in 2003 were used to study the growth trend of car ownership across Great Britain. The models were based on s-shaped growth curves. The data used to develop the models were obtained from Family

Expenditure Survey and the National Travel Survey undertaken by the Office for National statistics. The estimated models were applied using a methodology known as prototypical sampling. It allowed the application of models to 1203 zones to the year 2031.The percentage increase in the car ownership was estimated.

Ryan and Han (1999) performed a study and proposed a vehicle ownership model that describes an improved specification for discrete choice models. The source of data for the analysis was the 1995 Oahu household interview survey data from 4,060 households. Vehicle ownership models were estimated on the dataset. A variation of the model was tested on data for other American cities obtained from the 1990 US census. The results showed the models to be similar in structure and statistical significance. The multinomial logit formulation was used in vehicle ownership modeling. The variables found to influence how many vehicles a household owned included attributes of households and its members, and attributes of the residential location of the household. More specifically, the variables specified in the model were ownership cost, cost and inconvenience of parking at home, household income, vehicle availability for household members by age or work status, vehicle contribution to accessibility, number of workers, density of the residence zone, and geography.

The Southeast Michigan Council of Governments (SEMCOG) (Cambridge Systematics, Inc. (1997)) developed and applied disaggregate model to vehicle availability and trip generation forecasting. Data for the estimation were obtained from the US census's 1977 Annual Housing Survey. The models were developed using the curves for the fraction of households owning zero, one, two, and three or more vehicles as a function of household income level. The number of households in each category is estimated using conditional probabilities. The results were grouped into 20 categories based on household size and vehicle availability which are provided as inputs to trip generation forecasting.

The earliest logit auto ownership models were developed by the Metropolitan Transportation Commission for San Francisco Bay area in 1976 (Cambridge Systematics, Inc. (1997)). Data used to develop the models were a combination of travel survey data, highway and transit network level-of-service data, and zonal land use, population, and employment data. A wide range of variables such as transit and highway accessibility variables are included along with the household and socio-economic variables. Two models were developed in which one predicts the share of households vehicle availability without workers and the other predicts the share of households vehicle availability with workers. ALOGIT estimation package was used to provide statistical estimates of the model's parameters. Data preparation process for the MTC model was extensive and its incorporation in the travel forecasting process is also difficult.

One of the state of art disaggregate models is the LOGIT choice model in Oregon (Portland logit model), developed by the Metropolitan Service District in Portland in 1989(Cambridge Systematics, Inc. (1997)).Utility functions were developed to estimate the households vehicle availability. Data used were the household survey data supplemented with an accessibility variable (defined as number of employment opportunities which can be reached within 30 minutes of transit time from the residence zone), which incorporates zonal employment and transit level-of-service data. The model includes only household variables and simple accessibility variables and was applied at a market segment level in each traffic analysis zone. The model is simple to estimate and apply but it did not include the effects of differences in pedestrian amenities on vehicle availability. Later this model was modified and expanded in

1994 by including income and pedestrian environment variables. The results showed that auto ownership declines as pedestrian environment improves and the model is helpful in predicting the number of zero car households.

Another advanced Logit model is the Philadelphia ordered response logit model which is an extension of the Seattle and 1994 Portland models (Cambridge Systematics, Inc. (1997)). The features added are Extension of Portland model into two model structures as multinomial logit (MNL) which assumes that the decision of a household to own number of vehicle is one-time choice and ordered response logit (ORL) which assumes that households decision of current vehicle availability level follows a sequence of decisions, combining highway and transit accessibility variables with pedestrian environment variables and use of dummy variable that explains the choice of more than one vehicle per person. This model contains almost all the variables that are expected to affect the vehicle availability.

Mannering (1986) analyzed selectivity bias in discrete and continuous choice models and developed econometric techniques for correcting it. Mannering pointed out that a modeling system that has interrelated discrete and continuous choices will exhibit selectivity bias in its estimation of continuous equations. The corrective methods suggested involve joint estimation of a discrete model derived from a random utility theory and a continuous model estimated by regression procedures. To illustrate this joint estimation, household’s decision of type of vehicle to own and the extent to which it is utilized is assessed. For discrete choice of vehicle type and continuous choice of vehicle utilization, utility maximizing models are defined for each household. To eliminate bias and inconsistent parameter estimates, vehicle utilization is estimated as a part of joint model of discrete choice of vehicle type. The analysis considered single vehicle households that own low, medium and high accumulated mileage vehicles. The data used for estimation are a 364 household sample collected by the US department of Energy in the spring of 1980. Utilization of vehicles is defined over a 6 month period. Appropriate variables are included and coefficients are estimated which are reasonably statistically significant. It is observed that age variable has a strong negative effect on vehicle utilization and income variable was generally not significant. Residential location and number of drivers were found to be less significant. The estimated results for equations not corrected for selectivity bias and for those that are corrected for selectivity bias are compared and noticed that the difference is not large.

Abu-Eisheh and Mannering (1987) analyzed commuters' choices of routes and departure times using a discrete/continuous econometric models. They used the morning work trip data collected in state college Pennsylvania to estimate the models. Data collection is done by conducting a survey which includes a trip log with a variety of information on the most recent trip, socio economic and household characteristics of the respondent. In addition to this extensive traffic related data of the diverse routes connected to origin-destination pair were collected. A linear function is chosen to calculate utility provided by a route to the traveler. Route choice probabilities are calculated using multinomial logit - model and a continuous departure time model is specified taking in to account different factors influencing travel time of the commuters and delay cushion. As route and travel time choices are inter related decisions, selectivity bias occurs. To avoid this travel time is calculated conditioning on the choice of route and estimated probabilities from the route choice model. It is observed from the estimated results that higher income commuters find travel time to be more onerous than their lower income counterparts. Number of traffic signals and the probability of coordinated traffic signals had a positive influence on the probability of route selection. We can find improvement in the estimated coefficients of travel time between corrected and uncorrected models estimated using ordinary least squares. The model results were explained based on several other variables indicating selectivity bias is considerable.

A joint discrete and continuous choice model mentioned in the review of De Jong and Fox (2004) explain household car ownership and car use in an integrated micro economic frame work. De Jong developed two different models based on household private car ownership and car use. The first model is called statistical model and is used for demand predictions. It assumes that a household has a structural desired annual kilometrage depending on the attributes of household to own a car. Second model is the indirect utility model based on micro economic theory. It is based on the idea that households compare combinations of car ownership and car use and choose the one with higher utility. Further these models were improved by introducing additional explanatory variables.

De Jong and Fox (2004) reviewed and compared different car ownership models. The statistic disaggregate models reviewed include car ownership sub-model within the Dutch national model system (LMS) for transport, Bhatt and Pullugurta's model, Car Ownership model for Sydney, NRTF model by Whelan, Rich and Nielsen (2001). LMS model states that car ownership choices are conditioned on household license holding and are modeled as binary logit models. The models are explained including the variables like monthly income the household can freely spent and fixed car costs. Sydney Strategic Transportation Model models company and total car ownership of the households. Different test approaches in owning company and private cars are considered and the best structure is identified. Households choose number of private cars depending on the number of company cars owned. Car ownership is based on logarithm of net household income, license holders, parking costs, and job position and type. The 1997 NRTF model was improved to NRTF2001 model by including economic, environmental, and social impacts of traffic growth. The 1997 NRTF binary model is improved by introducing prediction of multiple car ownership and company cars dummies. Rich and Nielsen presented a nested- logit car ownership model conditioned on both residential and work location choice. These models are suited for short and medium run predictions. The review of De Jong and Fox also include static disaggregate car type choice models which deal with the choice of car type of the household, given car ownership.

2.3.3 Time-Series Studies

The state of art vehicle availability models includes vehicle type choice models. These models are combination of household vehicle availability and vehicle type choice. The model variables include information on characteristics of vehicles like price, physical dimensions, operating and repair costs and fuel efficiency, average utility of household's vehicle choice, and vintage choice. Dynamic models of vehicle availability predict the change in the vehicle availability of a household with time making use of the time-series data. Household panel survey data are required for these models. Extensive academic research is going on in this area. These models are complex and provide challenges in data collection but provides accurate estimates of household activities and travel behavior.

Roorda et al. (2000) conducted a study to collect data for developing dynamic models of car ownership. A longitudinal survey was conducted to obtain the data. A retrospective interviewing method was adopted in the research, which included(a) a letter of introduction to the study;(b) a primary telephone interview; and (c) secondary interviews. The interviews were conducted over a period of 7 weeks on a sample of 1741 households from different geographic locations. The relationship between characteristics of household and the occurrence of vehicle transactions, the choice of vehicle type, the duration a vehicle is held, and degree of consumer loyalty to different types of vehicles was analyzed. The study showed that the retrospective surveys capture in detail household/individual changes in characteristics that are associated with changes in vehicle ownership, hence are useful in forecasting vehicle ownership.

Even though the data collection method was satisfactory, its implementation had some drawbacks. These include, (1) The skipping of questions on account of interviewer haste, and the asking of wrong questions; (2) Errors occurring in the input of the data into a database program; and (3) the interviewee's inability to recollect all past information accurately. Only 53.7% of the total sample responded to the interview. The usable data represented 52.5% of the total sample, and this could have had negative implications for the reliability of the data. The authors indicated an improvement in the behavioral content of the model compared to existing models. The latter notwithstanding, the model specification was more complex than found in other studies, with far more explanatory variables. Further, in large metropolitan regions, the data collection effort to support such a model would involve very significant financial resources.

De Jong and Fox (2004) reviewed a broad range of car ownership models and classified them in to different categories. In aggregate time series model their review include the work done by Tanner and Button et al. using logistic function, Ingram and Liu's double logarithmic specification, the NRTF aggregate model, Dargay and Gately's Gompertz function to predict motorization rate on the basis of GDP/capita and some other explanatory variables for a large number of countries. They explained the trend of new products and income effects on car ownership in developed and developing countries. They also compared different types of functions with and without saturation levels. Another car ownership model type in the study is aggregate cohort model in which population is grouped in to cohorts and the behavior of car ownership is studied. Aggregate Cohort model studies explained the variation of car owning life style in different generations describing how cohorts acquire, keep, and lose cars with time and the influence of saturation level of license holding and income growth on car ownership.

Aggregate car market model studies were based on demand and supply of cars in the market. Models were developed considering car price, variation in income, and utility of using car. Some endogenous models were developed based on the prices of new and used cars. TREMOVE is a simulation model designed to analyze changes in transport flows across modes with changes in economic conditions. The model describes forecasts of transport flows, vehicle stocks and cost of vehicle usage and emissions. Another simulation model ALTRANS is developed for analyzing the environmental impact of car and public transport usage. Further FACTS and Zahavi Heuristic simulation models are developed for forecasting energy use and emissions. These models can be used to predict car ownership with some policy sensitivities.

2.3.4 Panel Studies

In the review conducted by De Jong and Fox (2004) several panel studies were analyzed. Kitamura developed a joint discrete and continuous model to determine car ownership and number of trips simultaneously using Dutch National Mobility Panel (LVO) data. The model is linear and it contained lag effects. An ordered response probit model for the number of cars in the household was developed using the same panel data. The variables specified included lagged variables and individual specific error components. Golob and Van Wissen developed an ordered response probit -model for car ownership combined with a standard tobit model for the continuous variable distance travelled. Meur's linear simultaneous equations of car ownership focused on effect of income. Nobile et al. estimated a random effects multinomial probit model of car ownership using longitudinal data collected in Netherlands which enables incorporation of both inter temporal and intra temporal dimensions. It is observed that most of the variability in choices could be attributed to between household differences rather than to within household random disturbances. Pseudo panel approach is a relatively new econometric approach to estimate dynamic demand models based on cohorts. Dargay and Vythoulkas developed a fixed effects model to estimate car ownership of a cohort using a pseudo panel dataset of 5-year cohort constructed from the UK family Expenditure Survey. Along with the socio economic characteristics and car purchase and running cost variables, generation effect and lagged dependent variable were included to estimate the dynamics of the model. It was noted that the earlier households tended to have a lower average car ownership rate over their lifetime than the ones born later and the number of cars of an average household depend on the number of cars in the previous year. This analysis was extended including area type. Hany and Dargay developed a quaternary ordered choice latent regression model. The model compared the state of car ownership a household was in last year with the state it is in this year and noted that last year's car ownership influences the current car ownership. Golounov developed a theoretical dynamic car ownership model based on lifetime utility. The utility in a period depends on the consumption of cars, other optional durables, long-term fixed purchases, and fixed day-to-day purchases in that period. An econometric model was estimated for the purchases and consumption of cars and other optimal durables considering financial assets and liabilities of the persons. The model used different brand-model-vintage combinations to explain car purchase behavior over time. Pseudo panel models are suitable to get short and long run policy sensitive forecasts of car ownership. Dynamic transaction models capture the changes in the car ownership states of the household. These models also suit for short and medium run policy sensitive forecasts.

CHAPTER 3

DATA AND DESCRIPTIVE ANALYSIS

3.1 Introduction

This chapter provides an exploratory analysis of the study database to identify variables within the study database that are associated with household trip generation and household vehicle ownership, respectively. The analysis consists of performing one-way and two-way cross tabulations of socio-economic, demographic, and spatial variables. The variables analyzed include household size, vehicle ownership, county of residence, lifecycle, land use density in origin and destination traffic zones, household income, total number of trips made by the members of a household, and number of vehicle-trips made by a household.

3.2 Survey

Good quality data are extremely important for all engineering investigations. The data for this research were drawn from the most recent travel survey conducted in the Nashville Metropolitan region. The survey was undertaken in1998, and the resulting database is referred to as the Nashville Travel Behavior Study database. The execution of the survey involved making prenotification calls, mailing advance letters, Computer–Assisted Telephone Interviewing program, and mailing diary packages. The survey obtained household travel behavior and travel related information from a stratified sample of households within the Nashville area and adjacent municipalities. Stratification was by county of residence. All told, 2,204 households were sampled from the 63,195 households in five counties. This research uses the data on household demographic, economic, and travel-related information for persons 5 years of age and older for the analysis.

3.3 Descriptive Analysis

Table 3-1 presents the household trip-rate by number of vehicles owned by household. A graphical representation of this information is presented in Figure 3-1.

Table 3-1: Trip rate by household Vehicle ownership

# of vehicles owned by household	Number of HH[1]	Trips made by all HH[1]	Trip-rate/ HH[1]
0	68	241	3.56
1	594	3045	5.12
2	872	7999	9.17
3	321	3210	9.99
4	108	1394	12.91
5+	37	502	13.72

HH ≡ Household(s)

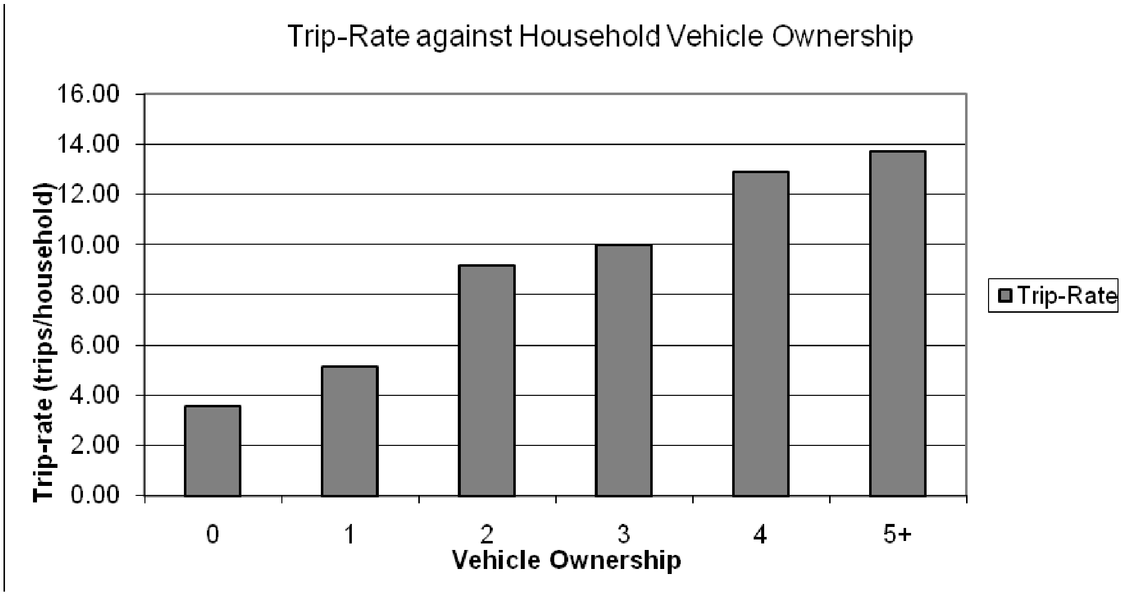

Figure 3-1: Trip rate against Household Vehicle Ownership

Vehicle Ownership is an important variable that influences each of the dimensions of travel modeled using the four step UTMS. The plot shows that households with higher levels of vehicle ownership have higher household trip-rate values. The general trend between trip-making and vehicle-ownership is the motivation for the primary hypothesis to be tested in this research; that is that households that have an expectation of making more trips tend to acquire more vehicles to be used in meeting these travel needs. Microeconomic consumer theory indicates that the consumer's budget serves as a constraint on how much the consumer can consume of vehicles and travel, and thus ultimately, the role of income would have to be factored into the analysis.

Household-size is another important household attribute that is associated with the travel behavior of the members of a household. Table 3-2 and Figure 3-2 show how both the household trip rate and person trip-rate vary with household size.

Table 3-2: Trip rate by Household Size

Household Size	Trips	Number of HH[1] in that category	Triprate/HH[1]	Trip-rate/person
1	1583	458	3.46	3.46
2	4871	762	6.39	3.20
3	3433	353	9.73	3.24
4	4009	295	13.61	3.40
5	1618	92	17.59	3.52
6	529	27	19.73	3.29
7+	349	14	25.65	

1. HH ≡ Household(s)

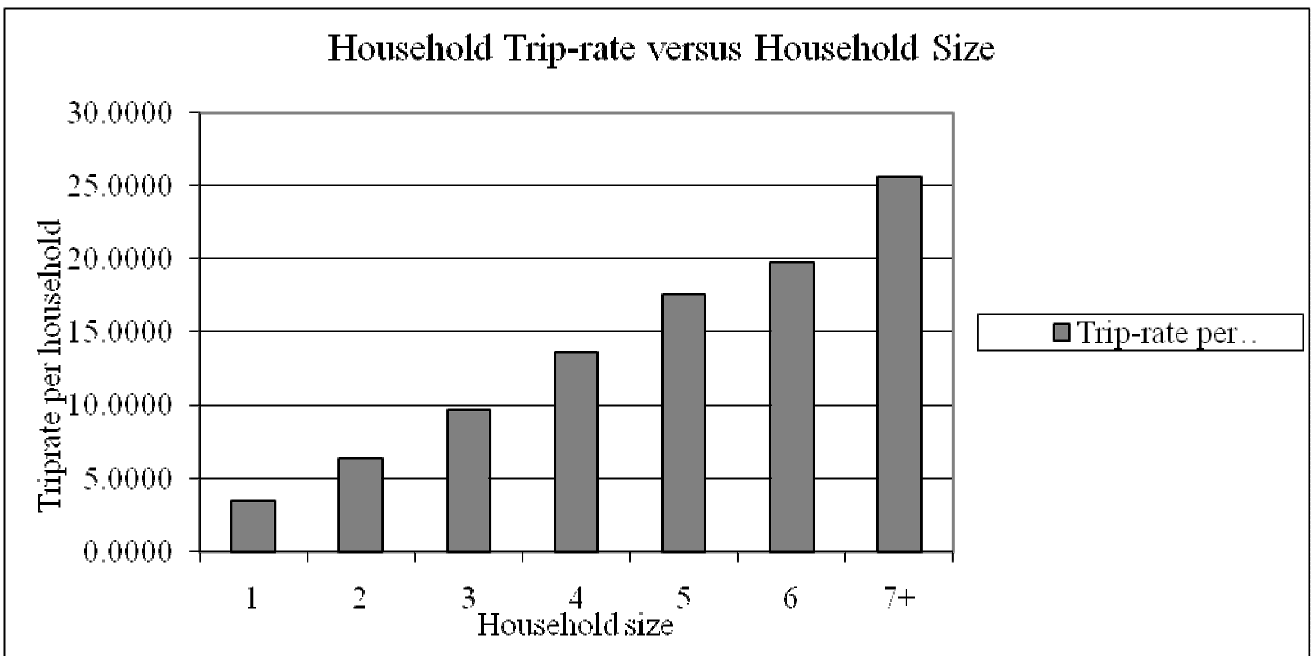

Figure 3-2: Household trip rate versus Household Size

A trend is observed between household trip rate and household size. The larger the household size the greater the number of trips on average that the household makes, rising from 3.46 trips /household for single person households to 25.65 trips/household for households consisting of seven persons or more.

The last column of Table 3-2 reports the person trip-rate for each household size. With the expectation of households with seven or more, the variation in the person trip-rate by household-size is relatively small, suggesting stability in the average number of non-home activities participated in by household-members irrespective of the size of the household.

Life cycle is another important variable that is associated with the travel behavior of the members of a household. Table 3-3 and Figure 3-3 show the household trip-rate by life cycle category.

Table 3-3: Household trip-rate by Life Cycle:

Life cycle group	Number of Households in group	Number of Trips	Trip-rate per household
Younger HH[1] with kids	198	2082	10.51
Younger HH[1] with no kids	233	1235	5.30
Older HH[1] with kids	526	7072	13.44
Older HH[1] with no kids	633	3931	6.21
Retired HH[1]	397	1980	4.99
Missing	12	91	7.44

1. HH ≡ Household(s)

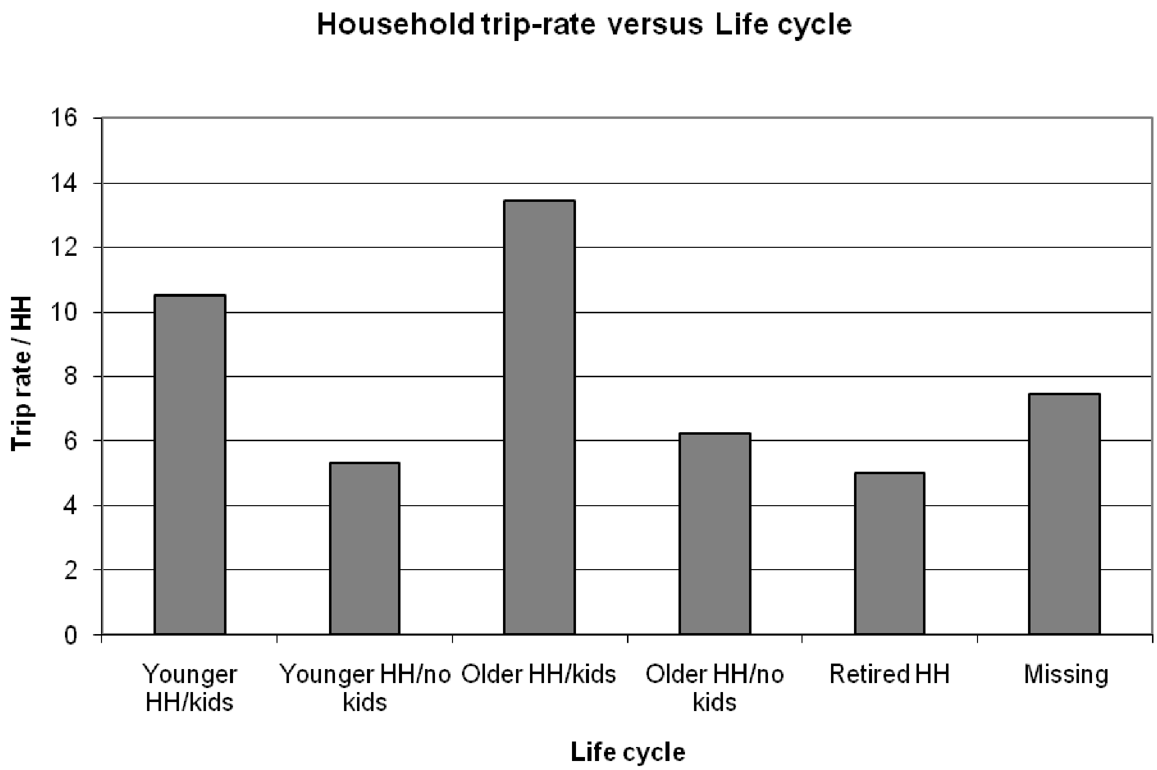

Figure 3-3: Household trip-rate versus life cycle

From the computed household trip-rates, it is observed that households with kids make on average more trips than households without kids. This observation is consistent with what one would expect based on intuition since kids attend school and participate in a variety of non-home activities resulting in more trips on average being made by such households with kids compared to households without kids. Depending on the age of the kids and the distance to be traveled, these trips could involve an adult member of the household, adding to the overall number of trips made.

Figure 3-3 also shows that households consisting of kids and older adult members make on average more trips compared to households comprising kids and with younger adult members.

This observation is consistent with what would be expected based on intuition since kids in households with older adults are also likely to be older and are therefore capable of generating additional trips themselves.

Households with retired members on average make the fewest trips daily. Again, this is to be expected since this group does not make mandatory trips (e.g., work trips) and, given the reduced household incomes that result from exit from the work force, on average do not participate in as many non-home activities.

The number of vehicles owned by a household is likely to depend on several factors including household size, population density, and vehicle prices. Very importantly, the decision of how many vehicles to own by a household is constrained by its budget. The expectation therefore is for households with higher income to own more vehicles on average than households with lower income. Table 3-4 shows the joint distribution of the number of vehicles owned by a household and household income category, while Figure 3-4 shows the frequency distribution of household vehicle ownership conditional on household income category.

As can be seen in the table, the percentage of households with no vehicle in general declines with increase in household income. There are no sampled households that had an income in excess of $75,000 and had no vehicle. The income frequency distribution of households conditional on vehicle ownership shifts to the right with increase in the number of vehicles owned. This is a direct result of (because) households with greater incomes on average owning more vehicles than households with lower incomes.

It is observed from Table 3-4 and Figure 3-4 that in the income groups less than $35k, the peak of the vehicle ownership frequency distribution occurs at one vehicle. The percentage of households though with two or more vehicles is higher the higher the income category. In the income groups greater than $35k, the peak of the vehicle ownership frequency distribution occurs at two vehicles. Again, as above, the percentage of households with two or more vehicles is higher the higher the income category.

The percentage of households with four or more vehicles as seen from the joint distribution is relatively small. This suggests that for modeling purposes, a plausible universal choice set of alternatives could have as elements 0, 1, 2, and 3 or more vehicles (denoted as 3+). An alternative plausible universal choice set, given the fairly steep drop in number of households in going from two vehicles to three or more vehicles, could have as elements 0, 1, and 2 or more vehicles (denoted as 2+).

Table 3-4a: Frequency distribution of households by number of vehicles owned and household income

# of Vehicles Owned by Household	<$10k	$10k-<$15k	$15k-<$25k	$25k-<$35k	$35k-<$50k	$50k-<$75k	$75k-<$100k	$100k-<$125k	$125-<$150k	$150k or more	Don't know	Refused
0	21	8	9	2	2	4	0	0	0	0	11	10
1	60	54	101	125	96	40	16	3	0	3	33	63
2	10	20	66	99	178	161	96	32	10	29	50	121
3	1	8	15	30	59	82	29	16	8	13	22	39
4	1	0	2	6	19	28	20	5	1	6	9	11
5	0	0	1	0	1	1	3	3	1	1	2	4
6	0	0	1	0	2	0	3	2	0	1	0	1
7	0	0	0	0	4	0	0	1	0	1	1	1

Table 3-4b: Trip rate of a household by number of vehicles owned and household income

# of Vehicles Owned by Household	<$10k	$10k-<$15k	$15k-<$25k	$25k-<$35k	$35k-<$50k	$50k-<$75k	$75k-<$100k	$100k-<$125k	$125-<$150k	$150k or more	Don’t know	Refused
0	2.05	3.74	5.74	7.82	16.58	4.51	0.00	0.00	0.00	0.00	2.41	1.94
1	4.55	4.44	4.38	5.73	6.67	5.74	5.44	8.98	0.00	3.00	3.82	3.99
2	4.74	5.75	7.44	8.94	9.71	10.18	11.05	9.56	10.90	10.96	8.05	7.40
3	13.00	12.15	6.20	9.11	10.44	10.84	10.75	9.63	12.01	10.96	7.63	9.40
4	7.16	0.00	8.09	12.21	10.73	15.40	13.37	9.53	22.27	10.16	16.04	11.13
5	0.00	0.00	6.45	0.00	15.00	9.50	13.15	21.97	11.00	22.50	9.84	15.91
6	0.00	0.00	0.00	10.00	9.43	11.00	15.98	35.60	0.00	16.22	10.00	9.00
7	0.00	0.00	0.00	0.00	10.21	0.00	0.00	5.00	0.00	15.50	6.00	6.00

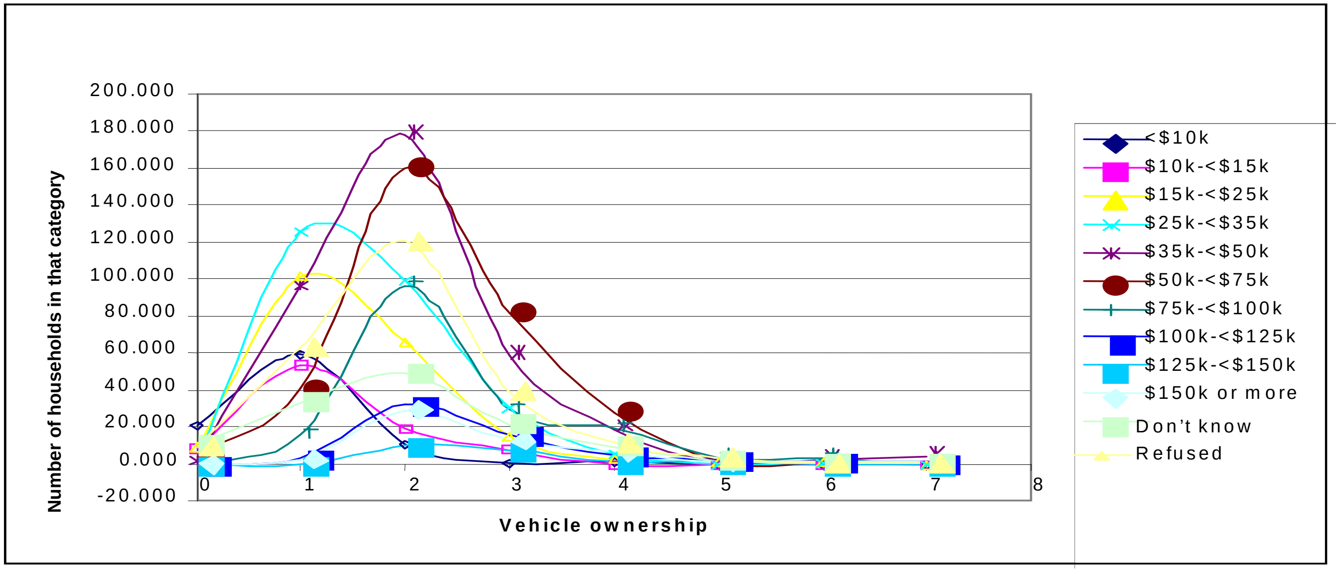

Figure 3-4: Frequency distribution of vehicle ownership conditional on household income.

County of residence of a household is a spatial factor that could be associated with socio- economic characteristics of households, the transit level of service experienced, and the average travel distance to activities. It therefore could be associated with trip generation and vehicle ownership level. Table 3-5 and Figure 3-5 show the distribution of the number of vehicles owned by county of residence.

Table 3-5a: Frequency distribution of households by vehicle ownership and county of residence

Number of vehicles owned by household	**County**				
	Davidson	**Rutherford**	**Sumner**	**Williamson**	**Wilson**
0	46	8	7	2	4
1	435	65	43	25	26
2	507	125	98	81	61
3	166	37	47	35	36
4	56	10	16	12	13
5	8	2	3	4	0
6	4	2	3	3	0
7	4	0	1	2	0

Table 3-5b: Household trip rate by vehicle ownership and county of residence

Number of vehicles owned by household	**Davidson**	**Rutherford**	**Sumner**	**Williamson**	**Wilson**
0	3.80	4.90	2.82	0.60	0.88
1	5.11	5.04	4.47	5.94	5.93
2	9.17	8.32	9.81	10.02	8.77
3	9.87	9.82	10.14	11.53	9.05
4	13.07	13.50	12.63	14.44	10.76
5	17.00	10.50	11.00	15.13	10.00
6	23.00	15.00	7.25	13.67	10.00
7	8.67	0.00	6.00	12.75	9.00

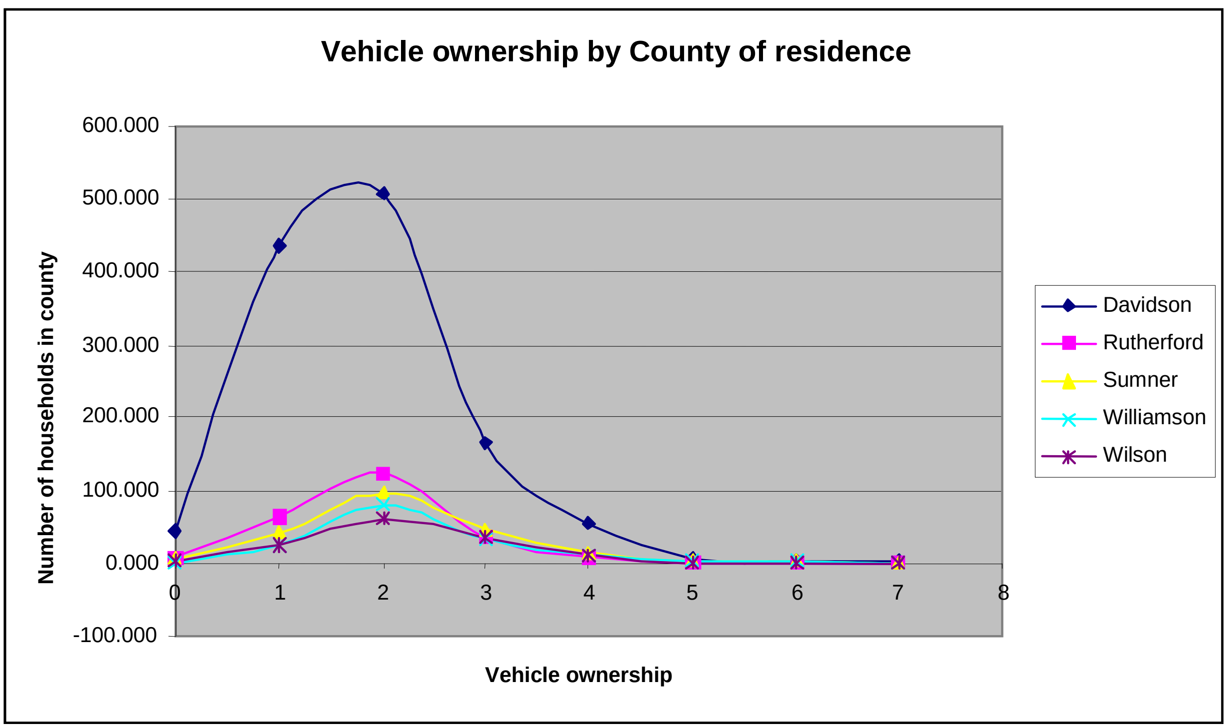

Figure 3-5: Vehicle ownership by county of residents

In Davidson and Rutherford Counties the majority of households own either one vehicle or two vehicles. These two counties happen to be the most urbanized in the metropolitan region. In the remaining counties, the pair of household vehicle ownership categories to which the majority of households belong are two and three vehicles. Though these counties also have substantial urbanized regions, their level of development is not to the extent as Davidson and Rutherford.

Davidson has the highest percentage of non-vehicle owning households. The comparatively better transit service in Davidson County would attract more of such households since they would have non-auto modal alternatives for accomplishing trips they need to make to participate in activities of interest to them.

Earlier analysis examined trip-rate by household size and vehicle ownership independently. Here, the distribution of households by size and vehicle ownership level is examined. Table 3-6 and Figure 3-6 show the frequency distribution of households by vehicle ownership and household size. The distribution shows that households of larger size tend to own more vehicles than households of smaller size. This observation appears to hold up to a vehicle ownership value of four beyond which the relatively small overall sample size results in cell frequencies that are low and hence do not permit any reliable statements to be made from the data. This is in part due to households of larger size on average having more members with a driver’s license leading to such households acquiring a greater number of vehicles for use in meeting the activity needs of these households. Clearly, household size alone is not the only determinant and, as has been discussed earlier, income and other variables influence the decision of how many vehicles to own.

Table 3-6a: Distribution of households by vehicle ownership and household size

Vehicle ownership	**Household Size**							
	1	2	3	4	5	6	7	8
0	39	16	6	5	1	0	0	0
1	360	156	48	22	3	4	1	0
2	41	453	159	162	43	9	5	0
3	13	104	106	66	25	5	2	0
4	3	23	27	29	15	7	4	0
5	0	5	3	5	2	1	1	0
6	1	2	2	3	3	1	0	0
7	2	2	0	2	0	0	0	0

Table 3-6b: Household trip-rate by vehicle ownership and household size

	Household Size							
Vehicle ownership	1	2	3	4	5	6	7	8
0	1.04	2.45	8.74	18.25	8.00	0.00	0.00	0.00
1	3.71	5.74	8.98	10.30	10.36	17.74	41.00	0.00
2	3.46	6.49	9.71	14.09	18.96	19.99	19.81	0.00
3	3.66	6.93	9.69	13.22	15.45	18.94	26.56	17.00
4	2.50	8.26	9.79	13.57	19.04	20.78	27.74	0.00
5	0.00	8.37	8.76	18.66	19.00	25.00	33.00	0.00
6	0.00	6.08	32.10	8.51	20.00	12.00	0.00	0.00
7	5.17	11.17	0.00	9.36	0.00	0.00	0.00	18.00

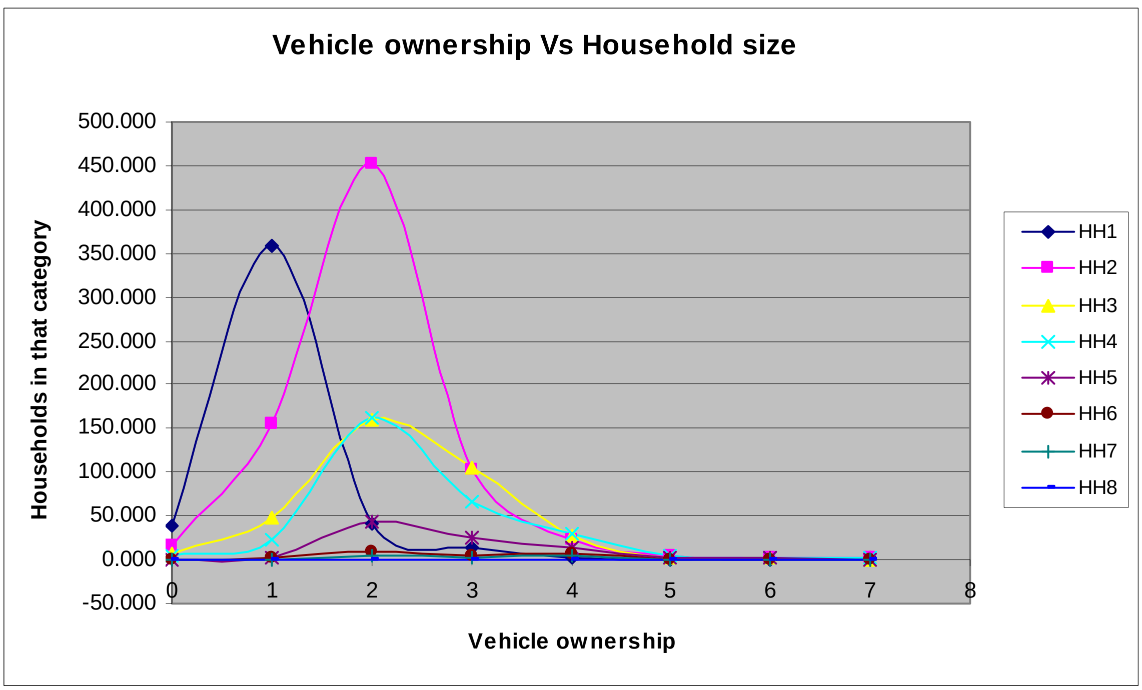

Figure 3-6: Frequency distribution of households by vehicle ownership conditional on household size

The joint frequency distribution of households by vehicle ownership level and life cycle is presented in Table 3-7 and Figure 3-7.

Table 3-7a: Vehicle Ownership by life cycle

Vehicle ownership	Younger HH/kids	Younger HH/no kids	Older HH/kids	Older HH/no kids	Retired HH	Missing
0	6	4	6	17	35	0
1	37	95	81	219	159	3
2	109	103	254	244	155	8
3	37	22	121	104	35	1
4	8	5	51	35	9	1
5	0	1	7	6	3	0
6	0	0	5	6	1	0
7	1	2	2	2	0	0

Table 3-7b: Household trip rate by vehicle ownership and life cycle

Vehicle ownership	Younger HH/kids	Younger HH/no kids	Older HH/kids	Older HH/no kids	Retired HH	Missing
0	11.17	2.33	14.05	2.34	1.24	0.00
1	7.98	4.76	9.37	4.24	3.74	4.00
2	10.93	5.63	14.13	6.63	6.23	8.13
3	11.43	6.41	12.99	7.50	7.84	11.00
4	12.69	4.22	16.62	10.57	6.32	6.00
5	10.00	12.00	18.19	11.99	13.00	0.00
6	0.00	0.00	15.91	15.84	10.00	0.00
7	6.00	4.68	14.60	12.26	0.00	0.00

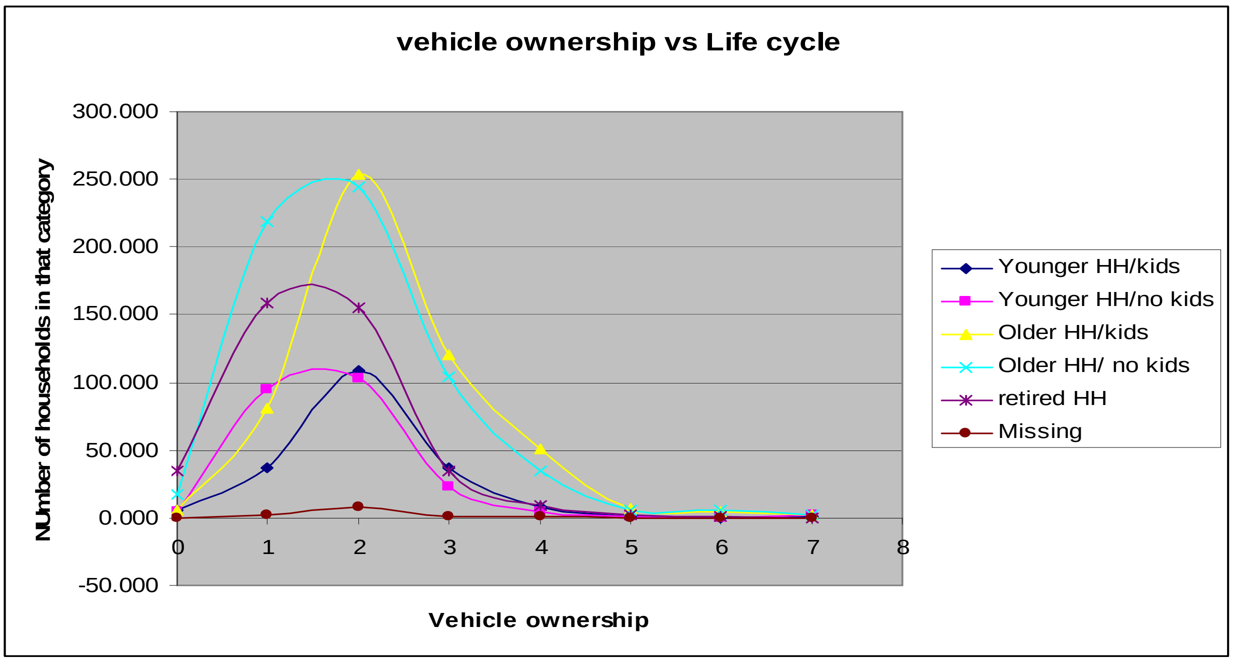

Figure 3-7: Frequency distribution of households by vehicle ownership conditional on life cycle of household

It is observed from the graph that there are many households with older members in the Nashville metropolitan region. Households with older members on average own more vehicles compared to households with younger members. As stated earlier, income is an important factor influencing consumption. The longer a person is in the workforce, the greater the experience they have, and on average, the better they are remunerated for their work. Thus, the expectation is for households with older members to own more vehicles than households with younger members.

Households with kids own more vehicles compared to households without kids. This is likely on account of the greater number of non-home activities households with kids participate in compared to households without kids. On an average most of the households with kids own two or three vehicles whereas the households without kids own one or two vehicles. Retired people in general own fewer vehicles, as the number of trips they make is also less.

Household vehicle ownership and household trip making are expected to be associated with household income. As explained earlier, this is because income serves as a check to consumption. Table 3-8 and Figure 3-8 provide a summary of the variation of vehicle ownership with household income.

Table 3-8: Household trip rate by household income

Income group	HH in that income group	Trips	Trip rate/HH
< $10K	93	380	4.10
$10 - <$15K	89	475	5.32
$15 - <$25K	196	1110	5.65
$25 - <$35K	264	1976	7.50
$35 - <$50K	361	3300	9.13
$50 - <$75K	317	3225	10.18
$75 - <$100K	168	1827	10.88
$100 - <$125K	62	665	10.73
$125K or more	74	821	11.15

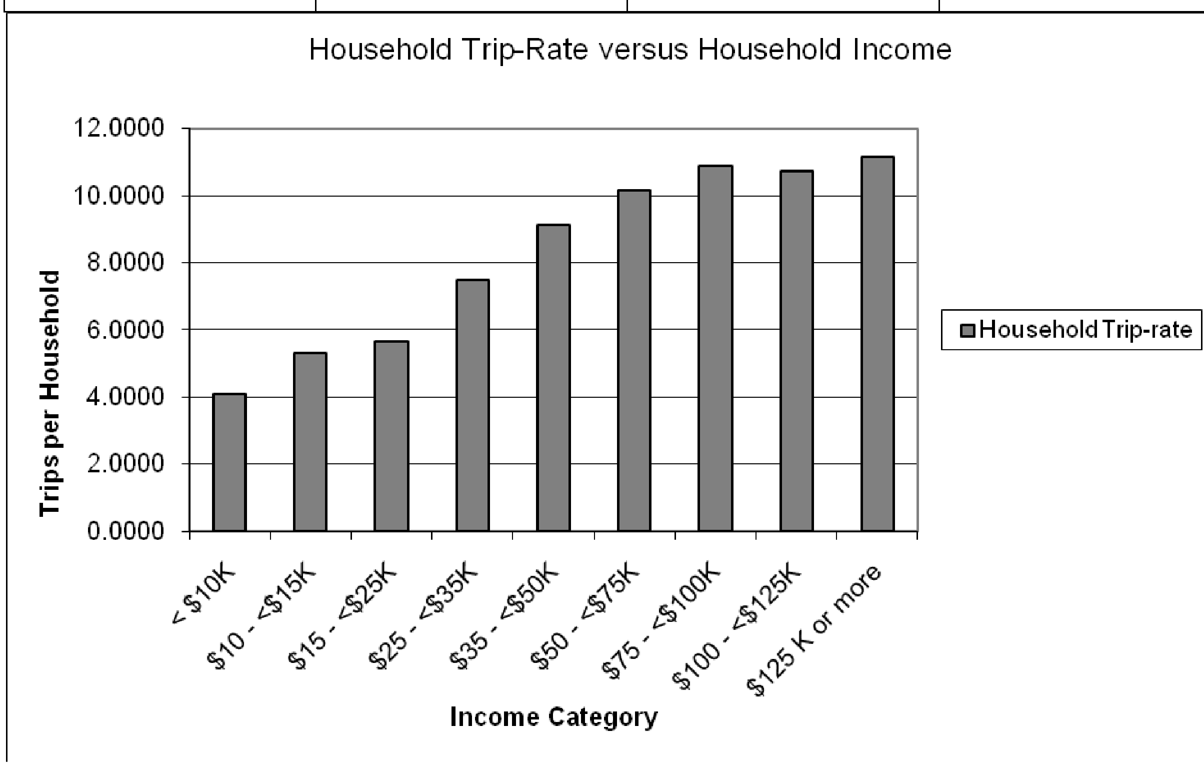

Figure 3-8 Household trip-rate by household income.

The computed trip-rates in the table show that on average, trip making increases with increasing household income. The lowest income group (less than $10k) had the lowest household trip rate. Beyond a household income of $75K, the household trip-rate became fairly stable. The survey sample had a disproportionate share of households with higher incomes. In the analysis, the weighting factors developed by the Nashville MPO are used to adjust the observations such that they reflect their actual shares in the population.

CHAPTER 4

METHODOLOGY

The primary objective of this research is to model trip generation with explicit accounting of its interrelationship with household vehicle ownership. The results of the descriptive analysis in Chapter 3 provide guidance on which variables in the dataset to consider in the modeling phase and, possibly, how these variables should be specified in the model to be developed. This chapter provides a description of the current approaches to modeling household trip generation and household vehicle ownership respectively, and a description of a model that accounts for the interdependency in the mobility and travel decisions.

4.1 Modeling Household Trip Generation and Household Vehicle Ownership Level

The conventional approach to modelling trip generation is premised on the supposition that characteristics of a household are the factors associated with the number of trips the households' members make daily. Thus, the procedure consists of first identifying the covariates, denoted by the vector $\mathbf{x}_h$, likely to be associated with y_h, the number of trips made for a given purpose by each household *h*. Mathematically, this relationship may be expressed as

$$y_h = f(\mathbf{x}_h) \quad [1]$$

The covariates, as stated above, typically comprise characteristics of the modelling unit, which in all the studies encountered in the literature review includes the number of vehicles owned by the household.

Then second, a linear statistical model between the number of trips made by the household (dependent variable) and the vector of explanatory variables (covariates) describing the modelling unit is postulated. Mathematically, this results in Equation [1] above being expressed as

$$y_h = a_0 + a_1 x_{ih} + ... + a_k x_{kh} + \varepsilon_h \qquad h = 1, \ldots\ldots, T \quad [2]$$

where

T is the number of households in the sample; and

a_j, $j = 0, \ldots, k$ are the unknown model parameters.

The method of least squares is used to obtain estimates of the unknown model parameters, yielding a linear model

$$\hat{y}_h = \hat{a}_0 + \hat{a}_1 x_{1h} + ... + \hat{a}_k x_{kh} \qquad h = 1, \ldots\ldots., T \qquad [3]$$

The estimated model in equation [3] above predicts the expected number of trips produced by each household h.

In the case of modeling household vehicle ownership level with cross-sectional data, the state of the art and state of the practice has been to use the multinomial logit model (Cambridge Systematics et al., 2007). The theory underpinning this model is in consumer theory in economics. The theory states that when a decision-maker has to select an alternative from a set of feasible alternatives, the decision-maker would first identify the characteristics of the alternatives that influence their choice. From these characteristics, the decision-maker would obtain an overall measure of attractiveness of each alternative. The decision-maker is then posited to select the alternative with the highest utility. By this theory then, when a household *h* is faced with selecting a number of vehicles to own, the household would first identify the characteristics associated with each vehicle ownership level *i*. Denote the vector of these characteristics for household *h* by $\boldsymbol{x}_{ih}$. Second, the household would then evaluate the attractiveness of each level of vehicle ownership. This measure of attractiveness of an alternative to a household is referred to as its utility. For a household *h*, let the utility derived from owning *i* vehicles be denoted by U_{ih}, where $i = 0, 1, 2^+$. Finally, the household is assumed to select the vehicle ownership level that gives the household the highest utility. It is noted here that it is the transportation analyst that seeks to predict the choice of vehicle ownership level of each household. The analyst, however, is not able to observe fully all the factors leading to a household's decision of alternative. Hence, the analyst is not able to state with certainty the choice of each household. On account of this, the utility is assumed to consist of two additive components: a systematic component denoted by V_{ih}, which is a function of the observed characteristics of alternative *i*, and a random component denoted by ε_{ih}. Mathematically, the latter can be expressed as

$$U_{ih} = V_{ih} + \varepsilon_{ih} \qquad i = 0, 1, 2+; \quad h = 1,\ldots,T \quad [4]$$

Given the definition of V_{ih} above, the indirect utility function in equation [4] above can be expressed as

$$U_{ih} = \boldsymbol{\theta}_i \mathbf{x}_{ih} + \varepsilon_{ih} \qquad i = 0, 1, 2+; \quad h = 1,\ldots,T \qquad [5]$$

where

U_{ih} = total indirect utility provided by vehicle ownership level i

$\mathbf{x}_{ih}$ = vector of characteristics of household h associated with vehicle ownership level i

ε_{ih} = a random term accounting for unobserved factors

$\boldsymbol{\theta}_i$ = a vector of estimable parameters

T = total number of households

If the ε_{ih} 's are assumed to be distributed according to Gumbel's Type I Extreme Value distribution, then the selection probabilities of the alternative vehicle ownership levels is given by the standard multinomial logit model

$$P_{ih} = \frac{e^{V_{ih}}}{\sum_{j \in C_h} e^{V_{jh}}} \qquad \text{where } i \in C_h, \qquad [6]$$

where P_{ih} is the probability of household h selecting vehicle ownership level i and V_{ih} is the deterministic component of the indirect utility U_{ih}.

4.1.1 Joint Discrete/Continuous Model

As discussed in the introductory section, the above model structure is premised on the assumption that the number of vehicles a household decides to own is determined independently of the total number of trips its constituent members expect to make daily, with the consequence that vehicle ownership level of the household is determined exogenously to the four-step UTMS.

Further, policies designed to reduce the number of trips a household makes by car would by the above model structure have no impact on the vehicles owned by the household since vehicle ownership is determined outside the travel-forecasting phase. This structure, as argued earlier, is not consistent with the decision-making process households in general follow to determine levels of consumption of these dimensions of mobility and travel.

The working hypothesis of this thesis is that the expected number of trips made daily by a household is dependent on the vehicle ownership level decision of a household, an endogenously determined variable. By this, the two decisions are interdependent. Therefore, an appropriate structural system of equations has to be formulated to reflect this interdependency in decisions. This is achieved by specifying the systematic component of the indirect utility to be a function of the expected number of trips to be made by a household *h* in addition to the characteristics associated with vehicle ownership level *i*. Mathematically, the latter leads to the total utility function being expressed as

$$U_{ih} = \boldsymbol{\theta}_i \mathbf{x}_{ih} + \phi y_{ih} + \varepsilon_{ih} \qquad i = 0, 1, 2+; \quad h = 1, \ldots, T \quad [7]$$

where

y_{ih} = expected number of daily trips made by household h with vehicle ownership level i

ϕ = coefficient of y_{ih}, the expected number of daily trips made by household h with vehicle ownership level i

All other variables are as previously defined.

The expected number of daily trips made by household *h* with a vehicle ownership level *i* can be expressed as

$$y_{ih} = \boldsymbol{\beta}_i \mathbf{z}_h + v_h \qquad [8]$$

where

y_{ih} = number of trips made by household *h* with a vehicle ownership level *i*;

$\mathbf{z}_h$ = vector of socioeconomic characteristics of household *h*;

v_h = unobserved characteristics of household *h* that influence trip generation; and

$\boldsymbol{\beta}_i$ = vector of parameters that vary by vehicle ownership level (i = 0, 1, 2^+).

The inclusion of y_{ih} in the indirect utility function reflects, logically, the significance of trip generation in determining vehicle ownership level. However, because y_{ih} is endogenous to the vehicle ownership choice process and observed for only the chosen vehicle ownership level, the reduced form of the indirect utility function is given by

$$U_{ih} = \boldsymbol{\theta}_i \mathbf{x}_{ih} + \phi \boldsymbol{\beta}_i \mathbf{z}_h + \phi v_h + \varepsilon_{ih} \qquad [9]$$

If the ε_{ih}'s are assumed to be distributed according to Gumbel's Type I Extreme Value distribution, then the selection probabilities of the alternative vehicle ownership levels is given by the standard multinomial logit model

$$P_{ih} = \frac{e^{V_{ih}}}{\sum_{j \in C_h} e^{V_{jh}}} \qquad \text{where } i \in C_h \qquad [10]$$

where

P_{ih} = probability of household *h* selecting vehicle ownership level *i* and

V_{ih} = the deterministic component of the indirect utility U_{ih}.

Note here that because the error term ϕv_h in Equation [9] does not vary across vehicle alternatives (*i*'s) it does not affect the assumption of a logit model structure.

With the choice model structure above, the selectivity bias corrected trip generation equations can be expressed as (Hay, 1980; Dubin and McFadden, 1984; Mannering, 1986),

$$y_{ih} = \beta_i z_h + \alpha_h \lambda_{ih} + \eta_i \qquad [11]$$

where

$\alpha_i = (-1)^{k+1}\left(\sigma 6 \rho_i / \pi^2 \right)$, a parameter estimable by OLS, and

$$\lambda_{ih} = \left(1/K \right) \sum_{k \neq i}^{K} \left\{ \left[P_{kh} \ln P_{kh} / \left(1 - P_{kh} \right) \right] + \ln P_{ih} \right\} \qquad [12]$$

Ordinary least squares estimates of Equation [11] has been proven to yield consistent parameter estimates (Mannering, 1986).

4.1.2 Joint Discrete-Continuous Model without Stratification

Alternatively, the modeling system can be developed without segmentation into the different vehicle ownership categories. The starting point of this method is Equation [7] above, which has the systematic component of the indirect utility function specified to include the number of trips to be made by a household *h*, y_h, in addition to the other characteristics associated with a vehicle ownership level *i*. This is modified such that trip generation is not conditional on household auto ownership level. Given that y_h will not differ with alternative, it is to be specified as alternative specific in the utility functions. Mathematically, the latter leads to the total utility function being expressed as

$$U_{ih} = \boldsymbol{\theta}_i \mathbf{x}_{ih} + \phi_i y_h + \varepsilon_{ih} \qquad i = 0, 1, 2; \quad h = 1,\ldots,T \qquad [13]$$

where

y_h = number of trips made daily by members of household *h*

ϕ_i = coefficient of y_h, the number of daily trips made by members of household *h*

All other variables are as previously defined.

The expected number of daily trips made by household *h* can be expressed as

$$y_h = \boldsymbol{\beta}' \mathbf{z}_h + v_h \quad [14]$$

where

y_{ih} = number of trips made by household *h* with a vehicle ownership level *i*;

$\mathbf{z}_h$ = vector of socioeconomic characteristics of household *h*;

ν_h = unobserved characteristics of household *h* that influence trip generation; and

$\boldsymbol{\beta}'$ = transposed vector of parameters that vary by vehicle ownership level (*i* = 0, 1, 2).

The inclusion of y_{ih} in the indirect utility function reflects the logical significance of trip generation in determining vehicle ownership level. Given the definition of y_h in Equation [14] above, the reduced form of the indirect utility function is given by

$$U_{ih} = \boldsymbol{\theta}'\mathbf{x}_{ih} + \phi_i \boldsymbol{\beta}'\mathbf{z}_h + \phi_i \nu_h + \varepsilon_{ih} \qquad [15]$$

The assumption of the ε_{ih}'s to be distributed according to Gumbel's Type I Extreme Value distribution leads to the selection probabilities of the alternative vehicle ownership levels to be determined by the standard multinomial logit model

$$P_{ih} = \frac{e^{V_{ih}}}{\sum_{j \in C_h} e^{V_{jh}}} \qquad \text{where } i \in C_h \qquad [16]$$

where

P_{ih} = probability of household *h* selecting vehicle ownership level *i* and

V_{ih} = the deterministic component of the indirect utility U_{ih}.

Since the actual vehicle ownership level of a household *h* is not known with certainty, the expected vehicle ownership level is computed using as probabilities the selection probabilities given by the logit model in equation [16]. Mathematically, this is expressed as:

$$E[i_h] = \sum_{i=0}^{2} iP_{ih} \quad [17]$$

The number of daily trips made by household *h* can then be estimated as a function of $E[i_h]$ and the other socioeconomic and spatial characteristics of the household. Equation [14] is modified to include the additional term of expected household vehicle ownership as follows

$$y_h = \boldsymbol{\beta}\mathbf{z}_h + \psi E(i_h) + v_h \quad [18]$$

where

y_h = number of trips made by household *h* with a vehicle ownership level *i*;

$\mathbf{z}_h$ = vector of socioeconomic characteristics of household *h*;

v_h = unobserved characteristics of household *h* that influence trip generation;

$\boldsymbol{\beta}$ = vector of parameters that vary by vehicle ownership level (i = 0, 1, 2); and

ψ = coefficient of the expected number of vehicles owned by household *h*;

4.2 Model Assessment

4.2.1 Goodness of Fit Measure for Vehicle Ownership (Discrete Choice) Model

The vehicle ownership model will be assessed by its ability to replicate the vehicle ownership level selected by each household in the estimation dataset. The index that indicates how well the model predicts at this disaggregate level is the likelihood ratio index, ρ^2.

Mathematically, it is expressed as

$$\rho^2 = 1 - \frac{L(\hat{\boldsymbol{\beta}})}{L(0)} \quad [19]$$

where

$L(\hat{\boldsymbol{\beta}})$ = the log-likelihood value with the model parameters at their converged values

$L(0)$ = the log-likelihood value with no model parameters specified

ρ^2 takes on values between 0 and 1. The closer the value is to 1, the better the model.

4.2.2 Goodness of Fit Measure for Household Trip Generation Model

The goodness of fit measure to assess how well the linear model of trip generation is able to predict the number of trips made by each household is the coefficient of determination, R^2.

$$R^2 = 1 - \frac{\sum_{\forall h} (y_h - \hat{y}_h)^2}{\sum_{\forall h} (y_h - \bar{y})^2} \quad [20]$$

where

y_h = the observed number of trips made by household h

$\hat{y}_h$ = the predicted number of trips made by household h

$\bar{y}$ = the average of the number of trips made by all the households

R^2 also takes on values between 0 and 1. All else being equal, the closer model's R^2 value is to 1 the better the model.

4.2.3 Signs and Statistical Significance of the Estimated Coefficients

The models will also be assessed on sign of each estimated parameter, and on their statistical significance. The latter will be accomplished by comparing the estimated t-values with critical t-values obtained from tables.

CHAPTER 5

RESULTS AND DISCUSSION OF RESULTS

This chapter presents and discusses the results of estimating a joint household vehicle ownership model and household trip generation model using household data collected in the Nashville region. The model structure, as discussed in Chapter 4, captures the interdependency between the decisions of how many vehicles to own and how many trips to produce.

5.1 Variables

Several potential explanatory variables were considered in the specifications of the various models developed. These variables are discussed in turn below.

Household size: All else being equal, households of larger size are expected to participate in a greater number of activities than households of smaller size. Thus, the expected sign of the coefficient estimate of this variable in the trip generation equations is positive. Its sign in the vehicle ownership model is also expected to be positive since the more activities there are to participate in, the greater the need would be for more vehicles to help accomplish all of the activities.

Number of workers in the household: All else being equal, a household with more workers will participate in more work trips than a household with fewer workers. This variable is also related to the household income in that a household with two workers would have greater disposable income than one that has one or no workers and so forth. The greater disposable income of multi-worker households makes it possible for them to participate in more non-home activities than households that have less income. The expectation is for the sign of this variable to be positive in the trip generation equations as well as the vehicle ownership models.

Household income: Income is a key socioeconomic characteristic that influences virtually all aspects of travel characteristics of the household. As discussed above, household consumption is tempered by the budget constraint. The greater the household income, the greater the number of vehicles the household can acquire, and the greater the number of non-home activities that the household can participate in. This variable though is likely to be strongly correlated with the variable “number of workers in a household.” This would mean that both might not be included in the same model specification. The expectation is for the sign of the coefficient estimate of this variable to be positive in both the vehicle ownership model and the trip generation model.

Location of household: The Nashville travel database does not have information on interzonal modal travel times and travel costs. This means modal accessibilities cannot be computed to determine what impact they have on these two dimensions of mobility and travel. A good surrogate variable under these circumstances is the distinction of the location of a household, that is, whether rural or urban. Urban areas generally tend to have a higher transit level of service than rural areas and therefore present the opportunity for some trips to be undertaken by transit. If the transit level of level service is of a very high quality, it could lead a household to acquire a smaller number of vehicles than would otherwise be the case. The quality of Nashville’s transit service cannot be described as “high” hence, the expectation is not for this variable to play a particularly strong role in influencing both trip generation and vehicle ownership.

<u>Life cycle status of a family</u>: All else being equal, families with kids on average participate in more non-home activities than families without kids. This latter statement is supported by the trip-rate analysis presented in Chapter 3. Further, families with older kids on average make more trips than families with younger kids. This variable has the potential to be associated with trip generation and vehicle ownership. However, its impact will have to be determined from simultaneous analysis of all the potential explanatory variables.

<u>Vehicle ownership of a household</u>: the number of vehicles a household owns is influenced by several variables including transport system characteristics, income, and the expected number of activities a household desires to participate in. The expected number of activities a household participates in is also dependent on the expected vehicle ownership level. Thus in this research, household vehicle ownership is an endogenously determined attribute.

<u>Number of household members licensed to drive</u>: this household variable could be associated with the number of vehicles a household acquires. The greater the number of persons in a household licensed to drive the more likely the household is to owning multiple vehicles. The actual role of this variable in influencing both vehicle ownership and trip generation will be determined from an analysis considering all the potential explanatory variables simultaneously.

<u>Rent or own household</u>: This variable was considered as a possible surrogate to the household income variable. Families that own their home were thought to be financially more stable than those that rented. Hence, the expectation was for households that own to make more trips and own more vehicles. Again, the actual role of this variable in influencing both vehicle ownership

level decisions and trip generation will be determined from an analysis considering all the potential explanatory variables simultaneously.

Each variable discussed above is given a name for modeling purposes. These names are defined in Table 5-1 below. In addition, the values that each of these variables can take is indicated in the definition of the variables.

Table 5-1: Definition of variables specified in the Vehicle Ownership Models

Variable	Definition
NWRKH	Number of household members that are employed. It is included in the utility function of the 2-vehicle alternative
HSIZE	Size of household - It is included in the utility function of the 2-vehicle alternative
DINC1	Dummy variable of household income; = 1 if income category of household exceeds 5 (household income is at least $50,000); =0 otherwise Variable is specified as an alternative-specific in the utility function of the 2-vehicle alternative
DINC2	Income of household variable; =1 if income category of household is between category 4 ($25,000) and category 5 ($50,000) inclusive =0 otherwise Variable is specified as an alternative-specific socio-economic variable in the utility function of the 1-vehicle alternative
DINC3	Income of household variable; =1 if income category of household is less than or equal to category 3 ($15,000) =0 otherwise Variable is specified as an alternative-specific socio-economic variable in the utility function of the 0-vehicle alternative
NLIC	Number of household members possessing a driver’s license. Variable is specified as an alternative specific socioeconomic variable in the utility function of the 2-vehicle alternative

DRURAL	Dummy household location variable = 1 if Urban area; = 0 Otherwise Variables is specified as alternative-specific in the utility function of the 0-vehicle alternative
DRENT	Dummy variable indicating home ownership or renting; =1 if Owning = 0 Otherwise Variable is specified as an alternative-specific in the utility function of the 2-vehicle alternative
INCP	Household income divided by household size. Specified as alternative specific in the 1-vehicle utility function (INCP1) and as alternative specific in the 2-vehicle utility function (INCP2)

5.2 Model Estimation Output – Model with Selectivity Bias Correction Term

The coefficients of the variables specified in the vehicle ownership model and their corresponding t-values are estimated using N-LOGIT 3.0 software (Greene, 2003). The results are presented in Table 5-2 below.

Table 5-2: Estimated coefficients of the variables in the model.

Variable	Coefficient	t-value	Coefficient	t-value
NWRKH	0.3822	3.331	0.24019	1.965
HSIZE	0.3902	4.563	0.3683	4.174
DINC1	1.2452	5.909	1.1037	6.559
DINC2	0.0521	0.309	-----	-----
NLIC	2.2295	13.572	2.2768	12.711
DRURAL	0.1081	0.241	-----	------
DRENT	0.7476	4.277	0.5657	3.002
A_ZVEH	2.6425	5.208	3.0119	8.794
A_ONEVEH	4.9982	17.430	5.2939	16.826
Likelihood Ratio Index- ρ^2	0.604		0.527	

The most basic test of the model estimation output is on the signs that the estimated variable-coefficients have. Based on travel behavior theory, a sign is expected for each estimated coefficient and these are compared with estimation results for consistency.

First, the estimates of the coefficients of the model that includes most of the potential explanatory variables are discussed. These coefficients are presented in the second column of Table 5-2.

The estimated coefficient of the total employment in a household variable (NWRKH) is positive, which is in agreement with intuitive expectation. This indicates that with increase in the number of working members of a household, there is an increase in the utility and probability of owning one or more vehicles.

The coefficient of the household size variable (HSIZE) in the estimated model is positive. This indicates that household of large size all else being equal find it more attractive to own more vehicles than households that have smaller size. As discussed in Section 5.1 above, this is consistent with what one would expect based on theory.

Two dummy income variables are specified in the model. One is for households with income of at least $50,000 (DINC1) and the second is a dummy variable for households with income between $25,000 and $50,000 (DINC2). They both have the appropriate positive sign, indicating that greater household income increases the probability of owning multiple vehicles. The magnitude of the coefficient of DINC1 is much larger than that of DINC2 and of greater statistical significance.

The estimated coefficient of the number of household members licensed to drive (NLIC) has the positive sign, which is consistent with intuitive expectation. This indicates that households with higher numbers of licensed drivers are likely to also own more vehicles.

The estimated coefficient of the spatial dummy variable indicating the location of a household, that is, whether it is located in rural area or urban area (DRURAL) is positive. This is consistent with what one would expect based on travel behavior theory. The coefficient estimate though is not statistically significant at the 5 percent level of significance. As discussed in Section 5.1, given that the quality of transit service in the Nashville metropolitan region is not uniformly high, this variable was not expected to impact significantly on vehicle ownership choice or trip generation.

The estimated coefficient of DRENT has the positive sign and is statistically significant at the 5 percent level of significance. It presumably captures an aspect of financial stability over and above the effect of household income category – it is noted that a dummy variable for household income is included in the specification of the utility function for the 2-vehicle alternative.

The two variables whose coefficient estimates are not statistically significant are DINC2 and DRURAL. They were therefore dropped from the final model specification whose model estimation results are presented in columns 4 and 5 of Table 5-1.

The estimated alternative specific constant terms, A_ZVEH and A_ONEVEH, represent the effect of factors influencing household vehicle ownership level that were not observed and therefore explicitly included in the utility functions. There is no prior expectation on the sign of these constant terms based on theory – each could emerge with a positive or negative sign. The p-values of these coefficients indicate that they are statistically significant, indicating the exclusion of important variables in the model specification.

The likelihood ratio index is an overall goodness of fit measure, which indicates how well the model is able to predict the choices of vehicle ownership level made by the various households. The final model has a likelihood ratio index of 0.527. In travel behavior modeling, this represents a very good model.

The household models of trip generation were estimated using linear regression. Given the possible bias that could occur on account of household self-selectivity to the sub-samples corresponding to the different vehicle ownership groups, a second trip generation model was estimated for each vehicle ownership level that had a selectivity bias correction term included in the model specification. The estimation results for the three sets of trip generation models are presented in Tables 5-3 to 5-5.

Table 5-3: Estimates of coefficients of trip generation variables for households owning zero vehicles

Variable	Uncorrected			Corrected		
	Coefficient	t-value	P-value	Coefficient	t-value	P-value
NWRKH	1.1955	1.2707	0.2107	1.5652	1.5966	0.1179
HSIZE	4.0794	5.8353	6.35E-07	4.3055	5.9989	3.986E-07
INCOME	0.2888	0.8085	0.4233	0.4892	1.2561	0.2160
NLIC	-0.2799	-0.3564	0.7233	0.8241	0.6991	0.4882
DRURAL	0.3822	0.2527	0.8017	0.3706	0.2466	0.8064
Intercept	-5.0414	-1.6882	0.0986	-1.5426	-0.3782	0.7072
R^2	0.6710			0.6829		
Adj R^2	0.6328			0.6376		

Table 5-4: Estimates of coefficients of trip generation variables for households owning one vehicle

Variable	Uncorrected			Corrected		
	Coefficient	t-value	P-value	Coefficient	t-value	P-value
NWRKH	0.7714	2.2526	0.0247	0.8335	2.3425	0.0196
HSIZE	2.3455	8.9552	7.711E-18	2.4322	8.2638	1.43E-15
INCOME	0.2488	2.0071	0.0453	0.2790	2.1054	0.0358
NLIC	0.1261	0.2582	0.7963	0.5703	0.6772	0.4986
DRURAL	-0.6879	-1.1797	0.2387	-0.7121	-1.2180	0.2238
Intercept	1.0559	0.8352	0.4040	0.6108	0.4243	0.6716
R^2	0.2832			0.2839		
Adj R^2	0.2757			0.2748		

Table 5-5: Estimates of coefficients of trip generation variables for households owning two vehicles

Variable	Uncorrected			Corrected		
	Coefficient	t-value	P-value	Coefficient	t-value	P-value
NWRKH	0.3403	1.3068	0.19151	0.3598	1.3787	0.16824
HSIZE	3.2111	19.0191	5.11E-71	3.2470	18.8916	3.4E-70
INCOME	0.3582	3.4616	0.0006	0.4039	3.6281	0.0003
NLIC	0.1630	0.4726	0.6366	0.3442	0.9024	0.3670
DRURAL	1.1728	2.5792	0.01006	1.14502	2.5147	0.0120
Intercept	-4.5185	-3.8157	0.00016	-5.4336	-3.7683	0.0002
R^2	0.3042			0.3049		
Adj R^2	0.3014			0.3015		

Table 5-3: Estimates of coefficients of trip generation variables for households owning zero vehicles

Variable	Uncorrected			Corrected		
	Coefficient	t-value	P-value	Coefficient	t-value	P-value
NWRKH	1.3798	1.7063	0.0948	1.7455	1.8681	0.0684
HSIZE	3.9867	6.3615	9.06E-08	4.2426	5.996079	3.42E-07
DRURAL	0.2402	0.1621	0.8719	0.2693	0.1809	0.8572
C-term	-----	-----	-----	1.1129	0.7912	0.4331
Intercept	-4.2880	-1.5322	0.1325	-2.4498	-0.6718	0.5052
R^2	0.6654			0.6701		
Adj R^2	0.6431			0.6401		

Table 5-4: Estimates of coefficients of trip generation variables for households owning one vehicle

Variable	Uncorrected			Corrected		
	Coefficient	t-value	P-value	Coefficient	t-value	P-value
NWRKH	0.7847	2.3235	0.0206	0.7976	2.2790	0.0231
HSIZE	2.3797	10.4009	5.78E-23	2.4051	8.2699	1.36E-15
INCOME	0.2510	2.0242	0.0435	0.2548	2.0073	0.0453
DRURAL	-0.6946	-1.1916	0.2340	-0.7010	-1.1978	0.2316
C-term	-----	-----	-----	0.0762	0.1420	0.8871
Intercept	1.1458	0.9447	0.3453	1.1344	0.9323	0.3517
R^2	0.2830			0.2831		
Adj R^2	0.2770			0.2755		

Table 5-5: Estimates of coefficients of trip generation variables for households owning two vehicles

Variable	Uncorrected			Corrected		
	Coefficient	t-value	P-value	Coefficient	t-value	P-value
NWRKH	0.4340	1.3426	0.1798	0.5231	1.5662	0.1177
HSIZE	3.3792	16.8153	2.24E-54	3.4718	15.8407	3.36E-49
INCOME	0.3691	2.8115	0.0051	0.4287	3.0016	0.0028
DRURAL	1.5678	2.6641	0.0079	1.5332	2.6014	0.0095
C-term	-----	-----	-----	-0.6817	-1.0589	0.2899
Intercept	-5.5246	-3.7522	0.0002	-6.4475	-3.7686	0.0002
R^2	0.2983			0.2993		
Adj R^2	0.2947			0.2948		

The results show that the sign of each of the estimated coefficients is in agreement with what one would expect based on travel behavior theory. Household size and number of working household members are amongst the important variables that influence the number of trips a household makes on a typical day. This indicates that being a worker has an effect on trip making that is over and above the effect of being a household member. This is important since most models of trip generation are usually based on household size only. Household income is also an important variable influencing the number of trips made daily by a household. However, its significance is diminished for households that do not own a vehicle. The spatial location of a household is also associated with the number of trips, with households in rural areas, all else being equal, generating more trips compared to households in the urban region.

The selectivity bias correction terms have relatively low statistical significance, suggesting that for this dataset the extent of bias is rather small. This could be due to the relatively small size of sample used in the study or the diversity of households in each household vehicle ownership group.

5.3 Model Estimation Output – Discrete-Continuous Model without Stratification

5.3.1 Vehicle Ownership Model

Table 5-6 presents the model estimation output for the discrete component of the discrete-continuous model of household vehicle ownership level and trip generation. As discussed in Chapter 4, household vehicle ownership is endogenously determined without a stratification of the sample into different vehicle ownership categories.

Table 5-6: Vehicle Ownership Model with Embedded Trip Generation Equation

Variable		Estimate of Coefficient of Variable	t-value
Zero vehicle constant term		4.881	4.920
One-vehicle constant term		4.750	9.757
HSIZE2 (2-vehicle alternative)		0.366	3.114
NWRKH1 (1-vehicle alternative)		0.330	1.178
NWRKH2 (2- vehicle alternative)		0.475	1.613
NLIC1 (1 vehicle alternative)		1.315	3.891
NLIC2 (2 vehicle alternative)		3.552	9.459
DINC1 (2 vehicle alternative)		1.263	5.799
DINC3 (0-vehicle alternative)		0.381	0.768
DRURAL (2-vehicle alternative)		0.316	0.655
INCP1 (2 vehicle alternative)		0.643	3.538
INCP2 (1 vehicle alternative)		0.611	3.026
Number of observations	1314		
Likelihood ratio index, ρ^2	0.652		

Each of the estimated coefficients has the sign consistent with what one would expect based on travel behavior theory. The t-values in column 3 indicate that the majority of the estimated coefficients are statistically different from zero at the 5 percent level of significance indicating that the majority of the variables specified in the choice model are associated with the choice of how many vehicles to own by a household. The likelihood ratio index value is very high (0.652), which indicates that the model predicts very well the choices made by households in the estimation dataset of how many vehicles to own.

The model shows that households of larger size (HSIZE2), all else being equal, find it attractive to own multiple vehicles. Having at least one member of the household working (NWRKH1, NWRKH2) makes it attractive to own a vehicle. Further, the magnitude of the coefficient of the number of working members of the household in the 2-vehicle alternative (NWRKH2) is larger than the magnitude of the coefficient in the 1-vehicle alternative (NWRKH1), indicating that the attraction of owning multiple vehicles is greater than the attraction of owning a single vehicle all else being equal.

Households with at least one member having a driver’s license (NLIC) all else being equal find it more attractive to own a vehicle. Further, the magnitude of the coefficient of the number of household members licensed to drive in the 2-vehicle alternative (NLIC2) is larger than the magnitude of the coefficient in the 1-vehicle alternative (NLIC1), indicating that the attraction of owning multiple vehicles is greater than the attraction of owning a single vehicle all else being equal.

Households in the higher income bracket (DINC1) find it more attractive to own multiple vehicles. This is consistent with what one would expect from travel behavior theory since the greater purchasing power of such households means they can afford to acquire more vehicles. Being in the lowest income categories (DINC3) did not prove to be statistically significant at the 5 percent level in terms of helping to determine whether a household would not acquire any vehicles. As stated above though, the sign of the estimated coefficient is correct.

The estimated coefficient of the spatial dummy variable indicating the location of a household, that is, whether it is located in rural area or urban area (DRURAL) is positive. This is consistent with what one would expect based on travel behavior theory. The coefficient estimate though is not statistically significant at the 5 percent level of significance. As discussed in Section 5.1, given that the quality of transit service in the Nashville metropolitan region is not uniformly high, this variable was not expected to impact significantly on choice of how many vehicles to own or trip generation.

Finally, the household income per household member has statistically significant coefficient estimates in both the 1-vehicle (INCP2) and 2-vehicle (INCP1) alternatives. The positive sign of the coefficient indicates that households with higher incomes per person find it attractive to own a vehicle. The magnitude of the coefficient in the 2-vehicle alternative is larger than in the 1-vehicle alternative indicating that all else being equal, a household would find it more attractive to own multiple vehicles compared to a single vehicle.

5.3.2 Trip Generation Model

Table 5-7 presents the estimation results for the trip generation model. The overall goodness of fit of the model, captured by the coefficient of determination, R^2, is from a travel demand modeling perspective respectable. All the estimated coefficients are statistically significant at the 5 percent level of significance. This means the variables of expected vehicle ownership (EVEH), household size (HSIZE) and household income per household member (INCP) are all important variables associated with the number of trips made by a household. The positive sign of each of the variable-coefficients indicate that higher values of each variable are associated with higher levels of trip making.

Table 5-7: Trip Generation Model with Expected Vehicle Ownership as Explanatory Variable

Variable		Estimate of Coefficient of Variable	t-value
Intercept		-3.223	-4.184
Expected number of vehicles owned by household (EVEH)		1.738	2.557
Household Size (HSIZE)		3.217	13.943
Household income per household member (INCP)		0.443	3.242
Number of observations	1314		
Coefficient of determination, R^2	0.358		
Adjusted coefficient of determination, R_a^2	0.357		

The above estimated models of vehicle ownership (Table 5-6) and trip generation model (Table 5-7) complete the process of estimating trip generation equations with vehicle ownership level being determined endogenously rather than exogenously.

CHAPTER 6

CONCLUSIONS AND FUTURE RESEARCH

This chapter presents a summary of the study and what the major conclusions are. It also provides some direction for future work.

6.1 Summary and Conclusions

This thesis made a case for modeling household trip generation treating household vehicle ownership level as an endogenously determined variable rather than as an exogenously determined variable. A model structure that would capture the interdependency in the household decisions of how many vehicles to own and the collective number of trips the household's members would make daily was sought. A discrete-continuous model relating to these two decisions was formulated and estimated. The models fit the base year data very well with estimated model-coefficients that were statistically significant at the 5 percent level of significance. Based on the study results, the following conclusions can be drawn:

1. Given the statistical significance of the coefficients of the trip generation variables that were specified in the vehicle ownership level choice model, the study results show that trip generation does indeed influence how many vehicles a household decides to own.

2. In the case of the trip generation model conditional on vehicle ownership level, corresponding parameter estimates in the trip generation models for the different vehicle ownership categories differ in magnitude. This indicates that the number of vehicles a household owns also influences the collective number of trips the household's members make daily.

3. In the case of discrete-continuous model without stratification, the statistical significance of the expected number of vehicles owned by a household indicates the importance of vehicle ownership to household trip making.

The above three conclusions lead to the conclusion that there is an interdependency between the decision of how many vehicles to own and the decision of how many trips a household expects to be make daily.

Socioeconomic variables such as household income, household size, and number of household-members who work strongly influence both vehicle ownership decisions and trip making decisions.

Spatial location of a household also influences both vehicle ownership decisions and trip making decisions.

6.2 Future Research

Given that the overall size of sample was not particularly large, it was not possible to set aside a validation sample from the parent sample. In future research, a larger sample must be used such that provision can easily be made for validation.

This study was not able to investigate what role transportation system characteristics play in the decisions of how many vehicles to own and how many trips to make on a typical day because transport system data were not available. Future research should investigate this.

REFERENCES

1. Ben-Akiva, M.E. and Lerman, S.R. (1985), *Discrete Choice Analysis*: *Theory and Applications to Travel Demand*, MIT Press, Cambridge, MA.

2. Cambridge Systematics, Inc. (1997). *Vehicle Availability Modelling*, Volume1, Final Report prepared for Federal Highway Administration, Washington, DC.

3. Cambridge Systematics, Inc., Nustats, McGunkin, N. and Ruiter, E. (2007), *NCHRP Report 588: A Guidebook for using American Community Survey Data for Transportation Planning*, Transportation Research Board, Washington, D.C.

4. Dubin, J. and D. McFadden (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, Volume 52, No. 2, pp. 345-362, March 1984.

5. Gerard De Jong, James Fox, Andrew Daly, Martis Pieters and Remko Smit (2004), Comparison of Car Ownership Models, Transport Reviews, Vol.24, No.4; pp.379-408.

6. Hay, J (1980). Occupational Choice and Occupational Earnings. Ph.D. Dissertation. Yale University, New Haven, Connecticut.

7. Heckman, J.J. (1979), Sample Selection Bias as a Specification Error, *Econometrica*, 47, pp. 153-161.

8. Mannering, F. L. (1986). Selectivity Bias in Models of Discrete and Continuous Choice: An Empirical Analysis. *Transportation Research Record* 1085, pp. 58 – 62. Washington, D.C.

9. McNally, M.G. (2000), “The Four-step Model,” in *Handbook of Transport Modelling* edited by Hensher, D. A. and K. J. Button, Elsevier Science, Pergamon Press, pp. 35 – 42.

10. Meyer, M. D. and E. J. Miller (2000), *Urban Transportation Planning: A Decision-Oriented Approach*, 2nd Edition, McGraw-Hill Publishers, Inc., New York.

11. Ortuzar, J. D. and L. G. Willumsen (2001), *Modelling Transport* 2nd Edition, John Wiley & Sons Book Publishers, West Sussex.

12. Roorda, M. J., Mohammadian, R.A., and Miller, E. J. (2000), Toronto Area Car Ownership Study A Retrospective Interview and its Applications, *Transportation Research Record* 1791, Transportation Research Board, National Research Council, Washington, D.C., pp. 69-76.

13. Ryan, J. M. and Han, G. (1999), Vehicle Ownership Model Using Family Structure and Accessibility Application to Honolulu, Hawaii, *Transportation Research Record* 1676, Transportation Research Board, National Research Council, Washington, D.C., 1999, pp.1-10.

14. Thakuriah, P. and Liao, Y. (2005), Analysis of Variation in Vehicle Ownership Expenditures, *Transportation Research Record* 1926, Transportation Research Board, National Research Council, Washington, D.C., 2005, pp.1-9.

15. Whelan, G., (2007), Modeling Car Ownership in Great Britain, *Transportation Research Part A*, Vol.41, pp. 205-219.

VITA

Kiranmai Chirumamilla was born on August 05, 1986, in Guntur, Andhra Pradesh, India. She grew up among a warm loving and caring family. After completing her elementary and high school education at Siddhartha Public School from 1989 to 2001, she joined Goutham Junior College for her intermediate studies. Kiranmai completed her undergraduate studies at Nagarjuna University with a Bachelor of Technology in Civil Engineering in 2007. While working towards a degree she finished a project entitled 'Use of Jute and Coir Geotextiles in Pavements on Expansive Soils'

In August 2007, Kiranmai entered the graduate program at Tennessee Technological University and is a candidate for the Master of Science Degree in Civil Engineering in December 2008. While working towards the degree of Master of Science, she held the position of Graduate Research Assistant in the Department of Civil and Environmental Engineering.