A Tour Level Stop Scheduling Framework and A Vehicle Type Choice Model System for Activity Based Travel Forecasting

by

Venu Madhav Garikapati

A Dissertation Presented in Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy

Approved November 2014 by the
Graduate Supervisory Committee:

Ram M. Pendyala, Chair
Xuesong Zhou
Yingyan Lou

ARIZONA STATE UNIVERSITY

December 2014

ABSTRACT

This dissertation research contributes to the advancement of activity-based travel forecasting models along two lines of inquiry. First, the dissertation aims to introduce a continuous-time representation of activity participation in tour-based model systems in practice. Activity-based travel demand forecasting model systems in practice today are largely tour-based model systems that simulate individual daily activity-travel patterns through the prediction of day-level and tour-level activity agendas. These tour level activity-based models adopt a discrete time representation of activities and sequence the activities within tours using rule-based heuristics. An alternate stream of activity-based model systems mostly confined to the research arena are activity scheduling systems that adopt an evolutionary continuous-time approach to model activity participation subject to time-space prism constraints. In this research, a tour characterization framework capable of simulating and sequencing activities in tours along the continuous time dimension is developed and implemented using readily available travel survey data. The proposed framework includes components for modeling the multitude of secondary activities (stops) undertaken as part of the tour, the time allocated to various activities in a tour, and the sequence in which the activities are pursued.

Second, the dissertation focuses on the implementation of a vehicle fleet composition model component that can be used not only to simulate the mix of vehicle types owned by households but also to identify the specific vehicle that will be used for a specific tour. Virtually all of the activity-based models in practice only model the choice of mode without due consideration of the type of vehicle used on a tour. In this research

effort, a comprehensive vehicle fleet composition model system is developed and implemented. In addition, a primary driver allocation model and a tour-level vehicle type choice model are developed and estimated with a view to advancing the ability to track household vehicle usage through the course of a day within activity-based travel model systems. It is envisioned that these advances will enhance the fidelity of activity-based travel model systems in practice.

DEDICATION

To my best half, Padmini Gudipudi

This dissertation would not have taken shape without the sacrifices you made.

To Amma, Naanna and Annayya

For the unflinching support, immeasurable affection and strength to persevere.

ACKNOWLEDGMENTS

The first shout out goes undoubtedly to the most intelligent, compassionate and creative individual I've ever met, my advisor Dr. Ram M. Pendyala who inspired and supported me all along the pursuit of my doctoral research. A thank you just seems so insignificant in return to the support he provided in bringing this work to fruition. I will forever be grateful for the guidance and creative freedom he has given me in honing my academic skills. I would like to extend my sincere gratitude to professors Yingyan Lou and Xuesong Zhou for taking the time to serve on my committee and provide their scholarly advice. Their suggestions added insight and value to this research work.

A special thanks to my colleague and dear friend Daehyun You who worked alongside me in the transportation research group at ASU. He was the sounding board for many of the ideas that form a part of this dissertation. His impeccable work ethic and inquisitive mind always encouraged me to think critically and work diligently. Apart for my dissertation, I cherished working with him on many research projects and journal papers. I also want to thank another colleague Ellie Volosin, from whom I have imbibed the quality of service in addition to academics.

I want to thank my dad Subrahmanyam Garikapati who said 'you can always earn back the money you spent, but not the time' and supported me in every possible way to pursue my doctoral dream. He always believed in me, even at times when I did not believe in myself, for which I am eternally grateful. This personal note would not be complete without mentioning my constant source of inspiration, my elder brother Bindu Madhav Garikapati whom I always look up to in matters both personal as well as academic. I have not met a person more righteous than him. He will forever be the corner stone for all my achievements and the guiding light in times of turmoil. A special thanks to my sister-in-law Alekhya who taught me how to cherish the ‘little things’ in life.

I want to thank the transportation systems modeling group at Maricopa Association of Governments (MAG) for their continuous support and for providing the data which is used to develop and test the model components of my dissertation. In particular, I want to recognize Dr. Vladimir Livshits and Dr. Kyunghwi Jeon from MAG for their valuable suggestions and insights. Last but not the least, I want to thank my friends in the transportation systems group Srivatsav Kandala, Mindy Kimball, Keith Christian, Sravani Vadlamani, Peiheng Li and Jiangtao Liu, for their cheerful company all along my journey here at ASU.

TABLE OF CONENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure | Page

Figure Page

Figure Page

Figure		Page

GLOSSARY

ABM	-	Activity-based Model
DTA	-	Dynamic Traffic Assignment
GHG	-	Greenhouse Gas
HBO	-	Home-based Other
HBW	-	Home-based Work
HMR	-	Heuristic Mileage Reallocation
HOV	-	High Occupancy Vehicle
id	-	Identifier
MAG	-	Maricopa Association of Governments
MC	-	Multivariate Count
MDCEV	-	Multiple Discrete-Continuous Extreme Value
MDCP	-	Multiple Discrete-Continuous Probit
MNL	-	Multinomial Logit
MNP	-	Multinomial Probit
NHTS	-	National Household Travel Survey
NL	-	Nested Logit
NM	-	Non-motorized
OP	-	Ordered Probit
PTR	-	Power Transformed Regression
SATC	-	Sequential Activity Type Choice
SOV	-	Single Occupancy Vehicle

TAZ	-	Traffic Analysis Zone
VFC	-	Vehicle Fleet Composition

CHAPTER 1

INTRODUCTION

Background

Travel is an integral part of our day-to-day life. We travel for a variety of reasons (work, leisure, chauffeuring a kid etc.) and use an array of modes. Until a few decades ago, the need to model/ forecast travel demand was mainly motivated by the necessity to evaluate the sufficiency of current infrastructure (roads, bridges etc.) for future travel demand. The assessment would provide policy makers with information regarding the infrastructural needs of a region in light of changing traffic patterns. The genesis of travel demand modeling began with trip-based methods which would predict the number of trips (vehicles) between a given set of origin-destination pairs. While the trip-based methods sufficed the need for evaluation of infrastructural needs back in the day, they were held down by a fundamental flaw in the way in which they viewed travel. Trip-based models view travel as trips going from one zone to another and fail to recognize the underlying reason behind travel – the necessity/want to participate in activities. Increasing infrastructure indefinitely is not the ideal solution to ever increasing congestion problems. The tip-based travel demand models were rather constrained to explore new avenues and test alternate congestion mitigation strategies (Kitamura 1988; Jones et al., 1990; Axhausen and Gärling, 1992). Thus came a paradigm shift in analyzing travel demand, from trip-based to activity-based methods. Jones et al. (1990) provide a comprehensive definition of activity-based travel analysis as:

"A framework in which travel is analyzed as daily or multi-day patterns of behavior, related to and derived from differences in life styles and activity participation among the population."

Activity-based microsimulation models have gained much attention in travel demand modeling field in the past few decades because of the sound behavioral background on which they are based. Activity-based models quantify travel as a derived demand that arises from the necessity of individuals to participate in various activities throughout the day (Kitamura and Fujii, 1998). Trip-based models fail to recognize this fundamental behavioral principle influencing travel decisions and model travel as trips going from one zone to another, thereby ignoring behavior of the travelers making these trips. Typically when one has to pursue an activity, he/she thinks about 'where' to go from a set of possible destinations and then decide on 'how' to reach there. In case of individuals who own a personal vehicle, that vehicle might be the most preferred mode, but for people who do not own a personal vehicle, the decision is whether to take transit or walk or so on. Activity-based models consider such nuances in travel decisions made by individuals and offer many distinct advantages in demand modeling in addition to realistic representation of personal travel. The behavioral realism incorporated in the foundations of activity-based modeling methods opened up a plethora of avenues to accurately represent travel demand, test a variety of policies and their impacts on the travel patterns in a region. For example, how would toll pricing impact the destination/route choices between a specific origin-destination pair (and across different social-demographic segments) or how the option of telecommuting would impact the work travel behavior of individuals? Instead of just increasing the infrastructure to meet the needs of growing demand, activity-based methods provided the flexibility to test demand management strategies such as congestion pricing, transit-oriented developments etc. in an intuitive and behaviorally realistic fashion.

The advent of activity-based modeling/activity-scheduling methods dates back to the 80's (Recker et al., 1986a; 1986b) and this area has seen tremendous progress since. Activity-based approaches provide the flexibility to test travel demand models under a wide variety of transportation and land-use policies to determine the extent of their effectiveness even before they are implemented. While activity-based models have proved quite effective in handling the finer nuances in modeling activity-travel behavior, this field like any other needs to be updated with state-of-the art techniques that are capable of handling the changes in travel behavior of people with time. In just the past decade, the world has seen major technological breakthroughs that have significant impacts in the way we view travel. Advancements such as self-driving cars, electric vehicles, connected vehicle technologies and autonomous taxi services show a promise of mitigating the constraints (physical, temporal and environmental) related to travel decisions. On the other hand, travel demand models are being utilized now more than ever to test the new-age demand modeling strategies such as dynamic tolling, eco-lanes, low emission zones etc. The broader scope of this dissertation is to learn from existing activity based models and develop frameworks to enhance them so that travel behavior modeling can be advanced one step closer to reality and provide activity-based models with the capability to handle challenges of the future. In a more specific sense, the research identifies some limitations in the extant activity-based model systems (in both research and practice), proposes frameworks to help fill these gaps and contribute to empirical literature in the field of activity-based modeling. The specific topical areas covered by this research are discussed next.

Tour-Based Representation of Activity-Travel Patterns

Trip chaining and tour formation is a topic of much interest in the activity-based modeling arena. Travel patterns of individuals consist of a set of activities bundled together into a tour. Within each tour, individuals participate in a multitude of activities and allocate different amounts of time to each of these activities. There have been extensive studies that emphasize on trip chaining/ tour formation patterns (Goulias et al., 1990; Hamed and Mannering, 1993; Bhat and Singh, 2000). Tour based models that are widely in practice today, form day level activity agendas for each individual and then simulate various tours and intermediate stops as governed by the individual's agenda. Most of the tour based models use simple/complex discrete choice models such as the multinomial/nested logit models to predict intermediate stops on the tours. The intermediate stops are then inserted into an individual's tour using time-of-day choice models. The temporal detail used in such models is limited to one-hour or half-hour time slots.

Despite significant advancements in the activity based modeling methods, most of the tour level activity-based models still operate in a discrete time representation paradigm, where each of the activities undertaken by an individual in a day are modeled disjointly and are sequenced using rule based heuristics. Recent advancements in econometric modeling theory (Bhat, 2005; 2008) allow for simultaneous prediction of the mix of activities undertaken by an individual in a tour along with the time allocated to each of these activities. Adopting this research, enhancements can be made to existing activity-based models in practice to follow a continuous time representation of activities. The current research effort utilizes state-of-the-art research in the profession and proposes a tour characterization framework capable of predicting the mix of activities pursued on a tour, time allocated each of the activities and the sequence in which these activities are undertaken. The framework provides a platform for tour level activity-based models to adopt an evolutionary continuous-time approach in modeling activity engagement that is at the heart of many activity scheduling model systems.

Vehicle Fleet Composition Modeling

Activity-based models provide an accurate account of travel which can then be assigned using network assignment models to quantify vehicle miles travelled (VMT) by all the households in a region. The calculated VMT is used as a determinant of greenhouse gas (GHG) emissions and fuel consumption. Emission calculations for region wide travel are usually carried out in emission modeling softwares such as EMFAC (2014) and MOVES (2014) that take the output of network assignment models as input. These emission modeling softwares have default distributions that represent the vehicle mix of a region. While these default values provide a quick and simple way to compute emissions, they are often not responsive to policy measures that might influence the vehicle fleet mix. For travel demand models to accurately predict emission footprint of a region, it is necessary to model the vehicle fleet composition at household level, which would help forecast the fleet mix (and corresponding emissions) accurately at the regional level.

Interest in modeling household vehicle fleet composition has been growing for several decades. The necessity to implement such models as an integral part of the activity based modeling framework in order to accurately quantify the emission footprint of a region has become all the more important today with increasing pollution levels, GHG emissions and global warming. While there has been tremendous amount of research in simulating activity-travel patterns of households using activity-based models, simulation of vehicle fleet and vehicle type choice at the trip/tour level has only seen light in the recent years in a handful of activity-based models. Such model systems are still in the research phase and have not yet fully made their way to be included in activity-based models in practice. The proposed research effort aims at developing a robust framework to predict the fleet mix of a household using a Multiple Discrete Continuous Extreme Value (MDCEV) model in conjunction with several other models that control and constrain the prediction of fleet mix such that it is representative of the observed fleet mix in the base year. This will impart much confidence in prediction of fleet mix made for any future year using the developed model system. The model system being proposed is designed as a self-contained package that can be integrated as a plugin to any activity-based model.

Beyond the Modeling of 'Mode' at the Tour Level

Most of the tour-based microsimulation model systems model different attributes of a tour such as the primary activity of the tour, stop frequency, tour accompaniment etc., but ignore modeling vehicle type choice at the tour level. The level at which an auto mode is represented in activity-travel decisions in most of the tour level activity-based models is either an SOV (single occupant vehicle) or a high occupancy vehicle (HOV). While this representation does provide some flexibility in terms of testing policies such as HOV or high occupancy toll (HOT) lanes, planners seek more disaggregate level of information in order to accurately represent complex activity-travel patterns. For example, in a household that owns a car and a Sports Utility Vehicle (SUV), which vehicle would be used more to make a work tour and which one would be utilized more to make a shopping tour. It is apt to model car as the SOV mode in the family and SUV as a HOV mode? Questions such as these cannot be answered with information regarding auto ownership information that most activity-based models simulate. A limiting reason to not identify the specific vehicle type used on a tour is unavailability of information about household vehicle fleet discussed in the previous section. Information regarding the type of vehicle utilized to undertake a tour/trip is critical for policy planners to accurately evaluate emission footprint of personal/regional travel. As a part of the current research effort, a framework is being proposed to predict the vehicle type choice for each tour undertaken by a household along the day. This model can include various tour level attributes as determined by the tour characterization framework and can span different body type x vintage classifications that are modeled using the fleet composition model system.

Objectives of the Dissertation

This dissertation aims to enhance the empirical literature in activity-based microsimulation modeling approaches along the following lines of enquiry.

- **Objective:** *Propose a framework and develop all model components required to enhance tour level activity-based microsimulation model systems to adopt evolutionary continuous time representation of activities.*

Research Contribution: In this research effort, a novel tour characterization framework is proposed that models tours in a continuous time domain. The proposed framework consists of an activity type mix model system that predicts the array of activities undertaken by an individual in a tour. Unit of analysis for this effort is considered in such a way that it facilitates representation of time in a continuous rather than discrete fashion. A stop sequencing model system takes the activities predicted and orders them using utility maximization methods that take into account the history of activity participation as well as anticipatory activity engagement decisions of individuals into consideration. At the end of application of this framework, one would be able to successfully simulate the mix of activities performed by an individual on a tour, the time allocated to each of the activities and the order in which the activities were undertaken. Results of the model components are presented for home-based work tours made by working individuals and home-based other tours undertaken by non-workers.

- **Objective:** *Develop an open source vehicle fleet composition simulator capable of predicting vehicle fleet mix owned by households classified by vehicle body type and age. The intent for development of this component is to provide a precise input of fleet mix at the household level to help in prediction of specific vehicle type choice at the tour level. The utility of such a module can be extended to a wide a variety of applications in activity-based model systems.*

  Research Contribution: A framework is proposed to simulate the vehicle fleet mix at the level of every household in a region. Adopting the framework proposed a vehicle fleet composition simulator is developed on open source coding platform 'R', which can be easily integrated into any of the extant activity-based microsimulation model systems. Estimation and validation results for all of the components of the model system are presented. The proposed framework takes the socio-demographic characteristics of a household as input (for any horizon year), simulates the array of vehicles owned by a household and annual mileage allocated to each of the vehicles. The information is simulated at the disaggregate classification defined by a cross between 4 vehicle body types (car, van, SUV and pick-up truck) and 3 vintage categories (new: 0-5 years, middle aged: 6-11 years, and old: ≥ 12 years). The vehicle fleet composition simulator developed as a part of this effort shows great promise in predicting the mix of vehicles at the household level.

- **Objective:** *Present a joint modeling framework capable of modeling the vehicle fleet mix (type of vehicles) and count (number of vehicles) dimensions together.*

  Research Contribution: This research effort builds on the previous objective, where a fleet composition simulator is developed to predict the fleet mix and count of vehicles in a household. In the previous effort, fleet mix and count components are estimated and applied separately using a novel approach to be able to predict the occurrence of multiple vehicles of similar body type and age classifications (for example, a household might own two cars, both belonging to age category 0-5 years old). Though the framework proposed in the previous effort has proved to be quite effective, it was felt prudent to model the fleet mix and count components in a joint framework as these dimensions are inextricably linked in a household's fleet. The models developed as a part of this effort will replace the fleet mix and count components with a joint model system in future incarnations of the vehicle fleet composition simulator.

- **Objective:** *Develop a tour level vehicle type choice modeling framework constrained by the vehicle fleet owned by a household for an accurate depiction of the activity-travel patterns of individuals in a household.*

  Research Contribution: This effort ties together the tour characterization and fleet composition frameworks to allocate the resources (vehicles owned by a household) to the activity-travel needs (tours undertaken by various members of the household) of individuals in a household. Unlike existing microsimulation models which only consider an aggregate auto mode (SOV/HOV) for modeling vehicle type choice at the tour level, the current effort attempts to model the choice dimension at the disaggregate level of body type and age classification of vehicles owned by a household. The vehicle choices available to an individual are constrained to the fleet of vehicles owned by the household to which the individual belongs to. This framework intends to propose a behaviorally consistent way of representing activity-travel decisions as observed in the real world. The proposed framework utilizes information provided by both the previous components to model the vehicle usage decisions of the household.

Objective: *Propose a conceptual framework for real-time vehicle allocation and tracking capable of accounting for temporal vehicular constraints.*

Research Contribution: While the tour level vehicle type choice modeling framework is a good starting point to introduce vehicular constraints in activity-travel decisions of individuals, it does not to take into consideration, the real-time availability of each of the vehicles owned by the household. A framework is proposed to mimic the real-time vehicle availability in the context of an integrated model system where an activity-based model and a dynamic traffic assignment model are tightly coupled. This framework assumes that the network status (active – meaning that the vehicle is currently on the network, inactive – meaning the vehicle is available) is provided to the activity based model by a dynamic traffic assignment model on a regular basis (say, every 5 minutes). Using the proposed framework, all household vehicles can be accounted at all times of the day, thus allowing for the flexibility to carry out a real-time vehicle allocation to activity-travel needs of the household. Separate frameworks are proposed for adults and children in a household so that the activity engagement patterns of the entire household are represented.

Dissertation Outline

The rest of the dissertation document is organized as follows. Chapter 2 presents a brief literature review on various activity-based modeling systems in research as well as practice, followed by a review of vehicle fleet composition and tour level vehicle type choice literature. In Chapter 3, frameworks for all the components developed as a part of this research are presented. Chapter 4 discusses the model estimation results of tour characterization framework followed by an assessment of the model performance in replicating observed activity-travel patterns. Chapter 5 is devoted to the model estimation/application results of the vehicle fleet composition framework followed by a proposed extension to the framework presented in Chapter 6. Chapter 7 covers the model components of tour level vehicle type choice framework and Chapter 8 discusses a conceptual framework for real-time vehicle allocation and tracking. Conclusions and directions of future research are presented in Chapter 9.

CHAPTER 2

LITERATURE REVIEW

In the past few decades, astounding progress has been made in the development of aggregate travel demand models, followed by increasing interest in activity-based (or agent based) modeling systems. The ultimate goal of these developments is to model travel behavior as closely as possible to get realistic estimates of outcomes in response to various policy measures. This will help planners/authorities weigh and choose policies that pave way for a sustainable future. In this section, a literature review on activity-based modeling systems is provided. A contrast is made between advancements in activity-based models in research and practice. This is followed by a brief account of existing literature on vehicle fleet characterization and tour level vehicle type choice. The chapter concludes with a brief overview of vehicle allocation literature.

Activity-Based Modeling Systems

Activity-based travel analysis traces its roots back to the works of Chapin (1974) who studied human activity patterns in an urban space and Hägerstraand (1970) whose work emphasized on constraints which limit an individual’s activity and travel choices. Hägerstrand’s constraints include coupling, authority and capability constraints. Coupling constraints refer to participation in joint household activities, while authority constraints are imposed by institutional measures such as work hours, store hours etc. Capability constraints are enforced by nature and technological limitations. Activity travel scheduling has been an area of considerable interest in the field of transportation including modeling activity participation and destination choice (Kitamura and Kermanshah, 1984), trip chaining (Adler and Ben-Akiva, 1979; Goulias and Kitamura, 1989; Shiftan, 1998; McGuckin et al., 2005) and choice of activity patterns (Hamed and Mannering, 1993; Bhat and Singh, 2000)

Based on these sound behavioral foundations, many activity based models were developed to estimate travel demand. There are different schools of thought followed in these models which can be broadly classified into: i) utility maximization principles (Recker et al., 1986a; 1986b; Kitamura and Fujii, 1998; Bhat et al., 2004), ii) heuristic or rule-based approaches (Gärling et al., 1989, 1994; Ettema et al., 1993; Kwan, 1997; Pendyala et al., 1998; Arentze and Timmermans, 2004), and iii) sampling from observed activity patterns in surveys (McNally, 1995; Barrett et al., 1999). In addition to the theoretical classification, activity-based models are separated by operational classification as i) activity scheduling modeling systems with continuous time evolutionary approach and ii) tour based model systems with tour/day level activity agenda approach. The activity scheduling systems provide a behaviorally intuitive representation of travel patterns using continuous time representation, but such models are in the development phase and are yet to find their way to be widely adopted in practice. On the other hand, tour level activity-based models that are widely in practice in the industry follow a discrete time representation of activities.

One of the earliest activity-based model systems is proposed by Recker et al. (1986a; 1986b) termed as STARCHILD. This modeling system simulates all possible sets of activity patterns in a given situation and then assumes that an individual will choose a pattern that maximizes his/her utility. STARCHILD however lacked the behavioral representation of travel in the sense that decision makers do not usually enumerate all possible patterns of activities before they make a choice. More often than not, individuals might make sub-optimal choices, which fit their needs. Also the enumeration of all possible travel patterns might become quite cumbersome and is an unnecessary computational burden.

Prism Constrained Activity Travel Simulator (PCATS) is an activity scheduling system that models the daily-activity travel patterns of individuals using the concept of Hägerstrand's time-space prisms (Kitamura et al., 1996; Kitamura and Fujii, 1998). For example, an individual might have access to a wide range of destinations to perform an activity early on in the morning, while the same individual might not have as many destination choices to perform the same activity late in the night. PCATS assumes that activity-travel patterns choices are sequential in nature, where each choice is conditional on a decision previously made (the choice of an activity influences the choice of destination, which might in turn influence the choice of mode).

Activity Mobility Simulator (AMOS) combines utility maximization and rule based heuristics into one microsimulation model that mimics an individual's decision making process (Kitamura et al., 1993; Kitamura and Fujii, 1998; Pendyala et al., 1998). AMOS consists of a host of submodels including a Household Activity Generation System (HAGS) and PCATS that is described above. HAGS generates synthetic population at household and person levels. It also generates the location choice of synthetic population (home and work locations) and mandatory activities (work/school) for all persons within in a household. PCATS then simulates non-mandatory activity-travel choices within a time-space prism corresponding to each open/free period for an individual. Recently, an updated version of AMOS named openAMOS was integrated with a dynamic traffic assignment model and a land use microsimulation model to form an integrated model of urban continuum dubbed SimTRAVEL (Simulator of Transport, Routes, Activities, Vehicles, Emissions, and Land). In addition to providing the conceptual framework, the authors successfully implemented an operational prototype of this system (Pendyala et al., 2012b).

The Comprehensive Econometric Micro-simulator for Daily Activity-travel Patterns (CEMDAP) institutes a suite of econometric models, coupled with several deterministic rules that help simulate the activity-travel patterns of individuals or households (Bhat et al., 2004; Pinjari et al., 2008b). The implementation schematic of CEMDAP is similar to that of openAMOS in the sense that it focuses on modeling the decisions of individuals/households to pursue different activities in a day and allocation of responsibilities (school escorting, other escorting etc.,) among household members. The scheduling model system determines how the generated activities are sequenced while accounting for time-space constraints imposed by mandatory activities such as school and work. Bhat et al. (2004) provide a detailed account of the mathematical structures of various econometric models involved. CEMDAP considers activity travel pattern of an individual as a three level structure: activity, stop and tour. Many of the activity scheduling model systems that are gaining popularity in the research arena follow an evolutionary continuous-time approach to model activity engagement.

Travel Activity Scheduling Model with Household Agents (TASHA) is prototype of a tour level microsimulation model of travel (Miller and Roorda, 2003) that is based on the concept of projects. Axhausen (1998) defines a project as "*a set of activities tied* *together by a common goal*". A common example for a project might be cooking dinner that involves shopping for the ingredients, preparing the food and the actual dinner itself. A project can encompass several sub-projects and a sub-project has its own unique objective. TASHA uses the concept of project in terms of travel, as different activity episodes are organized to form schedules. TASHA uses the concept of tours as the building block to model travel and proposes a rigorous framework to model household vehicle allocation and joint travel. Validation of a few model components in TASHA are presented recently (Roorda et al., 2008). The scheduling model system in TASHA follows a rule based method in which activities are added to 'project-agendas' based on a common purpose.

Another stream of activity based modeling systems are based on heuristics, which emphasize that individuals make a sub-optimal (yet satisfactory) decisions from a set of feasible solutions. Wilson (1998) introduced sequence alignment method (SAM) for analyzing activity patterns, which is based on a set of re-ordering rules. SAM however, is not sensitive to the position of an activity in a sequence. Joh et al. (2001) proposed a position sensitive sequence alignment method to address this shortcoming. A Learning-based Transportation Oriented Simulation System (Albatross) is a system that predicts various dimensions of activity-travel such as which activities are pursued, where, for how long, using what mode and accompaniment choice.

Albatross uses a decision tree mechanism that enumerates an exhaustive set of mutually exclusive rules for each of the decision dimensions mentioned (Arentze and Timmermans, 2004). Albatross assumes that choice behavior is based on rules that are formed and continuously adapted through learning while the individual is interacting with the environment (reinforcement learning) or communicating with others (social learning). Albatross builds the schedule of each individual's activities in a priority based approach, given an activity skeleton that consists of fixed activities (e.g., work or school). Sequence of activity stops are considered in order of activity priority. Janssens et al. (2003) explored the performance of Bayesian networks to extract decision rules from activity dairy data that can be used in rule-based transportation models (such as Albatross). SCHEDULER, developed by Gärling et al. (1989) models activity travel behavior on similar lines. Kwan (1997) developed a GIS based activity scheduling model termed as GIS-Interfaced Computational-process-model for Activity Scheduling (GISCAS) that combines rule based methods and spatial information to form activity schedules.

While the activity-based models in research arena continue to excel in depicting human activity-travel patterns as realistically as possible, they are often not very easy to be implemented on large scale networks in practice due to the heavy computational burden incurred. This lead to the development of another stream of activity-based models that are easy enough to implement on large networks while not compromising on the behavioral representation of travel. Most of the activity-based model systems in practice today operate at the tour level. Each tour may be characterized by a primary destination, where the primary destination may be defined based on an activity priority hierarchy, activity duration, travel duration, or any combination of these attributes. A tour may consist of multiple stops in which individuals devote varying levels of time to different activities. In addition to activity type choice and time allocation, there are a variety of characteristics that define a tour such as the tour mode, number of intermediate stops, sequence of the intermediate stops, and tour accompaniment. Most of the tour level activity-based models in practice resort to time-of-day choice models to sequence various activities undertaken by an individual.

First, a daily activity agenda is formed for every individual. Then, tours are generated for the individual along the day and the timing (start/end times) of the tour is determined using time-of-day choice models. Within each tour, a secondary stop frequency informs how many stops (other than the primary destination) are made on the tour. Activities from already generated agendas for the individual are then 'inserted' into the tours using time-of-day choice models. In the time-of-day choice models at both tour and stop level, time is represented as discrete bins to facilitate the estimation of discrete choice models. The temporal detail followed in most of these models is either one hour or half hour periods at the maximum, as more disaggregation poses dimensionality issues for model estimation/application. The discrete time representation in these models is followed for operational convenience and is not behaviorally realistic.

There is growing interest in enhancing tour level activity-based models in practice to continuous-time domain. Activity scheduling model systems in the research circles such as CEMDAP (Bhat et al., 2004) and openAMOS (Pendyala et al., 2012b) are well founded on evolutionary continuous time approaches. The proposed research effort learns and builds on the existing activity-based models in the research community, as well as recent advancements in econometric modeling literature and proposes a framework to bring the continuous-time evolutionary activity-based model systems and discrete-time tour-based model systems closer together in their representation of activity-travel schedules.

Vehicle Fleet Composition Modeling

Vehicle fleet composition modeling and its impacts on personal travel has been an area of significant interest in the travel demand modeling arena for the past couple of decades. In the recent years, there have been significant advancements in the field of activity-based modeling in generating various attributes of synthetic population and how characteristics of the individuals/households impact activity-travel patterns. Another important dimension that controls and constrains the travel of a household is the vehicle fleet composition. Without this information it is impossible to predict what 'type' of a vehicle will an individual choose to make a particular tour/trip. Modeling vehicle fleet mix has only seen light in the recent years and is making its way to being implemented in activity-based models in practice. Knowing the exact type of vehicle used for personal travel will help accurately quantify emissions at the household as well as regional level. Accurately quantifying emissions would provide planners/authorities with a powerful tool to evaluate pollution abatement strategies aimed at specific vehicle categories. Also, studying the sensitivity of fleet mix to changing land use dynamics will help build more compact and sustainable communities.

There has been considerable progress in the modeling of vehicle fleet composition and utilization in the recent past. This progress is motivated mainly by the increasing necessity to curb pollution arising from personal travel. United States accounts for 16% of all GHG emissions in the world (World Research Institute, 2014). In the US, transportation accounted for 28% of greenhouse gas emissions in 2011 (EPA, 2014) and 70% of all petroleum consumption (EIA, 2013). There is an alarming need to curtail the usage of polluting modes and shift to greener transportation. The problem however is that it is difficult to predict the outcomes of pollution abatement strategies aimed at specific vehicle types because such disaggregate level of information is usually not available at the household/regional level. This shortcoming gained the attention of several researchers in the profession to develop comprehensive fleet composition simulators that accurately predict the household fleet of a region for a given set of inputs, thereby helping to test a variety of pollution reduction strategies.

Several earlier studies examined auto ownership, in terms of the number of vehicles owned (Lerman and Ben-Akiva, 1976; Kain and Fauth, 1978; Golob and Burns, 1978; Bhat and Pulugurta, 1998), the type of vehicle (Lave and Train, 1979; Hocherman et al., 1983; Brownstone et al., 2000; Choo and Mokhtarian, 2004) or vehicle holdings of a household (Kitamura et al., 2000). A parallel stream of research efforts focused on joint modeling of different dimensions of auto ownership, such as household fleet size and composition (Hensher and Plastrier, 1985), vehicle body type and vintage (Berkovec and Rust, 1985; Mohammadian and Miller, 2003), vehicle make/model and vintage (Manski and Sherman, 1980; Mannering and Winston, 1985). Emphasis on estimating the annual miles travelled using vehicles owned by a household lead to studies that included usage of household fleet in addition to the dimensions mentioned above (Train, 1986; Golob and Wissen, 1989; de Jong, 1996; Golob et al., 1997). Predicting the annual mileage would not only help predict the annual emissions of the household disaggregated by vehicle type, but also help in evaluating policies that reduce the usage of specific type of vehicles.

Many of these studies provided invaluable insights into understanding the vehicle ownership and usage patterns of households, but a common constraint to most of these studies is that they were confined to modeling a few dimensions of vehicle ownership. This can be attributed heavily to the type of models (multinomial (MNL), nested logit (NL)) used in these studies. Traditional discrete choice models such as the MNL and NL models deal with situations where a single choice is made from a set of mutually exclusive alternatives. Also, the utility function of these models is linear in nature and is not flexible to capture the effect of diminishing marginal utility in the consumption of an alternative. The household fleet mix problem is not a single discrete choice problem as a household may own multiple vehicles and use them to varying degrees (because of the variety seeking nature) simultaneously. Using these models to address the fleet composition of a household limits the dimensionality of the modeling problem and restricts the analyst from modeling variety seeking behavior.

It should be identified here that modeling multiple discreetness using single discrete choice models is not impossible but computationally inefficient. For example, to model a discrete choice model of fleet mix encompassing 5 vehicle alternatives, a total of 31 (2^n-1) combinations of the 5 alternatives should be identified for each individual in the dataset. This number increases exorbitantly as the number of elemental alternatives increase. Bhat (2005; 2008) formulated the Multiple Discrete Continuous Extreme Value (MDCEV) model to overcome the limitations of single discrete choice models in such situations. The MDCEV model has since then been tested to model various multiple discrete choice situations such as decision to participate in different types of maintenance and leisure activities in a given time period (Bhat, 2005) and household vehicle fleet composition (Bhat and Sen, 2006). Paleti et al. (2011a) developed a comprehensive vehicle fleet composition, utilization, and evolution framework for integration in activity-based microsimulation models of travel demand. The framework consists of a vehicle selection module in which a joint discrete-continuous copula-based model predicts vehicle fleet composition and utilization.

Musti and Kockelman (2011) estimated a stated preference vehicle choice (MNL) model from a survey conducted in Austin, Texas that examines opinions of people on vehicle policy. They also developed a micro-simulator of vehicle transactions to simulate change of vehicle fleet composition in the household using the estimated models. The base year fleet characteristics of the data are considered as ‘given’ to the simulator, which will then evolve the fleet based on transaction and vehicle choice models. Pendyala et al. (2012a) apply a socio-economic model system for activity-based modeling to the region of Southern California. A component of the socio-economic model system is CEMSELTS, which includes a fleet composition module that simulates the vehicle fleet owned by a household and assigns a primary driver to each of the vehicles owned. Vyas et al. (2012) proposed a framework to model vehicle type holdings, usage and allocation of primary driver simultaneously at the household level. Such a model would be quite sensitive to transport policies that are aimed at bringing changes in the fleet mix at household as well as regional levels. The study considers a total of 54 vehicle body type – age categories to make sure that no household in the dataset owns multiple vehicles of same body type and age. While such disaggregation of choice dimensions is laudable, it is exhaustive and might make the estimation dataset sparse.

Despite significant advancements in the fleet composition modeling research in the past decade, none of the models developed have fully made their way to be included as an integral part of the activity-based microsimulation model systems in practice. Much of the fleet composition integration into activity-based modeling systems is still in the testing phases and it is safe to assume that none of the activity-based models in practice today house a comprehensive fleet composition model system. The current effort builds on the existing literature and develops an operational prototype of a robust vehicle fleet composition model system that predicts the fleet mix of households classified by body type and age. The model system is built on an open source platform and in a modular fashion such that it can be easily integrated into any activity-based microsimulation model.

Tour Level Vehicle Type Choice Models

Tour level activity-based models simulate a variety of attributes that are of interest to travel demand modelers such as number of stops, tour complexity, tour accompaniment etc. Many tour based modeling systems have been successfully implemented in the United States and elsewhere (Algers et al., 1996; Bowman and Ben-Akiva, 2001; Vovsha et al., 2001; Vovsha and Bradley, 2006). While these tour based model systems predict key dimensions that characterize a tour such as number of stops on the tour, their location, mode choice, sequencing and scheduling of stops, almost none of these model systems account for vehicle type choice at the tour level. Identifying vehicle type at the tour level has important use in policy evaluation. One example is testing low emission zone strategies where heavily polluting vehicle types are restricted from entering select few zones for certain times of the day.

A limiting reason for the tour based modeling systems to be constrained to modeling the ‘mode’ but not ‘vehicle type’ at the tour level is lack of information regarding the fleet mix of the household. Vehicle fleet composition and utilization patterns at the household level have been a topic of considerable interest in the travel demand modeling arena (Mohammadian and Miller, 2003; Bhat and Sen, 2006). Most of the existing literature aims at modeling the vehicle fleet and utilization patterns at the household level. While this information can be used as a good starting point to compute energy and emission footprint of travel in a region, it does not provide disaggregate detail at the level of individual tours that planners seek to quantify travel. Two important dimensions that need to be considered when modeling vehicle type choice at the tour level are: i) Fleet of vehicles owned by the household and ii) Inter-personal household constraints on vehicle usage.

Miller et al. (2005) present a microsimulation modeling framework that handles the household's allocation of resources (vehicles to drivers) based on maximizing the overall household utility. A key assumption made by Miller et al. (2005) is that if a vehicle is to be used on a particular tour, it would be used for the entire chain of activities comprising the tour and would be available to any other member in the household only after the vehicle returns home at the end of the tour. This is a behaviorally intuitive assumption in terms of the general vehicle usage patterns observed in day-to-day life. The authors of the study present a conceptual model and an operational prototype designed to be integrated with TASHA (Miller and Roorda, 2003). They also identify that the model could be used with any activity-based model that generates home-based tours. The authors provide an excellent account of vehicle allocation and joint household travel decision making, but the framework proposed still operates at the level of mode (such as auto, walk, transit), but not at the individual vehicle level (such as a car, van, pick-up truck etc.).

It is intuitive that different types of tours call for usage of different types of vehicles (subject to their availability in the household's vehicle fleet). For example, a tour consisting of only work activity, might be carried out using a small vehicle such as a car, but a tour involving all the members of a household might call for a much larger vehicle (such as a van), provided the household owns one. The vehicle type choice comes a sub-decision provided an individual in the household decides to use an auto mode on the tour.

More recently, Konduri et al. (2011) presented a framework that focuses on examining two dimensions of tours, the type of vehicle (in households that own multiple vehicles) chosen to undertake a tour and the tour length as these dimensions might have common unobserved factors, that drive the decision making process. The study examined the relationship between these two dimensions both ways i.e., does tour length affect vehicle type choice or vehicle type choice affect tour length and concluded that the latter specification performed better. The authors justify this finding with the fact that vehicle type choice is a longer-term choice that influences shorter-term tour length choices, than the other way around. The study is restricted to four body types (car, van, SUV and pick-up truck) for modeling vehicle type choice at the tour level and does not consider any vintage (old vs. new vehicles) classification. Wherein reality, tour length might also dictate the type of vehicle to be used. For example, when making long distance travel one would generally prefer to use newer (supposedly more fuel efficient) than older vehicles of the same body type, provided both vehicles are available in a household's vehicle fleet.

Paleti et al. (2011b) reports similar findings for a model that considers tour complexity, passenger accompaniment, vehicle type choice and tour length together. The authors use an integrated modeling framework to study a mixture of dependent variables within a single model system. While the methodological framework proposed by the authors is absolutely state-of-the-art, it is quite cumbersome from a model deployment and application standpoint to simulate choices using such a complex model system. In light of the multiple dimensions that need to be estimated using a single integrated modeling framework, the authors restrict the vehicle type choice to a similar level of aggregation as Konduri et al. (2011), thereby leaving out the vintage classification which is an intrinsic part of a household's vehicle fleet. The model estimation results provided by the authors are quite intuitive and are mostly in line with similar studies in the domain (Konduri et al., 2011).

While joint modeling systems discussed above bear great significance in understanding the inter-relationship between multiple endogenous variables, they are often not easy to apply from a practical standpoint. The current effort proposes a framework capable of simulating vehicle type choice at the tour level, which can fully leverage the information provided by the tour characterization framework with regard to types of activities on the tour, tour accompaniment and tour duration. The framework is designed with a view for it to be easily integrated into any existing activity-based microsimulation model systems. In application mode, the vehicle type choice model will be constrained to only include vehicles owned by the household. This information is available from the proposed vehicle fleet composition modeling framework at the level of body-type x age categorization.

Vehicle Allocation and Tracking

Household vehicle usage characteristics have an overbearing significance in quantifying the emission foot print of a household/region and also in accurately depicting activity-travel patterns of the household. Take the example of a two vehicle (a car and a van) household, with three licensed drivers (two workers, one non-worker). On an 'average day', if both the workers in the household have to go to work, how would the non-worker fulfill his travel needs? Between the workers in the household, which worker gets which type of vehicle? Do the workers carry out a joint commute to work and leave one of the vehicles for travel necessities of the non-worker? All of these are behaviorally valid questions and encountered by households with number of vehicles less than workers on a regular basis. To understand and accurately depict the activity-travel patterns of households in a region, it is therefore very important to

i. Have information regarding the entire range of activities carried out by all the members of the household
ii. Have knowledge regarding the different types of vehicles owned by a household
iii. Know the whereabouts (availability) of different vehicles in the household along various time periods in a day

Study of vehicle usage patterns in a quest to accurately quantify emission has been topic of significant interest in the travel behavior research community. This question becomes all the more important in the context of auto-deficient households (household with less number of vehicles than drivers). Early studies in this domain tried to model aggregate vehicle usage patterns of the household (relative preference of one vehicle over another) using household, person and vehicle attributes as explanatory variables (Hensher, 1985; Golob et al., 1996; 1997). For example, findings of these studies report that young people and males are likely to drive more (Mannering, 1983; Mannering and Winston, 1985; Hensher, 1985; Train, 1986). These studies are in part motivated by the necessity to forecast demand for alternative fuel vehicles. All of these studies assume household's vehicle fleet as a given and attempt to model the usage patterns of different vehicles based on the household as well as vehicle characteristics. Advancements in the profession now allow us to predict the vehicle fleet at the disaggregate level of body-type and age of the vehicle. Also, progress in the activity-based modeling arena led to modeling of household's travel decisions at the level of individual tours – thereby necessitating the modeling of vehicle type choice at the same level.

Despite significant advancements in the activity-based modeling arena, most of the tour level activity-based models in practice still operate at the level of a mode (auto vs. non-auto), but do not simulate the exact type of vehicle being used to undertake a particular tour. While the tour level vehicle type choice framework discussed in the previous section handles this issue partly, it is still static in nature meaning that it assumes continuous availability of the household's fleet to all drivers in the household. There is much necessity to handle the vehicle allocation and tracking problem as a dynamic one where the availability/unavailability of vehicles in a household are continuously updated at regular time intervals. This would require a tightly integrated model system where an activity-based model and a dynamic traffic assignment model communicate on a continuous basis (say every minute). Such models are not futuristic, but fully operational in the current day. An example of such a system is SimTRAVEL (Simulator of Transport, Routes, Activities, Vehicles, Emissions, and Land) where an activity-based model is tightly coupled with a dynamic traffic assignment model (Pendyala et al., 2012b) and information exchange between both model systems happens on a minute-by-minute basis. The current effort proposes a dynamic vehicle allocation and tracking framework that models vehicle type choice at the tour level subject to temporal – vehicular constraints of the household's fleet. The proposed framework can be implemented in any of the activity-based modeling systems. Some of the limitations of existing activity-based models (in both research and practice) that the research effort in this dissertation intends to address are:

- Virtually all of the tour level activity-based models in practice use a discrete time representation of activity-travel patterns.
- The sequence of activities undertaken by individuals along the day is determined using rule based heuristics in the most of the existing activity-based models in practice. In reality the sequence of activity participation is determined by an individual before he/she embarks on the journey.
- Most of the activity-based models (both in practice and the research community) only model auto-ownership and do not account for household vehicle fleet composition. This is a topic of growing interest in the profession.
- Almost none of the activity-based models in practice incorporate a vehicle type choice component in the modeling process which is of tremendous importance in calculating emission footprint of both personal as well as region wide travel. There are a few activity-based models in the research arena that are staring to implement fleet composition and vehicle type choice, but these efforts are still in the developmental stages.
- Dynamic vehicle allocation and tracking is garnering much interest in the research arena, which is indented to 'fine-tune' the behavioral representation of activity-travel patterns in extant microsimulation modeling systems.

The current research effort aims at addressing each of these issues by developing frameworks that explicitly model tour based travel in a continuous time domain and predict vehicle type choice at the tour level.

CHAPTER 3

METHODLOGICAL FRAMEWORKS

This chapter discusses in detail, the methodological frameworks proposed as a part of this dissertation research. Estimation and validation of components involved in each of these frameworks is presented in subsequent chapters. The proposed research effort focuses on three main themes aimed at enhancing the modeling methodology in activity-based models in practice as well as research domains.

- *Tour Characterization Framework*: This framework provides a methodology to enhance the current tour level activity-based models in practice to a continuous-time domain. Using this methodology, all the secondary activities undertaken by an individual in a tour could be predicted along with the sequence in which these activities are performed and time allocated to each of the stops.
- *Vehicle Fleet Composition Modeling Framework*: The proposed framework simultaneously predicts the fleet of vehicles owned by a household and the extent to which these vehicles are driven annually. This provides rich information to calculate emission footprint of a region. Information about fleet composition of a household will also help in modeling the tour level vehicle type choice.
- *Tour Level Vehicle Type Choice Modeling Framework*: This line of effort aims at building the model components required to allocate vehicles owned by a household to their activity-travel needs. This effort joins the tour characterization framework with the fleet composition modeling methodology to identify which vehicle amongst a household's fleet will be utilized to undertake a specific tour.

The schematic of each of these frameworks is explained in the following sections, along with an explanation of the components involved and the modeling process.

Tour Characterization Framework

A tour is defined as sequence of trips starting and ending at the same anchor point. If the anchor point is home, the tour will be called a 'Home-based Tour'. If it is any other location than home, then the tour will be called a 'Non-Home-based Tour'. Within a tour, a primary purpose is always identified which completes the definition of a tour. For example, if the primary purpose of a home based tour is work activity, then the tour is termed a Home-based Work (HBW) tour. If the primary purpose is any other activity than work, then the tour is termed as a Home-based Other (HBO) tour for the purposes of this research effort. Modeling a person's daily/tour schedule is behaviorally intuitive as people usually plan a set of tasks to be accomplished before embarking on a journey. For example, an individual might plan to drop the kids at school, go to work and on the way back home, he/she might plan on stopping at the grocery store and then head back home. All of these activities together constitute a tour. The concept of a tour or trip chain mimics the real world travel behavior of individuals.

The notion that an individual might pursue 'multiple' activities in a single tour is intuitive and this choice context is referred to as multiple discrete choice. Tour based travel demand modeling systems that are in practice do not consider multiple discreteness in predicting activity-travel patterns in a tour, but identify various dimensions of a an individual's tour such as primary destination, tour accompaniment and stop frequency using single discrete choice modeling methods. The tours are then 'filled' with activities that are in the individual's daily activity agenda. Multiple discrete choice situations such as participating in multiple activities as a part of the tour can be modeled using traditional single discrete choice modeling frameworks by building a choice set that encompasses all possible combinations of the elemental alternatives considered. The problem with this approach is that as the number of elemental alternatives increase, the bundle of composite alternatives to be considered explodes. Recent advancements in econometric modeling methodologies help us model such multiple discrete choice situations in a parsimonious yet efficient manner. In particular, the multiple discrete continuous extreme value (MDCEV) model structure proposed by Bhat (2005; 2008) is attracting much attention in modeling choice situations where individuals choose more than one alternative from a set of available alternatives.

The choice context of simultaneously modeling multiple activities in a tour lends itself aptly to the application of MDCEV modeling methodology. This also provides an elegant way to model tours in continuous-time domain. In addition to modeling the multiple discrete choice behavior, the MDCEV model also considers the effect of diminishing marginal utility (i.e., satiation effect) with increasing consumption of an alternative, which the traditional random utility maximization methods fail to accommodate for. The consideration of satiation coupled with flexibility to accommodate multiple discreteness allows us to model the different activities performed by an individual on a tour and the amount of time a person allocates to each of the activities chosen. The proposed framework is capable of characterizing tours in terms of the mix of activities pursued, the time allocated to various activities, and the timing and sequence in which activities chained in a tour will be pursued.

Figure 3.1 shows the structure of a HBW work tour, which has four stops on the outbound half tour (Home → Work) and four stops on the inbound half tour (Work → Home). In the context of HBW tours, home and work locations of an individual are known and the times at which the individual starts at home to go to work and the arrival/departure times of work are 'fixed'. The tour characterization framework predicts all the secondary stops on the tour. Whereas in the case of HBO tours, only the home location (or activity) of the individual is fixed and the framework predicts all activities performed on a tour including the primary activity.

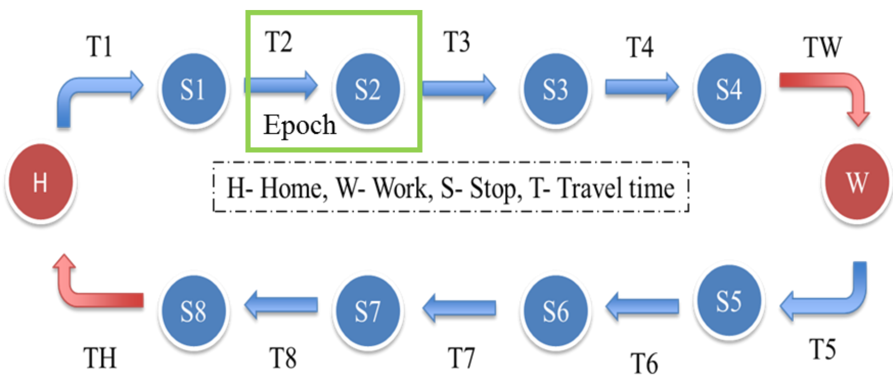

Figure 3.1. Sample composition of a home-based work tour.

The tour characterization framework considers the following information as exogenous input.

- Tour start and end times: The activity-based model system pegs start and end times of each tour undertaken by the individual. Knowledge about start and end times of the tour help to compute a 'time budget' for the tour which will then be allocated to different secondary activities on the tour.

- Tour primary purpose: The framework assumes that the activity-based model has already identified the primary purpose of the tour. If the primary purpose is work, the tour will be labeled as a HBW tour, else it will be labeled as a HBO tour. It is quite common for activity-based model systems in practice to simulate a primary activity for the tour as a part of the tour formation component.
- Household tour plan: The model system also has knowledge of joint tours made by individual in a household. While this component is not that critical to predicting secondary activities on the tour, this information can be utilized in allocating vehicles owned by a household to meet the travel needs of the household.

The overall tour characterization framework is depicted in Figure 3.2. The proposed methodology provides an elegant framework for moving tour-based model systems into continuous-time domain while also facilitating the characterization of tours in the daily activity-travel pattern. With the information discussed above as input, the framework identifies

- The multitude of secondary activities performed on a tour
- Time allocated to each of the secondary activities on the tour
- Placement of secondary activities with respect to primary activity on the tour (inbound or outbound half tours)
- If any half-tour that consists of multiple activities, the framework simulates the sequence (order) of activities before or after the primary activity along a continuous time axis

The tour characterization framework starts with computing a time budget for each tour from the exogenously provided tour start and end times. An MDCEV model takes the tour budget as input and simulates all the secondary stops that will be pursued as a part of the tour under consideration. The unit of analysis that is modeled for each secondary activity on the tour is termed as an 'epoch' (Figure 3.1), which is the summation of travel time and dwell time for an activity. This helps us allocate tour budget to various activities in continuous time.

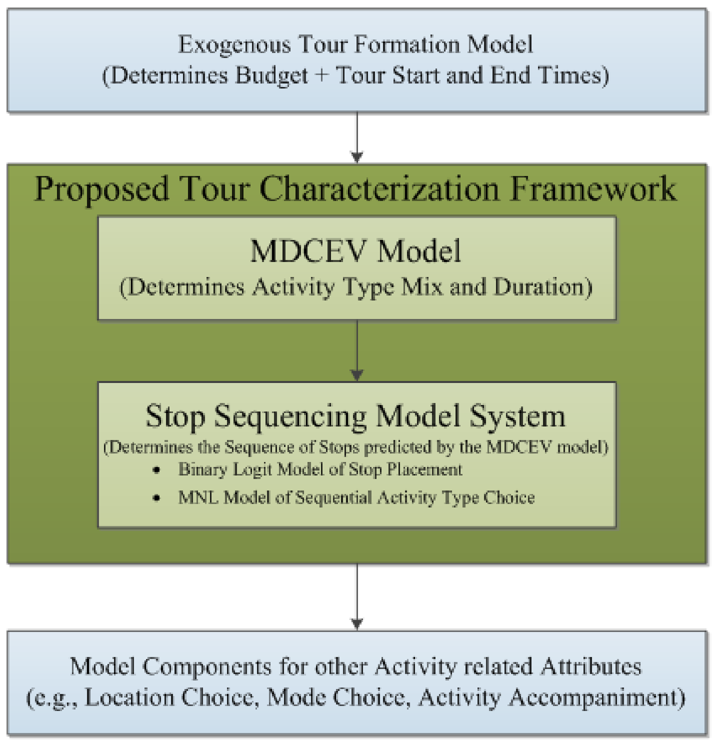

Figure 3.2. Tour characterization modeling framework.

Following the MDCEV model component, sequencing component of the framework is invoked. The first part of this component involves application of a binary logit model on all the secondary stops predicted by the MDCEV model to determine the placement of each of the stops relative to primary activity (inbound/outbound half tour). If the inbound/outbound half tour consist of only one secondary stop, no sequencing is required. In the event that multiple stops occur before or after the primary activity, a sequential activity type choice (SATC) model is invoked. The SATC model predicts the sequence of activities on a half tour subject to the constraints of the choice set.

The choice set for SATC model includes only those activities that are pursued on the half-tour under consideration (i.e., inbound/outbound half tour). Separate SATC models are estimated for inbound and outbound half tours as the impetus for organizing stops on the way from home and returning back home is varied. As the application of MDCEV model provides information about all the activities pursued, the SATC model can utilize information about activities that have been completed (earlier in the day/tour) as well as information about activities that are yet to be pursued. The SATC model is applied in sequence, starting from home location until the half tour is completed. The SATC model is applied '*m-1*' times, where *m* is the total number of secondary stops on tour, as knowing the order of '*m-1*' stops automatically positions the m^{th} stop on the tour. The tour characterization framework thus simulates a full tour schedule, with information about the secondary stops, time allocated to them and the order in which they are pursued.

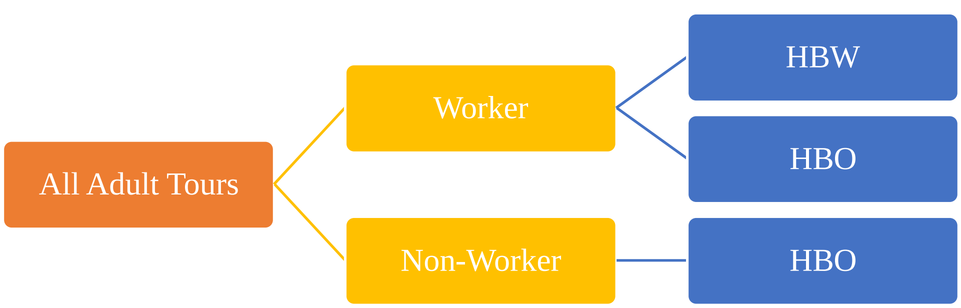

Figure 3.3. Segmentation for tour characterization framework.

Figure 3.3 shows the market segmentation considered for estimating model components in the tour characterization framework. At the top level, all adult tours are separated into tours made by workers and non-workers. For workers, separate models are estimated for HBW and HBO tours. For non-workers model components are estimated for HBO tours. For each of the market segments an MDCEV model is estimated that considers all of the secondary stops made by the individual on the tour. The models includes a host of socio-demographic attributes, zonal characteristics and accessibility measures. The estimated model components will be validated to see how well they predict the observed tour patterns. Checks are made to see how well the model can predict observed activity type frequency and duration.

The next section describes the fleet composition model framework that predicts household fleet composition and utilization classified by body-type and age categories.

Vehicle Fleet Composition Model Framework

Most of the activity-based model systems (both in practice and research) only model auto ownership at the household level without any consideration of the types of vehicles that the household owns. Very few activity-based models that are still in the research phase are beginning to incorporate fleet composition models as a part of their modeling framework. Generalizing the household vehicle ownership has important consequences in quantifying energy and emissions footprint at the household/regional level.

To overcome this issue, a fleet composition model framework is proposed in this research effort to simultaneously predict the number of vehicles owned by a household, body type of each of these vehicles, their vintage and the annual mileage put on each of these vehicles by the household. The system also predicts if the household owns multiple vehicles of the same body type–age category (for example a household might own two cars, both in the age category 0-5 years). The proposed framework is shown in Figure 3.4. Each of the models in the proposed fleet composition modeling framework are explained below.

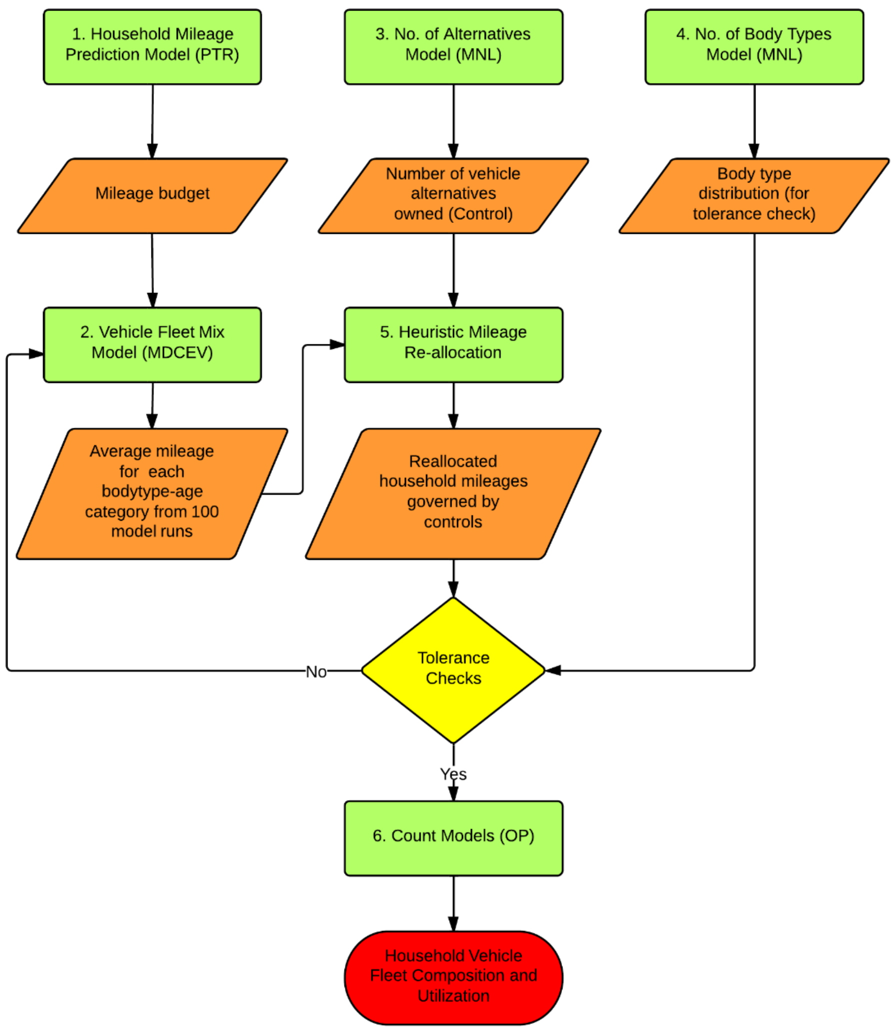

Figure 3.4. Proposed vehicle fleet composition model framework.

1. *Household Mileage Prediction Model*: The first element in the model system is a household mileage prediction model that predicts the annual motorized mileage consumption of households. This step is necessary as the subsequent component in the model system, the MDCEV model of fleet mix requires a mileage budget to allocate to the fleet owned by a household. The motorized mileage can be estimated using a simple linear regression model or some variation of it. Once the motorized mileage for each household is predicted, non-motorized mileage is computed using a preset formula (0.5 x household size x 365) as every household will inevitably have at least some amount of non-zero mileage consumption (walking from the parking lot, jogging etc.). The combined annual mileage is given as input to the MDCEV model, which will then predict the fleet mix owned by the household and allocate the household's mileage budget to all the vehicles owned by the household.

2. *Vehicle Fleet Mix Model*: The MDCEV model of fleet mix predicts the fleet composition of the households. Fleet mix for the MDCEV model is defined as a cross classification between 4 body types (car, van, SUV and pick-up truck) and 3 vintage categories (0-5 years old, 6-11 years old and ≥ 12years old). Motorbike is added as an alternative with no vintage categories. An additional alternative called the '*non-motorized vehicle*' is added to which non-motorized mileage of a household is allocated. MDCEV model is estimated and applied in such a fashion that every household in the dataset will consume at least some non-motorized mileage. Such an alternative is termed as an 'outside good' in econometric modeling jargon. The MDCEV model produces a different output each time a simulation is run. Which of the

simulations should be considered final? To answer this problem, the MDCEV model is applied on the data multiple times and mileage consumptions from each simulation are stored. After ‘n’ simulations of the MDCEV model are completed, an average mileage consumption is computed for each alternative, which will then be re-allocated using a mileage re-allocation algorithm.

Table 3.1

Average MDCEV Model Output After ‘n’ Iterations

Vehicle Alternative	**Mileage**
Car (Age > 0 & ≤ 5 Years)	5000
Car (Age > 6 & ≤ 11 Years)	8000
Car (Age ≥ 12 Years)	400
Van (Age > 0 & ≤ 5 Years)	2000
Van (Age > 6 & ≤ 11 Years)	7000
Van (Age ≥ 12 Years)	550
SUV (Age > 0 & ≤ 5 Years)	1100
SUV (Age > 6 & ≤ 11 Years)	1000
SUV (Age ≥ 12 Years)	300
Pick-up (Age > 0 & ≤ 5 Years)	200
Pick-up (Age > 6 & ≤ 11 Years)	500
Pick-up (Age ≥ 12 Years)	100
Motorbike	50
Total	***26200***

3. *Number of Vehicle Alternatives Model*: A sample output for a household from 50 simulations of the MDCEV model is shown in Table 3.1. Since each simulation gives a slightly different result, the average mileage consumption result from 50 runs show that the household owns almost all of the vehicle categories, whereas in reality this household might own only a couple of vehicles. The heuristic mileage reallocation algorithm does the job of reallocating this mileage distribution in such a fashion that it reflects the household vehicle fleet composition. But mileage reallocation algorithm

requires information about how many categories of vehicles does the household own. A multinomial logit model of number of alternatives predicts this information and provides it as an input to the mileage reallocation algorithm. Suppose, the household owns a car 0-5 years old and a SUV 0-5 years old, and a Van 6-11 years old, the number of alternatives model predicts the number of alternatives owned by this household as three.

4. *Number of Vehicle Body-types Model*: The structure of number of vehicle body types model is quite similar to that of the number of alternatives model, except this model predicts the number of different vehicle body types owned by a household, which provides marginal control totals for the mileage re-allocation model. While the vehicle body type distribution for the population is known in the base year (from survey data), this distribution is unknown for future years. The MNL model of vehicle body types predicts this distribution based on the projected synthetic population characteristics. This goes in as a control distribution that should be matched by the mileage re-allocation algorithm.

5. *Heuristic Mileage Re-allocation Algorithm*: The heuristic mileage re-allocation algorithm (HMR) takes outputs of MDCEV model and MNL model of number of alternatives as input, to re-distributes the mileage to number of alternatives owned by the household. The logic followed by the HMR algorithm is shown in Figure 3.5. The algorithm operates at the level of each household, where it reallocates the mileage using a choice occasion based approach. The output from MNL model of number of

alternatives provides information regarding how many body-type x age categories does the household own.

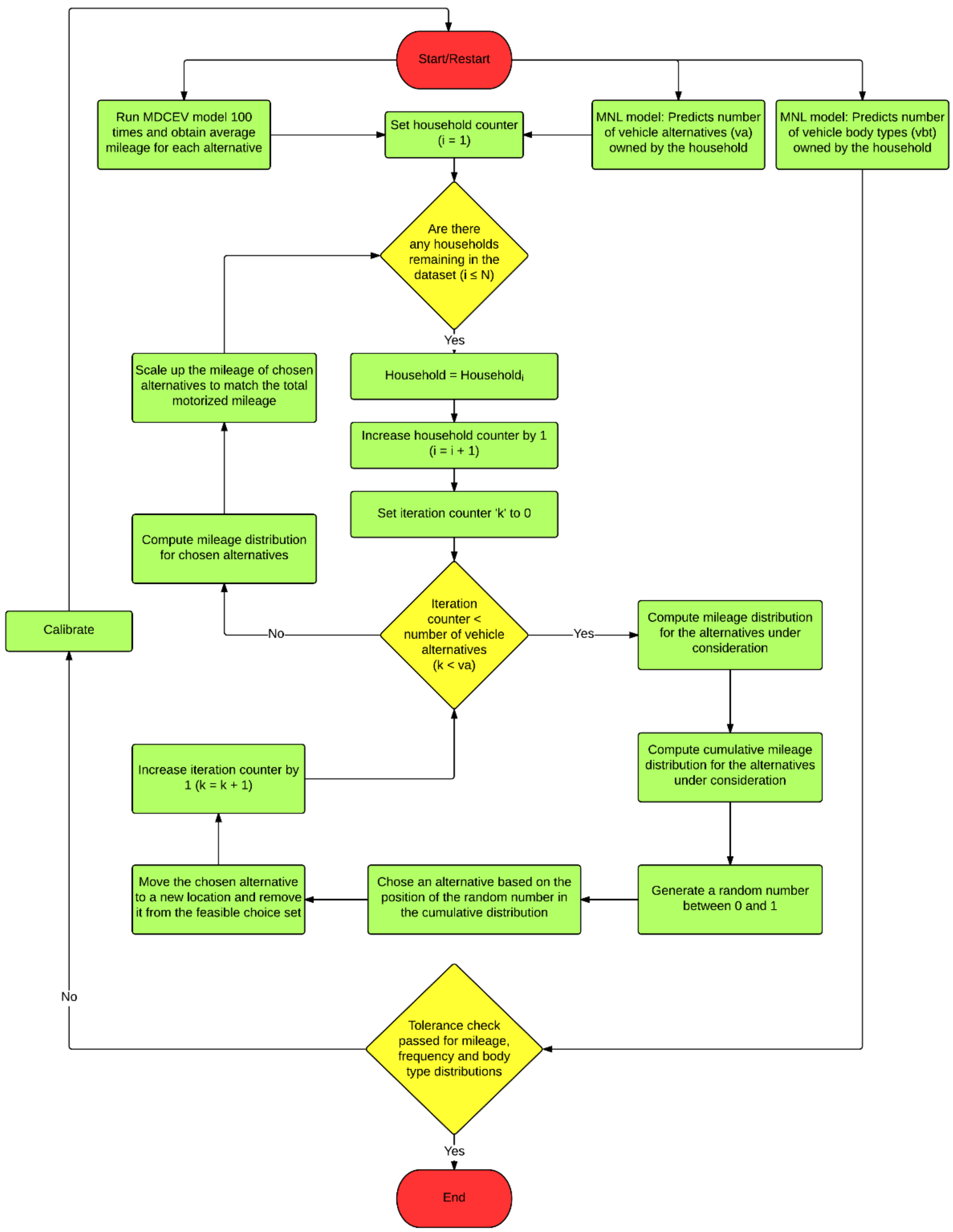

Figure 3.5. Heuristic mileage reallocation algorithm.

From the output of MDCEV model, cumulative mileage distribution of the household is computed. A random number is generated and based on location of the random number in the cumulative mileage distribution of the household, a vehicle is selected as owned by the household. The selected alternative is removed from the dataset thereby eliminating the possibility of choosing the same alternative multiple times. This process is carried out in a loop as dictated by the number of alternatives model. At the end of the loop, the HMR algorithm would select all the alternatives owned by the household. The mileage consumed by these alternatives is scaled up proportionally to account for the annual motorized mileage consumption of the household. Once, the HMR algorithm reallocates the mileages for all households according to the input provided by number of alternatives model, the predicted body type distribution of the entire population is compared against the body type distribution predicted by the MNL model of number of vehicle body types. The absolute percent difference is computed between both the distributions and checked against a pre-set tolerance limit. If the HMR algorithm passes the tolerance check, the output of HMR algorithm goes in as input to the count models. If not, the entire application is repeated after calibrating the model components as warranted.

This process is carried out repeatedly until the percent difference between the two distributions is within a set tolerance limit. The output of HMR algorithm provides the final fleet composition of every household in the dataset. The output of this algorithm would have successfully predicted the vehicle ownership of the household, body type and vintage composition of the vehicles owned. Sample output of the HMR algorithm for a household who owns 3 vehicle alternatives is shown in Table 3.2. This output goes as input to the count models.

Table 3.2

Output after HMR Algorithm

Vehicle Alternative	**Mileage**
Car (Age > 0 & ≤ 5 Years)	10000
Car (Age > 6 & ≤ 11 Years)	0
Car (Age ≥ 12 Years)	0
Van (Age > 0 & ≤ 5 Years)	0
Van (Age > 6 & ≤ 11 Years)	14000
Van (Age ≥ 12 Years)	0
SUV (Age > 0 & ≤ 5 Years)	2200
SUV (Age > 6 & ≤ 11 Years)	0
SUV (Age ≥ 12 Years)	0
Pick-up (Age > 0 & ≤ 5 Years)	0
Pick-up (Age > 6 & ≤ 11 Years)	0
Pick-up (Age ≥ 12 Years)	0
Motorbike	0
Total	***26200***

6. <u>*Count Models*</u>: Once the HMR algorithm re-distributes the mileage consumptions for all households such that they satisfy the marginal distributions provided by the body the distribution model, count models are applied for each household. The count models determine if all the mileage consumed by a household with a particular alternative belongs to one or multiple vehicles. Suppose, the output of HMR algorithm determines that a household uses a car 0-5 years old to travel 25000 miles annually, the count model determines if all of this mileage is put on just one car 0-5 years old or if the household owns multiple cars of 0-5 years of age. Ideally, a count model should be estimated for each of the 13 different vehicle categories defined for the MDCEV model,

but this might make the model system vulnerable because of too many components. So, it was felt prudent to estimate one count model for each of the vehicle body types, with vintage serving as an explanatory variable in the models. If the household has non-zero mileage consumption in any of the vintages of a vehicle body type, count model of that particular body type is applied for that household.

At the end of application of the entire model system, fleet composition of the household including body type, age and count of vehicles of each vehicle body type-age category is known along with their annual usage. Knowing the exact fleet composition is the first step toward accurate perdition of emissions. Each of the model components in the vehicle fleet composition model system are validated to test their predictive capability. Once each of the models are validated/calibrated to replicate observed distributions well, the model system is applied in its entirety to the data to see how well the model system as a whole would predict the observed fleet mix for the base year.

The MDCEV model of vehicle fleet mix is a comprehensive model that includes attributes at the household and zonal level. In addition to this, accessibility measures will computed for each zone to conduct sensitivity analysis to test changes in fleet mix with varying zonal accessibility. The hypothesis is that increasing zonal accessibility will propel lower auto-ownership levels and decrease pollution. With information about the tour composition for various types of tours undertaken by households (provided by the tour characterization framework) and information about the fleet composition of the household (as predicted by the fleet composition model system), the current effort proposes a framework that ties these two components together using a tour level vehicle type choice modeling framework described in the next section.

Tour Level Vehicle Type Choice Modeling Framework

Despite significant advancements in the activity-based modeling arena in the past decade or so, almost none of the existing models in practice house an extensive vehicle type choice model component that determines the particular household vehicle used by an individual to undertake a specific tour. Almost all of the activity-based microsimulation model systems only model the mode utilized (auto, transit, walk/bike) to make a trip/tour. It is important to identify the specific type of vehicle (among the vehicles owned by the household) used to make a specific trip/tour as this information has overbearing consequences in computing the emission footprint from person travel in a region.

A limiting reason for the activity-based models for not including vehicle type choice is lack of information regarding the fleet mix owned by a household. Without knowing 'what types' of vehicles are owned by a household, it is impossible to model 'which' vehicle (among the ones owned by household) would a person use to embark on a trip/tour. The vehicle fleet composition methodology discussed in the previous section provides a solution to this problem by predicting the fleet mix owned by a household classified by body-type and age. This information coupled with information regarding characteristics of a tour (solo/joint tour, tour accompaniment etc.) can be used to model the vehicle type choice of individuals at the tour level. The vehicle type choice models can be implemented as sub-models to a mode choice model, where a vehicle type choice model is applied to predict the type of vehicle use on a tour if and only if the mode choice model predicts the mode for the tour under consideration as auto.

Prior to entering this framework, but after the fleet composition framework, a primary driver allocation module identifies and allocates each vehicle in the household to a ‘primary driver’. This can be determined by allocating vehicle fleet to individuals in a household such that they maximize their utility. This behavior can be modeled using traditional discrete choice modeling methods or decided based on simple/complex heuristics involving one or more person level attributes (gender, income, age etc.). Once a vehicle is allocated to a primary driver, it is assumed that the vehicle is available to that driver throughout the day to travel to any activity.

With full information about all the stops on a tour from the tour characterization framework, and knowledge about fleet mix from the fleet composition module, the proposed tour level vehicle type choice framework is developed with an intent to allocate vehicles owned by the household to tours undertaken by them. The overall framework for tour level vehicle type choice modeling is shown in Figure 3.6. For each tour undertaken by the household, an exogenous joint tour formation component determines if the tour can be undertaken as a joint tour, depending on the travel dynamics of the household. For example, if a child needs chauffeuring, an adult in the household could perform a joint tour with the child. This can be a fully/partially joint tour depending on the activity agenda of the child. Similar joint tours can be formed between adults in a household. The joint tour formation component is an extensive component in itself and is left out in the current research effort. For the purposes of the proposed framework, it is assumed that the type of tour (single/partially joint/fully joint) is provided exogenously. The framework assumes that all household vehicles are available to drivers in the household at all times. This assumption is made for ease of model estimation and can be dealt with in a straight forward manner in application step, if real time vehicle availability information is known. A conceptual framework to this effect is proposed and discussed in a separate chapter (Chapter 8).

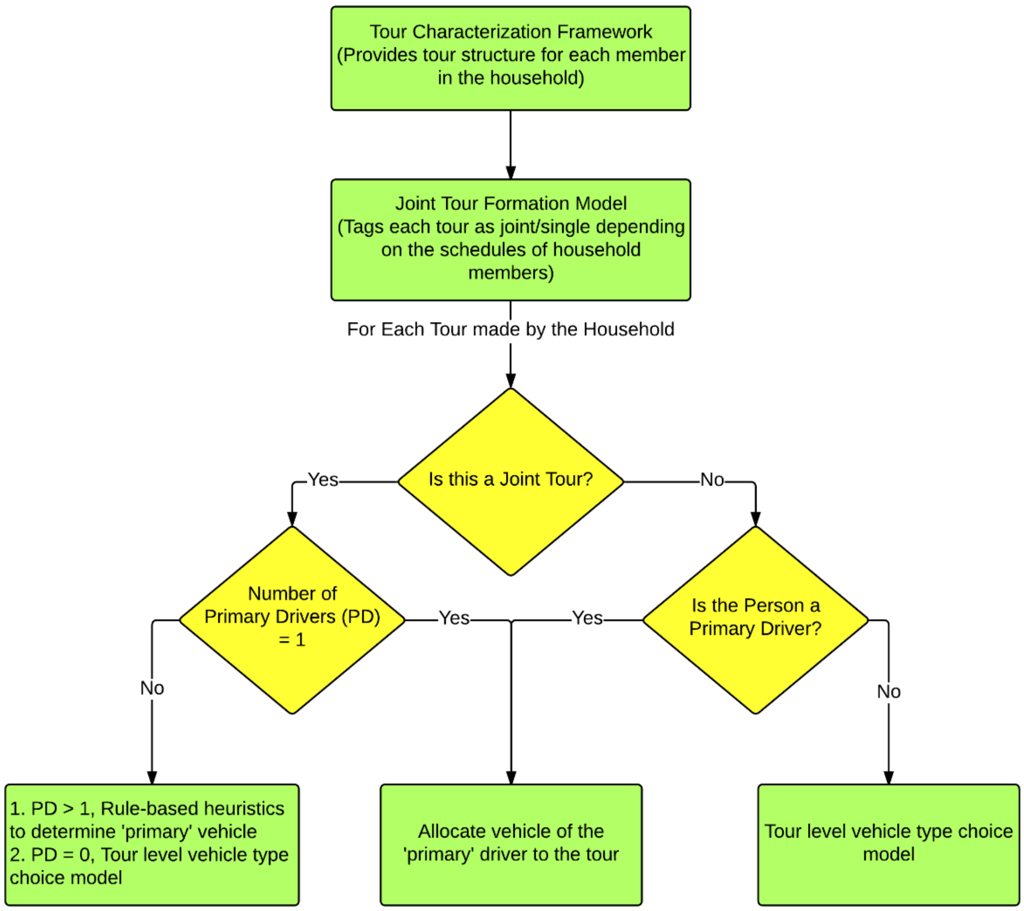

Figure 3.6. Tour level vehicle type choice framework.

Once the tour composition is identified, the proposed framework first checks if the tour is a solo/joint tour. For joint tours, a check is made to see if at least one individual on the tour is a primary driver that already has a vehicle allocated to him/her. If exactly one person on the tour is a primary driver, then vehicle allocated to the primary driver is assigned to the joint tour. If number of primary drivers on the tour is greater than one, rule based heuristics can be employed to see which primary driver's vehicle will be used for the joint tour under consideration.

In the case where there are no primary drivers on the joint tour, the vehicle type choice model determines which of the household vehicles will be utilized to carry out the tour under consideration. For solo tours, it is first checked whether the person undertaking the tour is a primary driver. If yes, the vehicle allocated to that particular person will be used to make the tour. If not, a vehicle type choice model including all the vehicles owned by the household is used to determine the vehicle type that would be used in the tour. The vehicle type choice model operates under the assumption that individual undertaking the journey has already chosen 'auto' mode to participate in the tour. This can be easily determined using mode choice models, which are well equipped in existing activity-based modeling systems.

The tour level vehicle type choice can be defined as a function of tour complexity as this information is made available by the tour characterization framework. Attributes such as number of stops on the tour, what activities are undertaken as the part of the tour as well as tour accompaniment can drive vehicle type choice in addition to individual attributes such as age income and gender to name a few. Tour composition is assumed to heavily influence the vehicle type choice. For example, a work tour with an escort stop will probably be undertaken using a 'bigger' vehicle in the household, whereas a tour that consists of only one shopping stop has the potential possibility to be undertaken by a car or an SUV. The tour level vehicle type choice model will be an MNL/NL model which will have the same vehicle classification as the MDCEV model of vehicle fleet mix. For each household, the dataset will be constrained to include only the fleet mix owned by them so that household tour level vehicle type choice can be depicted accurately.

The proposed model frameworks are aimed at enhancing tour level activity-based models in practice to a continuous-time domain with full representation of household fleet mix and tour level vehicle type choice. The research effort has a few important limitations identified below:

1. The research effort leaves out prediction of tour start and end times to an exogenous model component. This is not an operational shortcoming as activity-based modeling systems in practice are equipped to predict and provide tour start and end times (Parsons Brinckerhoff, 2010) to the tour characterization framework. The start and end times provided by such model systems will form the tour budget that MDCEV model allocates to various activities in the tour. The activity scheduling systems on the other hand might not be able to predict the tour start/end times. This is left for future research.
2. The current research effort does not include the joint tour formation component at the household level. This is an area of research that requires considerable exploration to come up with econometric modeling methods and rule based heuristics to develop joint tour formation at the household level.
3. This effort does not consider instantaneous vehicle availability where movements of vehicles in a household's fleet are tracked continuously through the course of a day. This is not so much of a modeling issue, but an implementation issue. The real-time vehicle accountancy framework would require constant exchange of information between the activity-based model and the dynamic traffic assignment model regarding

the usage of household vehicle fleet. The current effort proposes a conceptual framework that could potentially be implemented in such an integrated model system.

In summary, the proposed research effort aims to develop modeling frameworks that accurately depict the travel as well as fleet composition of households. This effort intends to enhance tour level activity-based model systems in practice with evolutionary continuous-time approaches followed by the activity scheduling models in research so that activity engagement patterns can be modeled as realistically as possible. A framework is proposed that ties the travel undertaken by household to vehicles owned by them. All of the model components in the proposed frameworks are estimated and validated for the 2009 NHTS add-on dataset for the Phoenix metropolitan region.

CHAPTER 4

TOUR CHARACTERIZATION FRAMEWORK

Tour level activity-based models that are widely adopted in practice follow a discrete representation of time (one hour or half hour time bins). In these models, activities performed by individuals in a day are modeled disjointly and are sequenced using rule based heuristics or discrete time-of-day choice models. Recent progress in econometric modeling arena (Bhat, 2005; 2008) allows for simultaneous prediction of the mix of activities undertaken by an individual in a tour. Adopting these advancements, a methodology is developed to enhance activity-based models in practice to follow a continuous time representation of activities. The proposed tour characterization framework is capable of predicting the mix of secondary activities pursued on a tour, time allocated to each activity and the order in which these activities are pursued. The proposed framework is discussed in detail in Chapter 3. This chapter presents the model estimation and validation results of various components in the tour characterization framework.

First, a brief account of the data preparation exercise is provided along with a discussion of the characteristics of different types of tours modeled. Model estimation results are presented for each of the components followed by the model's performance in replicating observed activity-travel patterns.

Data

The data used for estimating the model components of the tour characterization framework is from the latest wave of the National Household Travel Survey (NHTS) conducted in the year 2008-2009. NHTS data provides a wealth of information regarding travel patterns of individuals across the nation. A random sample of individuals representative of the population in each state are interviewed and data is collected regarding household/person level socio-demographics as well as travel characteristics of the individuals. Data collected from NHTS is organized into household, person, vehicle and trip files. Data collected at the household level includes details such as number of people/drivers/workers, household income etc. At the person level, information regarding the respondent's individual characteristics such as age, sex, driver status etc. are collected. A comprehensive travel diary is filled out by each respondent answering the survey that has detailed information regarding each trip (NHTS User's Guide, 2009) such as:

- Trip purpose (work, shopping etc.)
- Mode of transportation (e.g., car, bus, walk, light rail)
- Travel time
- Time of the day
- Travel day
- Trip composition
    - Occupancy
    - Driver characteristics
    - Vehicle attributes

This data can be used to understand the travel behavior at household as well as trip level (trip chaining patterns, modal usage etc.). Since the unit of analysis in the current research effort is a 'tour', but not an individual trip, data from NHTS is processed to convert trips reported by individuals into tours. A tour consists of a sequence of stops made by an individual that start and end the same anchor point. If a tour starts and ends at a home location, then the tour is called a home-based tour. If the anchor is not home, then the tour is called a non-home-based tour. Every tour is tagged with a primary purpose, which is the main motivation behind that particular journey. Tours that have work as primary purpose are called home-based work (HBW) tours, while tour that have other purposes (such as shopping, recreation etc.) are named accordingly. For the purposes of this effort, all the tours that have a primary purpose other than work are grouped together and labeled as home-based other (HBO) tours.

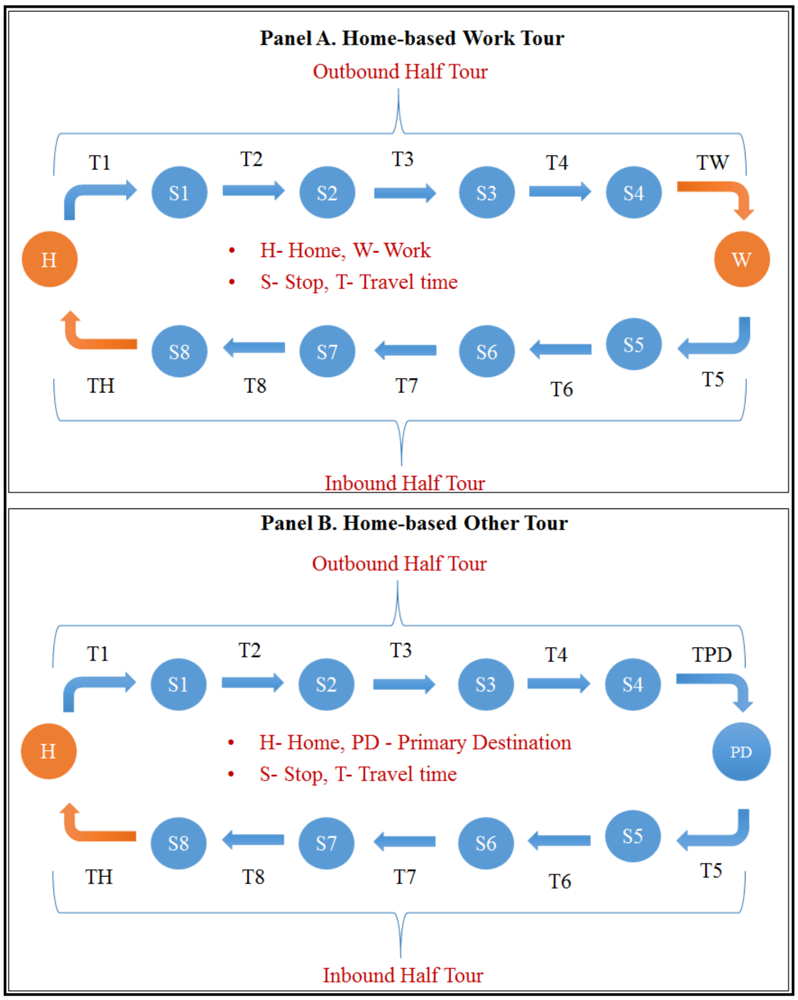

Figure 4.1. Structure of a tour.

Tours are constructed from trip level data by chronologically arranging all the trips made by an individual in a day. Travelers may undertake a variety of tours over the course of a day, and each tour may include a multitude of stops where individuals participate in various activities. The structure of home-based work and home-based other tours considered for this research effort is shown in Figure 4.1. Panel A presents the structure of home-based work tour and Panel B presents the structure of a home-based other tour. As mentioned earlier, each tour starts and ends at the same anchor point and is defined by a primary purpose. All the stops on the outbound journey from home to the primary destination constitute the outbound half tour and all the stops that are made on the return back home are together called as inbound half tour.

From preliminary analysis of the data, it was observed that maximum number of stops made on any half tour made by an individual is four. So, up to four stops on each half tour are taken into consideration for developing the tour characterization framework. All the stops other than the anchor points (home, primary destination) are called as 'secondary stops' made on the tour and are the main focus of this research effort. The intent of this research is to develop a framework to accurately predict the secondary stops made on a tour, time allocated to each of these stops and the sequence in which these stops are made in a continuous time domain. The framework proposed accounts for multiple stops of the same activity type made on a tour as dictated by the observations in the data.

Table 4.1 shows a brief sketch of the description of different types of tours considered for modeling. Statistics shown represent how many half tours in a specific tour type have 2, 3 and 4 stops on the outbound and inbound legs of the journey. Each panel represents a specific type of tour considered for modeling.

Table 4.1

Stop Making Patterns on Different Types of Tours

Panel A. Workers, HBW Tours				
	Outbound		Inbound	
Number of Stops on the Half Tour	Number of Half Tours	Percent (%)	Number of Half Tours	Percent (%)
2	42	63.60	162	72.60
3	16	24.20	40	17.90
4	8	12.10	21	9.40
Total	**66**	**100**	**223**	**100.00**

Panel B. Workers, HBO Tours				
	Outbound		Inbound	
Number of Stops on the Half Tour	Number of Half Tours	Percent (%)	Number of Half Tours	Percent (%)
2	77	69.40	76	65.50
3	21	18.90	30	25.90
4	13	11.70	10	8.60
Total	**111**	**100.00**	**116**	**100.00**

Panel C. Non-Workers, HBO Tours				
	Outbound		Inbound	
Number of Stops on the Half Tour	Number of Half Tours	Percent (%)	Number of Half Tours	Percent (%)
2	206	65.00	181	67.50
3	75	23.70	61	22.80
4	36	11.40	26	9.70
Total	**317**	**100**	**268**	**100.00**

From the table, it can be observed that among HBW tours, multiple stops are made predominantly on the inbound half tours, which is explained by the flexibility in activity participation for workers after their ‘regular’ work day. Whereas, HBO tours made by workers (tours made after or before work, which have a primary destination other than work), have more of an even spread of tours with multiple stops on both outbound and inbound half tours. For similar type of tours (HBO) made by non-workers, it was observed that a majority of outbound half tours have multiple stops. All these findings are intuitive and confirm to the usual activity-travel patterns expected from the respective market segments. Moreover, these patterns identify significant differences in stop making behavior among different types of tours and call for the necessity to develop separate model components for each tour type.

Further analysis was carried out to study the activity type distributions of secondary stops made on all the tours considered for analysis. Table 4.2 presents the activity type distributions observed on different types of tours. The total number of stops made on outbound and inbound half tours for all the tour types considered follow similar patterns observed in Table 4.1, as expected. There are some interesting similarities and differences in these different tour segments. Among all types of tours considered, outbound half tours have a greater proportion of maintenance stops, whereas inbound half tours tended to have more shopping stops, regardless of the segment under consideration. Other escort stops are more or less evenly distributed across outbound and inbound half tours across all segments. Within the worker segment, HBW tours have more maintenance stops on the outbound half tours, whereas HBO tours have an equal proportion of shopping and maintenance stops on the outbound half tours. Workers seem to push the shopping activity towards the return home journey of their HBW tour, which is intuitive. On HBO tours, which do not usually have a rigid temporal/spatial constraint, a more even placement of shopping stops is observed across outbound and inbound half tours. Another intuitive finding from this table is that HBW–outbound half tours have absolutely no social visit stops. This is finding is consistent with expectation, as individuals do not usually make a social visit stop on their way to work.

Table 4.2

Activity Type Distributions on Different Types of Tours

Activity Type	**Workers, HBW Tours**		**Workers, HBO Tours**		**Non-Workers, HBO Tours**	
	Outbound Half Tours (%)	**Inbound Half Tours (%)**	**Outbound Half Tours (%)**	**Inbound Half Tours (%)**	**Outbound Half Tours (%)**	**Inbound Half Tours (%)**
Other Escort	15.2	15.8	7.4	7.8	6.8	10.6
Shopping	26.2	34.4	33.8	34.8	31.9	45.8
Maintenance	32.9	25.1	33.8	27.0	35.9	23.1
Meal	17.1	11.2	13.4	19.1	14.0	12.8
Social Visit	0.0	4.2	3.0	4.3	2.7	3.2
Other Discretionary	8.5	9.3	8.6	7.1	8.8	4.5
Total No. of Stops	**164**	**526**	**269**	**282**	**781**	**649**

From observations in the data, it was felt prudent to estimate separate model components for worker and non-worker segments. Within tours made by workers, a classification is made between HBW tours and HBO tours as the stop making patterns on these tours are found to be significantly different.

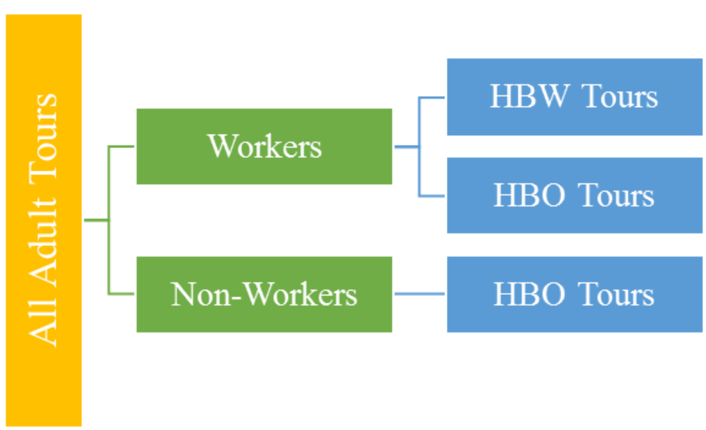

Figure 4.2. Segmentation considered for estimating components of tour characterization framework.

Segmentation considered for modeling exercise is depicted in Figure 4.2. In this chapter, results are presented here for HBW tours made by workers and HBO tours made by non-workers. Note that the model structure for HBO tours made by workers is similar to that of non-workers.

TCF Model Components for HBW tours made by Workers

Tours made by workers are classified as HBW tours and HBO tours depending on the primary destination and separate models components are estimated for each of the tour categories. Model components pertaining to HBW tours made by workers are presented in this section. All of the trips made by a worker (including secondary stops on the way to work and on the way back home) are synthesized to form HBW tours. Up to four secondary stops are considered on the inbound and outbound half tours. So, the total number of secondary stops that are considered on any HBW tour is eight. A total of 6 activity types are considered for classifying secondary stops made on a tour. A brief description of each of these activities is provided below.

- *Other Escort*: Activities such as dropping-off/picking up or chauffeuring an other household member. A different type of escort called as the school escort purpose is not considered in this modeling effort.
- *Shopping*: Includes activities such as shopping for apparel/electronics/gear etc. as we all as a quick grocery stop.

- *Maintenance*: Activities such as going to the bank, doctor or a beauty salon come under this category. This purpose is synonymous to personal business activity type seen in a lot of ABMs.
- *Meal*: Stop made to eat meal (lunch/dinner). A quick stop on the way to work to grab a cup of coffee also comes under this category.
- *Social visit*: Visiting friends, relatives etc. constitutes this purpose.
- *Other Discretionary*: A category defined to classify activities that do not belong to any of the purposes defined above. Examples of activities that fall under this category are jogging, going to the gym, recreational activities such as going to movie etc.

Multiple stops of the same activity are allowed on a tour. Table 4.3 presents the composition of HBW tours used for analysis. The last column of the table shows the percentage of tours that have at least 'n' number of stops of the activity under consideration. It can be observed that number of HBW tours that have secondary stops are quite few. Among, the tours that do have secondary stops, majority of them consist of shopping, maintenance and meal activities. This is expected behavior on HBW tours as workers are usually constrained by their work schedules and do not have a lot of flexibility to make secondary stops on the way to work. Further analysis of these tours revealed that most of the stops made on HBW tours are on the inbound half tours (work to home), corroborating the stated hypothesis. To model multiple stops of the same activity type, a criteria was set to consider activity types that have at least a 2% percent representation in the data set. According to this criteria, multiple stops of other escort and shopping are considered for estimating the MDCEV model of activity type mix, but the final model included 2 other escort stops. This model covers 93.3% of the tours considered for model estimation. Estimation and validation results of the activity type mix model are presented next.

Table 4.3

Composition of HBW Tours

Activity Type	Number of Episodes	Number of Tours	% of Total
Other Escort	1	165	5.62
	2	71	2.42
	3	14	0.48
	4	3	0.10
Shopping	1	369	12.56
	2	74	2.52
	3	12	0.41
	4	2	0.07
Maintenance	1	297	10.11
	2	51	1.74
	3	9	0.31
	4	2	0.07
Meal	1	208	7.08
	2	14	0.48
	3	1	0.03
	4	1	0.03
Social Visit	1	54	1.84
	2	2	0.07
Other Discretionary	1	143	4.87
	2	9	0.31
	3	1	0.03
Total number of tours in the data set			***2938***

HBW tours – MDCEV model of activity type mix. On any tour, an individual can participate in a multitude of activities and allocate different levels of time to each activity. This behavioral choice problem comprises of simultaneously modeling activity type mix (multiple discrete component) and the time allocated (continuous component) to each activity and lends itself nicely to the application of a Multiple Discrete Choice

Extreme Value (MDCEV) model. The MDCEV model has been successfully applied in the context of activity participation and time use process, and is proven to be quite effective in this context (e.g., Bhat, 2005; Bhat et al., 2006). In the proposed tour characterization framework, MDCEV model is utilized to identify stops (activity types) and durations of activity epochs within the tour. An epoch is defined as the sum of the activity episode duration and the travel time leading to the activity episode.

The MDCEV modeling methodology is presented in detail elsewhere (Bhat, 2008) and hence only a brief overview is presented here. The MDCEV model allows modeling of decision processes in which the choice makers are able to choose a mix of alternatives among the available alternatives, to maximize the utility of their consumption patterns. Individuals are allowed to select 'm' out of 'k' available alternatives and allocate varying levels of consumption to each of these alternatives. The functional form of utility proposed by Bhat (2008) is based on a generalized variant of the constant elasticity of substitution (CES) function:

$$U(x) = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \qquad (4.1)$$

$$\psi_k > 0, \gamma_k > 0, \alpha_k \leq 1$$

where $U(x)$ is a quasi-concave and continuously differentiable function with respect to consumption quantity vector x $(x_k \geq 0 \forall k)$. ψ_k represents the baseline marginal utility or the marginal utility at the point of zero consumption. α_k is the satiation parameter which governs the decrease in marginal utility with increasing consumption for good k. The translation parameter γ_k not only governs the level of satiation but also enables corner solutions (i.e., zero consumption of some goods). The baseline marginal utility ψ_k may be written in the following functional form:

$$\psi_k = \exp(\beta' z_k + \varepsilon_k) \tag{4.2}$$

where z_k is an *M*-dimensional column vector of attributes that characterize good k, β' is a corresponding vector of coefficients (of dimension M×1), and ε_k captures the effect of unobserved attributes. As both γ_k and α_k are parameters that incorporate the effects of satiation, it is difficult to uniquely identify and distinguish between them. For this reason, one of the two parameters is fixed and the other parameter is free to be estimated in most empirical model estimation efforts and the best model is chosen. In the current modeling context, γ-profile gave the best fit to the data.

The empirical context considered in this research effort is that of modeling activity engagement in the context of home-based work tours for workers who report going to work on the survey day. Home-based work tours will always include two activities – a work activity episode and a home activity episode. Any alternative that is always consumed by all decision-makers is considered an “outside good” in econometric modeling jargon and in the case of the home-based work tour characterization context, there are two outside goods – work and home. Thus, the MDCEV model is configured to accommodate two outside goods. With the presence of two outside goods, the specification of γ-profile of MDCEV model becomes:

$$U(x) = \sum_{k=1}^{2} \psi_{k,out} \ln(x_{k,out}) + \sum_{k=3}^{K} \gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right) \tag{4.3}$$

An individual maximizes his/her utility by optimally allocating consumptions to the k available goods, while always choosing the outside goods. Thus the constraint for the utility maximization problem is:

$$\sum_{k=1}^{K} t_k = T \tag{4.4}$$

where t_k is the budget allocated to good k and T is the total tour budget. Model estimation codes provided by Bhat (2008) were translated into open source coding language 'R' to estimate the MDCEV model and optimization algorithms within 'R' were used to search for parameter estimates that maximize the likelihood function. The MDCEV model estimation results for home-based work tour are presented in Tables 4.4-4.6. Table 4.4 presents the significant parameters in baseline utility equation of the MDCEV model. The model is found to offer intuitive results and reinforces the confidence for its use in estimating a model of activity type choice and time allocation. Among household attributes, it was found that individuals from households with more number of retirees are more likely to make other escort stops on their home-based work tours. Such individuals might usually drop-off/pick-up the retired household on the way to work or the way back home. Persons from households with high income tend to participate in maintenance and other discretionary activities on the HBW commute.

Table 4.4

HBW Tours, MDCEV Model of Activity Type Mix:

Significant Parameters in Baseline Utility

Activity Type	Explanatory Variable	Estimate	t-statistic
Other Escort Episode 1	Number of retirees in the household	0.31	1.27
	Number of vehicles per working individual in the household	-0.38	-2.63
	Multi-person tour made by a male person	0.35	2.06
	Tour mode is HOV	4.57	8.87
Other Escort Episode 2	Auto deficient household	0.97	3.27
	Count of total people on trip	0.21	2.08
	Tour mode is HOV	4.50	4.96
Shopping Episode 1	Respondent is a female person	0.41	3.78
	Number of household members between 0-5 years	-0.23	-2.04
Maintenance Episode 1	Age of the person ≥ 65	0.68	3.68
	Respondent is a male person	-0.61	-4.96
	Very high income household (> $100,000)	0.22	1.73
	Person is the driver on a multi-person tour	0.31	2.39
	Tour mode is any mode other than HOV and SOV	-0.72	-1.59
Meal Episode 1	Household size	-0.14	-2.53
	Respondent is a female person	0.28	1.85
	Tour mode is HOV	0.79	5.18
Social Visit Episode 1	End time of the tour 5pm - 7pm	-0.54	-1.72
	Number of vehicles in the household	-0.26	-1.71
	Respondent is a female person	0.82	2.74
Other Discretionary Episode 1	Very high income household (> $100,000)	0.67	3.84
	Auto deficient household	-0.50	-1.43

Workers from auto deficient households, tend to make other escort stops more, which is an intuitive finding as such households are more likely to engage in joint travel. Auto deficiency is also found to negatively influence other discretionary activity participation. Among person level variables, it was found that females tend to engage in more stop-making than males in the context of home-based work tours, which is consistent with findings reported in the literature (Mensah, 1995; Bhat, 1999). It was also found that older individuals tend to engage in more maintenance stops. Tour level attributes used in the model offer findings consistent with expectation. Other escort stops are found to be made using HOV mode and such stops are usually found to occur on multi-person tours. Tours made by non-auto modes are less likely to have maintenance stops (or any stops for that matter), as it might be inconvenient to make more stops when not using a faster and convenient mode (read auto). Meal stops tend to occur more on tours made using HOV mode. This is an intuitive finding as meal activity usually tends to be a joint activity. Table 4.5 presents the baseline constants and translation parameter values of the MDCEV model estimation result.

Table 4.5

HBW Tours, MDCEV Model of Activity Type Mix:

Baseline Constants and Translation Parameters

Activity Type	**Baseline Constant**		**Translation Parameter**	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
Work Episode (Outside Good 1)	-	-	0	NA
Home Episode (Outside Good 2)	-	-	0	NA
Other Escort Episode 1	-11.32	-21.1	18.31	3.95
Other Escort Episode 2	-13.33	-14.01	23.76	2.66
Shopping Episode 1	-8.03	-94.37	33.52	6.12
Maintenance Episode 1	-8.02	-81.67	39.08	6.73
Meal Episode 1	-8.46	-42.81	40.17	5.92
Social Visit Episode 1	-9.57	-21.06	87.94	2.19
Other Discretionary Episode 1	-9.07	-76.14	106.38	3.51

Baseline constants provide an indication of the inherent preferences for various alternatives and the marginal utility at zero consumption. These may be viewed as the preference for pursuing an activity relative to the outside goods (work/home in this case). From the results it was found that shopping and maintenance and activities have higher baseline preference while other escort activity has the lowest baseline preference. These findings line up nicely with trends observed in the estimation dataset and exhibit the capability of the model to accurately depict the observed activity-travel patterns. Translation parameters give an indication of the time allocated to an activity, once it is chosen. It is found that social visit and other discretionary activities have relatively larger translation parameters, suggesting that more time is allocated to these activities before satiation occurs. This finding is consistent with expectations and behaviorally intuitive.

It is interesting to see that the 2nd other escort activity episode has a higher translation parameter value that the first stop of the same activity. Upon further analysis, it was found that second episode of the other escort activity is more likely to happen on the inbound half tour (work to home journey), thereby revealing the reason behind this finding. Individuals usually have flexibility to spend a little more time to pursue any activity post work than on the way to work.

Table 4.6

HBW Tours, MDCEV Model of Activity Type Mix:

Goodness of Fit Measures

Statistic	**Value**
Log-likelihood of final model at convergence	-30728.4
Degrees of freedom of final model	36
Log-likelihood of base model at convergence	-31124.8
Degrees of freedom of base model	14
Likelihood ratio	792.7
$\chi^2_{22,0.001}$	48.27

The goodness of fit statistics of the MDCEV model are presented in Table 4.6. The model is found to offer an acceptable goodness of fit with a likelihood ratio of 729.7, which is substantially greater than the critical χ^2 value with 22 degrees of freedom at any level of significance. The estimated MDCEV model is applied on the entire dataset to see how well it can replicate the observed tour composition patterns. Procedures developed by Pinjari and Bhat (2011) are translated to open source coding language 'R' to apply the MDCEV model in forecasting mode. It should be noted that results presented do not constitute a true validation process. In a traditional validation process, a hold out sample (20-30% of the data) is kept aside and the model is estimated on the rest of the data. The estimated model is applied on the holdout sample to see how well it can replicate the observed patterns. However, in the current context, it was necessary to use the entire dataset for model estimation process to have adequate sample size for all the activity types considered. The estimated model is applied on the entire sample to see how well it can replicate the observed activity-travel patterns. This is more of a replication process than a validation. Comparisons are made across two distributions to test the predictive capability of the model.

- Activity Frequency Distribution: This represents the percent of home-based work tours that have each of the different stop types.
- Average Epoch Duration (excluding zero epoch durations): Average epoch duration is computed as the summation of all epoch durations of the stop type under consideration divided by total number of stops that have non-zero epoch durations.

Figure 4.3 presents the comparison of observed and predicted activity-travel patterns of uncalibrated version of the MDCEV model of activity type mix. It can be observed that the model is able to replicate the overall activity frequency distribution patterns quite well (from a qualitative standpoint). Both work and home epochs are observed and predicted as 100% as these are considered as outside goods in the current empirical context and hence have to be chosen on every tour considered for model estimation. Slight calibration is warranted to exactly match the observed distributions. It can be observed that the model represents the relative abundance of some activity types (shopping and maintenance) and paucity of others (other escort, social visit) quite well.

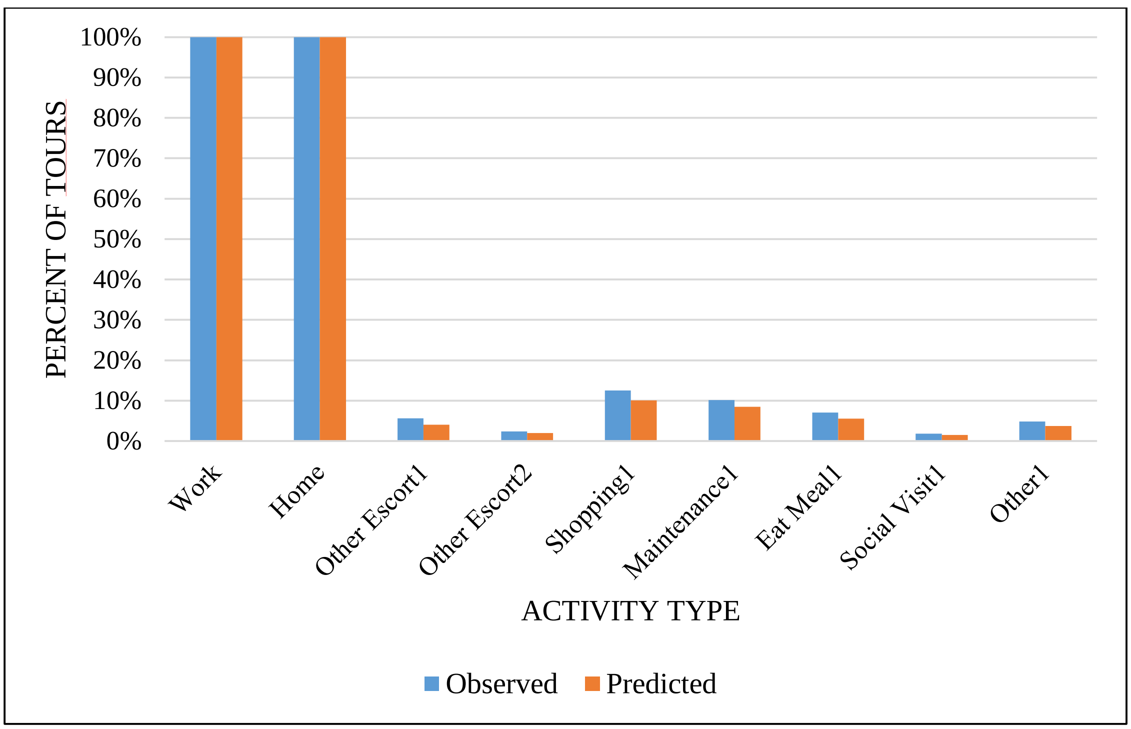

Figure 4.3. HBW tours, MDCEV model:
observed vs. predicted activity frequency distributions.

Figure 4.4 shows the comparisons of observed vs. predicted average epoch durations. From an initial glance it feels like the model is heavily overestimating the average epoch durations of all activity types. But, careful investigation revealed that it is not a systematic deficiency in predictions of the model but rather a problem with small sample size of stops for most activity types considered. For example, among all tours considered for estimation, there were a total of 200 meal stops made and the total duration of all of the meal epochs is 9000 minutes. This means that the average observed duration of meal epoch is 45 minutes. If the MDCEV model predicted that the total number of meal stops is only 150, with total meal epoch duration of 9000 minutes, the average predicted meal epoch duration will be close to 60 minutes for each meal epoch. The discrepancy of 15 minutes per an average meal epoch is more a manifestation of the small sample size than a systematic deficiency in model predictions. As mentioned before, any estimated model will require some amount of calibration to exactly match the observed patterns. The intent of this effort is to exhibit the ability of the proposed tour characterization framework to accurately represent the 'trends' observed in the data. Calibration of the model parameters is not undertaken as a part of the current research effort. It can be observed that the model qualitatively represents the observed epoch duration patterns quite well.

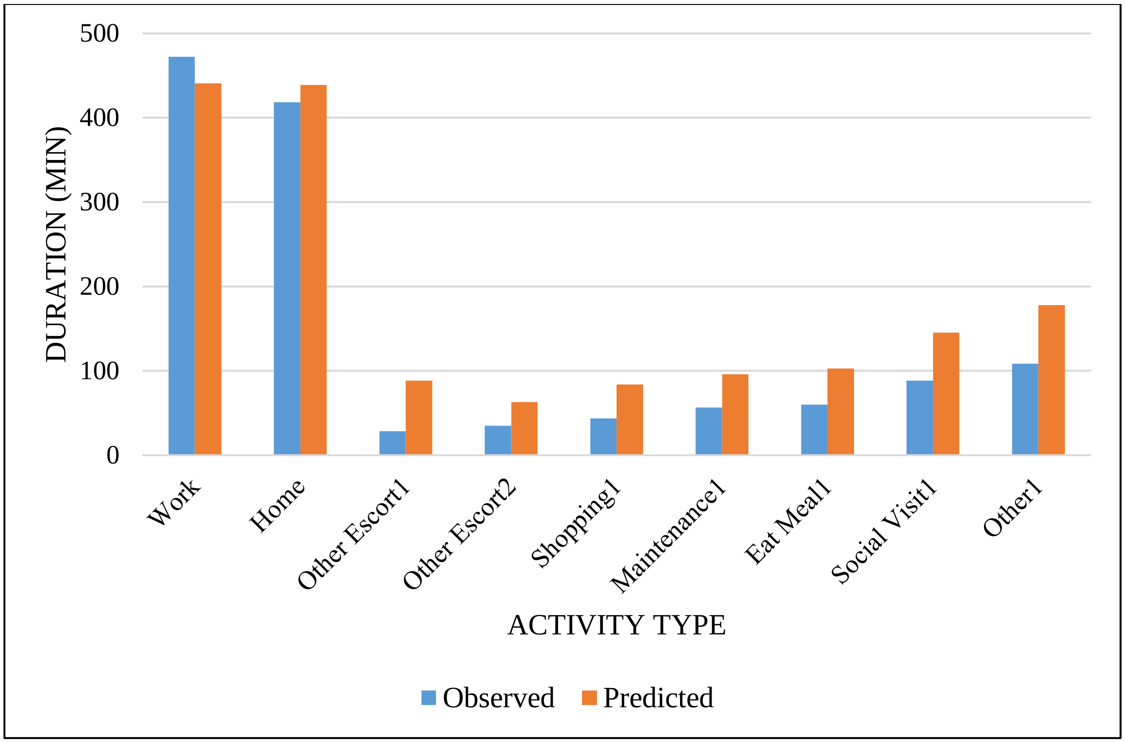

Figure 4.4. HBW tours, MDCEV model: observed vs. predicted average epoch durations.

Next component in the tour characterization framework is the stop sequencing model system that predicts the location of each stop on a tour with respect to the primary destination and then determines the relative position of each stop on a both inbound and outbound half tours.

HBW tours – Binary logit model of stop placement. The binary logit model of stop placement determines the location of each stop with respect to primary destination (i.e., either the inbound or the outbound half tour). To estimate this model, each stop in the dataset is categorized using a binary indicator where the stop gets a value of '1' if the stop is on the inbound half tour, '0' otherwise. Estimation results of binary logit model for HBW tours is presented in Table 4.7. A positive value for a coefficient in the model estimation result increases the chances of the stop occurring in the inbound half tour than on the outbound one.

Table 4.7

HBW Tours, Binary Logit Model of Stop Placement

Explanatory Variable	Coefficient	t-statistic
Constant	0.31	2.70
Respondent is a female person	0.28	2.41
Start time of the tour (7am - 9am)	0.43	3.41
Start time of the tour (9am - 11am)	-0.56	-3.19
Duration of Other Escort1 epoch on the tour	0.01	1.88
Duration of Other Escort2 epoch on the tour	-0.01	-2.61
Duration of Shopping1 epoch on the tour	0.01	4.08
Duration of Eat Meal1 epoch on the tour	0.003	1.87
Duration of Social Visit1 epoch on the tour	0.01	2.88
Goodness of Fit		
Log-likelihood at convergence for the full model		-876.61
Log-likelihood at convergence for the restricted model		-914.22
Likelihood ratio		75.21
$\chi^2_{8,0.001}$		26.13

From the results it was found that tours starting earlier in the day are more likely to have stops on the inbound half tours than tours starting later in the day. This is a very intuitive finding and refers to the schedules of different types of individuals. The early birds, who go to work quite early in the day, tend to make more subsistence stops (grocery shopping, going to the bank etc.) on the way back to home i.e., inbound half tour, While those starting their travel later in the day might take care of such activities before the start of a ‘work day’. Activities with longer durations have a greater proclivity to happen on the inbound half tour. This result is consistent with expectation as individuals usually tend to schedule activities with longer duration on the way back home as they have a greater time flexibility on the home to work commute. Tours undertaken by female respondents are found be more likely to have stops on the inbound half tour than the outbound half tour. The likelihood ratio of the model is greater than the critical χ^2 value at the 99% confidence level.

The estimated model is applied on the entire dataset to see how well it can replicate observed stop making patterns. Results of the comparison are depicted in Figure 4.5. First litmus test for the efficiency of the model is to predict the correct proportion of stops on the outbound and inbound half tours. From the figure it can be seen that model predicted the proportion of stops quite close to the observed data. An interesting observation here is that number of stops made on the inbound half tours is twice as many as on the outbound half tour. This finding is in line with observations from extant literature (Bhat, 1997) and it was encouraging to observe that the model predicts this pattern quite accurately. In addition to identifying the location of stops with resepct to primary destination, it was found that the model is able to predict the profile of activities on outbound and inbound half tours quite well. The statistic depicted in this graph is the percent of stops on each half tour that pertain to a specific activity. Panel A presents the result of observed and predicted activity type distributions on the outbound half tours and Panel B presents the results for inbound half tours. It can be seen from the comparison charts that the model is able to predict activity distribution patterns quite nicely on both legs of a HBW tour. The next section presents results of sequential activity type choice (SATC) model of outbound half tours.

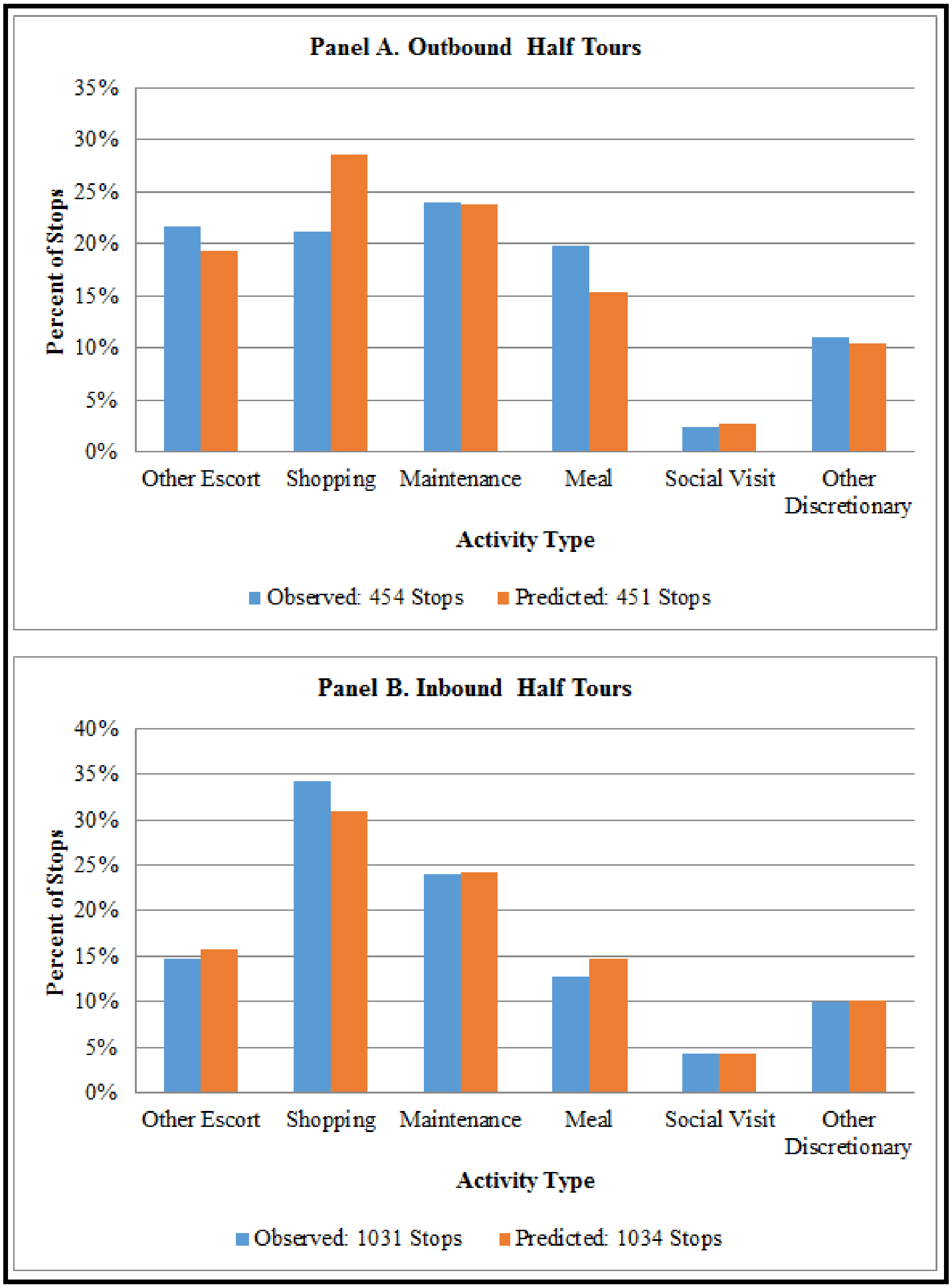

Figure 4.5. HBW tours, binary logit model:
observed vs. predicted activity type distributions.

HBW tours – SATC model for outbound half tours. Sequential activity type choice models are estimated with an intent to determine the 'impending activity' on a half tour. If there is only one stop other than the primary destination on any half tour, no sequencing is required as that stop is automatically positioned between home and primary destinations. If there are multiple stops on a half tour, then the sequence of is simulated using the SATC model. The SATC model is an MNL model for which every stop on a tour had a separate entry in the estimation data set. Information about all the activities pursued by the individual prior to the stop under consideration (both on the current tour as well as any other tours made earlier in the day) are included in the data set. This information entered the models in the form of binary variables indicating whether the individual did or did not perform an activity earlier in the day or on the tour under consideration.

For outbound half tours, prospective information about activities planned on subsequent portions of the tour (i.e., activities that are not yet scheduled) was included. Choice set for estimating the SATC model is constrained to only stops made on the outbound half tour. In application mode, the MDCEV would predict all the secondary stops on a tour and binary logit model would place all of these stops on either the outbound or the inbound half tours. Analysis of the estimation dataset revealed that only one social visit stop was made among all the stops considered for model estimation on the outbound half tour. As making a social visit stop on the outbound half tour is not a phenomenon observed often, this activity was not considered in the set of feasible activities that can be performed on an outbound half tour. Other discretionary activity was considered as the base alternative. The dataset used for estimating SATC models for outbound half tours consisted of a total of 164 stops. Model estimation results are furnished in Table 4.8.

Table 4.8

HBW Tours, SATC Model for Outbound Half Tours

Activity Type	Explanatory Variable	Coefficient	t-statistic
Other Escort	Constant	0.63	0.76
	History of other escort activity participation in the day	-3.74	-2.59
Shopping	Constant	1.32	1.45
	History of shopping activity participation in the day	-1.81	-1.89
	Anticipated shopping activity participation on the tour	-1.46	-1.46
Maintenance	Constant	1.38	1.44
	History of maintenance activity participation in the day	-4.25	-3.67
	Anticipated maintenance activity participation on the tour	-4.12	-3.65
	Anticipated shopping activity participation on the tour	2.78	2.70
Meal	Constant	1.83	2.03
	History of meal activity participation in the day	-4.36	-3.57
	Anticipated meal activity participation on the tour	-2.95	-3.14
Goodness of Fit Statistics			
Sample size (number of stops)		164	
Adjusted ρ^2		0.80	
Likelihood ratio		399.46	
$\chi^2_{(8,0.001)}$		26.13	

From the model results, it was found that history of activity participation earlier in the day, in general has a negative effect on the immediate occurrence of a similar activity on the outbound half tour. The behavioral interpretation of this result is that if an individual has already performed a specific type of activity earlier in the day (on the current tour or any other tour made earlier in the day), it is less likely that the individual will schedule a similar activity on the outbound half tour. Anticipatory activity participation has a similar impact on stop making patterns. In tours with both shopping and maintenance stops, an anticipated shopping activity on the tour positively influences the propensity to execute a maintenance stop as the next stop (on the outbound half tour). This finding is consistent with generally expected activity-travel patterns in that individuals usually take care of maintenance activities (such as going to the doctor/bank etc.) early on the tour, while taking care of shopping activities on the return home journey.

The estimated model was applied on the entire dataset to see how well it can predict the observed stop making patterns and the comparison is presented in Figure 4.6. Panel A presents the activity type distribution as observed in the data compared to the predicted distribution. The graph shows what percent of stops made on outbound half tours belong to each of the activity types considered in the SATC model. The model is able to replicate the observed activity type distribution very well. The performance of uncalibrated version of the model is quite appealing both from a qualitative and a quantitative standpoint.

The SATC model is aimed at predicting the 'next activity' on the half tour. So, in addition to accurately depicting the overall activity type distributions, it is important that the model also represent the activity type distributions at different stop levels on a half tour. This translates to the models ability to forecast the activity sequences observed in the data. Panel B of Figure 4.6 presents the comparison of observed and predicted activity type distributions at the stop level. Up to four stops were considered on each half tour, but the figure shows comparisons only until stop 3, as only 8 out of the 164 stops considered for model estimation were made as the 4th stop on the outbound half tour. For this reason, it was felt prudent not compare observed vs. predicted distributions on such meager sample sizes. The observed and predicted distributions at each stop level are presented side-by-side for easier comparison. Similar color coding scheme was maintained between observed and predicted activities. The legend at the bottom of the graph identifies observed activities categories with a suffix '_O' and predicted ones with a suffix '_P'. From the figure, it can be observed that the model is able to replicate activity type distributions at the stop level reasonably well.

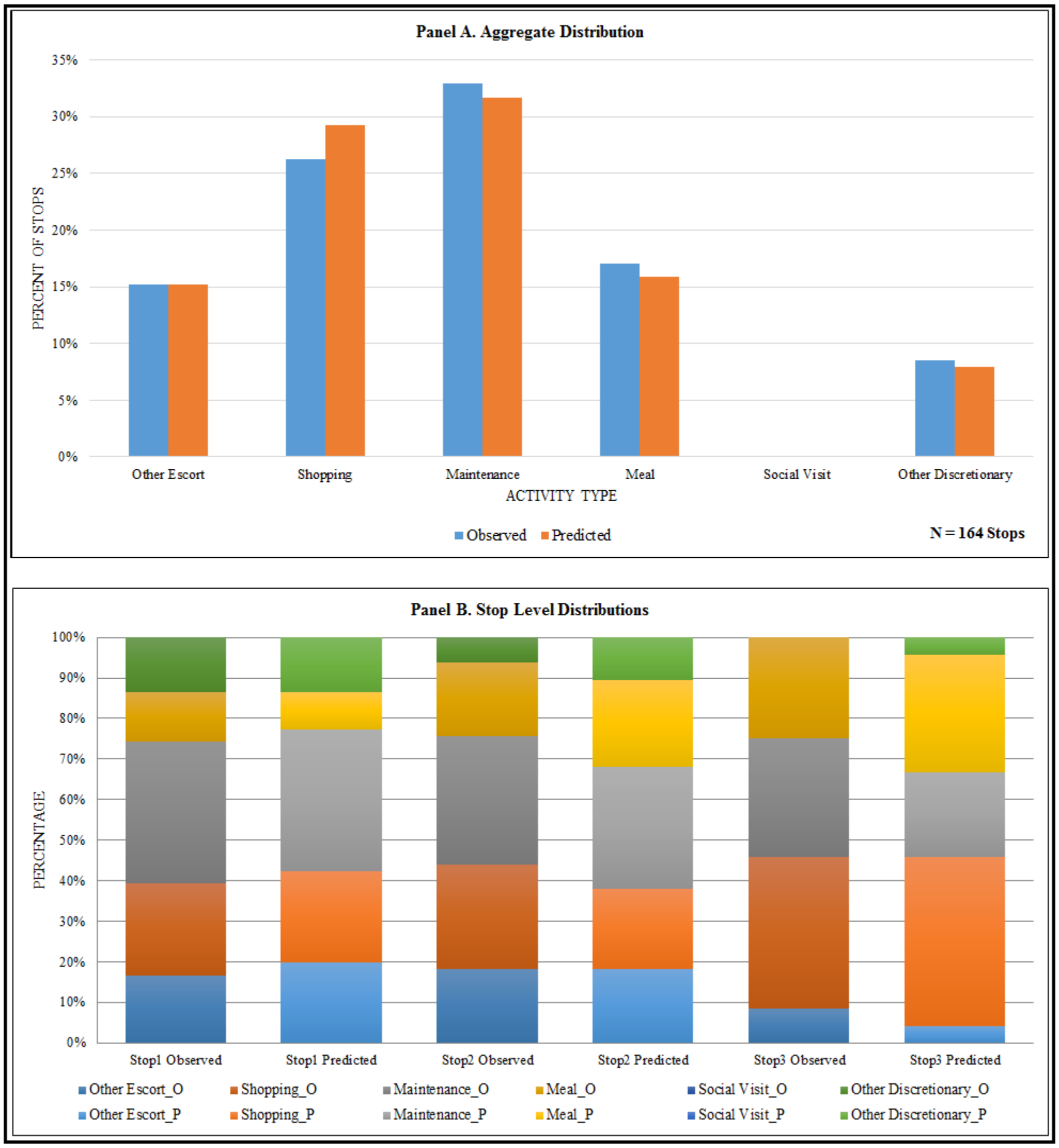

Figure 4.6. HBW tours, SATC model for outbound half tours: observed vs. predicted activity type distributions.

HBW tours – SATC model for inbound half tours. The structure of SATC model for inbound half tours is similar to that of the outbound half tours. To model the sequence of stops on the inbound half tours, information regarding history of activity participation earlier in the day as well anticipatory activity participation on the current half tour are taken into consideration. From the data used for model estimation, it was observed that a total of 526 stops were made on the inbound half tours as opposed to a mere 164 stops made on the outbound half tours. This is an intuitive observation, consistent with the notion that individuals usually have greater flexibility on the inbound journey (evening commute) of a HBW tour than that of the morning commute. Model estimation result of SATC model for inbound half tours is presented in Table 4.9.

From the results it was found that pursuit of 'other escort activity' as the next activity is positively influenced by the presence of a planned meal or social visit activity on the inbound half tour. This finding has an intuitive behavioral interpretation as meal and social visit activities usually tend to be joint activities, thereby necessitating the occurrence of an other escort activity (where the worker is probably picking up a family member), before meal or shopping activities. Similar pattern is observed for placement of meal and shopping activities, where occurrence of a shopping activity is found to precede the occurrence of a meal activity. This behavior is consistent with expectation as individuals often participate in a meal activity after a shopping activity. This is finding is nicely complemented by observation that history of shopping activity participation positively influences the occurrence of a meal stop as the impending activity.

Table 4.9

HBW Tours, SATC Model for Inbound Half Tours

Activity Type	Explanatory Variable	Coefficient	t-statistic
Other Escort	Constant	-0.16	-0.19
	History of other escort activity participation in the day	-0.83	-1.4
	Anticipated other escort activity participation on the inbound half tour	-1.64	-2.82
	Anticipated meal activity participation on the inbound half tour	2.92	2.91
	Anticipated social visit activity participation on the inbound half tour	4.19	2.73
Shopping	Constant	1.18	1.44
	History of shopping activity participation in the day	-2.27	-3.71
	Anticipated shopping activity participation on the inbound half tour	-2.68	-4.18
	Anticipated meal activity participation on the inbound half tour	3.38	4.07
	Anticipated social visit activity participation on the inbound half tour	3.22	2.72
	End time of the tour (3pm - 5 pm)	-0.94	-1.72
Maintenance	Constant	1.25	1.64
	History of maintenance activity participation in the day	-3.7	-5.7
	Anticipated maintenance activity participation on the inbound half tour	-3.12	-4.62
	Anticipated shopping activity participation on the inbound half tour	1.28	2.09
	Anticipated meal activity participation on the inbound half tour	3.33	3.33
Meal	Constant	0.38	0.58
	History of meal activity participation in the day	-3.13	-4.81
	History of shopping activity participation in the day	1.03	1.48
Social Visit	History of social visit activity participation in the day	-3.21	-2.26
	Anticipated meal activity participation on the inbound half tour	3.39	1.93
	End time of the tour (5pm - 7 pm)	1.93	1.42
Other Discretionary	Constant	-0.23	-0.35
	History of other discretionary activity participation in the day	-2.26	-3.1
	History of maintenance activity participation in the day	-1.22	-1.75
Goodness of Fit Statistics			
Sample size (number of stops)		526	
Adjusted ρ^2		0.83	
Likelihood ratio		1381.85	
$\chi^2_{(20,0.001)}$		45.32	

Tours ending later in the day (5-7 p.m.) had a greater propensity for occurrence of a social visit stop as the next activity in the inbound half tour, which is expected behavior as people usually tend to participate in social visit activities after completing all other activities in their agenda (duration of social visit activities in general is observed to be longer, which would logically place them after the completion of all other activities on a tour). Similar to outbound half tours, history of activity participation of an activity (earlier in the day) had a negative impact on pursuing a similar activity as the 'next' activity. The likelihood ratio statistic of the model is substantially greater than the critical χ^2value at 99% level of significance, reinforcing the confidence in statistical validity of the model.

The estimated model is applied on the entire dataset to see how well it can replicate the observed stop sequencing patterns on inbound half tours and results of the comparison are presented in Figure 4.7. From the figure, it can be observed that majority of the stops made on inbound half tours are for shopping and maintenance activities. Panel A presents the comparison of observed and predicted overall activity type distribution of stops made on the inbound half tours. The uncalibrated model performed exceedingly well in replicating the observed activity type distributions. Panel B presents a similar comparison, but at the disaggregate stop level, which tests the efficacy of the model to accurately depict the stops sequencing patterns observed in the data. The observed and predicted distributions at each stop level are presented side-by-side for easier comparison. A total of four stops are considered for modeling on the inbound half tour.

Figure 4.7. HBW tours, SATC model for inbound half tours:
observed vs. predicted activity type distributions.

From the figure it can be observed that shopping stops are usually kept until the end of the tour (final stops before heading home), whereas maintenance stops are usually completed early on in the inbound half tour. This finding makes intuitive sense as maintenance activities such as visiting a doctor or going to a bank, salon etc. are more time constrained than shopping activities which can be carried out at a person's own convenience. The proportion of social visit activity is quite small across all stops and this is expected behavior, as one would not usually prefer to make a social visit stop after a tiring workday (as a part of the HBW tour). From pairwise comparisons, it can be seen that the model is able to replicate the observed stop making patterns very well, giving the necessary confidence to use it to determine the sequence of stops in multi-stop half tours.

Next section presents the estimation results of model components of the tour characterization framework estimated for HBO tours made by non-workers. As discussed before, significant differences were observed in tour compositions between worker and non-worker segments. A readily apparent reason for this is the daily schedules of both these segments. While workers on the one hand are constrained by work start/end times and tend to plan their day around the 'work' activity, non-workers on the other hand have more flexible schedules to participate in different types of activities throughout the day. It was observed that HBW tours had majority of the stops made on inbound half tours (the return home journey), whereas HBO tours made by non-workers have a greater proportion of stops on the outbound half tours. HBO tours made by workers are more or less similar to HBO tours made by non-workers. To be able to identify the difference in tour compositions across different types of tours, the next section presents the model estimation and validation results of tour characterization framework for HBO tours made by non-workers.

TCF Model Components for HBO Tours made by Non-Workers

This section presents the estimation and replication results of HBO tours made by non-workers. The major difference in model structure between HBW tours and HBO tours is that while HBW tours have two outside goods (work and home epochs), HBO tours have only one outside good or one alternative that is consumed by every individual in the dataset, which is home epoch (travel home + home sojourn). For HBW tours, the MDCEV model of activity type mix simulates all secondary stops (other than work and home), while for HBO tours, the model simulates all activities including the primary destination. The distribution of primary purpose for HBO tours is taken as an exogenously provided information and the MDCEV model is run multiple times until it matches the observed primary purpose distribution. If the MDCEV model simulates multiple epochs for the activity tagged as primary purpose, the epoch with longer duration is considered the primary purpose, thereby making all the other epochs of that activity, secondary stops on the tour. The stop sequencing system however works exactly in the same way for HBW and HBO tours, where a binary logit model first simulates the half tour to which each secondary stop on the tour belongs to and then the sequential activity type choice model simulates the sequence of activities on each half tour.

Table 4.10 presents the composition of HBO tours made by non-workers. HBO tours made by non-workers had more stops on them relative to the previous tour type (HBW tours made by workers), which is explained by the flexibility in daily schedules for this segment. Representation of multiple episodes of same activity type on a tour gradually decreases, which is an intuitive observation. It is rare to find a tour with 4 separate shopping stops (or any other stops for that matter), as individuals usually bunch activities together, than make multiple stops for the same activity. Following the criteria set before, any stop that has at least a 2% representation in the dataset is included in model estimation for MDCEV model of activity type mix. According to this criteria up to 2 episodes of other escort and maintenance, 3 shopping episodes and one stop of all other activities is considered for model estimation. In addition to these, a second stop of other discretionary activity is also included in the final model. This model configuration accounts for 95% of HBO tours in the dataset. Model estimation results for HBO tours made by non-workers is presented next.

Table 4.10

Composition of HBO Tours (Non-Workers)

Activity Type	Episode Number	Number of Tours	% of Total
Other Escort	1	528	10.89
	2	100	2.06
	3	24	0.50
	4	7	0.14
Shopping	1	1771	36.54
	2	405	8.36
	3	125	2.58
	4	33	0.68
	5	7	0.14
	6	3	0.06
Maintenance	1	1561	32.21
	2	241	4.97
	3	59	1.22
	4	12	0.25
	5	2	0.04
Meal	1	739	15.25
	2	22	0.45
Social Visit	1	311	6.42
	2	12	0.25
Other Discretionary	1	1236	25.50
	2	49	1.01
	3	3	0.06
	4	2	0.04
Total Number of Tours			***4847***

HBO tours (non-workers) – MDCEV model of activity type mix. As discussed before, the MDCEV model of activity type mix for HBO tours consists of only one outside good i.e., only one activity that is consumed by everyone in the dataset, the home epoch. Start and end times of all tours made by an individual in a day are assumed as exogenous inputs to the tour characterization framework. From this information tour budget can be computed (start time – end time of the tour) which goes as input to the MDCEV model. Home sojourn is computed as the difference between end time of the tour and end of day if the individual made only one tour in a day. For individuals who made multiple tours in a day, home sojourn is computed as end time of the current tour minus start time of the next tour. Unit of analysis considered foe MDCEV model of activity type mix is an 'epoch' which is the summation of travel time to an activity and the duration of activity participation. Choosing epoch as the unit of analysis facilitates the modeling of tours in continuous time domain. For MDCEV model with a single outside good, the utility function is given in equation 4.5. Model estimation results are provided in Table 4.11.

$$U(x) = \psi_{out} \ln(x_{out}) + \sum_{k=2}^{K} \gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right) \tag{4.5}$$

The model offered intuitive and behaviorally consistent results. Among household attributes, it was found that individuals from households with more young adults (6 – 17 years) are more likely to make other escort stops on their HBO tours. This might point to the chauffeuring necessities of children in such households. Lowest income households are found to have lesser proclivity to make meal stops, as such households might usually return home for meal activity. Larger households are also found to have a lesser probability to make meal stops.

Table 4.11

HBO Tours (Non-Workers), MDCEV Model of Activity Type Mix:

Significant Parameters in Baseline Utility

Activity Type	Explanatory Variable	Coefficient	t-statistic
Other Escort Episode1	Number of retirees in the household	-0.47	-7.45
	Number of household members between 6-17 years	0.31	6.07
	Tour modes is HOV	1.84	13.86
Other Escort Episode 2	If person is driver on the multi-person tour	2.71	7.75
	Respondent is a female person	0.79	3.29
	Count of total people on tour	0.36	3.99
Shopping Episode 1	Start time of the tour 1pm - 3pm	0.47	6.90
	Respondent is a female person	0.15	2.68
	History of shopping activity	-0.35	-4.52
	Tour modes is SOV	0.41	7.07
Shopping Episode 2	Respondent is a female person	0.42	3.14
	History of shopping activity	-1.05	-4.83
	Tour modes is HOV	0.25	2.25
Shopping Episode 3	Respondent is a female person	0.52	2.20
Maintenance Episode 1	Person is the driver on a multi-person tour	-0.13	-1.72
	Count of total people on tour	-0.18	-3.31
	Age of the person $\geq$ 65	0.09	1.71
	Tour modes is SOV	0.13	1.63
Maintenance Episode 2	History of maintenance activity	-0.40	-1.91
	Very high income household (> $100,000)	-0.52	-2.16
	Tour modes is SOV	0.56	4.04
Meal Episode 1	Respondent is a female person	-0.19	-2.38
	Household size	-0.25	-5.94
	Lowest income household (< $25,000)	-0.35	-3.35
	Count of total people on tour	0.40	9.64
	Tour mode is any mode other than auto	-0.66	-3.28
Social Visit Episode 1	Start time of the tour 5pm - 7pm	0.31	1.77
	Respondent is a female person	0.37	2.93
	Number of vehicles per person in the household	0.28	2.30
Other Discretionary Episode 1	Household size	-0.05	-1.67
	Number of vehicles per person in the household	0.30	5.49
	Tour mode is any mode other than auto	1.42	17.54

Auto sufficient households (households with more number of vehicles per person) are found to make more other discretionary stops on the tour, as such households do not have vehicular constraints and are free to undertake discretionary travel. Among person level attributes, older individuals (≥ 65 years) tended to make more maintenance stops, while female individuals had a greater probability to make shopping stops on HBO tours. Tour level attributes used in the model show intuitive signs. It was found that higher the number of people on the tour, greater is the propensity to have an other escort and meal stops on the tour, consistent with expectation as these activities usually tend to be joint activities. Having more number of people on the tour had a negative influence on the occurrence of a maintenance activity on the tour, meaning that such activities are usually performed on solo tours, rather than joint tours. It was found from the model results, that most of the activities undertaken as a part of the HBO tour had SOV as the tour mode, which indirectly points to the fact that HBO tours are predominantly solo tours. Other escort activities however had a greater probability of happening on tour with mode as HOV. History of shopping activity participation earlier in the day had a negative impact on shopping activity participation in the current tour, which is another finding consistent with expectation. Other discretionary activities are found to have high probability of occurring on tours which are made using non-auto modes. This might point to recreational activities such a walking/jogging/biking etc. Table 4.12 presents the baseline constants and translation parameter values of the MDCEV model estimation result.

Baseline constants may be viewed as the preference for pursuing an activity relative to the outside good (home epoch in this case). Since majority of the time on HBO tours is allocated to home epoch (travel home + home sojourn), all of the coefficients are negative meaning they are less preferred relative to the home epoch. From the model results, it was found that shopping and maintenance activities are the most pursued activities on HBO tours (values of baseline constants for these activities are then that of others). This finding represents the models ability to depict inherent preferences observed in the data as shopping and maintenance are indeed the most chosen stops on HBO tours with 36% and 32% representation in the data (see Table 4.10) respectively. The least preferred activity undertaken on HBO tours is other escort activity, again a finding that lines up with the observations from estimation data. Among activity types for which multiple stops are considered, the preference for making 'another' stop of the same activity type gradually decreases (for example, baseline preference for shopping episode 1 > shopping episode 2 > shopping episode 3), which is consistent with expectation.

Table 4.12

HBO Tours (Non-Workers), MDCEV Model of Activity Type Mix:

Baseline Constants and Translation Parameters

Stop Type	**Baseline Constant**		**Translation Parameter**	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
Home Episode (Outside Good)	-	-	0	NA
Other Escort Episode 1	-8.95	-65.14	26.04	9.27
Other Escort Episode 2	-12.74	-25.97	24.99	3.47
Shopping Episode 1	-6.76	-115.65	33.22	14.44
Shopping Episode 2	-8.48	-65.17	38.88	5.62
Shopping Episode 3	-9.76	-49.09	36.97	3.41
Maintenance Episode 1	-6.47	-52.29	39.49	16.42
Maintenance Episode 2	-8.89	-85.30	32.47	5.87
Meal Episode 1	-7.52	-65.17	60.84	9.13
Social Visit Episode 1	-9.03	-54.89	177.79	6.41
Other Discretionary Episode 1	-7.48	-74.10	96.70	12.91
Other Discretionary Episode 2	-10.38	-70.00	123.70	2.38

Translation parameters give an indication of the time allocated to an activity, once it is chosen. In other words, greater the value of the translation parameter, higher is the time allocated to the activity. Though social visit is not the most preferred activity, it has the highest translation parameter value, which means that individuals who do participate in a social visit activity, pursue it for longer durations, which is generally observed behavior. Other escort activity has the least translation parameter value, as escort activities (such as drop-off/pick-up) usually do not last very long. Table 4.13 presents the goodness of fit statistics of the estimated model. The estimated model has a likelihood ratio of 1547.36, substantially greater than the critical χ^2 at any level of significance which indicates a good model fit.

Table 4.13

HBO Tours (Non-Workers), MDCEV Model of Activity Type Mix:

Goodness of Fit Measures

Statistic	**Value**
Log-likelihood of final model at convergence	-54939.25
Degrees of freedom of final model	54
Log-likelihood of base model at convergence	-55712.93
Degrees of freedom of base model	22
Likelihood ratio	1547.36
$\chi^2_{32,0.001}$	62.49

The estimated model is applied on the entire estimation dataset to see how well it can replicate the observed activity type and duration distributions. Figure 4.8 shows the comparison of observed and predicted activity-travel patterns of uncalibrated version of the MDCEV model of activity type mix. The patterns shown in the graph depict what percent of tours in the dataset have at least one stop of the activity episode under consideration. From the figure it can be observed that tours with multiple stops of the same activity type are fewer than those with only one stop of that activity.

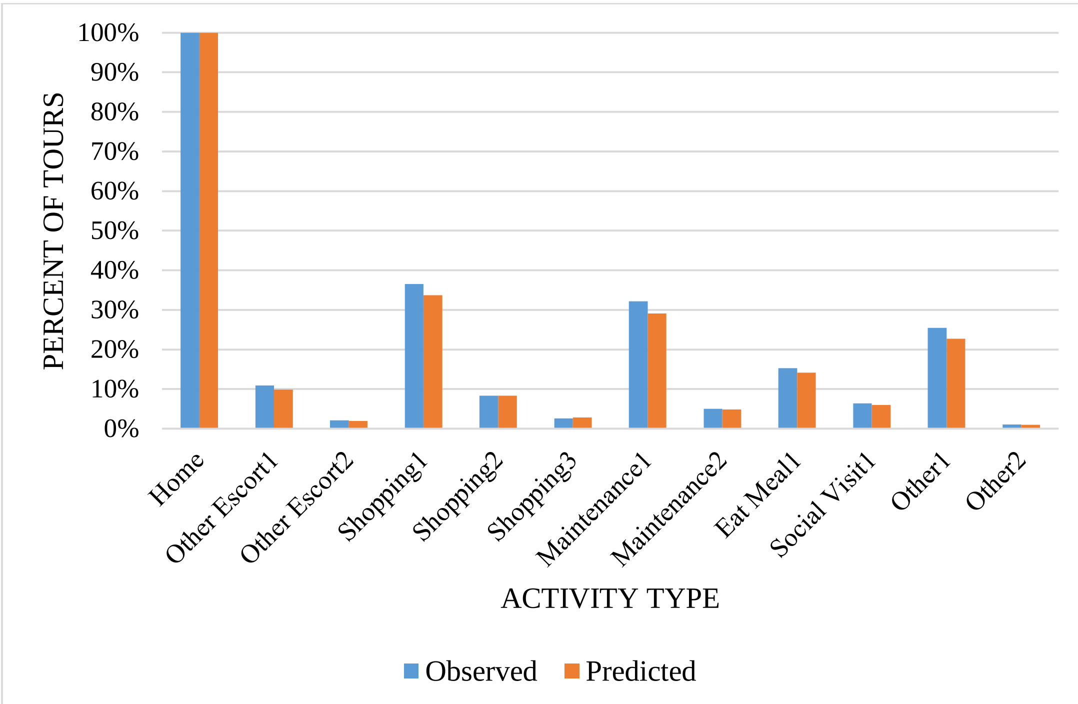

Figure 4.8. HBO tours (non-workers), MDCEV model:
observed vs. predicted activity frequency distributions.

The model accurately predicts this pattern within each activity type and also predicts the overall distribution of activity types quite well. Slight calibration of the model is warranted to match the observed distributions better. In the figure it can be seen that home episode is predicted on each and every tour as this activity is considered as an outside good in this model specification. Majority of the HBO tours made by non-workers have shopping, maintenance and other discretionary activity stops (in that order) in them and the model captures this nuance in the observed data very well. Overall, the predictions of the uncalibrated version of the model line up well with the observed patterns in the data.

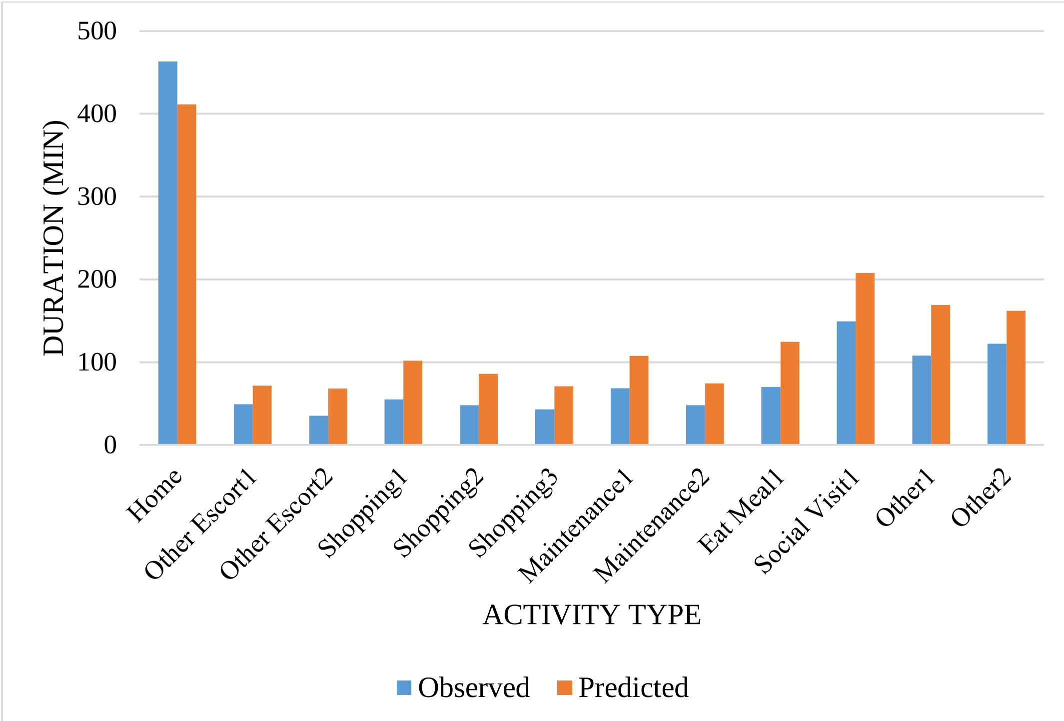

Figure 4.9. HBO tours (non-workers), MDCEV model:
observed vs. predicted average epoch durations.

Figure 4.9 presents the comparisons of observed vs. predicted average epoch durations. As it was observed in the case of HBW tours, the model seems to over predict the average epoch durations but it should be kept in mind that comparisons shown only consider stops with non-zero durations. So the over prediction of epoch durations is actually a manifestation of under prediction of stops on the tours (see Figure 4.8) coupled with small sample size issues. But, more importantly, it is to be observed that the model is able to predict the duration distribution trends quite accurately. For example, among all the activity types considered social visit stops have the longest epoch durations which is consistent with expectation. The model is able to predict this pattern quite well. The decreasing epoch durations for second stop of the same activity type (across all activities) are also captured by the model.

HBO tours (non-workers) – Binary logit model of stop placement. The binary logit model of stop placement determines the location of each stop with respect to primary destination (i.e., either the inbound or the outbound half tour). To estimate this model, every stop in the dataset is identified using a binary indicator where the stop gets a value of '1' if the stop is on the inbound half tour, '0' otherwise. Estimation results of binary logit model for HBO tours made by non-workers is presented in Table 4.14. Since the primary purpose of HBO tours is varied (and is known beforehand), a primary purpose dummy is used in the model to explain the stop placement.

An interesting deduction is made from the value of constants in the binary logit models of stop placement for HBW tours (workers) and HBO tours (non-workers). The value of the constant signifies the probability of a stop occurring on the inbound half tour, ceteris paribus. While the sign of the constant for HBW tour model is positive which means that all else being equal, a stop has a higher probability of occurring on the inbound half tour, the sign on the constant for HBO tour (non-worker) model is negative which means that the probability of stop occurring on the inbound half tour of HBO tours is lesser. This is a pertinent finding in that it explains the difference in activity-travel patterns on different types of tours under consideration and hence the necessity for separate models. Workers on the one hand are more constrained by work schedules in the morning time and tend to make more stops on the evening commute which can be loosely dubbed as the inbound half tour (explained by the +ve sign of the constant in binary logit model for this segment, see Table 4.7). Non-workers on the other hand have flexible schedules and tend to front load the activities on a tour and get as much as done on the outbound half tour (explained by the -ve sign of the constant in binary logit model for this segment).

Table 4.14

HBO Tours (Non-Workers), Binary Logit Model of Stop Placement

Explanatory Variable	**Coefficient**	**t-statistic**
Constant	-0.50	-7.41
Number of children in the household	0.17	3.69
Primary purpose of the tour is Other Escort	0.60	4.05
Primary purpose of the tour is Maintenance	1.35	11.37
Primary purpose of the tour is Meal	0.49	2.04
Primary purpose of the tour is Social Visit	-0.53	-2.30
Primary purpose of the tour is Other Discretionary	0.99	5.36
Duration of Maintenance2 epoch on the tour	-0.01	-4.27
Duration of Other Discretionary1 epoch on the tour	-0.002	-1.95
Start time of the tour (7am - 9am)	0.25	2.35
Start time of the tour (1pm - 3pm)	-0.26	-2.24
End time of the tour (5pm - 7pm)	0.37	2.89
Goodness of Fit		
Log-likelihood at convergence for the full model	-1749.07	
Degrees of freedom for the full model	12	
Log-likelihood at convergence for the restricted model	-1850.7	
Degrees of freedom for the restricted model	1	
Likelihood ratio	203.26	
$\chi^2_{11,0.001}$	31.26	

Activities with higher durations have lesser probability of occurring on the inbound half tour, which means that individuals might undertake ‘exerting’ activities early on in the tour (the outbound half tour). Tours with most purposes as primary activity are likely (to varying degrees) to have a stop on the inbound half tour. Tours with social visit as primary activity however have are less likely to have a stop on the inbound half tour. This finding complements quite nicely with observations from the MDCEV model, where social visit activity had a greater probability of occurring on tours starting late in the evening (see Table 4.11) and also the fact that social visit activities usually tend to be longer (from the translation parameter values in Table 4.12). Putting these observations together, tours with social visit activities tend to run late into the night thereby making the probability of stop occurrence on the inbound half tours bleak. Likelihood ratio statistic of the model is 203, which is significantly greater than the critical χ^2 value at 99% level of confidence.

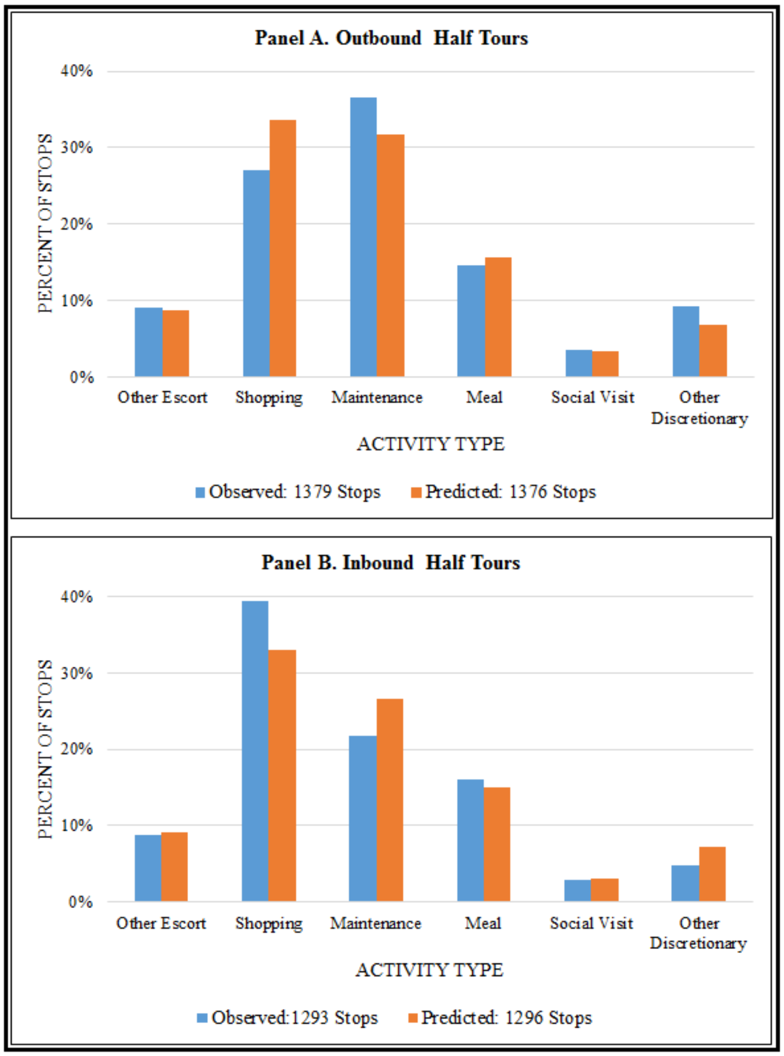

Figure 4.10. HBO tours (non-workers), binary logit model: observed vs. predicted activity type distributions.

Figure 4.10 shows the comparison between observed activity type distributions, vs. the ones predicted by the binary logit model of stop placement. The model accurately predicted 'number of stops' on the outbound and inbound half tours, which is the first test for the effectiveness of the model. But the intent of this model is to simulate which types of activities are more likely to happen on the outbound and the inbound half tours. To check this, activity type distributions from observed data are compared against the distributions predicted by the model. Panel A of the figure shows this comparison for outbound half tours. The uncalibrated version of the model performed quite well in predicting the activity distribution patterns in the outbound half tours. Shopping stops on outbound half tours are slightly over predicted, but the percent difference is not significant. Some calibration effort is warranted to take care of such finer details dictated by patterns observed in the data.

Panel B shows similar comparison for inbound half tours. It can be observed that shopping stops are slightly under predicted in this segment which is a direct manifestation of over prediction of the same activity in the outbound half tours. Again, it is important for the model to capture the activity-travel patterns observed in the data and it is quite difficult to match exactly match the observed distribution without some calibration effort. Majority of the stops made on HBO tours by non-workers are for shopping, maintenance and meal activities. This phenomenon is depicted accurately by the MDCEV model of activity type (from baseline constant values in Table 4.12) as well as from the results of the binary logit model of stop placement. This reinforces the confidence in effectives of the model system to accurately simulate the activity-travel patterns observed in the data.

HBO tours (non-workers) – SATC model for outbound half tours. Sequential activity type choice models are estimated to determine the 'impending activity' on any half tour. If a half tour has only one stop other than the primary activity, it is automatically positioned between the anchor points (home and primary destination). For tours with multiple stops on either the outbound or the inbound half tours, the SATC model simulates the sequence of activity participation. Inputs to the SATC model come from the binary logit model of stop placement which tags each stop on a tour as occurring on either the outbound of the inbound half tour. The STAC model estimates the impending activity as a function of history of activity participation (in the day/on the tour), anticipated activity participation on the tour and tour attributes. Model estimation results of SATC model for HBO tours made by non-workers is presented in Table 4.15.

History of activity participation earlier in the day or anticipated activity participation on the tour had a negative influence on the same type of activity to be the 'next stop' on the outbound half tour. For example, if an individual has already participated in a shopping activity early on in the day (on the a different tour) or plans to participate in a shopping activity later on in the tour, then it is less likely that the next stop on the tour is also a shopping activity. A planned meal activity (or an already finished meal activity) positively influences the next stop on the tour to be shopping activity. This finding is consistent with expectation as shopping and meal activities are often coupled together. Moreover, this finding is complemented by the observation that completion of a shopping activity early on in the day positively influences the occurrence if a meal stop as the next stop on an outbound half tour. Tours ending mid-day (1-3 p.m.) are more likely to have meal activity as the next stop on the outbound half tour.

Table 4.15

HBO Tours (Non-Workers), SATC Model for Outbound Half Tours

Activity Type	Explanatory Variable	Coefficient	t-statistic
Other Escort	Constant	4.96	3.53
	History of other escort activity participation in the day	-5.65	-4.17
	Anticipated other escort activity participation on the tour	-4.87	-3.60
	Anticipated meal activity participation on the tour	2.61	2.74
Shopping	Constant	0.18	0.39
	Anticipated shopping activity participation on the tour	-1.55	-4.94
	History of meal activity participation in the day	0.86	1.46
	Anticipated meal activity participation on the tour	4.44	8.51
	Start time of the tour (1pm - 3pm)	-0.77	-1.76
Maintenance	Constant	2.24	3.62
	History of maintenance activity participation in the day	-4.65	-9.44
	Anticipated maintenance activity participation on the tour	-3.71	-7.78
	Anticipated meal activity participation on the tour	3.99	7.33
	Primary purpose of the tour is shopping	0.54	1.54
Meal	Constant	0.96	1.80
	History of meal activity participation in the day	-4.27	-5.66
	History of shopping activity participation in the day	1.39	2.72
	End time of the tour (1pm - 3pm)	1.29	2.02
	End time of the tour (5pm - 7pm)	-1.04	-1.59
Social Visit	Start time of the tour (5pm - 7pm)	2.35	1.53
Other Discretionary	Constant	2.00	3.61
	History of other discretionary activity participation in the day	-4.90	-6.61
	History of maintenance activity participation in the day	-1.35	-1.99
	Anticipated social visit activity participation on the tour	2.00	2.22
	Primary purpose of the tour is other escort	4.12	3.05
Goodness of Fit			
Sample size (number of stops)		781	
Adjusted ρ^2		0.83	
Likelihood ratio		1900.85	
$\chi^2_{(23,0.001)}$		45.32	

Tours starting later in the evening (5-7 p.m.) are more likely to have a social visit stop as the next stop on the tour, which is consistent with the times during which social visit activities usually occur. In tours involving both maintenance and meal stops, an anticipated meal activity on the tour positively influences the propensity to execute a maintenance stop as the next stop (on the outbound half tour). This is logically intuitive behavior as people may take care of maintenance activities (such as going to the doctor or bank) earlier in the day while participating in meal activities later. The model has a robust likelihood ratio which is much greater than critical χ^2 value at any reasonable level of significance. Replication results from the model are compared against observed activity type distribution and the results are presented in Figure 4.11. As expected, the activity type distribution on outbound half tours is not very different from the activity type distributions observed/predicted on the outbound half tours from the previous (binary logit) model.

The model performs exceedingly well in predicting the observed patterns on the outbound half tours. Though the activity type distributions depicted by these two models are the same, the purpose of these two models is entirely different. While the binary logit model of stop placement determines to which half tour does each stop predicted by the MDCEV model belong to, the SATC model takes all stops on the outbound half tour and sequences them with respect to each other. The functional utility of the SATC model of outbound half tours is to predict the 'impending activity' on the half tour, using information regarding activity participation of the individual earlier in the day or planned activity participation later on in the tour. Thus the true test for the model's efficacy is to see how well it would be able to predict the activity duration distributions at the stop level.

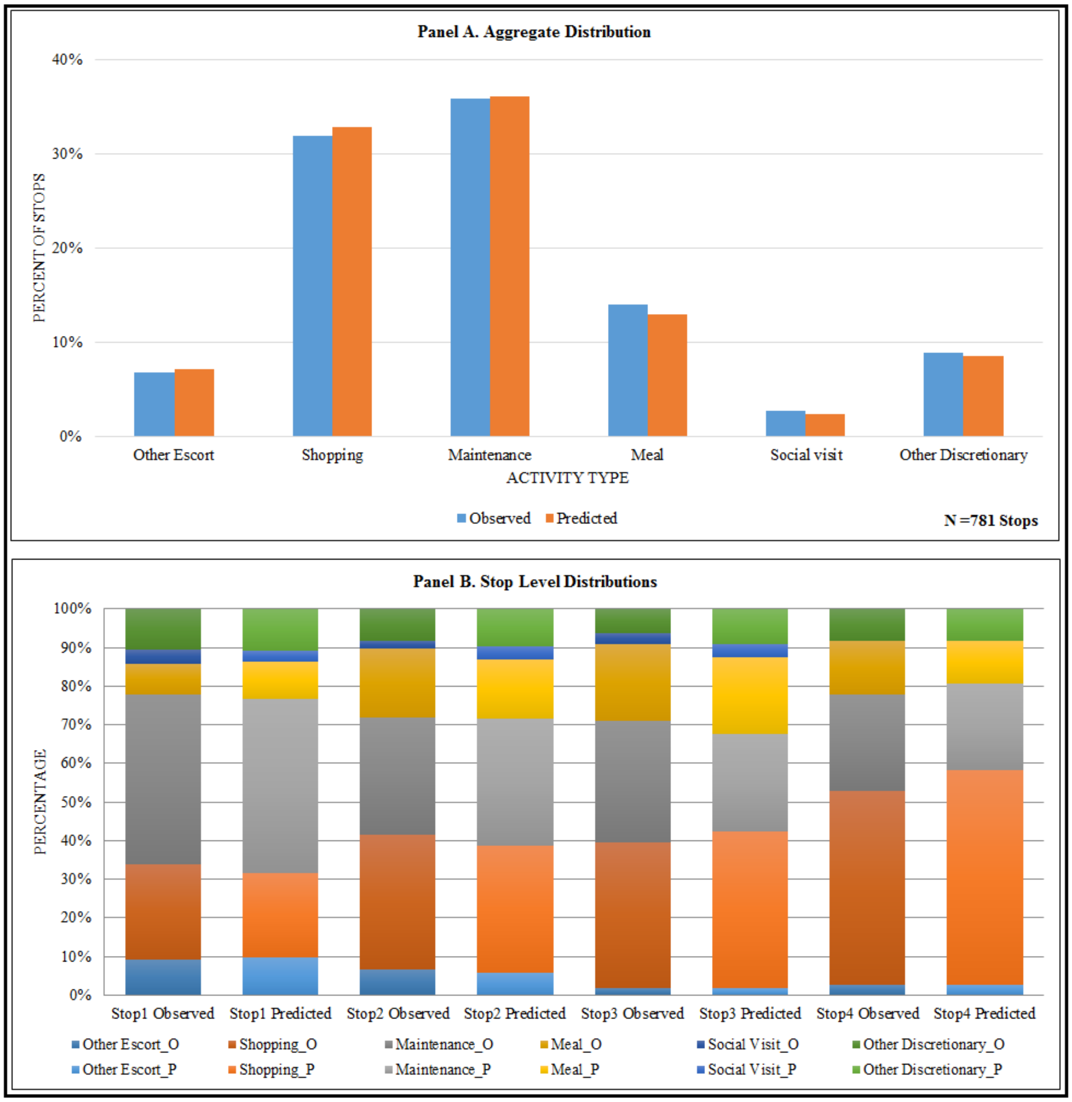

Figure 4.11. HBO tours (non-workers), SATC model for outbound half tours: observed vs. predicted activity type distributions.

Panel B of Figure 4.11 shows this comparison between observed and predicted distributions. Each stacked bar in the chart represents the activity type distribution at that stop level. The observed and predicted distributions are juxtaposed side-by-side for each stop level on the half tour for easier comparison. The pairwise comparison of these stacked bars reveals that, in addition to performing well at the aggregate level, the SATC model is able to replicate the activity type distributions exceedingly well at the disaggregate stop level. Interesting travel patterns are revealed from observing the distributions in this figure. While the share of other escort and maintenance activities is relatively higher in the first two stops of the half tour, shopping activity is predominant on later stops.

Meal activity is safely located mostly on the middle stops (stop no's 2 and 3) on the half tour. These findings depict how people usually organize their activities once they start their journey from home. Individuals usually take care of other escort activities (dropping off/ picking up someone) first, followed by meal activities and keep the shopping activity till the later part of the half tour. Social visit stops are more sporadically distributed amongst various stops (in very meager proportions) on the half tour, while other discretionary stops are evenly distributed across different stops. It is quite heartening to see the model almost mirror the patterns observed in the data, which speaks to the statistical as well as behavioral fidelity of the model.

HBO tours (non-workers) – SATC model for inbound half tours. The structure of SATC models for inbound half tours is same as that of the outbound half tours, except that the choice set for this model is constrained to stops made on the inbound half tours. While the HBW tours had a higher representation of stops in the inbound half tours (76 % of the stops), the HBO tours made by non-workers are more evenly distributed among outbound (55%) and (45%) inbound half tours. This finding identifies the differences in stop making patterns between these segments and corroborates the necessity to model them separately. The explanatory variables in the SATC model for inbound half tours include history as well as anticipatory activity participation dummies, household/person level attributes and tour characteristics. Table 4.16 presents the estimation results for this segment.

Table 4.16

HBO Tours (Non-Workers), SATC Model for Inbound Half Tours

Activity Type	Explanatory Variable	Coefficient	t-statistic
Other Escort	Constant	3.45	4.13
	History of other escort activity participation in the day	-2.46	-3.23
	Anticipated other escort activity participation on the inbound half tour	-3.64	-5.23
	Anticipated meal activity participation on the inbound half tour	3.93	3.01
	Start time of the tour (7am - 9am)	2.17	2.97
Shopping	Constant	0.95	2.55
	Anticipated shopping activity participation on the inbound half tour	-2.35	-0.70
	Anticipated meal activity participation on the inbound half tour	3.58	5.13
	History of meal activity participation in the day	0.71	1.97
	Respondent's age between 18 and 25	-2.88	-2.6
Maintenance	Constant	3.08	5.01
	History of maintenance activity participation in the day	-2.08	-4.32
	Anticipated maintenance activity participation on the inbound half tour	-2.68	-5.62
	End time of the tour (1pm - 3pm)	-1.25	-2.65
	Presence of children in the household	-1.34	-2.70
Meal	Constant	1.05	1.99
	History of meal activity participation in the day	-3.26	-5.57
	History of shopping activity participation in the day	1.07	2.08
	Anticipated social visit activity participation on the inbound half tour	1.92	1.62
Other Discretionary	Constant	2.81	2.45
	History of other discretionary activity participation in the day	-3.44	-2.82
	Anticipated other discretionary activity participation on the inbound half tour	-4.9	-3.52
	Start time of the tour (9am - 11 am)	1.50	1.79
	End time of the tour (3pm - 5pm)	2.83	1.99
Goodness of Fit			
Sample size (number of stops)		649	
Adjusted ρ^2		0.81	
Likelihood ratio		1460.53	
$\chi^2_{(19,0.001)}$		43.82	

The model offers intuitive results and has a robust likelihood ratio that is significantly greater than the critical χ^2value at any level of significance. The findings from this model are in line with results of similar models estimated for other segments. It was found that history as well as anticipatory activity participation of particular negatively influences the probability of a similar activity being the ‘next’ stop on an inbound half tour. This finding is behaviorally consistent as individuals usually bunch activities together and hence presence of multiple stops of same activity is not frequently observed in a person’s daily schedule. History/anticipatory meal activity participation increases the propensity of occurrence of next stop to be a shopping stop.

HBO tours starting earlier in the day tended to have a greater probability of having an other escort stop as the immediate stop on the tour. Anticipated meal activity participation had a positive influence on the impending stop being an other escort stop. This translates to logical ordering of stops in the real world where individuals pick up a family member (kids/spouse) before proceeding to a meal activity. Tours starting just after usual work start times (9am in the morning) and ending just before the usual work day’s end (before 5 pm in the evening) have a greater probability of having an other discretionary stop as next stop on the inbound half tour. As the name of the activity aptly suggests, people engaging in other discretionary activities have greater flexibility in their schedule and would organize their activities in a way that would avoid the morning or evening traffic rush. The model is able to identify and depict this interesting nuance in travel behavior of non-workers.

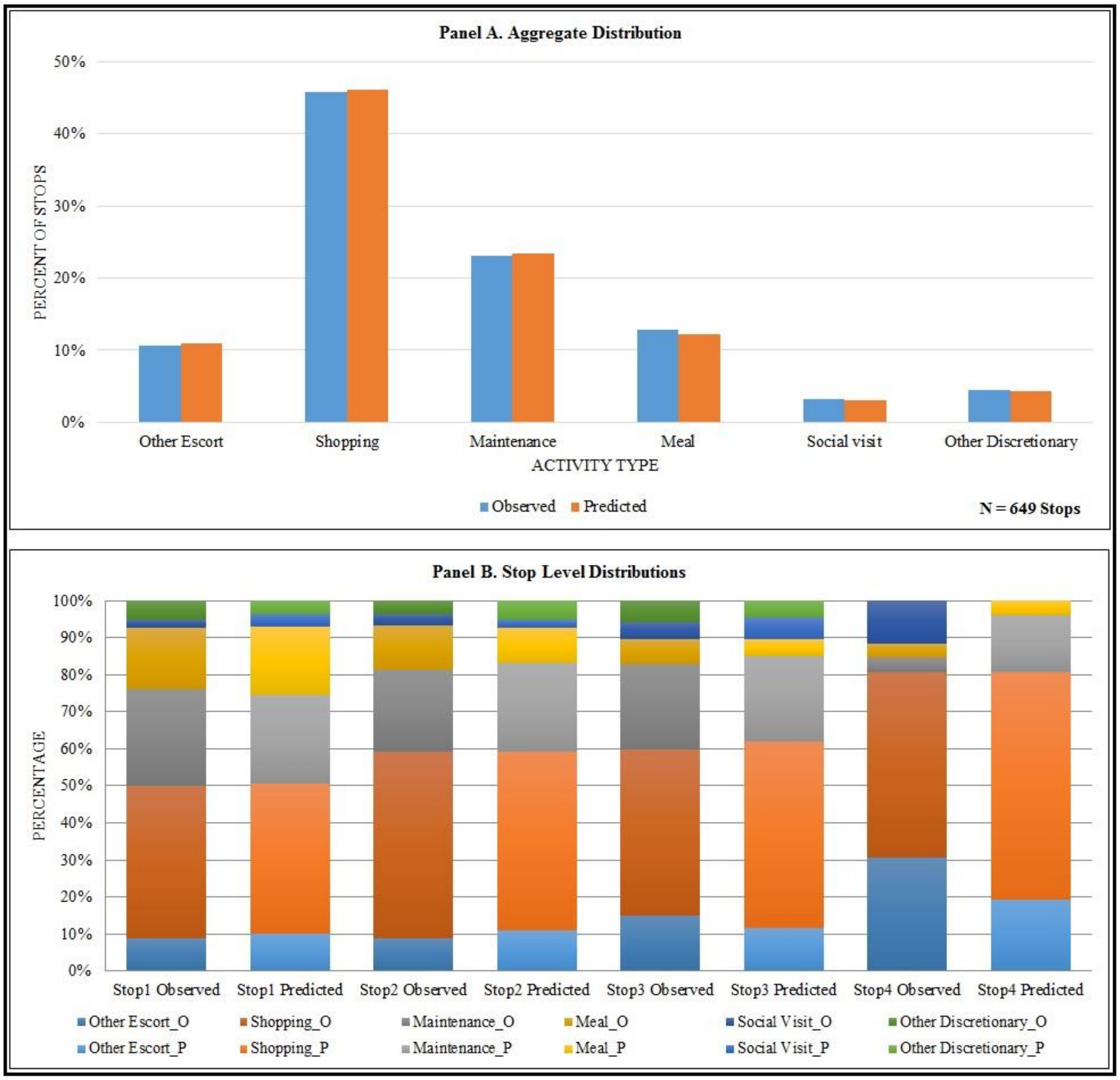

Figure 4.12. HBO tours (non-workers), SATC model for inbound half tours: observed vs. predicted activity type distributions.

Comparison of observed and predicted patterns from the model replication result are shown in Figure 4.12. Shopping activity dominates the stops made on inbound half tours, and this observation complements well with a similar finding from the outbound half tours where individuals are found in general to postpone the shopping activity until later part of the half tour which might spill over onto to the return home journey (read inbound half tours). Maintenance and meal activities have a fair share of stops on inbound half tours, which is a similar finding as the one found on outbound half tours. Overall, stops on HBO tours made by non-workers are majorly for shopping, maintenance and meal activities, a finding that is in line with the results of the MDCEV model of activity type mix that consider all stops made on the entire tour. Panel A of Figure 4.12 presents the results of observed and predicted activity type distributions. Without any calibration, the model performs astoundingly well in replicating the patterns observed in the estimation dataset.

Panel B of the figure presents similar distributions at the stop level. From visual inspection, it can be seen that the activity distributions patterns on inbound half tours are different from the patterns observed on outbound half tours (see Figure 4.11), necessitating separate SATC models for outbound and inbound half tours. While shopping stops tended to be more in proportion toward the end of outbound half tours, they have a more even presence across different stops on the inbound half tours. In the overall tour composition, this would mean that individuals are taking care of other activities before engaging in a shopping activity. Other escort stops gradually decreased with increasing stop number on the outbound half tours and the exact opposite pattern is seen on the inbound half tours, where the proportion of other escort stops increases with the stop number on the tours. This directly correlates to the two types of escort activities, where individuals are dropping of a household member at the beginning of the tour and picking them up toward the end of the tour. The presence of social visit stops is higher toward the later portion of the inbound half tour, an observation consistent with expectation. Other discretionary stops are evenly distributed across different stops on the inbound half tour similar to outbound half tours. All these observations are behaviorally consistent with the activity-travel patterns observed in the real world. The uncalibrated SATC model for inbound half tours is able to replicate these patterns quite well. Some calibration is warranted for exactly matching the observed patterns.

Model components of the tour characterization framework estimated for HBO tours made by non-workers are significantly different from those estimated for HBW tours made by workers corroborating the necessity of a separate model system for this segment. Similar model components are estimated for HBO tours made by workers.

Summary and Conclusions

This chapter presented the estimation results for components of the tour characterization framework for HBW tours (made by workers) and HBO tours (made by non-workers). All of the components developed show great promise in depicting the stop making patterns observed in the data. The MDCEV model of activity type mix takes tour budget as the input and simulates all secondary stops made on a tour. The binary logit model determines the half tour (outbound/inbound) to which each stop simulated by the MDCEV model belongs to. The SATC models developed for each half tour, take half tours that have more than one secondary stop and order the stops by predicting the next stop on the half tour under consideration. The unit of analysis considered for the modeling effort is the summation of travel time to an activity and duration of engagement in the activity, termed as an ‘epoch’. The reason for choosing an epoch as the unit of analysis is to develop a framework capable of modeling tours in a continuous time domain. The motivation behind this effort is to enhance the discrete time representation adopted in tour based models currently in practice with an evolutionary continuous-time approach that is capable of leveraging both history of activity participation as well as anticipatory activity engagement details and determine the sequence of stops undertaken on a tour.

The MDCEV model of activity type mix utilizes a variety of household, person and tour level attributes (mode, tour accompaniment etc.) to simulate the array of secondary of activities performed on a tour. For HBW tours, the MDCEV model simulates all secondary activities except home and work (primary purpose) activities, whereas for HBO tours it simulates all activities including the primary purpose of the tour. Estimated models are applied to replicate the observed activity-travel patterns. Comparisons are made between observed and predicted i) Activity frequency distributions and ii) Average epoch durations. The estimated models are found to offer intuitive findings and perform quite well (without calibration) in replicating the observed activity frequency distributions. The average epoch durations predicted by the model seemed to be higher than the observed durations, but on further analysis it was found that this a manifestation of slight under prediction of activity frequencies as well as the small sample sizes of secondary stops made on the tours. Slight calibration of the model parameters should be able to accommodate these differences without compromising the behavioral integrity of the model.

Binary logit model of stop placement is developed to determine the location of a stop on a tour relative to the primary purpose (inbound/outbound half tours). The models presented for HBW tours and HBO tours have simple yet elegant specifications. The estimated models are able to accurately predict the proportion of stops on outbound and inbound half tours. In addition to this, the models are also able to represent the activity type distributions on outbound and inbound half tours quite effectively with no necessity for calibration. Sequential activity type choice (SATC) models of stop placement are developed to determine the placement of stops relative to each other on multi-stop tours. Separate models are developed for outbound and inbound half tours as the stop making patterns on these different halves of the journey were observed to be significantly different. SATC models take information regarding history of activity participation (on the tour/in the day) as well as anticipatory activity engagement decisions (on the tour) and determine the 'next stop' on the tour. Suppose there are 3 stops on a half tour, the SATC model is applied twice to determine the first and second stops on a tour, automatically positioning the third stop. The SATC models estimated for different tour segments (HBW/HBO) and different half tours, were successful in replicating the activity type distributions at the aggregate as well as stop level.

The tour characterization framework developed as a part of this effort is tested for potential application in the Maricopa Association of Governments CT-RAMP activity-based travel demand model system. With very minor calibration efforts, the proposed model system could improve the activity agenda-based approach (with discrete time representation) in the current tour-based models in practice with the evolutionary continuous-time activity type choice modeling process embedded in scheduling-oriented models.

CHAPTER 5

VEHICLE FLEET COMPOSITION MODEL SYSTEM

This chapter presents the estimation and validation results of an open source vehicle fleet composition simulator that can be integrated into any existing activity-based microsimulation model systems. First, description of the data used for model estimation is provided. This is followed by model estimation results coupled with a sample replication result from a sequential application process of the model system. The process is continuous, in the sense that output of each component serves as input to the subsequent component in the model system. The logic followed by the fleet composition model system is discussed in detail in Chapter 3.

Data

Data used for estimating various components of the vehicle fleet composition model system is from the latest wave of National Household Travel Survey (NHTS) conducted in the year 2008-2009. NHTS collects data regarding socio-economic, demographic, vehicle ownership and personal travel characteristics of a random sample of households across the nation. Data collected from the survey is organized into four different files namely

- *Household File*: Contains information regarding the household level socio-demographic characteristics such as household size, income, vehicle ownership, presence of children etc.

- *Person File*: Contains information regarding person level characteristics such as age, gender, worker status, driver status, etc. Each respondent from the household has a separate entry in this file. All the respondents in a household are grouped by the same household id.
- *Trip File*: This file has information regarding all trips made by a person in the day. Each trip made by the person gets a separate line entry grouped by the same person id. Trip level characteristics such a trip purpose, length, duration etc. are stored in this file.
- *Vehicle File*: This file has information regarding each vehicle owned by a household. Information regarding year, make, model etc. are collected.

In addition to data collected for each state at the national level, metropolitan planning organizations (MPOs) have the opportunity to purchase add-on samples for model development purposes. Maricopa Association of Governments (MAG), the MPO of Greater Phoenix Metropolitan Region purchased an add-on sample for development of travel demand models for the region. This research effort uses the MAG add-on sample from 2008-09 NHTS for model estimation purposes. The vehicle fleet composition model system operates at the household level. Hence, the household and vehicle files are predominantly used for estimating components of the model system. A brief sketch of the household level socio-demographics of the data set is provided in Table 5.1.

From the table, it can be observed that the average number of vehicles owned by a household is about the same as average number of drivers in a households. This tells us that the data set under consideration is quite mobile and that households indeed own multiple vehicles. The intent of this research effort is to explicitly identify the body type, age and annual mileage consumption of each of the vehicles owned by a household in the dataset. Majority of households in the dataset reside in urban areas and in single family dwelling units. The income distribution of the dataset is uniform, with slightly higher representation of medium income households. This lines up with the income profile of the data collected for the entire nation (National Household Travel Survey, 2009).

Table 5.1

Data Description: Household Level

Characteristic	**Mean**	**Standard Deviation**
Number of vehicles in the household	1.95	1.054
Number of persons in the household	2.43	1.333
Number of adults in the household	1.90	0.708
Number of children in the household	0.53	1.016
Number of workers in the household	0.97	0.889
Number of drivers in the household	1.83	0.771
Population density (sq miles)	4401.23	2557.65
Employment density (sq miles)	1164.92	1548.96
% of Households residing in urban area	83.80%	0.369
% Single family housing units	95.80%	0.201
% Households with income < $25,000	17.80%	0.383
% Households with income ≥ $25,000 & < $50,000	28.40%	0.451
% Households with income ≥ $50,000 & < $75,000	18.80%	0.391
% Households with income ≥ $75,000 & < $100,000	15.00%	0.357
% Households with income > $100,000	20.00%	0.4
Sample Size, N	***4,262 Households***	

For the purposes of this research effort, vehicles from the NHTS data are categorized by a cross classification between four body types (car, van, sports utility vehicle (SUV), pick-up truck) and three vintages (0-5 years old, 6-11 years old, ≥ 12 years). A motorbike category is also considered (with no vintage classification) bringing the total number of motorized alternatives to thirteen. In addition to the motorized alternatives, a non-motorized vehicle alternative was considered to capture the walk/bike travel undertaken by each household in the dataset. This will allow for modeling the total annual mileage consumption for a household irrespective of the type of mode used for travel. While the NHTS data has information regarding the estimated annual mileage of each vehicle owned by a household, non-motorized mileage is not readily available. Annual non-motorized mileage is computed from the walk and bike trips reported by all individuals in a household.

The non-motorized alternative is the one that is consumed by every household in the dataset and is considered as an outside good. An outside good is an alternative that is chosen by every choice maker in the dataset in the econometric modeling perspective. In the current context, every household invariably undertakes some amount of non-motorized travel such as walking from the parking lot, walking to the bus station or jogging etc. and hence this alternative is considered as an outside good. To compute the annual non-motorized mileage of a household, (weighted) walk/bike trips reported by all of the household members are aggregated. If none of the household members reported walk/bike trips, annual non-motorized mileage for that household is computed as '0.5 (miles/person/day) x 365 (days/year) x number of persons in the household'. Previous studies have successfully incorporated this formulation in a similar context (Vyas et al., 2012). In total the model system consists of 14 alternatives (4 body types x 3 vintage categories + motorbike + non-motorized mileage).

An alternate vintage classification was also tested (0-3 years old, 4-9 years old, ≥ 10 years) and it was found that the fleet composition model system is robust to the vintage classification considered. The vintage classification (0-5 years old, 6-11 years old, ≥ 12 years) was finalized based on the observation that most car manufactures offer a five year power train warranty (My Car Stats, 2010). Also, this classification provided a healthy sample size for all the 13 motorized alternatives considered for model estimation. The vintage classification can be further disaggregated to include an alternative for each year for a vehicle body type (bringing the total number of vehicle alternatives to 50), but this level disaggregation would make the dataset sparse for model estimation and also increase the computation burden in model application process. Table 5.2 provides a description of vehicle fleet characteristics of the NHTS dataset considered for this research effort.

Table 5.2

Data Description: Vehicle Level

Panel A. Vehicle Body Type					
	Car	**Van**	**SUV**	**Pick-up**	**Motor Bike**
Average Age	8.55	7.46	6.52	9.52	9.21
Average Mileage	10204.4	11317.7	11296.6	10723.0	3838.9
Number of Vehicles	3,997	635	1,537	1,376	240
Panel B. Vehicle Body Type vs. Annual mileage					
Annual Mileage					
0 - 4,999	27.5%	18.4%	21.1%	24.9%	71.3%
5,000 - 9,999	30.6%	31.3%	28.6%	29.4%	15.8%
10,000 - 14,999	21.4%	26.9%	26.2%	22.8%	7.9%
15,000 - 19,999	11.3%	13.9%	12.8%	12.5%	2.9%
≥ 20,000	9.1%	9.4%	11.3%	10.5%	2.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

From Panel A of the table, it can be observed that households prefer relatively newer SUVs and older pick-up trucks consistent with expectation. This finding is corroborated by the body type and age distribution shown in Figure 5.1, where it can be observed that more than half of the SUVs are in the 'newer' vehicle category, whereas pick-ups have relatively lower representation in this category.

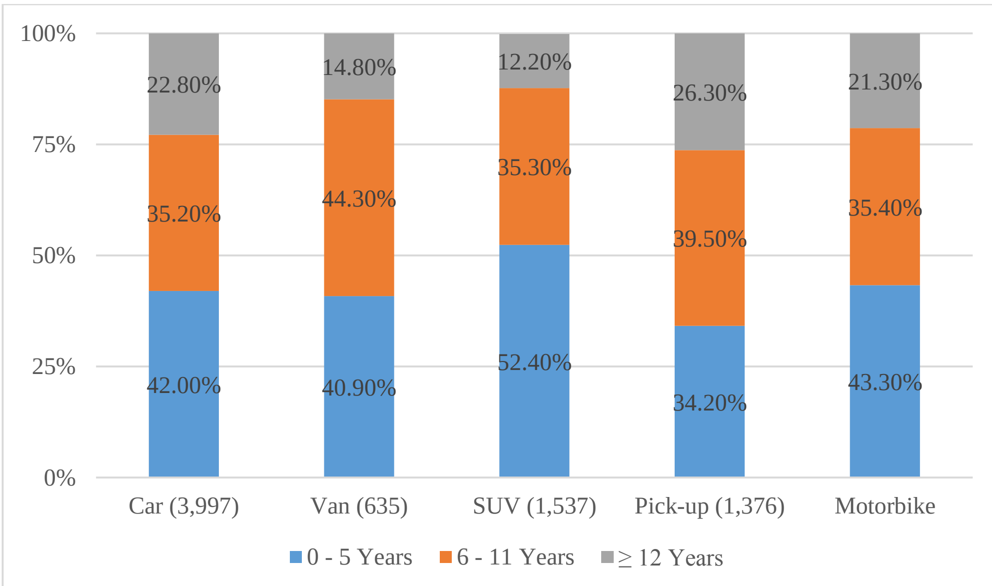

Figure 5.1. Vehicle body type and age distribution.

Table 5.3 provides the distribution of vehicle body types for each household income category. It can be observed that, while lowest (< $25,000) and low income ($25,000 - $49,999) households tend to own more cars, medium and high income household tend to own a mix of vehicles. It can also be observed that with increasing household income, the ownership of SUVs gradually increases. One possible reason for this might that high income households can afford sports utility vehicles more. Another reason could be that while low income households usually own fewer vehicles and utilize them for all travel needs, households with high income might own a mix of vehicles and use them to varying degrees for specific purposes (refers to a combination of affordability and variety seeking nature of the segment). The total number of vehicles owned by individuals from each income category is shown in parenthesis in the first column of the table.

Table 5.3

Vehicle Body Type Distribution by Household Income

	Vehicle Body Type					
Household Income	**Car (%)**	**Van (%)**	**SUV (%)**	**Pick-up (%)**	**Motor Bike (%)**	**Total (%)**
< $25,000 (852)	61.30	9.50	11.60	16.70	0.90	100
$25,000 - $49,999 (1,898)	54.90	9.00	14.60	19.00	2.50	100
$50,000 - $74,999 (1,547)	48.80	7.60	21.40	18.30	3.90	100
$75,000 - $99,999 (1,418)	48.00	8.50	20.10	19.50	3.90	100
≥ $100,000 (2,070)	48.20	7.10	26.30	15.20	3.30	100

Model Estimation and Application Results

This section provides the estimation results of all the components in the vehicle fleet composition model system coupled with comparisons between observed and predicted patterns from a sample replication exercise. It should be identified that the process adopted here does not constitute a true validation exercise. In the traditional validation process, the data would be split (say in the proportion of 80:20) and the larger sample is used for model estimation. The estimated models are applied on the holdout sample to test the predictive capability of the model. In the current context, the number of different components included in the model system and level of disaggregation of vehicle alternatives warranted the use of entire survey sample (4,262 households/7,785 vehicles) for model estimation. The estimated models are applied to the entire survey sample to compare predicted patterns against the observed patterns in the data. In order to ensure the efficacy of the model system, the model was tested on specific market segments (different income categories, urban/rural residents) and the predictive performance of different model components was tested against observed data. A detailed sensitivity analysis exercise was carried out to predict the effect of changes in land use dynamics on vehicle fleet composition patterns.

Each of the model components was estimated and validated separately to ensure the predictive capability of the models in replicating observed vehicle ownership patterns. The model system is then applied in its entirety to the estimation dataset to test its efficacy. Estimation results as well as results of the sequential model application process are provided here.

Motorized mileage prediction model. The first element in the vehicle fleet composition model system is the household mileage prediction model that predicts the annual motorized mileage consumption of households. The motorized mileage is estimated using a power transformed linear regression model. Once the motorized mileage for each household is predicted, non-motorized mileage is computed using a preset formula (0.5 x household size x 365) as every household will inevitably have at least some amount of non-zero mileage consumption. The combined annual mileage is provided as input to the MDCEV model, which will then predict the fleet mix owned by the household and allocate the mileage budget to all the vehicles owned by the household.

To fit the observed annual motorized mileage distribution, several model structures were explored and the power transformed linear regression model fit the data best. In practice, the activity-based model to which the vehicle fleet composition simulator is integrated will provide the annual mileage budget as an input. Since this is a standalone model application process, a separate mileage prediction model is estimated. Use of a power transformed linear regression model avoids the possibility of negative mileage predictions that a regular linear regression model may provide. The model structure of mileage prediction model is shown in equation 5.1.

$$Motorized\ Mileage^{0.3} = \beta_0 + \beta_i' x_i + \varepsilon \quad (5.1)$$

Where, $\beta_0$is a constant, β_i is the array of coefficients to be estimated and $x_i$is the array of socio-demographic characteristics included in the model. The error term 'ε' is normally distributed with mean zero and standard deviation of the dependent variable. Estimation results of the mileage prediction model are presented in Table 5.4. Various socio-demographic characteristics, lifestyle variables and TAZ characteristics of the household's residential location were used to estimate the motorized mileage consumption patterns. Household income was observed to be a significant variable in explaining the annual motorized mileage consumption. Households in the lowest income category are likely to have low motorized mileage consumptions, while households in the highest income category are likely to have higher motorized mileage consumptions. This finding is directly related to the number of vehicles owned by respective income categories (presented in Table 5.3) where it was seen that lowest income category households own approximately 10% (852 vehicles) of the vehicles in the dataset, while highest income households own about 26% (2,070) of the vehicles. This relates to the proportional higher mileage consumption of highest income households.

Households with more number of drivers were observed to consume higher mileages, an observation consistent with expectation. Similarly, household with more number of children had higher mileage consumptions. Possible reason for this might be due to chauffeuring associated with children's activities in such households. Retired households with no children tended to have lesser mileage consumption. This observation is behaviorally intuitive as such household might not engage in a lot of activity. Households residing in TAZs that had higher proportion of affluent households tended to have higher motorized mileage consumptions. This finding couples nicely with the higher mileage consumptions for the highest income (≥ $100,000) category. Households residing TAZs that have a lot of employment accessibility within 10 minutes of auto travel have lesser motorized mileage consumptions. These TAZs probably refer to locations in the urban core, with mixed-use development where discretionary travel can be easily undertaken by non-motorized modes (walk/bike). Residential self-selection also has a possible role to play in this finding as people residing the urban core might be more environment friendly and are willing to opt out of motorized modes of travel.

Table 5.4

Motorized Mileage Prediction Model: Estimation Results

Explanatory Variable	Coefficient	t-statistic
Constant	13.03	35.00
Number of drivers in household	2.01	12.80
Count of adult household members at least 18 years old	0.37	2.28
Household resides in rural area	0.87	5.30
Lowest income household (< $25,000)	-1.05	-6.03
Highest income household (≥ $100,000)	1.33	7.13
Number of children in the household	0.52	5.04
Zero worker household	-1.51	-8.78
Two worker household	0.83	5.52
Household size = 4 or more	-0.67	-2.48
Single family housing unit (owned)	0.57	3.14
Retired household (one/two person) with no children	-0.88	-5.15
Proportion of households in the highest income quintile	1.36	2.95
Proportion of single family housing units in the TAZ	0.69	1.99
TAZ with high regional employment accessible within 10 minutes by auto (1st Quartile)	-0.29	-2.20
R^2	0.404	

The model is applied on the estimation dataset to see how well it could replicate the observed mileage consumption patterns. Results of this comparison are presented in

Figure 5.2. The model is able to replicate the observed patterns quite well. Results shown are for the calibrated model where the constant in the regression equation was slightly adjusted to better match the observed patterns. Each bar in the chart represents the percentage of households in the data set that pertain to the mileage bin under consideration.

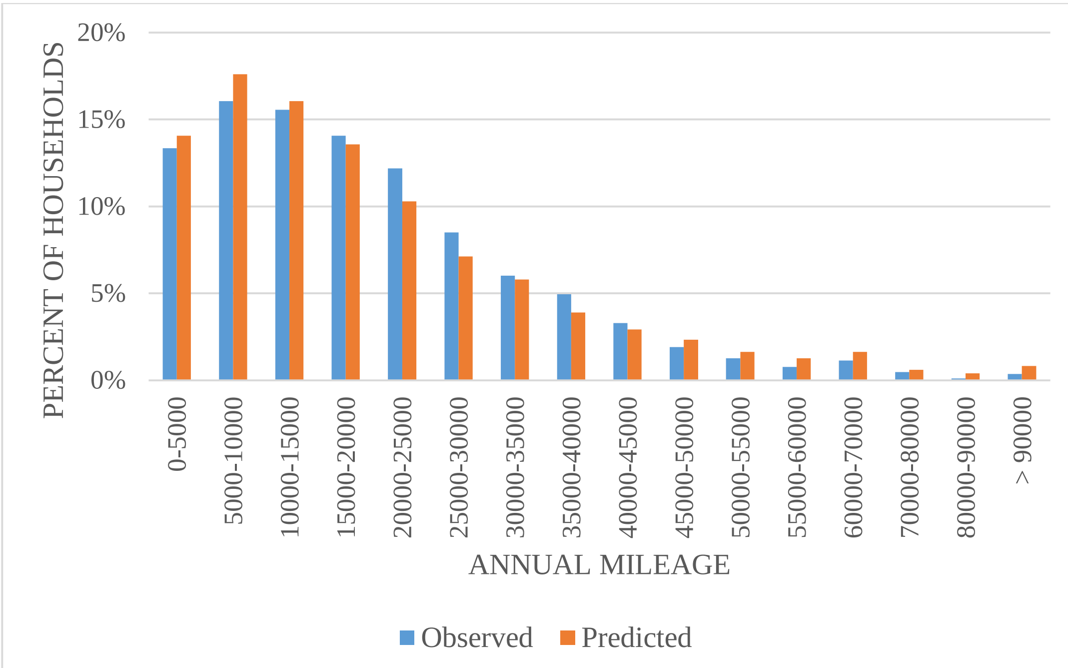

Figure 5.2. Observed vs. predicted mileage distributions.

MDCEV model of vehicle fleet mix. The next component in the model is system is the MDCEV model of vehicle fleet mix which takes the mileage predicted by the previous component as input, predicts the vehicle fleet mix owned by the household and allocates the mileage budget to different vehicles owned by the household. The fleet mix model system is an MDCEV model which is capable of simultaneously predicting the array of vehicles owned by a household. The MDCEV model is ideally suited to model the vehicle fleet mix and utilization patterns due to the multiple discrete (ownership of multiple vehicles) and continuous (mileage allocation to vehicles owned) nature of the problem. The MDCEV model was proposed by Bhat (2005; 2008) to efficiently model multiple discrete choice behavior, addressing the short shortcomings of single discrete choice models. A number of recent studies used the MDCEV model to estimate vehicle fleet mix at the household level (Bhat and Sen, 2006; Eluru et al., 2010; Vyas et al., 2012). Notable features of the MDCEV model include consideration of diminishing marginal utility with increasing consumption of an alternative and its capability to collapse to the standard MNL model structure, given every behavioral unit in the dataset chooses only one out of ‘k’ available alternatives.

As discussed earlier, the vehicle classification for the fleet mix model system consists of a total of 14 alternatives (4 vehicle body types x 3 vintage categories + motorbike + non-motorized alternative). In order to account for household with no vehicles at all, the MDCEV model specification with presence of an outside good is adopted in the current empirical context. An outside good is an alternative that is chosen by every household in the data set, which in this case would be the non-motorized alternative. After the mileage prediction model predicts the motorized mileage consumption of the household, non-motorized mileage is computed using a preset formula (0.5 x household size x 365) and added to the motorized mileage to determine the ‘total’ mileage consumption of the household. The MDCEV model takes the total mileage consumption of the household as input and distributes it to different vehicles owned (as predicted by the model) by the household. The formulation of the MDCEV model allows for selection of ‘m’ alternatives out of ‘k’ available alternatives, while definitely choosing the outside good (non-motorized alternative) for each and every household in the dataset. The functional form of the utility expression of the MDCEV model proposed by Bhat (2008) for a case with the presence of an outside good is:

$$U(x) = \frac{1}{\alpha_{out}} \psi_{out} x_{out}^{\alpha_{out}} + \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \tag{5.2}$$

Where baseline marginal utility for the outside good, $\psi_{out} = \exp(\varepsilon_{out})$ and baseline marginal utility for the rest of the alternatives $\psi_k = \exp(\beta' z_k + \varepsilon_k)$, is a function of various parameters that capture the observed and unobserved attributes of the alternative. z_k is a set of attributes that define and alternative 'k' and ϵ_k captures the effect of unobserved attributes. $U(x)$ is a quasi-concave and continuously differentiable function with respect to consumption quantity vector x ($x_k \geq 0 \; \forall \; k$). ψ_k represents the baseline marginal utility or the marginal utility at the point of zero consumption. α_k is the satiation parameter which governs the decrease in marginal utility with increasing consumption for good k.. The translation parameter γ_k not only governs the level of satiation but also enables corner solutions (i.e., zero consumption of some goods).

As both γ_k and α_k are parameters that incorporate the effects of satiation, it is difficult to uniquely identify and distinguish between them. For this reason, one of the two parameters is fixed and the other parameter is free to be estimated in most empirical model estimation efforts and the best model is chosen. In the current modeling context γ-profile gave the best fit to the data. A household maximizes its utility by optimally allocating consumptions to the k available goods (vehicles), while always choosing the outside good (non-motorized alternative). Thus the constraint for the utility maximization problem is:

$$\sum_{k=1}^{K} t_k = T \tag{5.3}$$

The MDCEV model is estimated with 14 alternatives and the estimation results of the MDCEV model system are presented in Tables 5.5-5.7. Table 5.5 presents the significant parameters in the baseline marginal utility equation of the MDCEV model. From the model estimation result, it was found that high income households are more likely to own newer vehicles and also tend to prefer cars and SUVs over other types of vehicles. Households with children tend to own vans more than cars. It can be observed that number of children in the household has a negative impact on owing cars, meaning such households would rather prefer a vehicle that would help them attend to the child's necessities (such as a van).

Households with more number of workers tend to prefer newer cars and SUV vehicles, which is intuitive as the number of workers in a household could probably act as a proxy characteristic for affluence of the household. While the impetus for owing vans is explained best by the presence/number of children in the household, the ownership patterns for car, van and SUV body types are explained very well by the income categorization. Within the car body type, high income households prefer to own newer cars (0-5 years), low income households tend to own medium aged cars (6-11 years) while the lowest income households are more likely to own older cars (≥ 12 years). Similar patterns are observed in the van and SUV categories as well. Thus the model is able to represent the vehicle ownership patterns of different income categories, where high income households usually change their vehicle fleet more often but less affluent households do not have such flexibility. Lager households prefer to own vans as they offer the flexibility of accommodating greater number people, which facilitates joint travel in such households. Households living in rural areas are more likely to own pick-up trucks, another finding consistent with expectation.

Table 5.5

MDCEV Model Estimation Results:

Significant Parameters in Baseline Utility: Cars and Vans

Vehicle Type	Explanatory Variable	Coefficient	t-statistic
Car 0-5 years old	High income household ($75,000 - $99,999)	0.16	2.16
	Number of children in the household	-0.19	-5.91
	Three or more worker household	0.17	1.38
	Proportion of households in the lowest income quintile	-1.01	-3.61
	Percent of regional employment within 10 minutes of auto accessibility from the TAZ	-13.83	-3.24
Car 6-11 years old	Two worker household	-0.16	-2.21
	Low income household ($25,000 - $49,999)	0.13	1.9
Car 12 years or older	Lowest income household (< $25,000)	0.57	5.84
	Household has one/two retired adults and no children	0.2	2.59
	Proportion of households in the lowest income quintile	0.57	1.74
Van 0-5 years old	Number of children in the household	0.38	8.23
	Two worker household	-0.39	-2.53
	TAZ with high density (1st Quartile)	-0.31	-1.89
	Percent of regional employment within 30 minutes of auto accessibility from the TAZ	-1.72	-1.78
Van 6-11 years old	Number of children in the household	0.33	7.15
	TAZ with high density (1st Quartile)	-0.26	-1.87
	Low income household ($25,000 - $49,999)	0.27	1.87
Van 12 years or older	Count of HH members	0.14	1.98
	TAZ with high density (1st Quartile)	0.68	3.1
	Lowest income household (< $25,000)	0.66	2.57

Table 5.5 (Continued)

MDCEV Model Estimation Results:

Significant Parameters in Baseline Utility: SUVs, Pick-ups and Motorbikes

Vehicle Type	Explanatory Variable	Coefficient	t-statistic
SUV 0-5 years old	Lowest income household (< $25,000)	-1.01	-5.59
	Two worker household	0.17	1.92
	Household has one/two retired adults and no children	-0.17	-1.78
	Proportion of households in the lowest income quintile	-1.48	-3.85
	Percent of regional employment within 10 minutes of auto accessibility from the TAZ	-17.83	-2.82
	TAZ with low density (3rd Quartile)	0.27	2.48
SUV 6-11 years old	Medium income household ($50,000 - $74,999)	0.26	2.41
	Household size = 4 or more	0.33	3.3
	Single family housing unit (owned)	0.86	4.37
	TAZ with high regional employment accessible within 30 minutes by auto (1st Quartile)	-0.30	-2.85
SUV 12 years or older	High income household ($75,000 - $99,999)	0.36	1.89
	Presence of children in the household	0.31	1.99
	Household in a single family housing unit	-0.74	-2.38
	TAZ with medium density (2nd Quartile)	-0.36	-2.28
Pick-up 0-5 years old	Highest income household (≥ $100,000)	0.24	2.28
	Household size = 1	-0.98	-4.64
	Household resides in rural area (from variable URBRUR)	0.24	1.95
	Proportion of single family housing units in the TAZ	0.74	2.44
Pick-up 6-11 years old	Household resides in rural area	0.15	1.3
	Household has one/two retired adults and no children	-0.35	-3.41
	High income household ($75,000 - $99,999)	0.16	1.44
	TAZ with high regional employment accessible within 10 minutes by auto (1st Quartile)	-0.23	-2.17
Pick-up 12 years or older	Proportion of households in the lowest income quintile	1.28	2.95
	Low income household ($25,000 - $49,999)	0.35	2.86
	Presence of children in the household	-0.25	-1.95
	TAZ with high regional employment accessible within 10 minutes by auto (1st Quartile)	-0.34	-2.6
Motorbike	Household resides in rural area	0.71	4.41
	Single family housing unit (owned)	0.75	2.36
	Household size = 1	-0.57	-2.11

Among TAZ characteristics, it was found that households living in TAZs with high proportion of households in the lowest income quintile are not likely to own newer cars. This finding is corroborated by another observation from the model that households in such TAZs are more likely to own older cars. Households in TAZs with high density were less likely to own vans. It is possible that spatial and social dependency effects play a role in vehicle ownership and this finding is consistent with such a notion (Paleti et al., 2013). Households in TAZs with high density are less likely to own newer vehicles. Households in such TAZs might have accessibility to alternative modes of transportation and also use walk/bike to satisfy their mobility needs, which might in turn prompt them to just keep their older vehicles in the fleet mix. This finding is nicely coupled by the observation that households in lower density TAZs tend to own newer vehicles and are more likely to have larger vehicles in the fleet mix such as SUVs.

Table 5.6 presents the model estimation results of baseline constants and translation parameters in the MDCEV model. A baseline constant provides an indication of the inherent preferences for various alternatives and the marginal utility at zero consumption. The values of the baseline constant reveals the preference for a particular type of vehicle for an 'average' user in the dataset. For cars and vans, it was found that newer vehicles have a greater baseline utility than older ones, suggesting that households would rather own newer cars, all other things being equal. In general baseline utility decreases with age of the vehicle, although this trend is not seen consistently for SUVs and pick-up trucks. For SUVs , there is lower baseline utility for middle aged SUVs suggesting that households tend to acquire newer SUVs and hold on to their SUV for a long time, which is fairly expected behavior. For pick-up trucks, there is a lower baseline utility for newer pick-up trucks, suggesting that households hold on their middle aged and older pick-up trucks more and do not see a necessity to own the ‘newest’ pick-up truck. This finding is consistent with general behavior, where pick-up trucks have a very slow turnover rate. All of the findings from model estimation results line up well with actual vehicle ownership patterns observed in the dataset.

Table 5.6

MDCEV Model Estimation Results: Baseline Constants and Translation Parameters

Vehicle Type	**Baseline Constants**		**Translation Parameters**	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
Non-motorized vehicle (Outside Good)	NA	NA	0	NA
Car 0-5 years old	-5.98	-83.27	23668	10.07
Car 6-11 years old	-6.51	-140.27	18621	10.37
Car 12 years or older	-7.29	-93.28	12164	9.46
Van 0-5 years old	-8.13	-61.11	29431	3.41
Van 6-11 years old	-8.43	-82.8	22248	4.21
Van 12 years or older	-10.04	-41.51	12691	3.22
SUV 0-5 years old	-6.65	-64	25172	6.7
SUV 6-11 years old	-8.32	-42.25	16717	6.71
SUV 12 years or older	-7.88	-25.94	8397	5.1
Pick-up 0-5 years old	-8.29	-30.87	20610	5.69
Pick-up 6-11 years old	-7.35	-95.28	14758	6.92
Pick-up 12 years or older	-8.06	-70.37	9542	6.7
Motorbike	-9.24	-29.1	2223	7.67

Translation parameters in the MDCEV model represent the diminishing marginal returns with increasing consumption of an alternative. A higher value for the translation parameter pertaining to a specific vehicle means that households are less satiated with the use of that vehicle and are likely to drive that vehicle alternative more. For all of the vehicle body types, the translation parameters show a consistent pattern where newer vehicles have a higher translation parameter than an older vehicle in the same body type. This finding is behaviorally intuitive as households usually tend to drive newer vehicles more than the older ones. Amongst all the alternatives, new vans have the highest translation parameter. A detailed exploration of the data revealed that households owning vans use these as multipurpose vehicles to meet the regular household travel necessities as well as the chauffeuring needs of the children. Motorbikes have the lowest translation parameter, which is expected as motorbikes are used mostly for pleasure/hobby travel but not as a primary vehicle in the household. Table 5.7 shows the goodness of fit statistics of the estimated model. The likelihood ratio of the estimated model is 645.36 which is substantially greater than the critical χ^2 value with 50 degrees of freedom at 99% level of confidence.

Table 5.7

MDCEV Model Estimation Results: Goodness of Fit Measures

Statistic	Value
Log-likelihood of final model at convergence	-77020.49
Degrees of freedom of final model	75
Log-likelihood of base model at convergence	-77343.17
Degrees of freedom of base model	25
Likelihood ratio	645.36
$\chi^2_{50,0.001}$	86.66

The estimated MDCEV model is applied on the entire data see how well the model can predict observed fleet composition patterns. Gauss codes made available by Pinjari and Bhat (2011), were translated to open source coding language 'R' to implement the MDCEV forecasting procedure. The model results were compared to match the following observed patterns

- Average annual mileage excluding zero mileage households: Average mileage is computed as the total mileage of each alternative divided by total number of households that have non-zero mileage consumption for the alternative
- Vehicle type distribution: Frequency of vehicle ownership for each vehicle alternative is computed as the total number of households who own a particular type of vehicle, divided by the total number of households in the dataset
- Body type distribution: Vehicle body type distribution of the observed data is compared against the body type distribution derived from the output of MDCEV model. This is an important check that should be passed by the MDCEV model, in order to impart necessary confidence in the model specification to be used for predicting fleet composition for a given (future) horizon year data. The body type distribution is not a factor that is inherently modeled in the MDCEV model specification. If the model is able to accurately predict this uncontrolled distribution, it would instill required confidence in the forecasts done using this model for any future year.

Figure 5.3 shows the comparison of observed and predicted average annual mileage consumption patterns across different alternatives for the dataset. The uncalibrated MDCEV model performed quite well in replicating the observed mileage consumption patterns. The model slightly over predicts the average mileage distributions for car body type. Some calibration of the model coefficients is warranted to exactly match the observed patterns. Figure 5.4 presents the observed versus predicted vehicle type distribution of the dataset.

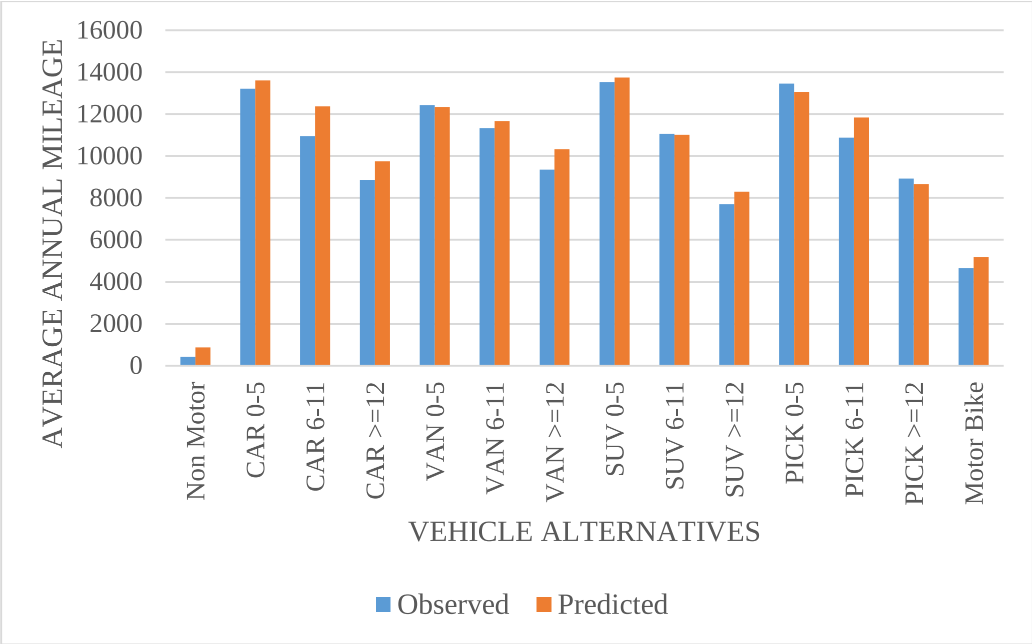

Figure 5.3. Observed vs. predicted average annual mileage consumption patterns: Uncalibrated MDCEV model.

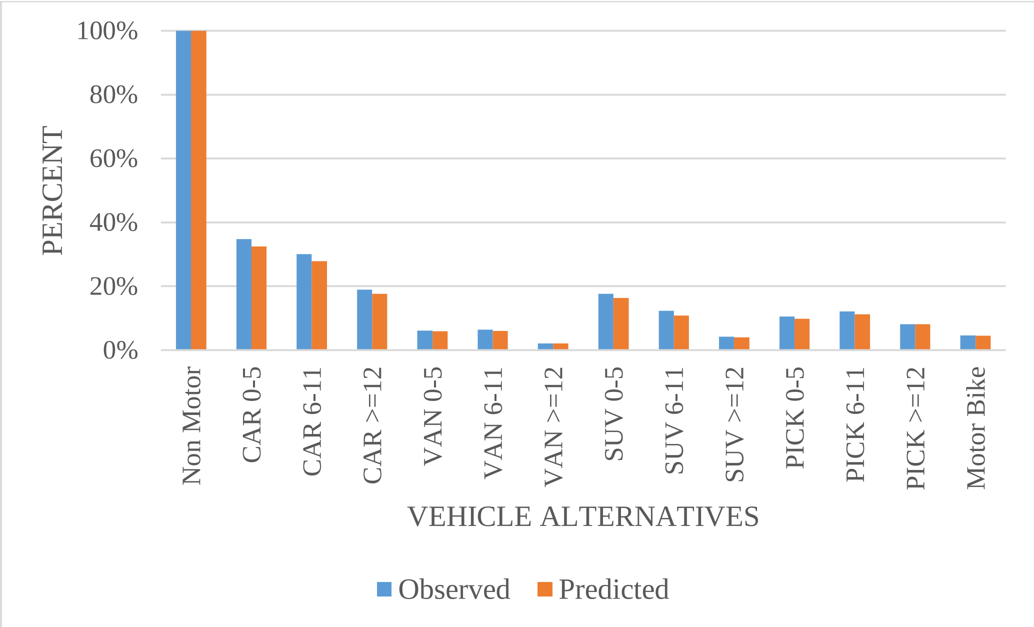

Figure 5.4. Observed vs. predicted vehicle type distribution: Uncalibrated MDCEV model.

The uncalibrated MDCEV model predicts the vehicle type distribution reasonably well, except it was observed that the model is slightly under predicting ownership across almost all body types. The reason for this behavior of the model is not readily apparent. In order to investigate this at a greater detail, the body type distribution of the observed data is compared against the body type distribution derived from the predictions of the MDCEV model. The results of this comparison (see Figure 5.5) elucidate the under predictions of vehicle ownership across the board by MDCEV model. The figure shows percentage of households in the dataset who own vehicles from distinct body type categories. Suppose a household owns a car 0-5 years old and a van 6-11 years, the household would be categorized as owing 2 vehicle body types. If a household owns a car 0-5 years old and a car 6-11 years, the household would be categorized owning only 1 vehicle body type. The comparison of observed and predicted categories of this 'implied' distribution, that is not directly controlled or modeled by the MDCEV specification sheds more light on the predictive capability of the model.

The MDCEV model is estimated and applied in such a fashion that every household in the dataset will consume at least some non-motorized mileage. The MDCEV model gives a different output each time a simulation is run. Which of the simulations should be considered final? It was observed that a single simulation of the MDCEV model predicts fleet mix quite well, but the implied body type distribution from the MDCEV model result almost always over predicts the proportion of households owning a single body type and under predicts all of the other categories. This finding answers the observation form Figure 5.4 where the MDCEV model is found to under predict the ownership of vehicles across all categories. A plausible explanation for this phenomenon is that as soon as a vehicle is selected as owned by the household, the MDCEV model apportions almost all of the mileage consumption for that household to the selected vehicle instead of choosing another vehicle alternative and apportioning a portion of the mileage to it. A calibration exercise was carried out to see if this behavior can be controlled by adjusting a few model coefficients, but the adjustments provided little improvements to the model predictions. To address the issue, and aid in the vehicle fleet composition modeling process, a framework is proposed (discussed in details in Chapter 3) which involves components that abet MDCEV model and provide better fleet composition predictions.

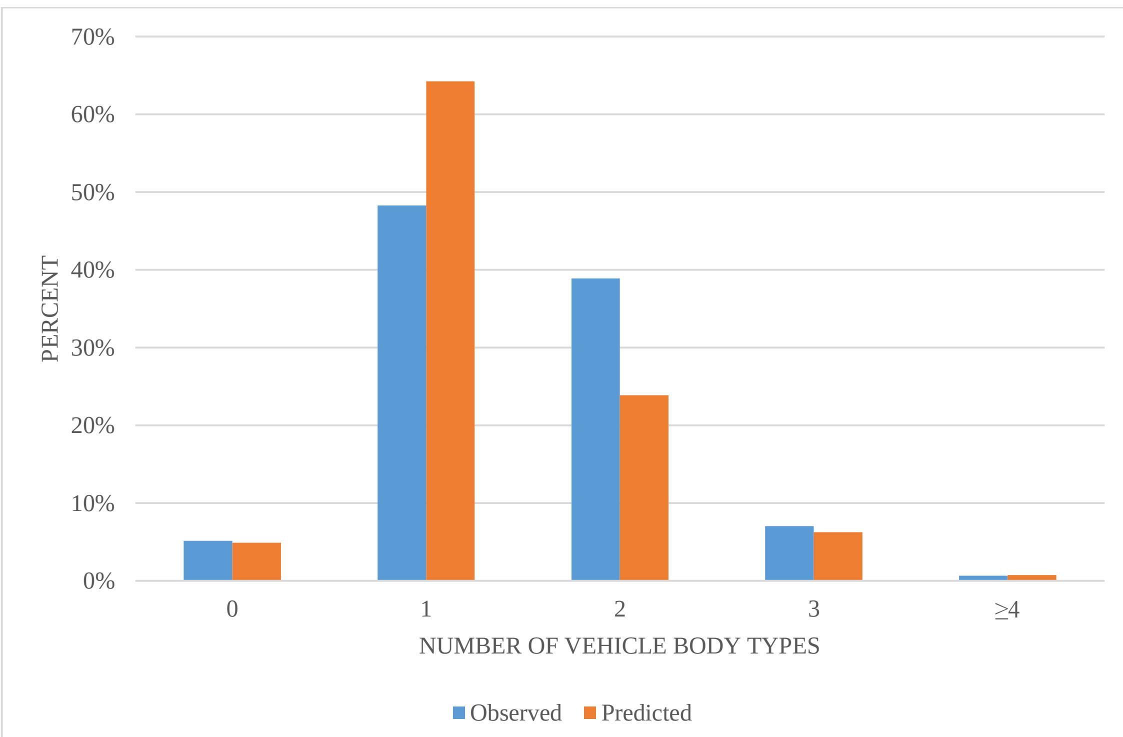

Figure 5.5. Observed vs. predicted vehicle body type distribution:
Uncalibrated MDCEV model.

In the proposed framework, MDCEV model simulation is carried out multiple times and mileage consumptions from each simulation are stored. After 'n' (say 100) simulations of the MDCEV model are completed, an average mileage consumption is computed for each alternative, which is then redistributed using a mileage reallocation algorithm. Since each simulation of the MDCEV model gives a slightly different result, the average mileage consumption result from 'n' MDCEV model runs show that a household owns almost all of the vehicle categories, whereas in reality the household might own only a subset of the vehicle categories considered by the MDCEV model of vehicle fleet mix. The heuristic mileage reallocation algorithm does the job of reallocating this mileage distribution in such a fashion that it reflects the household's vehicle fleet composition. The mileage reallocation algorithm requires information about how many distinct categories of vehicles does the household own. MNL model of number of vehicle alternatives predicts this information and provides it as an input to the mileage reallocation algorithm. Suppose, a household owns a car 0-5 years old and a car 6-11 years old, and a van 6-11 years old, the number of alternatives model is supposed predict the number of alternatives owned by this household as three.

MNL model of number of vehicle alternatives. The purpose of this model is to provide input to the heuristic mileage reallocation algorithm regarding the number of distinct vehicle alternatives owned by a household. Since the choice phenomenon at hand is a single discrete choice case (every household owns a unique number of vehicle alternatives), an MNL model structure is opted. Ideally, this model should have a total of 14 alternatives in accordance with the number of motorized alternatives in the MDCEV model structure. It is not required to consider the non-motorized alternative for this model or any subsequent models after the MDCEV model, as this is an outside good that 'should' be consumed by every household and hence need not be modeled separately. Observations from the estimation dataset revealed that the maximum number of distinct alternatives that any household in the dataset own is five. So, an MNL model is estimated with six categories (0-4, ≥ 5 vehicle alternatives). The final category (≥ 5 vehicle alternatives) served as the base alternative. Model estimation results are presented in Table 5.8.

From the model results, it was observed that lowest income households are likely to own fewer vehicle alternatives, while medium and high income households tended to own multiple vehicle alternatives which is consistent with expectation. Higher income households in general have the financial flexibility to own a mix of vehicles to cater for specific purposes. Single person households are more likely to own fewer vehicles (zero/one) which is an intuitive finding. Households living in TAZs with high population density tended to own zero vehicles. This might represent the category of households, who self-select themselves into mixed urban use TAZs (environmentally proactive households). Larger households are found to own multiple vehicle alternatives, another finding consistent with expectation as such households usually sport a vehicle (such as van) for family travel in addition to a vehicle to cater for regular travel necessities. Similar behavior was found in households with children. Highest income (≥ \$100,000) households were found to own 4 vehicle alternatives, which is an intuitive observation. The likelihood ratio of the model is substantially higher than the critical χ^2 value at any reasonable level of significance.

Table 5.8

MNL Model of Number of Vehicle Alternatives

Number of Vehicle Alternatives	Explanatory Variable	Coefficient	t-statistic
Zero	Constant	1.73	4.20
	Lowest income household (< $25,000)	2.37	9.70
	Low income household ($25,000 - $49,999)	0.82	3.23
	Housing unit owned	-1.65	-9.95
	Household size = 1	2.12	10.83
	Zero worker household	1.21	6.42
	Population density of the TAZ that the household resides	0.00011	4.16
One	Constant	4.27	13.72
	Lowest income household (< $25,000)	1.43	11.59
	Low income household ($25,000 - $49,999)	0.96	10.39
	Household size = 1	2.22	18.38
	Proportion of multi-family housing units in the TAZ	0.4	2.05
	Two worker household	-0.86	-6.37
Two	Constant	5.31	17.14
	Household with 2+ adults, youngest child 0-5	0.42	3.76
	Medium income household ($50,000 - $74,999)	-0.21	-2.29
	Two worker household	-0.34	-2.95
	Households in lowest income quintile	-0.00034	-2.03
Three	Constant	1.69	3.36
	Housing unit owned	1	3.01
	Count of adult HH members at least 18 years old	0.48	5.4
	Three or more worker household	1.02	4.52
	Population density of the TAZ that the household resides	-0.000088	-3.49
	Presence of children in the household	0.38	3.25
	Household with 2+ adults, youngest child 16-21	0.50	2.50
Four	Constant	1.18	3.19
	Highest income household (≥ $100,000)	0.72	3.00
	Household with 2+ adults, youngest child 16-21	1.63	5.47
	Household size = 4 or more	1.51	6.25
	Two worker household	-0.70	-2.59
Goodness of fit			
Sample Size (Number of Households)		4,262	
Likelihood ratio		2099.1	
$\chi^2_{25,0.001}$		52.62	

Figure 5.6 shows the comparison of observed and predicted vehicle alternative distributions. The results shown are for the uncalibrated version of the model and it can be observed that the model replicates the observed patterns exceedingly well with no necessity for calibration.

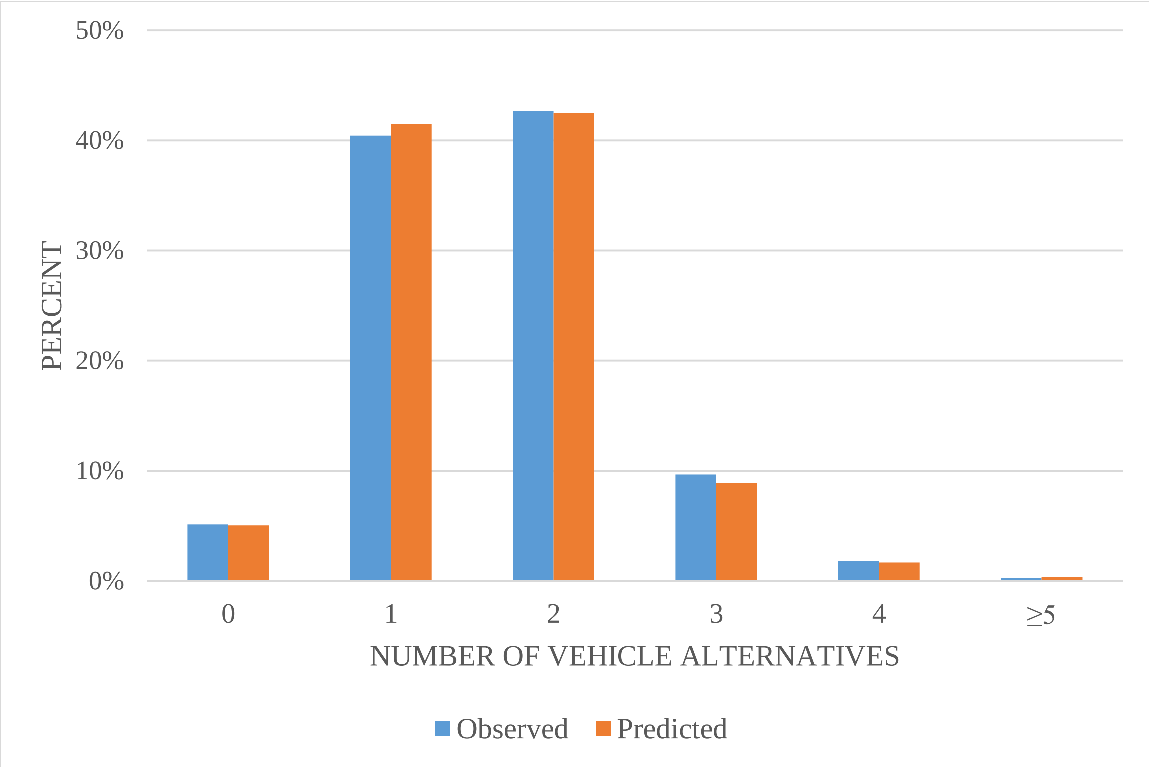

Figure 5.6. Observed vs. predicted vehicle alternative distribution: Uncalibrated model.

In addition to comparing the overall fit of the model for the entire dataset, comparisons were made across different income categories to see how well the model could replicate the vehicle alternative distributions of different income segments. Results of this comparison are presented in Figure 5.7. Panels A-E present the comparisons by income category. Sample sizes for each income category are given on the bottom right corner of the graph. Panel F depicts the synthesis of results from all the panels, where the income segments are juxtaposed within each vehicle alternative category. For example, in the 0 vehicle alternative category, the (pairs of) bars from left to right correspond to income categories starting from lowest to highest income. Similar presentation is followed for other vehicle alternative categories. The height of y-axis was kept the same across all panels for easier comparison. The blue bars always represent the observed patterns and the orange ones show predicted distributions. From the comparison charts, it can be observed that the model estimated on aggregate data performs quite well in predicting distributions across different income categories. This signifies that the model specification is robust enough to represent the difference in vehicle ownership patterns across different income segments. A high proportion of lowest income households (≤ \$25,000) correspond to the zero vehicle alternative category and with increase in income segment the proportion of households in this category slowly decreases with almost no households in the zero alternative category for highest income household segment. This finding is behaviorally intuitive, in the sense that as the household's income level increases, so does the financial flexibility to own more number of vehicles. It is quite heartening to see the model predict the same phenomenon.

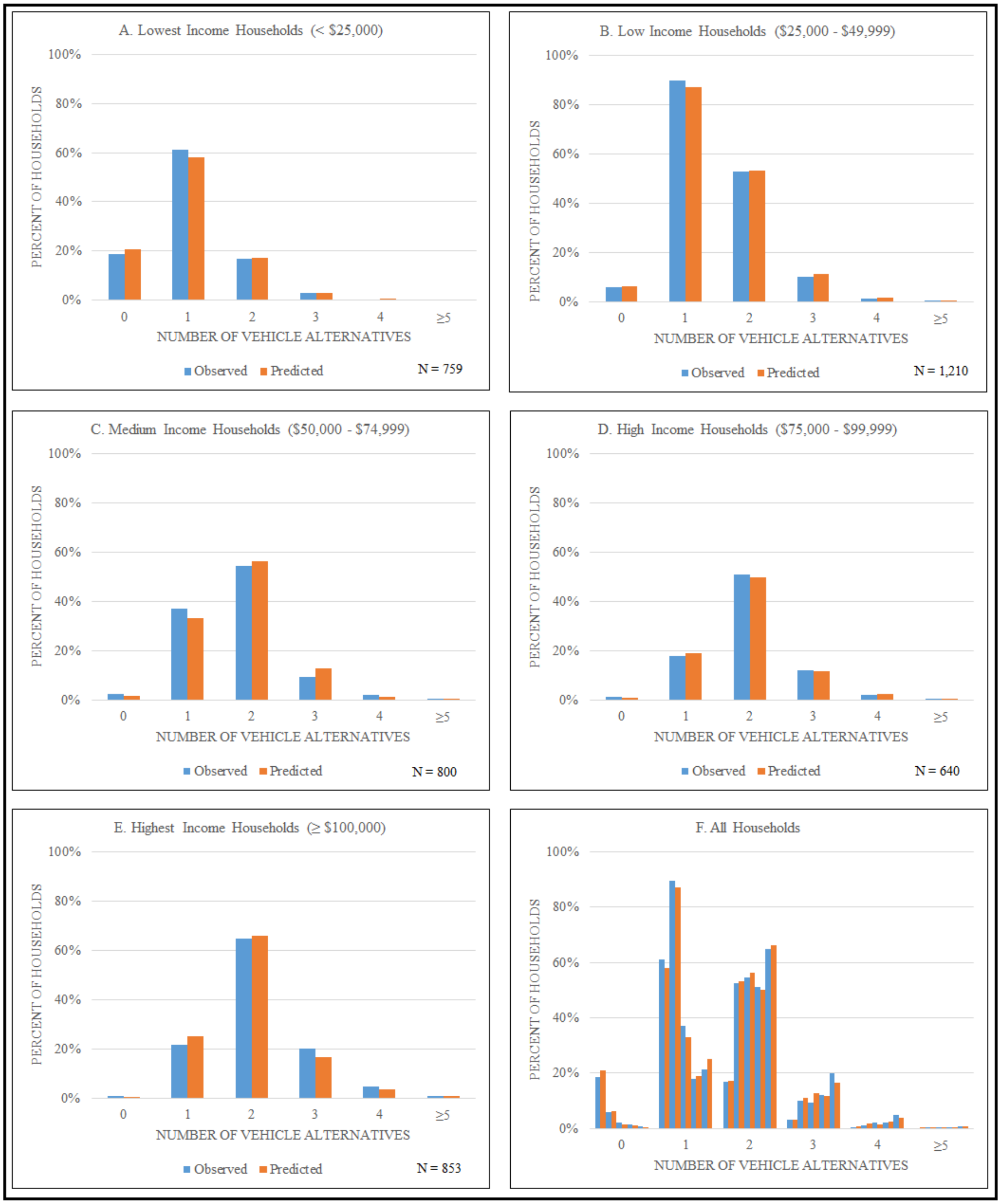

Figure 5.7. Observed vs. predicted vehicle alternative distributions by income category: Uncalibrated model.

Low income households majorly own one vehicle alternative (about 90%). Lower representation of medium through high income households in this category is supplemented their stronger presence in the higher vehicle alternative categories. For vehicle alternative categories 3 and 4, a gradually increasing representation can be observed from lowest to highest income households. This exceedingly good performance of the model to predict subtle nuances in the dataset imparts necessary confidence to use the output of this model as a governing distribution in the heuristic mileage reallocation algorithm.

MNL model of number of body types. For every household in the dataset, the heuristic mileage reallocation algorithm takes the output of the MDCEV model and redistributes the mileage as governed by the number of vehicle alternatives model. After this process is carried out for every household in the dataset, comparisons are made across observed and predicted distributions of average annual mileage, vehicle type and body type. While the data to calibrate the model on these three grounds is readily available for base year, how can one be sure that the predicted vehicle fleet composition distributions are representation of the actual fleet composition for a future year input data? To ensure this consistency, a vehicle body type MNL model is estimated and calibrated for the base year. The model predicts this distribution based on the projected synthetic population characteristics for any horizon year. This goes in as a control distribution that should be matched by the heuristic mileage reallocation algorithm.

The structure of the MNL model of number of body types is very similar to that of the previous model, except in this case the number of distinct vehicle body types owned by a household are modeled instead of the number of distinct vehicle alternatives. Going of the example from previous section, if a households owns a car 0-5 years old and a car 6-11 years old, and a van 6-11 year old, the household is said to own a total of 2 vehicle body types (car and van). There are a total of 6 body types considered in the context of the current research effort namely car, van, SUV, pick-up, motorbike and non-motorized alternative. The MNL model should ideally include 6 alternatives, but observations from the estimation dataset revealed that the maximum number of body types owned by any household in the dataset is 5, with very few households owning 4 or more vehicle body types. So, the vehicle body type count is truncated at 4, thereby providing a total of 5 alternatives (0-3, ≥ 4 vehicle body types). The final category (≥ 4 vehicle body types) served as the base alternative. Model estimation results are presented in Table 5.9.

From the model results, it was found that lowest and low income households are most likely to own zero or one vehicle body type. This finding traces back to the question of affordability of multiple vehicle types for this segment. Also, this finding couples nicely with the result of the MNL model of number of vehicle alternatives where households in this segment appeared in the zero and one vehicle alternative category, which automatically positions them in the same category in the vehicle body type model. Single person households are found to own none or one vehicle body types at the most, which is consistent with expectation as such household do not need more than one vehicle for their travel needs in general. Larger households as well as households with more number of drivers are likely to own more vehicle body types. This translates to the variety seeking nature of different individuals in such households.

Table 5.9

MNL Model of Number of Vehicle Body Types

Number of Vehicle Body Types	Explanatory Variable	Coefficient	t-statistic
Zero	Constant	0.68	2.09
	Lowest income household (< $25,000)	2.09	8.42
	Low income household ($25,000 - $49,999)	0.61	2.38
	Housing unit owned	-1.43	-8.58
	Household size = 1	2.11	10.45
	Zero worker household	1.14	6.15
	Population density of the TAZ in which the household resides	0.00011	4.08
One	Constant	3.67	19.69
	Lowest income household (< $25,000)	1.02	7.95
	Low income household ($25,000 - $49,999)	0.65	6.86
	Household size = 1	1.99	15.62
	Proportion of multi-family housing units in the TAZ	0.26	1.39
	Presence of children in the household	-0.33	-3.2
Two	Constant	2.96	13.01
	Household resides in rural area	0.19	2.01
	High income household ($75,000 - $99,999)	0.35	3.34
	Highest income household (≥ $100,000)	0.23	2.37
	Household size = 4 or more	0.3	2.73
	Housing unit owned	0.98	6.47
Three	Housing unit owned	1.42	5.9
	Count of adult HH members at least 18 years old	0.55	6.4
	Three or more worker household	0.46	1.96
	Population density of the TAZ in which the household resides	-0.00008	-2.82
	Presence of children in the household	0.36	2.55
Goodness of fit			
Sample Size (Number of Households)		4,262	
Likelihood ratio		1658.30	
$\chi^2_{21,0.001}$		46.80	

Presence of children is found to negatively influence owning a single vehicle body type. Chauffeuring needs of children require owning a bigger vehicle (such as a van) in addition to owning another vehicle body type for usual travel in such households. This finding is corroborated by the significance of same variable in the three vehicle body type category. The likelihood ratio test statistic for the model is 1658.30 which is substantially higher than the critical χ^2 value with 21 degrees of freedom and 99% level of signifance. This confirms the presence of exogenous variable effects in the model specification.

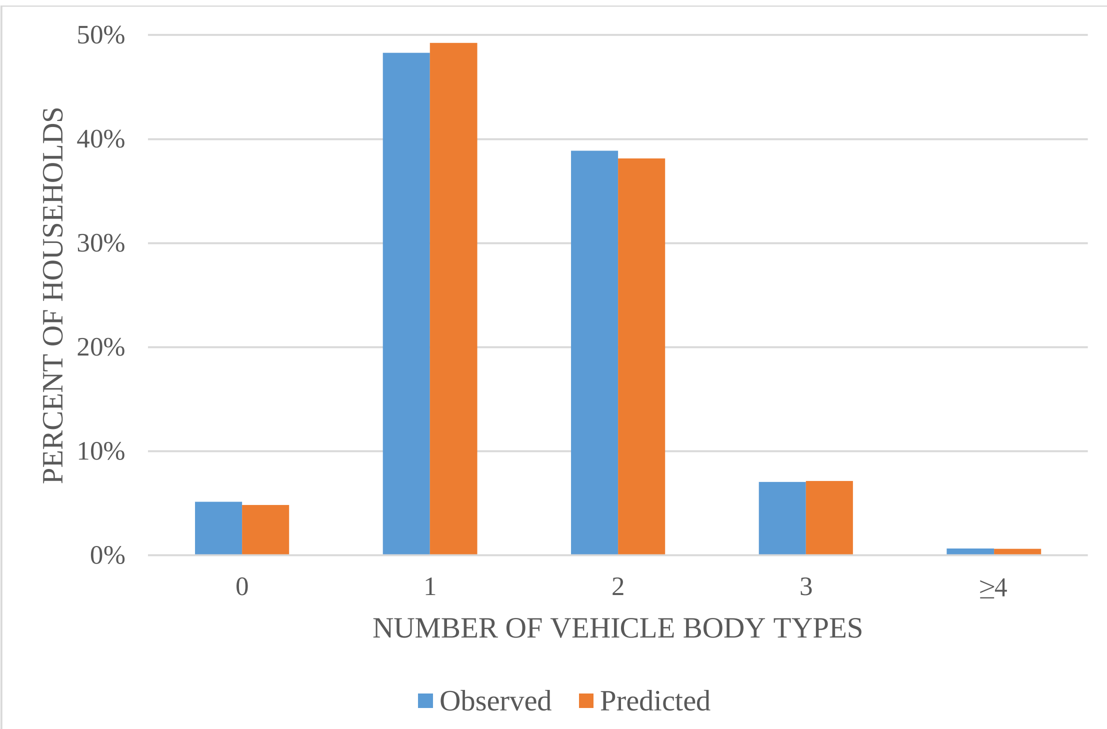

Figure 5.8. Observed vs. predicted vehicle body type distribution: Uncalibrated model.

Figure 5.8 presents the comparison of observed and predicted body type distribution of the dataset. It can be observed that the model replicates the observed vehicle body type distributions quite well. As expected the households owning more vehicle body types (3, ≥ 4) are very few in the dataset and the model predictions line up quite well with the observed distributions. Majority of the households in the dataset are found to own one vehicle body type, which doesn't necessarily mean that these households own a single vehicle. Even households owning multiple vehicles might fall under this category, if all of their vehicles happen to be the same body type (multiple cars, vans etc.). The results presented are for uncalibrated version of the MNL model.

In addition to the comparison of aggregate distributions, the model was tested for its efficiency in replicating the body type distributions at disaggregate income level classification. Results of the comparison are presented in Figure 5.9. The comparisons are made pairwise, where each set of bars (observed and predicted) corresponding to a particular income category are juxtaposed side-by-side. The pairwise comparison of different income segments reveal that the model performs quite well in replicating the vehicle body type distribution of different income segments. A more interesting observation comes from looking at the vehicle body type distribution of different income segments together (left to right in the figure). The percentage of households owning zero vehicle body types (or no vehicles at all) is higher in the lowest income category and this percentage slowly reduces as we move across the higher income segments. Similarly, the percent of households owning one vehicle body type is high in lowest and low income categories and this percentage gradually decreases as the household income increases. This observation speaks to the affordability combined with variety seeking nature of households in respective income segments. Though the observed and predicted distributions are presented side-by-side, the propagation of vehicle type distribution across different household income segments seems rather continuous because the model is able to predict the vehicle body type distributions across all the market segments quite well. This imparts much confidence to use the model prediction as control distribution that should be match by the heuristic mileage reallocation algorithm.

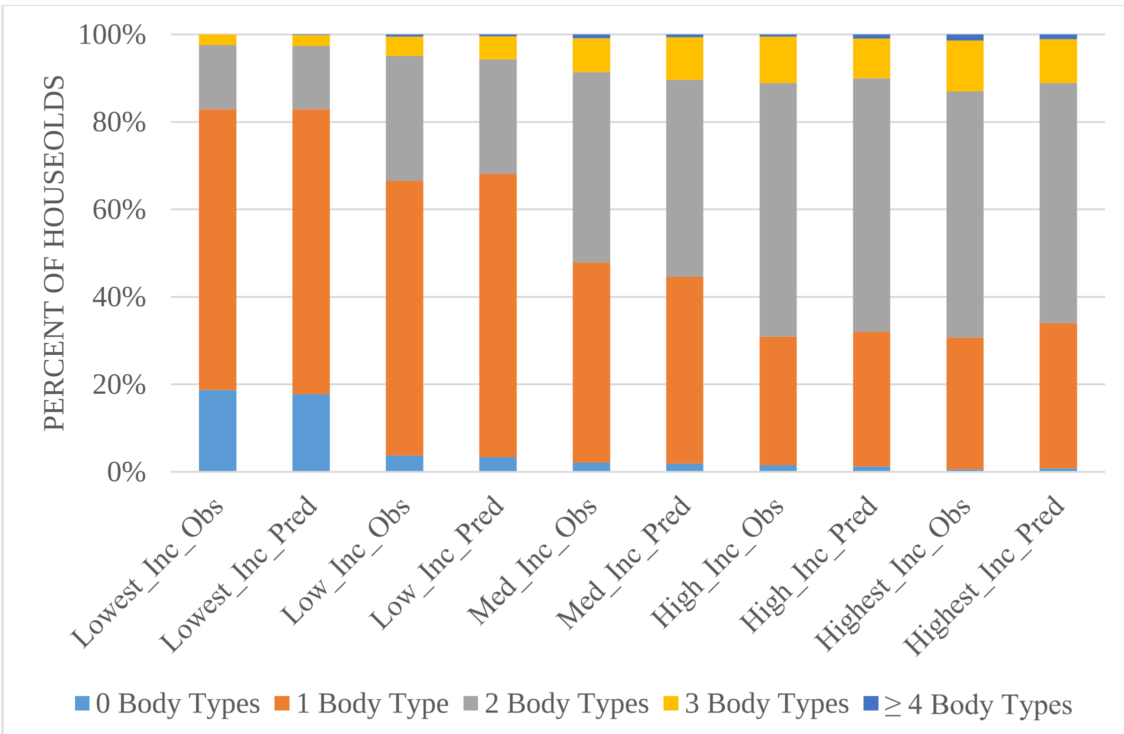

Figure 5.9. Observed vs. predicted vehicle body type distributions by income category: Uncalibrated model.

Heuristic mileage reallocation algorithm. The heuristic mileage re-allocation algorithm takes output of MDCEV model and MNL model of number of alternatives as input, to redistribute the mileage to number of alternatives owned by the household. Several alternative approaches were tested to effectively predict the fleet mix as well as body type distributions and the best approach among the once tested is presented. The logic followed by the heuristic mileage reallocation algorithm is discussed in detail in Chapter 3.The mileage reallocation algorithm operates at the level of each household, where it reallocates the mileage output of the MDCEV model. The output from MNL model of number of vehicle alternatives gives information about how many distinct vehicle body-type x age categories does the household own. From the output of MDCEV model, cumulative mileage distribution of the household is computed for the household. A random number is generated and based on location of the random number in the cumulative mileage distribution of the household, a vehicle is selected as 'owned' by the household. The alternative chosen (and the corresponding mileage) is removed from the dataset thereby eliminating choice of the same alternative multiple times. This process is carried out '*k*' times, where *k* (number of vehicle alternatives) is predicted by the number of alternatives model.

At the end of the iteration for a particular household, the HMR algorithm selects all the alternatives owned by the household. Now, the mileage consumed by these alternatives is scaled up proportionally to account for the annual motorized mileage consumption of the household. Once, the HMR algorithm reallocates the mileages for all households according to the input provided by number of alternatives model, the predicted body type distribution of the entire sample is compared against the body type distribution predicted by the MNL model of number of body types. Absolute percent difference is computed between both the distributions and this is checked against a pre-set tolerance limit selected by the user (say 5%).

If the HMR algorithm passes the tolerance check, the output of HMR algorithm goes in as input to the count models. If not, the entire application process is repeated after calibrating the model components as necessary. This process is carried out repeatedly until the percent difference between the two distributions (from MNL model of number of body types and the output of HMR algorithm) satisfies the tolerance criteria. In the context of the current modeling effort, a few coefficients in the MDCEV model specification are calibrated/asserted to match the observed vehicle fleet composition patterns better. The calibration/assertion exercise was carried out with due caution regarding any unexpected consequences such changes might bring about. The output of HMR algorithm gives us the final fleet composition of every household in the dataset. The output of this algorithm would have successfully predicted the vehicle ownership of the household, body type composition and vintage composition of the vehicles owned. As discussed before, comparisons between observed and predicted distributions are made for

- Average annual mileage
- Vehicle fleet mix distribution
- Aggregate vehicle body type distribution

Comparison of observed and predicted fleet composition patterns from the output of the HMR algorithm are presented in Figure 5.10. Average annual mileage distribution comparisons are presented on the bottom axis whereas the top axis presents the vehicle type distributions (the axis is inverted for ease of presentation).

It can be observed from the figure that the model replicated the observed fleet composition as well as mileage consumption patterns very well. It should be noted that an exact match between observed and predicted distributions is quite difficult to achieve and might require extensive calibration of the model system. It was felt prudent to rather capture the fleet composition patterns with slight calibration of the model system, than exactly match both of these distributions. It can be observed from the figure that within each body type, households prefer to drive newer vehicles more than that of older ones.

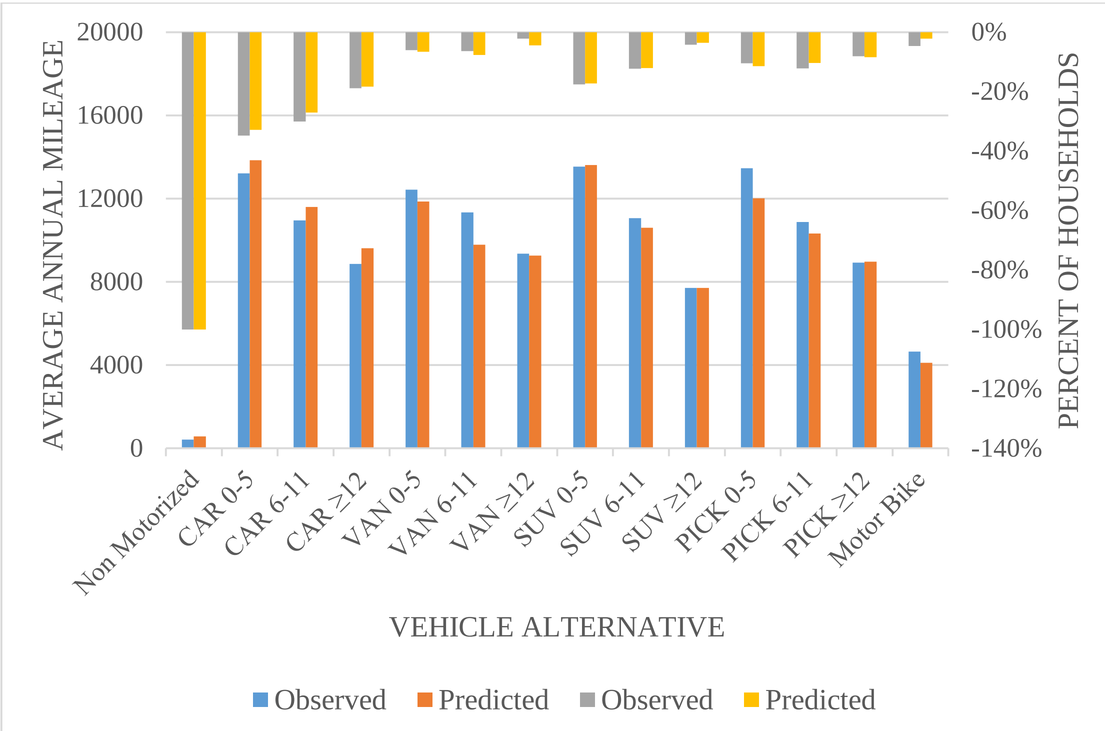

Figure 5.10. Observed vs. predicted distributions: Output from HMR algorithm.

Another distribution that the output of the HMR algorithm is expected to match is the aggregate vehicle body type distribution. This comparison is presented in Figure 5.11. Note that the comparisons made in this figure are not against the observed vehicle body type distribution, but are against the vehicle body type distribution as predicted by the MNL model of body type count. While the base year data for this comparison is readily available for making this comparison, horizon years for which the model system would have to predict the vehicle fleet composition will not have this data readily available. Only socio-economic data will be provided for the model as input from which the vehicle fleet composition model system should predict the fleet mix. Keeping this in mind, comparisons are made against predicted data and not the observed, as this is how the model would be used for any horizon year prediction. Compared to the vehicle body type distribution without the HMR algorithm (see Figure 5.5), the predicted distribution replicates the observed vehicle body type distribution quite closely. The output from HMR algorithm matches the observed patterns across all the comparisons made.

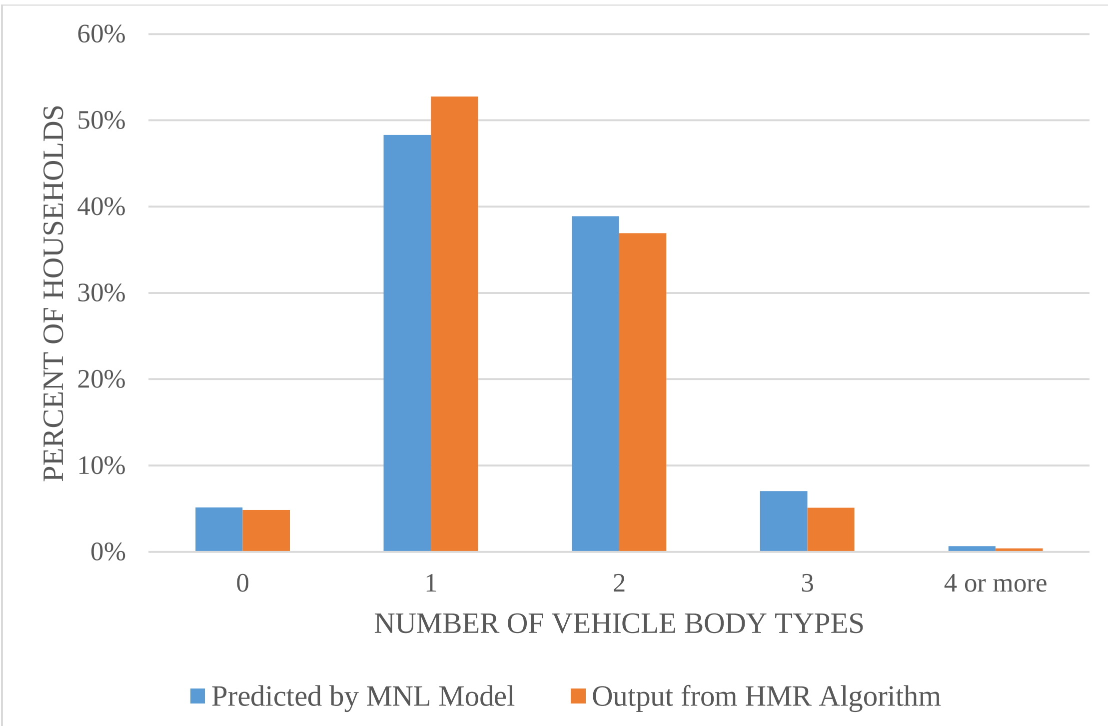

Figure 5.11. Observed vs. predicted vehicle body type distribution: Output from HMR algorithm.

In addition to comparisons at the aggregate level, the model system is tested on specific market segments to see how well it replicates the fleet mix and mileage consumptions of specific types of households. One such comparison is presented in Figure

5.12 for households residing in urban vs rural areas as these households might have significantly different mileage consumption as well as fleet mix patterns.

Figure 5.12. Observed vs. predicted distributions for urban and rural residents: Output from HMR algorithm.

The representation of households living in urban areas is slightly higher in the dataset (84%) than the national average of about 70%. This finding is consistent with the geography of Greater Phoenix Metropolitan Region which is predominantly urban. An important observation from the comparison chart shown in Figure 5.12 is that rural residents have greater average annual mileage consumptions across almost all vehicle types. This finding is behaviorally intuitive as rural residents indeed tend to have greater annual mileages as a result of travel to/from the adjacent city to engage in various activities. It is also observed that households residing in rural areas have a higher proportion of pick-ups in their vehicle fleet (across all vintage categories), than their urban counterparts. Similar comparison is shown for a low and high income households in Figure 5.13.

The main takeaway from the comparison of vehicle fleet composition patterns of low income households is the representation of number of vehicles in the 'newest' vintage category (0-5 years old) across all vehicle body types. High income households are found to own and use newer vehicles more. This is generally expected behavior as such households usually have a faster turnover of vehicles in their fleet. The vehicle fleet composition model system is able to replicate the ownership and utilization patterns of both these market segments reasonably well. The model does not 'exactly' match the observed distributions, but in general captures the patterns in observed data quite well.

Figure 5.13. Observed vs. predicted distributions for low and high income households: Output from HMR algorithm.

The output of HMR algorithm replicated the observed vehicle fleet composition patterns quite well. The output of HMR algorithm goes as input to the count models. The count models determine if all the mileage consumed by a household within a particular vehicle alternative belongs to one or multiple vehicles. The count models are necessary because vintage classifications are aggregated into 3 categories for ease of estimation and application of the fleet composition model system. Suppose, the output of HMR algorithm determines that a household uses a car 0-5 years old to travel 25000 miles annually, the count model determines if all of this mileage is consumed using just one car 0-5 years old or if the household owns multiple cars of 0-5years of age. Ideally, a count model should be estimated for each of the 13 different vehicle categories defined for the MDCEV model, but this will heavily increase the number of individual components in the model system while decreasing the data available to estimate each of the individual count models. So, it was felt prudent to estimate one count model for each of the body types, with vintage serving as an explanatory variable in the models. If the household has non-zero mileage consumption in any of the vintages of a particular body type, count model of that body type is applied for the household. The output of count models is a test to the efficacy of the entire model system as this is a sequential application process. Each of the count models and their performance in replicating observed patterns is presented next.

Count models. Count models take the mileage output from HMR algorithm and determine if all of that mileage is consumed by a single vehicle or multiple vehicles in the alternative under consideration. Ideally, 13 different count models should be estimated (one for each of the body type – age classification), but this would heavily increase the number of individual components in the model system. It was felt prudent to estimate one count model for each body type with vintage categories serving as explanatory variables in the models. The effectiveness of this simplification is tested and the results were satisfactory across all the vehicle body types. Ordered probit count models are estimated for car, van, SUV and pick-up body types. Model estimation results of car count model are presented in Table 5.10.

From the model estimation results, it was observed that highest income households usually sport multiple cars in their fleet while low income households on the other hand are not likely to multiple cars. It was also observed that three or more worker households tend to own multiple cars. This finding is behaviorally consistent as such households might usually require more than one car for daily commute travel for multiple workers in the households. Households living in an owned housing unit have a greater propensity of owning multiple cars. This variable might act as a proxy for the income of the household. The likelihood ratio statistic of the model (830.75) is significantly greater than critical χ^2 value at 99% confidence level.

Table 5.10

Car Count Model Estimation Results

Explanatory Variable	Coefficient	t-statistic
Constant	-2.74	-17.09
Indicator for car 0-5 years old	0.95	14.44
Indicator for car 6-11 years old	1.05	16.21
Low income household ($25,000 - $49,999)	-0.10	-1.63
Highest income household (≥ $100,000)	0.19	3.10
Three or more worker household	0.41	3.44
Count of adult HH members at least 18 years old	0.47	9.15
Household size = 1	-0.42	-3.93
Household size = 4 or more	-0.26	-3.64
Housing unit owned	0.28	2.62
Threshold parameters for index		
μ(1)	1.45	28.22
Goodness of Fit		
Likelihood ratio	830.75	
$\chi^2_{9,0.001}$	27.88	

The car count models are applied on only those households for whom the HMR algorithm allocates at least some non-zero mileage in any of the car vintage categories. Suppose the HMR algorithm allocated a mileage of 10000 miles for car 0-5 years old category and 5000 miles for car 6-11 years old category, car count model is applied on each of these categories to identify if the households own multiple cars in the category 0-5 years old and/or car 6-11 years old. If the car count model predicts multiple cars for any of the categories, mileage for that alternative is evenly distributed among the number of predicted alternatives. In the above example, if the model predicts that the households owns two cars in vintage category 0-5 years, each vehicle is assigned a mileage of 5000.

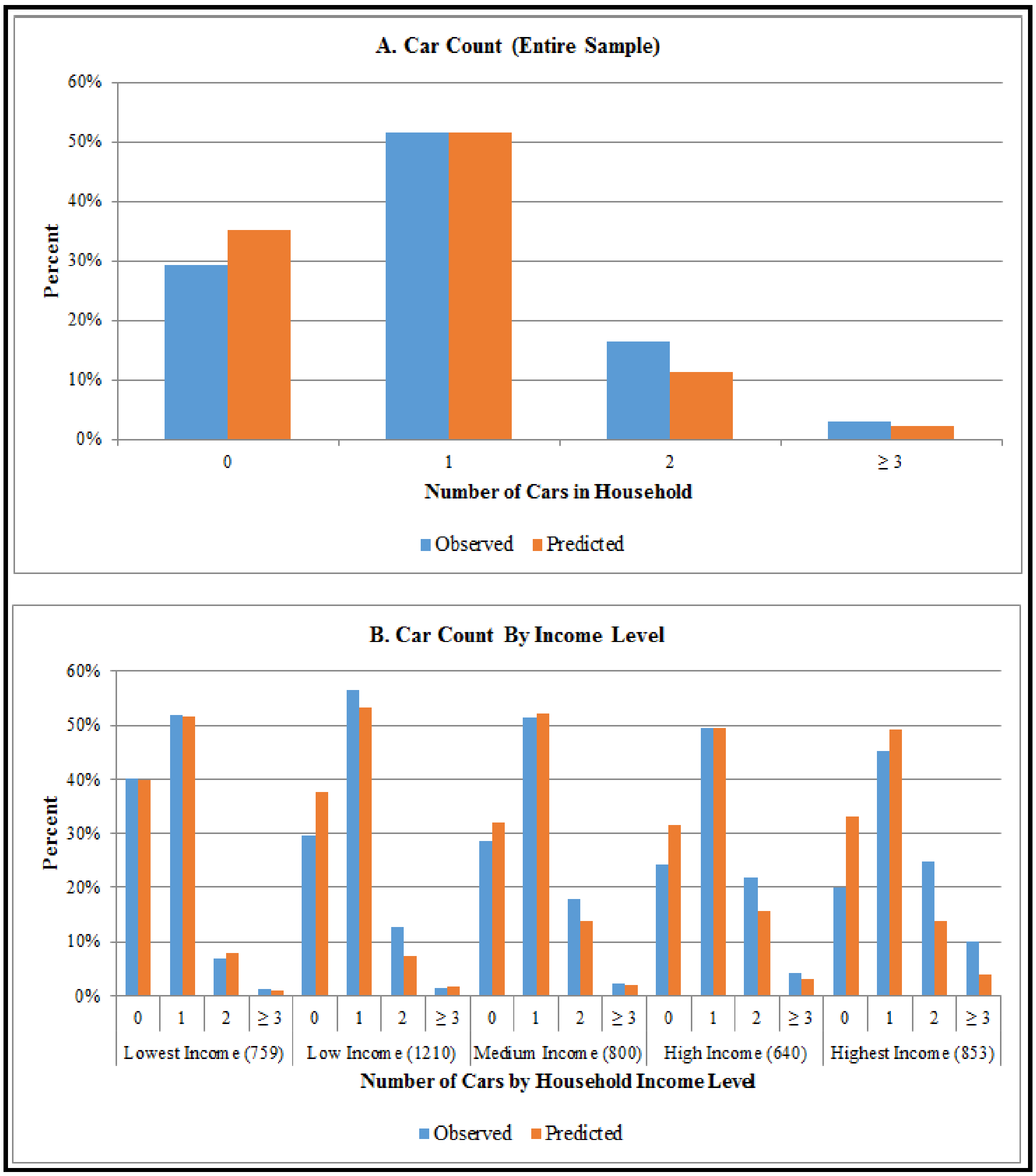

Figure 5.14. Observed vs. predicted distributions for car count.

The comparison of observed and predicted car counts is presented in Figure 5.14. Panel A presents the comparison between car count distributions of the entire dataset. The results presented are for the uncalibrated version of the count model. It can be observed that model replicates the car count distribution quite well with slight under prediction in the 2 car category. Some calibration of the model is warranted to better match the observed distributions. Panel B of the Figure 5.14 presents the car count distribution comparisons by income level. The model performs reasonably well in replicating the observed car count distributions for most of the income levels. The only segment which might require some further investigation is the highest income segment where the count model seems to under predict multiple car households.

Table 5.11

Van Count Model Estimation Results

Explanatory Variable	**Coefficient**	**t-statistic**
Constant	-4.62	-7.75
Indicator for Van 6-11 years old	1.95	3.92
Indicator for Van 12 years or older	2.57	4.97
Low income household ($25,000 - $49,999)	-0.59	-1.82
Annual mileage of Van trips	$6*10^{-5}$	5.46
TAZ with low density (3rd Quartile)	0.60	1.76
TAZ with high regional employment accessible within 10 minutes by auto (1st Quartile)	0.87	2.99
Employment density of the TAZ that the household resides	$-2*10^{-4}$	-1.47
Goodness of Fit		
Likelihood ratio	103.14	
$\chi^2_{7,0.001}$	24.32	

Table 5.11 presents the estimation results for van count model. From the model result, it was found that households tend to own multiple vans of older vintages than new

ones. Also, the mileage consumption variable is positive and significant. If the HMR algorithm allocates a high mileage to any of the van categories the count model will be able to identify and allocate that mileage to multiple vehicles of the same category. Low income households had a lower propensity to own multiple vans (or multiple vehicles of any category for that matter), which is intuitive. Households residing in TAZs with high employment density (mixed use zones) tended not to own multiple vans.

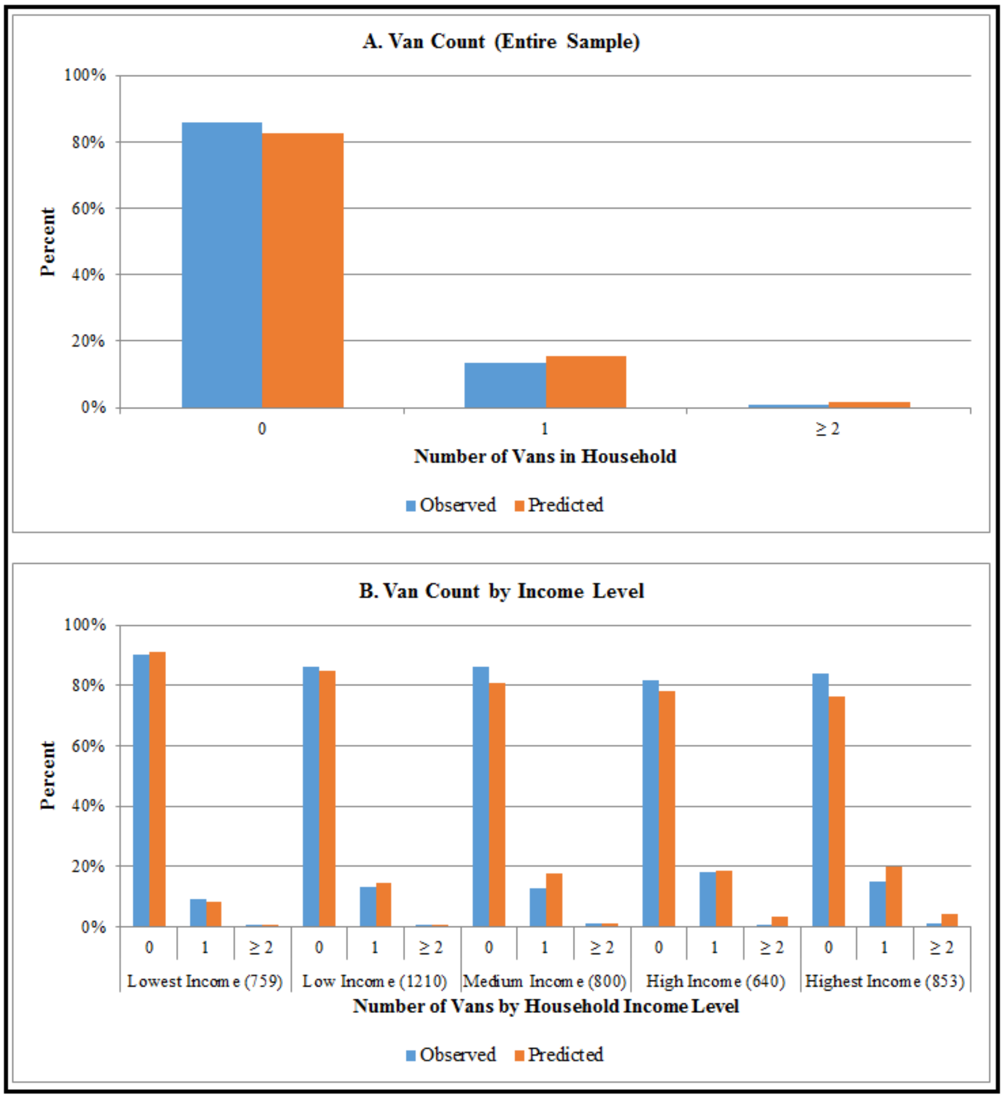

Figure 5.15. Observed vs. predicted distributions for van count.

Figure 5.15 presents the comparison between observed and predicted van count distributions. Panel A presents the results for the entire data and van count distributions for various income segments is shown in Panel B. It can be observed that the model performs consistently well overall, as well for specific income segments. Form the figure it can be observed that the van count distributions are not very different across the income categories. Possible reason for this might be that the impetus for owning a van is not solely based on the affluence of the household but rather on the composition of the household (such as presence of a child or more number of people in the household) and the corresponding travel dynamics.

Table 5.12 presents the model estimation results for SUV count model. Unlike the van count model, SUV count model has a positive and significant coefficient for the newer vintage category (0-5 years), which specifies that household who own and drive SUVs, like them rather new than old. Income categories used in the SUV count model show very intuitive findings. From model results, low income households in general have a lower probability of owning multiple SUVs, whereas high income households on the other hand have a greater propensity to multiple vehicles of this type. Usage of annual mileage variable in the model ensures distribution of high mileage predictions for any alternative to multiple vehicles in this category.

An interesting observation with respect to the SUV count model is the magnitude of the coefficient on annual mileage. For SUV count model the value of this coefficient is $4*10^{-5}$, where as in the van and pick-up count models, the value same coefficient is greater ($6*10^{-5}$). This finding translates to the fact that households who own multiple SUVs drive them for relatively lower annual mileage than that of vans and pick-up truck. This finding is behaviorally consistent in that SUVs are generally used for leisure travel.

Table 5.12

SUV Count Model Estimation Results

Explanatory Variable	Coefficient	t-statistic
Constant	-3.64	-14.06
Indicator for SUV 0-5 years old	0.90	6.16
Indicator for SUV 12 years or older	1.50	8.61
Count of adult HH members at least 18 years old	0.46	5.23
Low income household ($25,000 - $49,999)	-0.53	-2.84
Highest income household (≥ $100,000)	0.40	3.57
Annual mileage of SUV trips	$4*10^{-5}$	9.92
Three or more worker household	-0.46	-1.93
Percent of regional employment within 30 minutes of transit accessibility from the TAZ	-34.16	-2.07
Goodness of Fit		
Likelihood ratio	309.57	
$\chi^2_{8,0.001}$	26.13	

Figure 5.16 presents comparison of observed and predicted distributions for SUV count model. Panel A depicts the comparison of SUV count distributions for the entire dataset and Panel B presents the results by income level. The model predicts presence of multiple SUVs across different income categories. As the category of household income increases, it can be observed that the presence SUVs in the household (1, ≥ 2) slowly increases and the model is able to predict this pattern quite well. It should be noted that the results presented are for uncalibrated version of the count model. Slight calibration is warranted to match the observed distributions better.

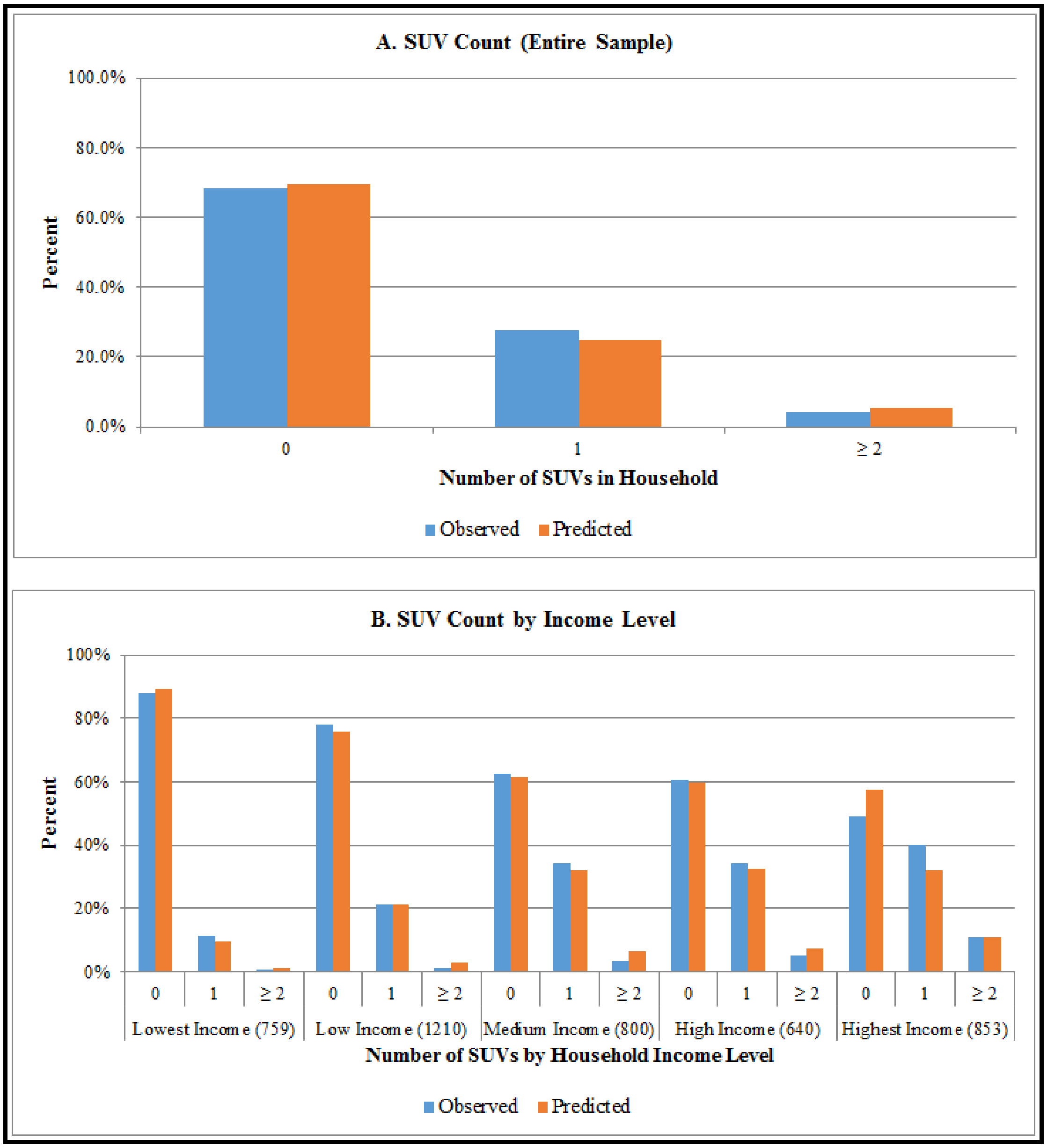

Figure 5.16. Observed vs. predicted distributions for SUV count.

Model estimation results for pick-up count models are presented in Table 5.13. Annual mileage variable is positive and significant in the model specification which avoids allocation of greater mileages to a single pick-up truck category. Households residing in

TAZs with low density are found to own multiple pick-up trucks. These might refer to households residing in rural areas.

Table 5.13

Pick-up Count Model Estimation Results

Explanatory Variable	Coefficient	t-statistic
Constant	-3.95	-8.40
Indicator for pickup truck 0-5 years old	0.87	5.16
Indicator for pickup truck 6-11 years old	1.02	6.21
Annual mileage of pickup truck trips	0.00006	9.88
Highest income household (≥ \$100,000)	-0.29	-2.08
Three or more worker household	0.46	2.32
Housing unit owned	0.81	1.96
Employment density per 10 square miles	0.01	1.58
TAZ with low density (3rd Quartile)	0.30	1.88
TAZ with high regional employment accessible within 30 minutes by auto (1st Quartile)	0.27	2.09
Goodness of Fit		
Likelihood ratio	240.62	
$\chi^2_{9,0.001}$	27.88	

Highest income households (≥ \$100,000) had a lower propensity to own multiple pick-up trucks. Figure 5.17 shows the comparison of observed and predicted distributions of pick-up counts. Percentage of households who own multiple pick-trucks are quite low in the dataset (only 3 % households own ≥ 2 pick-up trucks). The model is able to replicate observed patterns quite well across different income categories as well as the overall distribution. On the whole, results of the fleet composition model system are quite encouraging. Some calibration of the model system is warranted to better match the observed patterns. Information at such disaggregate level regarding the fleet composition patterns of households in a region help in accurate emission predictions.

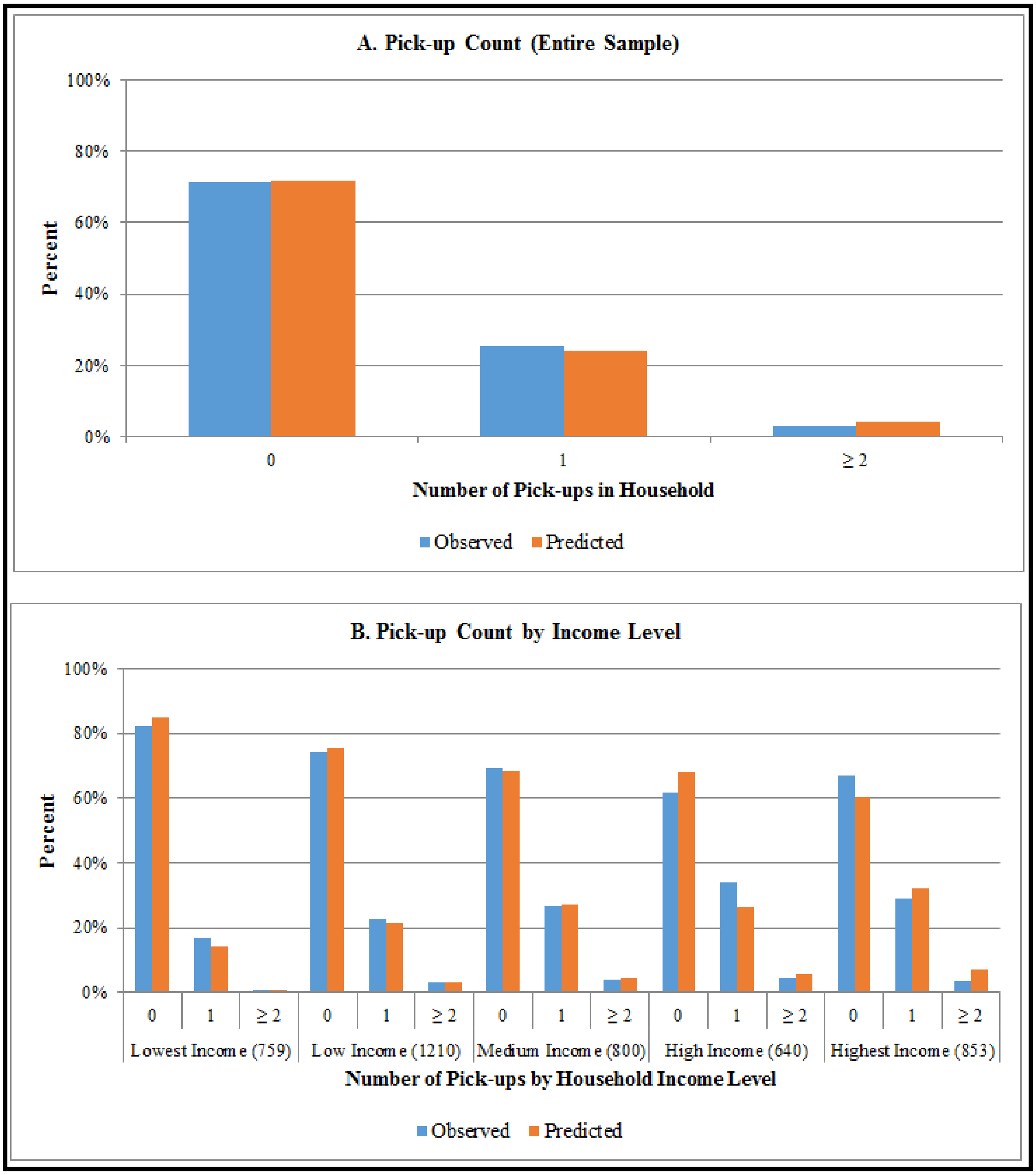

Figure 5.17. Observed vs. predicted distributions for pick-up count.

Sensitivity Analysis Exercise

Once the model was found to replicate the observed vehicle ownership patterns satisfactorily, a sensitivity analysis exercise was carried out to examine the ability of the model system to respond in a meaningful way to changes in input test conditions. First, a baseline scenario was established by applying the model system on the entire sample (4,262) households. Five scenarios were created by varying the percent of regional employment accessible from a households TAZ location by auto mode. The regional employment accessible was increased incrementally by 10%, 20%, 30%, 50% and 100% from the baseline. To build these scenarios, auto skims were used to select the percent regional employment accessible within a set travel time (10 minutes and 30 minutes) in the baseline and this employment was increased by respective percentages for each of the scenarios. This translates to increasing the accessibility of a household's TAZ location by enhancing the percentage of regional employment accessible from it.

It was observed in the modeling exercise that accessibility has a negative impact on vehicle ownership patterns i.e., households living in denser developments usually tended to own fewer vehicles. The reason for this behavior is twofold. First, households who are more environmentally proactive and already own fewer vehicles might self-select themselves into dense mixed-use urban locations. This phenomenon is called as residential self-selection and plays in an important role in auto ownership as well as travel demand in general. This topic has been the focus of many earlier studies (Cao et al., 2006; Bhat and Guo, 2007; Pinjari et al., 2008a; Bhat et al., 2013) and is not dealt with in the current research effort. The second reason is the fundamental causality between built environment and travel-behavior, which explains why dense urban developments tend to be more walk/bike/transit friendly than sparse suburban neighborhoods (Frank and Pivo, 1994; Cervero and Seskin, 1995; Cervero and Kockelman, 1997; Ewing, 2008). Results of the sensitivity analysis test are presented in Figure 5.18. The figure depicts the changes in vehicle ownership patterns with varying accessibility measures. The results are aggregated by vehicle body type for easier understanding.

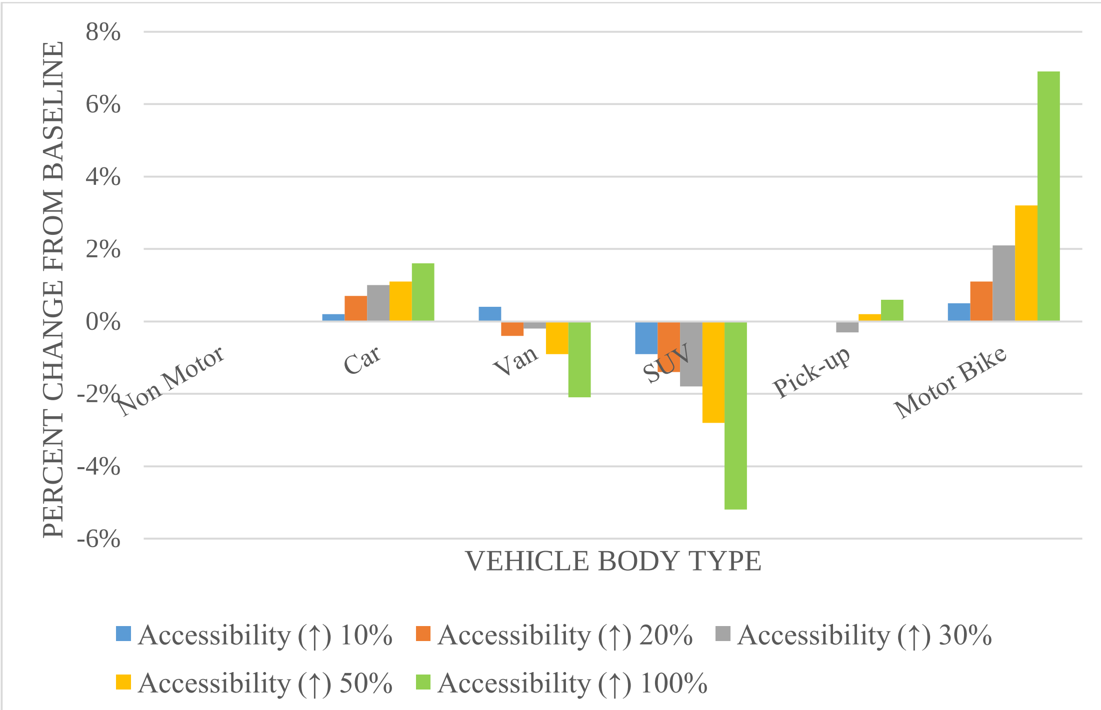

Figure 5.18. Change in vehicle ownership patterns in response to accessibility enhancement.

In general, the model system provides intuitive predictions for changes in vehicle ownership and utilization patterns in response to increasing accessibility measures. It was found that increasing accessibility positively influences ownership of smaller vehicles (such as cars) while the percent of households owning larger vehicle types (such as SUVs) gradually decreases. The percentage of households owning vans drops as well, but the changes are not as large in magnitude as that of SUVs, suggesting that households are more inclined to hold onto multipurpose vehicles in their fleet than the ones that are used mostly for luxury travel. Percent of households owning pick-up trucks remains largely unchanged.

The reason for this might be the type of TAZs in which households owning pick-up trucks usually reside in. It was observed from the results of the vehicle fleet mix model that households residing in rural localities usually tend to own pick-up trucks more. Such TAZs do not have a lot of regional employment accessible within 10 minutes of auto travel to begin with. And for this reason, even doubling the percent of regional employment does not have a significant effect on the ownership of this specific vehicle body type.

The changes in the vehicle ownership patterns with increasing accessibility are largely consistent with the notion that with increasing accessibility, households need to drive smaller distances to fulfill their daily travel needs (going to a grocery store, a mover etc.). Similar studies have found that larger vehicles are preferred for long distance travel (Konduri et al., 2011) and are less preferred in light of increased accessibility. The percent of households owning motorbikes heavily increases with increasing accessibility. While this is an intuitive finding in the sense that such vehicles are more convenient to make short trips, the magnitude of this change should be interpreted with caution. The changes are amplified owing to a lower level of motorbike ownership in the baseline conditions (only 190 households owned a motorbike). Even an increase of this category by 6% from the baseline only translates to 12 more motorbikes.

The average annual mileage patterns in the dataset with varying accessibility levels is shown in Figure 5.19. The average annual mileage is computed by summing up the mileage attributed to a vehicle type and dividing it by the total number of vehicles in the fleet (for each scenario). As expected, the average annual mileage value for all motorized body types decreases as accessibility increases. Car mileage gradually decreases with increasing accessibility. This is an interesting finding in that, though the market share of this body type has seen an increase with increasing accessibility, there is an associated decrease in the usage of cars. This translates to the convenience of owning smaller vehicles and driving them to lower degrees with increasing accessibility.

Vans and SUV categories also show a decreasing trend in annual mileage patterns though the pattern is not as consistent as in the case of cars. One plausible reason for this might be that as the number of vehicles in the fleet drops, even a slight change in the denominator in the calculation of average annual mileage might contribute to a more modest increase/decrease in per vehicle mileage. A closer look at both the market share and average mileage graphs for the van and SUV categories reveals that decrease in market share for these categories is more pronounced than the decrease in average annual mileage. It is possible that while fewer households own these vehicles in light of increased accessibility, the households who own such vehicles continue to drive these vehicles (on a per vehicle basis) for about the same number of miles. This points to the households who own and use larger vehicle type for long distance travel; such usage is therefore not impacted by changes in local accessibility. The non-motorized vehicle on the other hand shows a consistent increase in average annual mileage with increasing accessibility. This change is readily explained by the fact the decreased mileage consumption of the motorized alternatives translates to a corresponding increase in non-motorized mileage. Again, the percent increase in non-motorized mileage seems amplified due the smaller magnitude of non-motorized mileage in the baseline scenario. The increase in non-motorized mileage in the extreme accessibility scenario is about 25 miles per household per year (6% increase). While this may not seem all that significant, it should be noted that this is complemented by a corresponding decrease in motorized mileage. For the dataset under consideration this comes out to about 100,000 lesser vehicle miles driven per year. An increase in accessibility is in general associated with increased levels of walking and bicycling (Ewing and Cervero, 2001; Krizek, 2003) synonymous to the travel characteristics of mixed use urban neighborhoods.

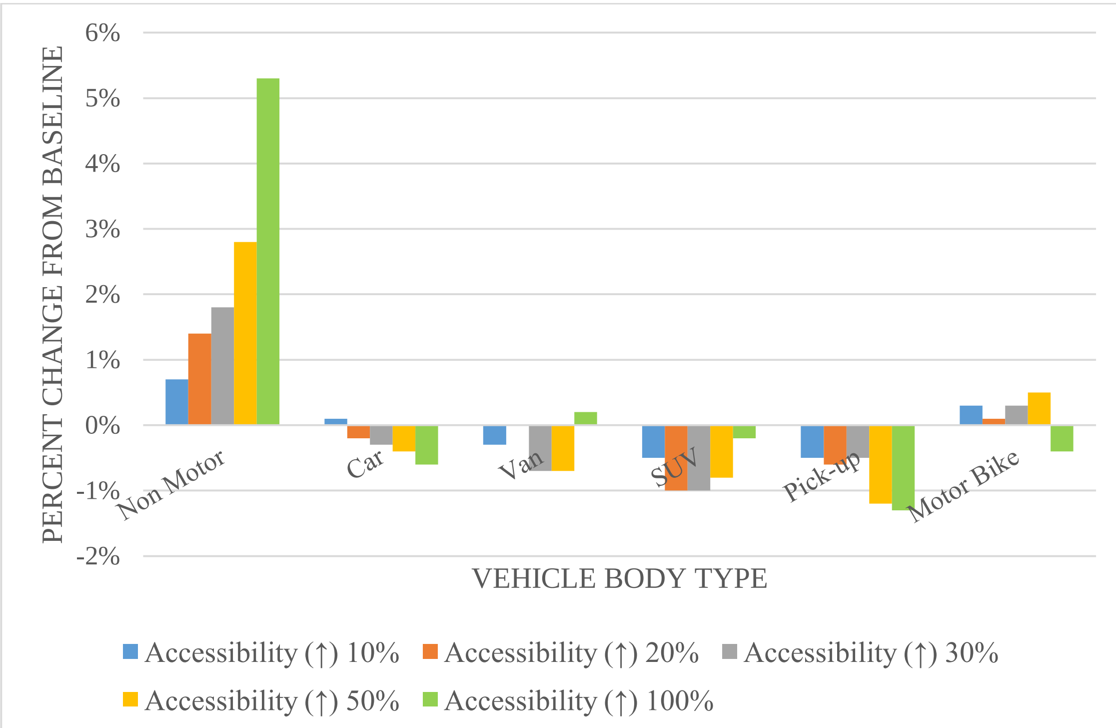

Figure 5.19. Change in annual mileage consumption patterns in response to accessibility enhancement.

In addition to replicating the observed fleet composition patterns quite well in the base year, the model system is found to respond in a behaviorally intuitive way to changes in inputs provided. This reinforces the confidence in using the developed model system as a plug-in fleet composition module to any of the existing activity-based models in practice. For ease of integration, the model system is completely coded on open source coding platform ‘R’. A brief description of the fleet composition package developed as a part of this effort is provided in the next section.

Vehicle Fleet Composition Package

The vehicle fleet composition package is developed on open source coding platform ‘R’ for ease of integration to any exciting activity-based model systems. The system takes socio-economic data as input and outputs the fleet composition for every household in the dataset that meets set tolerance criteria. There a two parts to the fleet composition model package

i. *Model Estimation:* This part consists of all the codes required for estimating various components of the vehicle fleet composition model system. This includes codes for estimating MNL, OP, MDCEV and transformed linear regression models. The intent for providing estimation codes is so that the modelers are provided with the ability to readily update the models with newly available data in the future.

ii. *Model Application:* This part consists of the codes required to apply the fleet composition model system for any given dataset. The application of each of the model components is coded following the logic explained earlier in this section. Figure 5.20 presents a screenshot of a sample application (in an add-on package called R-studio) of the model system. The top left window houses the code for estimation/application of the model system while the bottom left window prints the results for each of the components so that the analyst can perform intermediate checks as necessary. Any graphical results for comparing observed and predicted fleet composition patterns is outputted to the bottom right window. The application file has two main components that control and run the code.

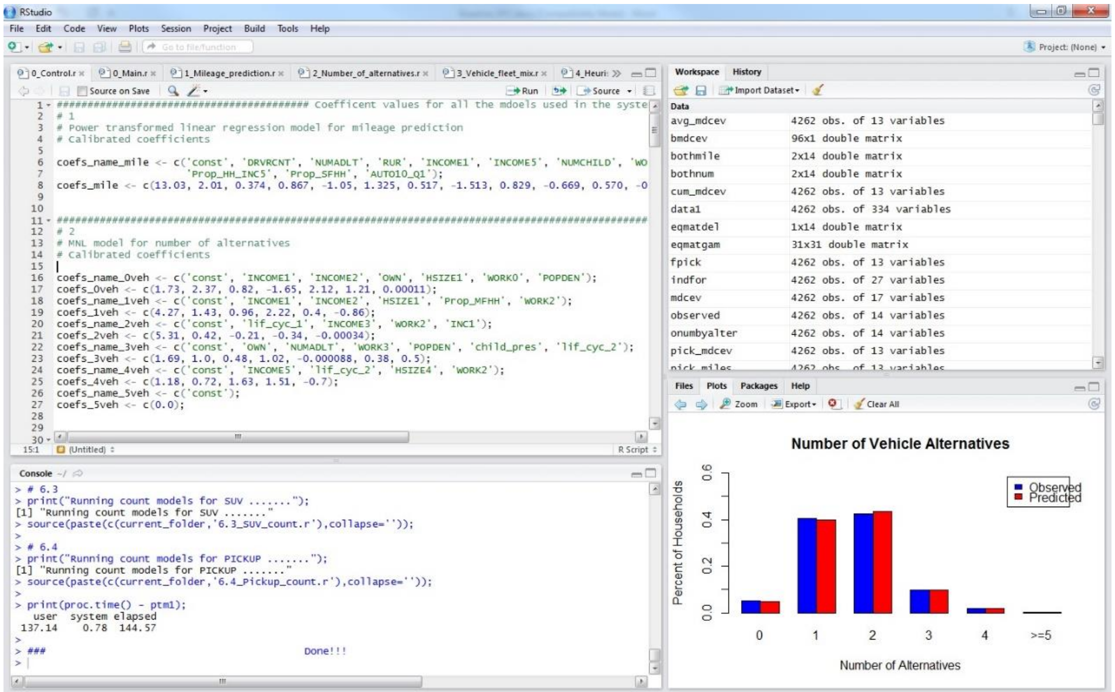

Figure 5.20. Vehicle fleet composition model system in ‘R’.

- Control file: This file contains the coefficients of all the estimated model components of the vehicle fleet composition package. If any of the models are re-estimated with new data, The new coefficients need to be updated in this control file for them to take effect and impact the fleet composition patterns.
- Main file: This files reads the socio-economic data provided and applies all of the model components in sequence. An output file is written with fleet composition as well as mileage allocation for each of the households in the dataset.

Summary and Conclusions

The vehicle fleet composition simulator developed as part of this research effort is quite comprehensive and shows promise in depicting the snapshot of vehicle fleet composition in the observed data along multiple dimensions. The model system developed is parsimonious in the sense that number of model components are kept to a minimum, yet quite effective in predicting vehicle ownership patterns accurately. The model system takes any horizon year data comprising of the socio-demographic characteristics of the households as well as built environment characteristics of household's residential location and predicts the vehicle owned by a household classified by body type, age and count. The model system is tested for its sensitivity to changing land use characteristics and it provided logically intuitive results.

A shortcoming of the model system that future research should focus on is including a vehicle turnover module in addition to a fleet composition module that can determine the vehicular transactions (adding a new vehicle, selling a vehicle, scrapping etc.) at the household level over a period of time. Another enhancement to the model system is joint estimation of the body type and count components in the proposed framework. These two dimensions in a household's vehicle fleet are inextricably linked and hence arises the necessity to tie a count model to the multiple discrete-continuous model so that counts of vehicles within each type may be accurately predicted in a joint modeling framework. Next section discusses the model estimation results of a joint household vehicle fleet composition and count model system intended to enhance the fleet composition simulator proposed in this chapter.

CHAPTER 6

INTEGRATED MODEL OF FLEET COMPOSITION AND COUNT

The material in this chapter is drawn substantially from the following paper, accepted for publication:

Garikapati, V. M., R. Sidharthan, R. M. Pendyala, and C. R. Bhat, "Characterizing Household Vehicle Fleet Composition and Count by Type in an Integrated Modeling Framework," forthcoming, *Transportation Research Record.*

This chapter presents the results of a proposed model system that could potentially replace several components of the vehicle fleet composition model system with an integrated model design framework. Previous chapter presented the model estimation results of a vehicle fleet composition model system proposed as a part of this research effort. One of the key components in the model system is an MDCEV model of vehicle fleet mix, which predicts the array of vehicles owned by a household. Although the MDCEV modeling methodology constitutes a promising development in the modeling of vehicle fleet composition and utilization, it is not without its limitations. One of the key limitations of MDCEV model is that the model does not return the exact count of vehicles that a household owns within each vehicle type category. Suppose a vehicle type category is defined by a combination of body type and age group as “cars 0-5 years old”. While the MDCEV model is able to indicate whether a household consumes (owns) cars 0-5 years old and the total miles that vehicle(s) in that category are driven (utilized), the model is not able to return the exact count of vehicles within the category. For this reason, separate count models are estimated and tied to the MDCEV model system to predict the exact count of vehicles for each vehicle alternative predicted by the MDCEV model.

Though this methodology seems to work reasonably well in the context of the current study, it was felt prudent to enhance the framework to overcome this problem. One way is to define the vehicle type categories in such fine disaggregation that it is virtually impossible for a household to own multiple vehicles in any of the categories. However, this may lead to the definition of a prohibitively large number of discrete alternatives in the MDCEV model. There is, essentially, a critical need for the ability to tie a count model to the multiple discrete-continuous framework so that counts of vehicles within each type may be accurately predicted. In addition to this key limitation, the MDCEV model has drawbacks similar to those of the traditional single discrete choice multinomial logit model including violations of the IIA property in the presence of correlated alternatives and the inability to reflect random taste variations in the behavioral choice phenomenon under investigation.

To overcome these limitations of the MDCEV model, Bhat et al. (2013) recently formulated and developed a multiple discrete-continuous probit (MDCP) model that can be tied together with a multivariate count model in an integrated modeling framework. Just as the multinomial probit (MNP) model offers a methodology to overcome the limitations of the logit model, the MDCP model offers a methodology to overcome the limitations of the MDCEV model. The joint MDCP-multivariate count modeling methodology (Bhat et al., 2013) is applied in this research effort to model vehicle fleet composition and utilization, and the number of vehicles (vehicle count) within each vehicle type alternative, so that the fleet mix of a household can be characterized in its entirety. This methodology is proposed to replace the MDCEV and count components with a joint modeling framework in future incarnations of vehicle fleet composition model system. The joint model system is estimated on the same data used in previous chapter, the 2008-2009 National Household Travel Survey sample drawn from the Greater Phoenix metropolitan area in Arizona.

Modeling Methodology

This section presents a brief overview of the multiple discrete-continuous probit (MDCP) – multivariate count (MC) modeling methodology employed in this effort. The complete details of the model formulation and methodology are provided in Bhat et al. (2013) and hence only a brief synopsis is provided here.

The use of the MDCP model in the current effort, rather than the multiple discrete-continuous extreme value (MDCEV) model (Bhat, 2005; 2008), is motivated by the need to tie the multiple discrete-continuous (MDC) model component (which caters to modeling the fleet composition dimension) with the multivariate count (MC) model (which handles the number of vehicles within each vehicle class dimension). For the MC model, a latent variable representation with normal error terms is used, and this facilitates the linkage with the MDCP model which is also based on a multivariate normal characterization of the error distribution. The model components are described further in this section.

The multiple discrete-continuous probit (MDCP) model. The utility equation proposed by Bhat (2008), where a consumer maximizes his/her utility subject to a binding budget constraint is:

$$\max U(\mathbf{x}) = \frac{1}{\alpha_1}\psi_1(x_1+\gamma_1)^{\alpha_1} + \sum_{k=2}^{K}\frac{\gamma_k}{\alpha_k}\psi_k\left(\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k}-1\right) \tag{6.1}$$

$$s.t. \qquad \sum_{k=1}^{K} p_k x_k = E \ ,$$

where $x \geq 0$ is the consumption quantity (vector of dimension K×1 with elements x_k), and γ_k, α_k, and ψ_k are parameters associated with good *k*. In the linear budget constraint, E is the total expenditure (or income) of the consumer $(E > 0)$, and p_k is the unit price of good *k* as experienced by the consumer. The utility function form in Equation (6.1) assumes that there is an essential outside good consumed by all behavioral units. α_k $(\alpha_k \leq 1)$ and γ_k capture satiation effects and hence it is difficult to disentangle and uniquely identify the effects of both parameter vectors. Bhat (2008) suggests estimating both a γ-profile and α -profile model specification (i.e., specifications in which only one of the parameter vector is free to be estimated, and the other vector is restricted) and choose the one that fits the data best. In addition to explaining satiation effects, γ_k also enables corner solutions (zero consumption) for alternatives, and hence is often preferred in empirical application contexts. ψ_k $(\psi_k > 0)$ represents the stochastic baseline marginal utility; it is the marginal utility at the point of zero consumption. To complete the model structure, stochasticity is added by parameterizing the baseline utility as follows (see Bhat, 2008 for a detailed discussion):

$$\psi_k = \exp(\beta' z_k + \xi_k), \tag{6.2}$$

where z_k is a *D*-dimensional column vector of attributes that characterize good *k*, β is a corresponding vector of coefficients (of dimension *D*×1), and ξ_k captures the idiosyncratic (unobserved) characteristics that impact the baseline utility of good *k*. Bhat et al. (2013) assumes that the error terms ξ_k are multivariate normally distributed across goods *k*: $\xi = (\xi_1, \xi_2, ..., \xi_K)' \sim MVN_K(\mathbf{0_K}, \mathbf{\Lambda})$, where $MVN_K(\mathbf{0_K}, \mathbf{\Lambda})$ indicates a *K*-variate normal distribution with a mean vector of zeros denoted by $\mathbf{0_K}$ and a covariance matrix $\mathbf{\Lambda}$.

The multivariate count (MC) model. Let y_k be the index for the count (say, of vehicles) for discrete alternative *k*, and let l_k be the actual count value observed for the alternative. Castro et al. (2012) recast the count model for each discrete alternative using a special case of the generalized ordered-response probit (GORP) model structure as follows:

$$y_k^* = \eta_k, \; y_k = l_k \; \text{ if } \; \psi_{k,l_k-1} < y_k^* < \psi_{k,l_k}, \; l_k \in \{0,1,2,......\} \;, \tag{6.3}$$

$$\psi_{k,l_k} = f_{k,l_k}(s) = \Phi^{-1}\left[e^{-\lambda_k} \sum_{r=0}^{l_k} \left(\frac{\lambda_k^r}{r!} \right) \right], \text{ where } \lambda_k = e^{\boldsymbol{\varsigma}_k' \mathbf{s_k}}.$$

In the above equation, y_k^* is a latent continuous stochastic propensity variable associated with alternative *k* that maps into the observed count l_k through the ψ_k vector, which is itself a vertically stacked column vector of thresholds $(\psi_{k,0}, \psi_{k,1}, \psi_{k,2}, ...,)'$. This variable, which is equated to η_k in the GORP formulation above, is a standard normal random error term. $\boldsymbol{\varsigma_k}$ is a vector of parameters (of dimension $\tilde{C} \times 1$) corresponding to the conformable vector of observables s_k (including a constant).

The η_k terms may be correlated across different alternatives because of unobserved factors. Formally, define $\eta = (\eta_1, \eta_2, \eta_3, \ldots, \eta_K)'$. Then η is assumed to be multivariate standard normally distributed: $\eta \sim MVN_K(0_k, \Gamma)$, where Γ is a correlation matrix.

Joint model system and estimation approach. An important feature of the proposed joint model system is that y_k (the count corresponding to discrete *k*) is observed only if there is some positive consumption of the alternative *k* as determined in the MDC model. That is, y_k is observed only if $x_k^* > 0$, and $y_k > 0$ in this case (y_k is not observed if $x_k^* = 0$). Thus, the proposed model resembles the hurdle model used in the count literature, albeit with the flexibility that the error components of the MDC model (ξ) and the MC model (η) can be correlated. As a result, the estimation approach involves the joint estimation of the MDC and MC model components. For details on the derivation of the likelihood expression, and the estimation procedures, please see Bhat et al. (2014).

Data

The data set used in this study is derived from the 2008-2009 National Household Travel Survey (NHTS), which is a survey of the nation's travel behavior conducted by the US Department of Transportation on a periodic basis. In the 2008-2009 version of the survey, individual jurisdictions were provided the option to purchase additional samples for their region to aid in model development and travel behavior analysis at the local and regional level. The Maricopa Association of Governments (MAG), the planning agency for the Greater Phoenix metropolitan area, purchased more than 4,400 such add-on sample households, thus obtaining a large sample household travel survey data set that could be used for model development and estimation purposes.

Each respondent household was geolocated within a traffic analysis zone. Using secondary traffic analysis zone (TAZ) and network skim data provided by MAG, the data set is augmented with an extensive set of built environment and accessibility variables. The built environment variables characterized the density and development patterns within the residential location TAZ of each household. The accessibility variables served as measures of the amount of employment in the region that could be accessed from the household’s residential TAZ within certain travel time bands by auto (10 and 30 minute bands) and transit modes (30 and 60 minute bands). As built environment and accessibility variables are likely to be important predictors of vehicle fleet composition and utilization, it was considered important to augment the travel survey data set with such secondary variables.

Vehicle type choice was characterized by five distinct body type alternatives, namely, car, van, sport utility vehicle (SUV), pick-up truck, and motorbike. Further disaggregation of vehicle type alternatives can be done. For example, it is possible to consider a disaggregation of vehicle body types by age category and fuel type category. While such a disaggregation of vehicle type classification is appealing, the number of households with non-zero consumption in the various categories could easily become too thin to support model estimation of such a joint model system. Moreover, with the inclusion of a multivariate count model within the modeling framework, it is not necessary to try and create a disaggregate categorization where households consume (choose to own) only one vehicle. An annual mileage value (continuous dimension) is associated with each vehicle in the estimation data file. Future work in this area should focus on enhancing the model specification to include a couple of vintage categories (perhaps classifying old and new vehicles) within each bod type.

In addition to the five vehicle choices, an outside good that is consumed by all households is introduced in the choice set to account for zero-vehicle households. This outside good is the non-motorized vehicle mileage. All households have to walk (and/or bike) for at least some non-zero distance over the course of an entire year. For households that report walk and bicycle trips in the survey, the reported non-motorized distance is scaled up to compute an annual non-motorized vehicle mileage. For households that report absolutely no walk and bicycle trips in the survey, a value for this consumption is estimated as 0.5 miles/person/day x 365 days/year x household size. This approximation is found to be reasonable and model parameter estimates are robust to alternative mileage computation schemes for the outside good (Vyas et al., 2012). The result of the exercise is creation of a data set where every household has six alternatives, one of which is consumed by each and every household.

The data set was subjected to an extensive quality check and cleaning process to ensure that the data would be able to support the model estimation effort. The final cleaned data set includes 4,262 households owning 7,785 vehicles. The socio-economic and demographic characteristics of the sample are same as the ones provided in the last chapter (see Table 5.1) and are hence not repeated here. On average, households owned 1.95 vehicles per household. Average household size is 2.43 persons per household. There are 1.9 adults per household, and just about one worker per household (on average). There are 1.83 drivers and 0.53 children per household. A vast majority (95.8 percent) live in single family dwelling units and just about 85 percent of the respondent sample owns the home in which they live. An examination of the income distribution shows that 47 percent of the households have incomes that fall within the band of $25,000 to $75,000 per year. About one-fifth of the households have incomes greater than $100,000.

Table 6.1 presents the vehicle ownership profile in the survey sample. The average age of pick-up trucks is larger than other vehicles in the fleet. Sport utility vehicles tend to be newer relative to the other vehicle types. A majority of the vehicles (3,997 of 7,785) are cars. The number of vans is less than 10 percent of all vehicles at just 635. Pick-up trucks and sport utility vehicles are prevalent in larger numbers than vans. The number of motorbikes is rather modest at 240. Panel B shows the age distribution for the vehicles in each body class. An examination of the age distribution suggests that SUVs tend to be newer vehicles, while cars and pick-up trucks tend to be older vehicles in the fleet.

Table 6.1

Vehicle Fleet Mix and Mileage Characteristics

Panel A. Vehicle Body Type					
	Car	**Van**	**SUV**	**Pick-up**	**Motorbike**
Average Age	8.55	7.46	6.52	9.52	9.21
Average Mileage	10204.41	11317.66	11296.57	10722.98	3838.92
Number of Vehicles	3,997	635	1,537	1,376	240
Panel B. Vehicle Type vs Age					
Age					
0 - 5 Years	42.00%	40.90%	52.40%	34.20%	43.30%
6 - 11 Years	35.20%	44.30%	35.30%	39.50%	35.40%
≥ 12 Years	22.80%	14.80%	12.20%	26.30%	21.30%
Total	100.00%	100.00%	99.90%	100.00%	100.00%

Model Estimation Results

This section provides a summary of the model estimation results. The model estimation effort involved a systematic attempt at including explanatory variables such that the model offered behaviorally intuitive and statistically significant interpretations. Some variables were retained in the model specification even if they were statistically insignificant for considerations of behavioral sensitivity and intuitiveness.

MDCP component. Estimation results for the MDCP component are furnished in Tables 6.2 and 6.3. A γ-profile of MDCP model was estimated with one outside good. A baseline utility equation is estimated for each vehicle type. The values of the coefficient estimates indicate whether a certain characteristic or variable positively or negatively contributes towards ownership (consumption) of that vehicle type. Cars tend to be owned by smaller households evidenced by the negative coefficients on child presence and household size. It is to be expected that larger households, and households with children, have a higher baseline preference to own SUVs and vans – this is indeed supported by the model estimation results as single person households show a negative propensity to own larger vehicle types.

Households at all income levels show a proclivity towards owning SUVs (presumably at different vintage levels), with the highest positive coefficient exhibited by the high income household category. Retired households with no children and those renting their single family housing unit have a lower preference to own SUVs. Home ownership, on the other hand, is positively associated with SUV ownership. It is found that high income households shy away from owning vans, a finding that is consistent with expectations. Higher income household are more likely to prefer luxury vehicles and vans generally do not come in luxury models.

It is found that multi-worker households are less likely to own vans, a finding that merits further investigation. Again, it is likely that these households are higher income households who prefer to own luxury vehicles. In the case of pick-up trucks, high income households tend to show a lower inclination to own trucks. Retired households, single person households, very large households, and households in single family dwelling units all have a low preference to own pick-up trucks. As pick-up trucks tend to be more specialized and likely to be used as utility and work-related vehicles to haul cargo, it is not surprising that there is a general disinclination to own pick-up trucks across the board. Households that own their residence and households with two workers show a positive inclination to own trucks.

Single person households and large households are less likely to own motorbikes (similar to pick-up trucks). It is likely that motorbikes and pick-up trucks are owned by two or three person households (not single person and not too large). Households in all come income categories show an inclination to own motorbikes with those in the middle range exhibiting larger coefficients. Retired households, as expected, are less likely to own motorbikes. Households in rural area, households that own their single family dwelling unit, and two worker households are more likely to own motorbikes (again similar to pick-up trucks). As both pick-up trucks and motorbikes tend to be rather specialized vehicles, they appear to exhibit common traits.

Table 6.2.

MDC Component Estimation Results: Significant Parameters in Baseline Utility

Vehicle Body Type	Explanatory Variables	Coefficient	t-statistic
Car	Constant	1.20	9.29
	Child presence	-0.13	-2.82
	Household size	-0.19	-12.67
	High income household ($75,000 - $99,999)	0.06	1.50
	Retired household (one/two person) with no children	-0.14	-3.81
	Single family housing unit (owned)	0.29	7.03
	Household residing in TAZ with low density	0.60	4.88
	Household residing in TAZ with medium density	0.63	5.18
	Household residing in TAZ with high density	0.61	4.97
Van	Constant	0.31	2.53
	Household size = 1	-0.11	-1.74
	Highest income household (≥ $100,000)	-0.16	-2.97
	Single family housing unit	-0.26	-1.84
	Single family housing unit (owned)	0.39	5.33
	Two worker household	-0.12	-2.45
	Three worker household	-0.34	-3.58
	Household residing in TAZ with medium density	0.08	2.05
SUV	Constant	0.93	8.94
	Household size = 1	-0.16	-13.71
	Lowest income household (< $25,000)	0.12	2.11
	Low income household ($25,000 - $49,999)	0.36	6.26
	Medium income household ($50,000 - $74,999)	0.29	4.59
	High income household ($75,000 - $99,999)	0.43	7.17
	Retired household (one/two person) with no children	-0.22	-5.59
	Single family housing unit	-0.21	-2.05
	Single family housing unit (owned)	0.51	8.20
	Three worker household	-0.10	-1.12
	Proportion of households in the lowest income quintile	-0.39	-3.38
PickUp	Constant	1.24	12.10
	Child presence	-0.17	-3.04
	Household size	-0.14	-7.14
	Household size = 1	-0.24	-4.33
	Highest income household (≥ $100,000)	-0.13	-3.38
	Retired household (one/two person) with no children	-0.30	-6.96
	Single family housing unit	-0.42	-4.05
	Single family housing unit (owned)	0.58	9.12
	Two worker household	0.13	3.52
Motorbike	Constant	-0.04	-0.21
	Household size	-0.27	-8.93
	Household size = 1	-0.24	-2.15
	Low income household ($25,000 - $49,999)	0.24	1.83
	Medium income household ($50,000 - $74,999)	0.42	3.04
	High income household ($75,000 - $99,999)	0.39	2.78
	Highest income household (≥ $100,000)	0.24	1.68
	Retired household (one/two person) with no children	-0.38	-4.83
	Household in rural area	0.24	3.80
	Single family housing unit (owned)	0.49	4.28
	Two worker household	0.11	1.62

The translation parameters from the model estimation result are furnished in the Table 6.3. Translation parameters represent the diminishing marginal returns with increasing consumption of an alternative. Van has the highest translation parameter, suggesting that it tends to be driven most. Vans tend to be multipurpose family vehicles, and are used for long distance family vacation trips. Thus this finding is consistent with expectations. SUV has the next highest translation parameter, once again consistent with expectations. These vehicles are more likely to be driven longer distances. Cars and motorbikes show lower translation parameters, presumably because these vehicles are driven shorter distances.

Table 6.3

MDC Component Model Estimation Results: Translation Parameters

Vehicle Body Type	Coefficient	t-statistic
Non-motorized vehicle	0	0
Car	22.75	74.8
Van	73.11	27.89
SUV	36.54	53.41
Pick-Up	29.14	52.43
Motor	10.9	26.66

Multivariate count component. Estimation results for the multivariate count model component are furnished in Table 6.4. The parameters in table refer to the elements of the ς_k vector (k=1,2,3,4) embedded in the threshold functions.

Table 6.4

Multivariate Count Component Model Estimation Results

Vehicle Body Type	Explanatory Variables	Coefficient	t-statistic
Car	Constant	-0.12	-0.66
	Child presence	-0.4	-5.88
	Household size = 1	-1.26	-12.37
	Household size = 2	-0.29	-4.75
	Zero worker household	-0.23	-3.38
	Three or more worker household	0.43	6.08
	Lowest income household (< $25,000)	-0.14	-1.65
	High income household ($75,000 - $99,999)	0.18	2.85
	Highest income household (≥ $100,000)	0.27	5.11
	Proportion of single family housing units in the TAZ	0.28	2.31
	Retired household (one/two person) with no children	0.07	1.14
	Household in rural area	-0.07	-1.29
	Single family housing unit	-0.51	-3
	Single family housing unit (owned)	0.36	3.13
Van	Constant	-2.49	-12.94
	Three or more worker Household	0.74	1.45
SUV	Constant	-2.25	-6.47
	Household size = 1	-2.02	-3.73
	Three or more worker household	0.47	3.82
	Lowest income household (< $25,000)	-0.44	-1.95
	Low income household ($25,000 - $49,999)	-0.66	-3.81
	Highest income household (≥ $100,000)	0.51	5.74
	Proportion of single family housing units in the TAZ	0.72	2.91
	Retired household (one/two person) with no children	-0.34	-3.16
	Household in rural area	-0.13	-1.47
	Single family housing unit (owned)	0.56	2.2
Pick-up	Constant	0.23	1.25
	Child presence	-0.51	-3.99
	Household size = 1	-1.24	-5.06
	Household size = 2	-0.58	-4.64
	Zero worker Household	-0.19	-1.15
	One worker household	-0.29	-3.01
	Lowest income household (< $25,000)	-0.45	-2.36
	Retired household (one/two person) with no children	-0.41	-2.85
	Household in rural area	0.14	1.52
	Single family housing unit	-1.46	-4.52
	Single family housing unit (owned)	0.71	2.53

The constant coefficient in the ς_k vector does not have any substantive interpretation. For the other variables, a positive coefficient in the ς_k vector for a specific vehicle type *k* shifts all the thresholds toward the left of the count propensity scale for that vehicle type, which has the effect of reducing the probability of one vehicle of type *k* if the household decides to own vehicles (which is determined by the MDC component). That is, the household has a higher probability of owning multiple vehicles of type *k*, should it hold any vehicles at all of that type. On the other hand, a negative coefficient shifts all the thresholds toward the right of the count propensity scale, which has the effect of increasing the probability of one vehicle of type *k* (or decreasing the probability of multiple vehicles of type *k*), conditional on owning a vehicle of type *k* .

It is found that households with children are less likely to own multiple cars or pick-up trucks. This is consistent with expectations as such households are likely to own larger SUV and van type vehicles, thus resulting in a lower propensity to own cars and pick-up trucks. Single person households are less likely to own more than one vehicle of any type. This is behaviorally intuitive as single person households would not generally own more than one vehicle. Households with three or more workers are more likely to own multiple cars, presumably because these households need multiple cars to meet their commuting needs. These household may also have a higher income, making it possible for them to own multiple cars. They appear to choose multiple cars or multiple vans, presumably because these vehicles offer better fuel economy.

An examination of the income dummy variables shows that lowest income households are less likely to own multiple cars, pick-up trucks, or SUVs, while high income households are more likely to own multiple vehicles – particularly in the car and SUV categories. These findings are consistent with expectations. Households in rural areas are more likely to own pick-up trucks and less likely to own multiple cars or SUVs. As these households tend to own a pick-up truck, they are likely to own just one (if any) of the other vehicle types. Retired households are less likely to own multiple SUVs, a finding that is consistent with expectations. These households would not have the need for multiple large vehicles; likewise, these households are less likely to own multiple pick-up trucks. On the other hand, retired households are more likely to own multiple cars. Households that own their single family dwelling unit are more likely to own SUVs. Households of small size (one or two person) are less likely to own multiple pick-up trucks; as these tend to be specialized vehicles, it is unlikely that small households would need to own multiple vehicles of this category. The negative coefficients on these variables are indicative of this.

Table 6.5

Joint Model System: Error Correlation Matrix

	Panel A					Panel B					
	Car	**Van**	**SUV**	**PickUp**	**Motor**	**NM**	**Car**	**Van**	**SUV**	**PickUp**	**Motor**
Car	1	0.5	0.5	0.5	0.5	0	0.454	-0.257	-0.311	-0.231	0
Van	0.5	1	0.5	0.5	0.5	0	-0.202	0.296	-0.192	-0.198	0
SUV	0.5	0.5	1	0.5	0.5	0	-0.314	-0.092	0.468	-0.165	0
PickUp	0.5	0.5	0.5	1	0.5	0	-0.216	-0.144	-0.203	0.506	0
Motor	0.5	0.5	0.5	0.5	1	0	0.055	-0.305	0.074	0.057	0
	Panel C					**Panel D**					
NM	0	0	0	0	0	1	0	0	0	0	0
Car	0.454	-0.202	-0.314	-0.216	0.055	0	1	-0.394	-0.34	-0.182	0
Van	-0.257	0.296	-0.092	-0.144	-0.305	0	-0.394	1	-0.066	-0.131	0
SUV	-0.311	-0.192	0.468	-0.203	0.074	0	-0.34	-0.066	1	-0.083	0
PickUp	-0.231	-0.198	-0.165	0.506	0.057	0	-0.182	-0.131	-0.083	1	0
Motor	0	0	0	0	0	0	0	0	0	0	1

Table 6.5 presents the error correlation matrix for the joint model system. The top left matrix (Panel A) refers to the MNP error differences in the MDC component of the joint model and the bottom right matrix (Panel D) corresponds to the error correlation matrix for the count propensities. The off diagonal block of (Panels B and C) the matrix is the covariance matrix between MNP errors and count propensities, which is the main focus

of this effort. All of the error correlations presented in the table are found to statistically significant.

In general, it is found that within body type correlations are positive while cross body type correlations are negative. For example, the error correlation for car across the two model components is positive. This suggests that unobserved factors that contribute to car consumption in the MDCP component also contribute to owning more cars in the MC component. Such positive correlations are seen for all vehicle body types. This is consistent with expectations; it is very likely that unobserved attributes that contribute to greater mileage of a certain vehicle type will also contribute to a higher vehicle count for this class. A household whose members appreciate and desire comfortable and roomy vehicles are likely to choose and drive larger vehicles (such as vans and SUVs), and the same unobserved factors (desire for comfortable and roomy vehicles) will also contribute to such households owning multiple large vehicles. Across vehicle categories, error correlations are generally found to be negative, suggesting that there is an inherent inverse effect across body types. In the above example, the unobserved factors (desire for comfortable and roomy vehicles) are the very same factors that will negatively impact the choice of smaller vehicles such as cars or vehicles with harsher rides such as pick-up trucks. Thus, the correlation between cars and vans (or SUVS and pickup trucks) is negative, both within the model component and across model components.

A rather interesting finding from the table is correlation between the model components for motorbike category. It can be observed from the table that error terms for motorbike have a positive correlation across all body types except vans. The behavioral interpretation of this finding is that ownership and usage of motorbikes (which is a specialized vehicle category, thereby marking its presence across all type of households) is in general positively correlated to ownership and usage of any other body type. Only households owning vans have unobserved factors that portray a disinclination towards owning motorbikes (negative correlation of error terms). Observations in the data as well as significant variables in the model specification suggest that ownership of vans is positively influenced by presence as well as number of children in the household. The indisposition of such households to own and use motorbikes is intuitive. It should however be noted that error correlations for motorbike category are quite modest compared to other vehicle body types.

Model Goodness of Fit and Assessment

Goodness of fit measures of the joint model are furnished in Table 6.6. The model system is found to offer a good fit with the log-likelihood of the final model at convergence equal to -20989.96.

Table 6.6

Joint Model System: Goodness of Fit Measures

Statistic	Value
Log-likelihood of final model at convergence	-20989.97
Degrees of freedom of final model	112
Log-likelihood of base model at convergence	-22191.08
Degrees of freedom of base model	14
Likelihood ratio	2402.23
$\chi^2_{98,0.001}$	146.99

The log-likelihood of the base model with only constants in the baseline utility, translation parameters, and constants in the count models is -22191.08. The likelihood ratio for the estimated model is 2402.23, which is significantly larger than the critical χ^2 value with 98 degrees of freedom at any level of significance. In addition to examining model goodness-of-fit statistics, an assessment of the efficacy of estimating a joint model system (MDCP-MC) was performed. A simple assessment can be made by comparing the fit and indications offered by the joint model against those offered by an independent MDCP-MC model system where error correlations across the discrete-continuous and count components of the model system are ignored. The latter is akin to estimating two model components separately and then applying them in forecast mode in a sequential fashion – first, apply the MDCP model to predict the vehicle fleet mix by body type, and second, given the predictions of this model component, apply the count model for each body type consumed by a household to estimate the number of vehicles owned in each class. An independent MDCP-MC model system was estimated by setting all error correlations equal to zero. The coefficient estimates and goodness of fit statistics were compared between the two models. It was found that the independent model system offered coefficient estimates that were considerably different from those provided by the joint model and the goodness of fit was inferior to the joint model. This comparison offered the first indication that the joint modeling approach is critical to modeling vehicle fleet composition, utilization, and count in a holistic framework.

An examination of the error covariance matrix (Table 6.5) shows that there are a number of significant error correlations across the alternatives in the MDCP model component and the MC (count) model component. The large number of significant error correlations lead to two noteworthy considerations. First, the joint model is capable of accounting for error correlations that may exist across choice dimensions. Ignoring such error correlations, when in fact they exist, will lead to inconsistent parameter estimates unsuitable for forecasting applications. Second, it points to the presence of unobserved factors that affect behavior and yet remain accounted in the model specifications. Qualitative research methods should be employed to identify these factors, and survey designs should be enhanced to measure these variables so that they may be included as observed covariates in the model specifications.

Table 6.7

Comparison of Measures of Fit - Per Household: Log-likelihood by Subsample

Sample details	Number of households	Joint Model	Independent Model
Full Sample	4262	-4.9249	-5.0585
Household Size			
Household size = 1	992	-1.8513	-1.8384
Household size = 2	1830	-4.9422	-5.0938
Household size greater than 2	1440	-7.0204	-7.2320
Household income			
Lowest income household (< $25,000)	759	-2.6603	-2.6880
Low income household ($25,000 - $49,999)	1210	-3.9486	-4.0106
Low income household ($50,000 - $74,999)	800	-5.3207	-5.4750
High income household ($75,000 - $99,999)	640	-6.4789	-6.6743
Highest income household (≥ $100,000)	853	-6.7878	-7.0514
Number of workers in household			
Zero worker Household	1496	-5.8713	-6.0516
One worker household	1597	-4.9777	-5.0945
Two worker household	1011	-6.7623	-7.0083
Three or more worker household	158	-9.2027	-9.6031
Household TAZ density			
Lowest density	16	-8.5189	-8.6719
Household in TAZ with low density	620	-5.4100	-5.5464
Household in TAZ with medium density	2158	-5.1054	-5.2616
Household in TAZ with high density	1468	-4.4155	-4.5146
Single family housing unit			
No	179	-3.4535	-3.5164
Yes	4083	-4.9894	-5.1261
Single family housing unit (owned)			
No	642	-3.0915	-3.1396
Yes	3620	-5.2501	-5.3988
Retired household (one/two person) with no children			
No	2504	-5.7956	-5.9667
Yes	1758	-3.6848	-3.7650
Household in rural area			
No	3570	-4.6936	-4.8194
Yes	692	-6.1182	-6.2921

In addition to an examination of the error correlations, a comparison of the joint and independent model systems was performed by computing the log-likelihood value on a per household basis for a number of subsamples in the survey data set. If the log-likelihood in one model is higher than that in the other model, then the model with the higher log-likelihood may be considered superior from a statistical perspective. If the improvement in log-likelihood per household is seen across all (or nearly all) subsamples, then it indicates that such a model is likely better able to predict vehicle ownership, fleet composition, count, and utilization patterns for all socio-economic and demographic market segments. This comparison is presented in Table 6.7.

The comparison in Table 6.7 suggests that the joint model is consistently performing better than the independent model across all socio-economic and demographic market segments. The log-likelihood value per household is consistently higher (and therefore better) in the joint model relative to the independent model. There is only one subsample for which this does not hold true – household size=1. For single person household subsample, it is found that the independent model system is very marginally better. For every other market segment depicted in the table, the joint model offers a stronger fit as evidenced by the higher likelihood value.

Summary and Conclusions

The motivation for this effort stems from the growing interest in modeling household vehicle fleet composition and utilization behavior so that richer predictions of vehicle fleet mix and miles of travel by vehicle type can inform energy and environmental analysis. Recent work in this domain has focused on the development and application of techniques that recognize the multiple discrete-continuous nature of the vehicle fleet composition and utilization modeling problem. Recent work involving the use of the multiple discrete continuous extreme value (MDCEV) model has provided a promising approach to model vehicle fleet composition and utilization behavior. However, the MDCEV model is not able to offer predictions of the count of vehicles within each vehicle class, thus necessitating the statistically inefficient and behaviorally counter-intuitive stitching of a separate count model system (to the MDCEV model) capable of predicting vehicle counts. Such an approach ignores the presence of possible common unobserved factors affecting both the consumption of alternative vehicle types and the number of vehicles owned within each vehicle type. In order to overcome this limitation and account for such presence of common unobserved factors, this research effort employs a joint model that incorporates a multiple discrete-continuous probit (MDCP) model component and a multivariate count (MC) model that takes the form of the generalized ordered probit model structure. The use of the probits in the two model components allows the use of multivariate normal distribution to characterize the error covariance structure accommodating correlations between the MDCP component (that models vehicle type choice and mileage) and the MC component (that models the number of vehicles or vehicle count within each chosen vehicle class).

The model is estimated on a household travel survey data set of 4,262 households drawn from the Greater Phoenix region of Arizona in the United States. The model system is found to offer plausible parameter estimates with a host of socio-economic, demographic, and built environment variables affecting both the MDCP model of vehicle type choice and mileage, and the MC model of vehicle counts. The model is found to fit the data well, and a comparison of the goodness of fits between the joint model presented and an independent model that ignores error correlations across the choice dimensions shows that the joint model consistently outperforms the independent model system. The comparison involved an examination of the per-household log-likelihood value between the two model systems; the model with the larger log-likelihood value offers the better fit to the data. The joint model is found to offer a better fit for all socio-economic and demographic market segments of interest. In addition, it was found that there were a number of significant error correlations across the two choice dimensions in the joint model. The presence of significant error correlations implies that there are common unobserved factors that affect both the MDC dimension (vehicle type choice and mileage) and the count of vehicles. For example, a person who is fun-seeking and gregarious in nature may like to own and drive sports cars. The unobserved attitudinal trait (being fun-seeking and gregarious) is likely to influence both the mileage (this person will likely drive more miles, thus representing a higher level of vehicle consumption/utilization), and the count of cars (as this individual might purchase additional sports cars that are fun to drive). There are likely to be a number of such attitudinal and contextual factors that are unobserved and yet influence both the multiple discrete continuous and multivariate count components of the model system.

The modeling of vehicle fleet composition, utilization, and counts by vehicle type is critical to performing energy and environmental impact analysis for a variety of policy, market, and technology scenarios. The introduction of vehicle fleet composition and utilization model systems is particularly made possible by the implementation of microsimulation-based activity-based travel demand model systems in practice. By accurately modeling vehicle fleet composition and usage patterns, planning agencies will be able to address energy sustainability and environmental concerns and implement policy actions that promote a more sustainable and energy friendly fleet mix and vehicle utilization pattern in the region. The MDCP-MC model system presented can be used to fill this modeling need. Future work in this domain should focus on including additional explanatory variables to make the model sensitive to policy, pricing, and market/technology changes. The data set used in this study did not support the inclusion of such variables. Household travel survey should be designed to collect such data so that model systems capable of responding to a wide variety of scenarios can be estimated and deployed in practice. Future research efforts also should be aimed at reporting results of model validation and sensitivity analysis to demonstrate the ability of the model system to replicate base year conditions and respond in behaviorally intuitive ways when subjected to changes in input variables. Also, the model system can be enhanced by classifying the vehicle body types into different vintage categories.

CHAPTER 7

TOUR LEVEL VEHICLE TYPE CHOICE

There has been remarkable progress made in the past few decades in the field of activity-based modeling in depicting activity-travel patterns of individuals in a behaviorally realistic way. The traditional trip-based methods model travel as trips going from one zone to another, whereas activity-based models are founded on the behavioral paradigm that travel is 'derived demand' which arises from the necessity of individuals to participate in various activities along the day. This behavioral representation of travel in activity-based models is not only intuitive but also helps in providing accurate outcomes in response to various policy measures. While the cause for travel (activity participation) is depicted quite well in these models, the means (mode) of travel is still represented at the aggregate level. Most of the activity-based modeling systems in research as well as practice still operate at the level of 'mode' to represent personal travel (auto, transit, walk, bike etc.). There is much need for progress in this domain to model 'vehicle type' instead of just the mode used to travel so that emission foot print from personal travel can be estimated accurately. The tour level vehicle type choice modeling framework proposed in this effort aims at developing a practical methodology that can be implemented in an activity-based modeling system to predict the exact type of vehicle (from the fleet of vehicles that the household owns) that a person would choose to make a tour.

The tour level vehicle type choice modeling framework is discussed in detail in Chapter 3. The framework begins with a primary driver allocation model that assigns all the vehicles owned by a household to the drivers in the household. It is assumed that all primary drivers in the household will adhere to the vehicle assigned to them to travel along the day. In case a person in the household who is not a primary driver (has no household vehicle assigned to him/her) needs to embark on a journey, a tour level vehicle type choice model determines which vehicle amongst the household's fleet will be used to make the journey. Tour level vehicle type choice is modeled as a function of tour attributes such as primary purpose of the tour, tour composition, type of tour (solo/joint) etc. Both the primary driver allocation model and the tour level vehicle type choice model are discussed in detail this chapter. For each model, first the data preparation method is discussed along with a description of the data. This is followed by model estimation and validation results for each of the components.

Primary Driver Allocation Model

Data. The data used for estimating primary driver allocation model is from the 2008-09 National Household Travel Survey (NHTS) add-on data for the Greater Phoenix Metropolitan Region. In the NHTS, a random sample of individuals are selected and data is obtained regarding their socio-demographic and travel characteristics. The NHTS provides data segmented into four types of files (NHTS Codebook Browser, 2014).

- Household file: This file contains data regarding the characteristics of the households responding to the survey. Data such as household size, income, number of workers in the household, number of drivers in the household etc. is available from this file.

- Person file: This file contains information regarding the characteristics of all individuals in a household responding to the survey. Data such as age, gender, employment status, education level etc. is available from this file.
- Trip file: This file has data regarding each and every trip made by a person responding to the survey on the most recent travel day. Information such as purpose and duration of the trip, vehicle used, number of people participating in the trip etc. are collected in this file.
- Vehicle file: This file contains data regarding characteristics of all vehicles owned by households responding to the survey. Each vehicle owned by a household is assigned a unique identifier and data regarding the make/model, fuel economy average annual miles put on the vehicle etc. are available from this file. Each vehicle reported in this file has a primary driver assigned to it among the drivers in the household. The person number (unique id) assigned to a vehicle in this file is same as the person id from the person file.

To estimate the primary driver allocation model, person characteristics (from the person file) of the primary driver are added to each vehicle reported in the vehicle file. All the vehicles owned by the household are then added as available choices to the person record. This choice context translates to a situation where a vehicle amongst the household fleet is being assigned to a person based on his/her characteristics (such as age, gender, educational level etc.). The vehicle fleet composition of the household is constructed from the vehicle file for the fleet composition modeling effort described in Chapter 5. In application mode, the vehicle fleet composition model system predicts the array of vehicles owned by the household and provides them as an input to the primary driver allocation model.

After extensive cleaning and analysis of the data, the estimation sample included 6,842 vehicles allocated to 5,992 primary drivers from 3,870 households. An immediate observation from these statistics is that there might be situations where multiple vehicles are assigned to one primary driver in the household, but such occurrences are very less in the data. The tour level vehicle type choice modeling framework assumes that each person in the household is assigned a unique vehicle from the household's fleet. Any unassigned vehicles in the household are assumed to be equally accessible to all the drivers in the household. A brief sketch of the household level socio-demographics of the data set is provided in Table 7.1.

Table 7.1

Primary Driver Allocation Model: Data Description

Characteristic	Mean	Standard Deviation
Households (N = 3870)		
Number of vehicles in the household	2.05	1.013
Number of adults in the household	1.91	0.699
Number of drivers in the household	1.87	0.729
Number of workers in the household	1.02	0.887
Number of children in the household	0.55	1.023
% Households with income < $25,000	15.14%	0.358
% Households with income ≥ $75,000	37.31%	0.483
Person (N = 5992)		
Age of the respondents	54.07	16.113
% of Male respondents	46.50%	0.499
% of Respondents that are workers	60.10%	0.490
% of Respondents with an associate's degree	33.44%	0.472
% of Respondents with an bachelor's degree	25.58%	0.436
% of Respondents with a graduate or professional degree	17.36%	0.378

The characteristics of the dataset used for estimating primary driver allocation model are found to be similar to the characteristics of the dataset used for estimating components of the vehicle fleet composition model system. This ensures consistency between components estimated across different model systems. From the table, it can be observed that number of vehicles to number of drivers ratio of the dataset is just above one which is corroborated by the earlier observation that some individuals are assigned as primary drivers to multiple vehicles in a few household. The vehicle fleet composition module does the job of identifying the body type and age of each vehicle owned by a household. The primary driver allocation model determines which driver in the household will be assigned as a primary driver to each of the vehicles owned by the household. Average age of the primary driver is about 54 years. Though this number appears to be on the higher side, it should be kept in mind that the dataset comprises only of driving age individuals (≥ 16 years). The data set has almost an equal proportion of male and female respondents. About 60% of primary drivers are workers which is an intuitive finding as workers in the household would need a vehicle for commute purpose on a daily basis and hence have a greater proclivity of getting assigned as primary drivers. 76% of the primary drivers have an educational attainment of at least an associate's degree.

Table 7.2 presents fleet composition profile of the dataset used for model estimation which is in line with the fleet composition observed in the estimation of components for vehicle fleet composition model system (see Figure 5.1). The body type and age classification used in vehicle fleet composition framework is carried through to this component to ensure consistency between different frameworks proposed in this research work. It can be observed that majority of vehicles in the dataset are cars, which is consistent with expectation. Within car and SUV body types a consistent ownership pattern was observed where households preferred to own a greater percentage of newer vehicles than older ones. This speaks to the greater turnover rates of these body types. For van and pick-up truck categories however, it was observed that there are a greater proportion of middle aged vehicles (6-11 years) than newer (0-5 years) or older (≥ 12 years) vehicles. All of these patterns are consistent with observations from the fleet composition model dataset (Chapter 5).

Table 7.2

Primary Driver Allocation Model: Vehicle Profile

	Body Type				
Age	**Car (%)**	**Van (%)**	**SUV (%)**	**Pick-Up (%)**	**Motorbike (%)**
0-5 years	41.9	39.9	53.7	33.9	
6-11 years	35.5	45.4	34.6	39.6	100.0
≥ 12 years	22.6	14.7	11.7	26.5	
Total No. of Vehicles	***3518***	***557***	***1338***	***1185***	***244***

Figure 7.1 depicts the cross classification between different vehicle alternatives and gender of the primary driver assigned to the vehicle. A few interesting observations can be made from this figure regarding the primary driver allocation characteristics. Across all body types, females tend to be primary drivers for newer vintages than older ones. The reason for this might be the household's decision to allocate safer vehicles to females in the household or that males are relatively less particular about the vintage classification (old vs. a new vehicle) than females. It can also be observed that the van body type is predominantly allocated to female primary drivers, while males are predominantly assigned as primary drivers to pick-up truck body type. These are intuitive findings and

consistent with general expectation. Females often take up the duty to chauffer children to various activities thereby inherently preferring a roomier vehicle such as a van. Pick-up trucks whose build is considered masculine tend to be preferred more by male individuals in the household.

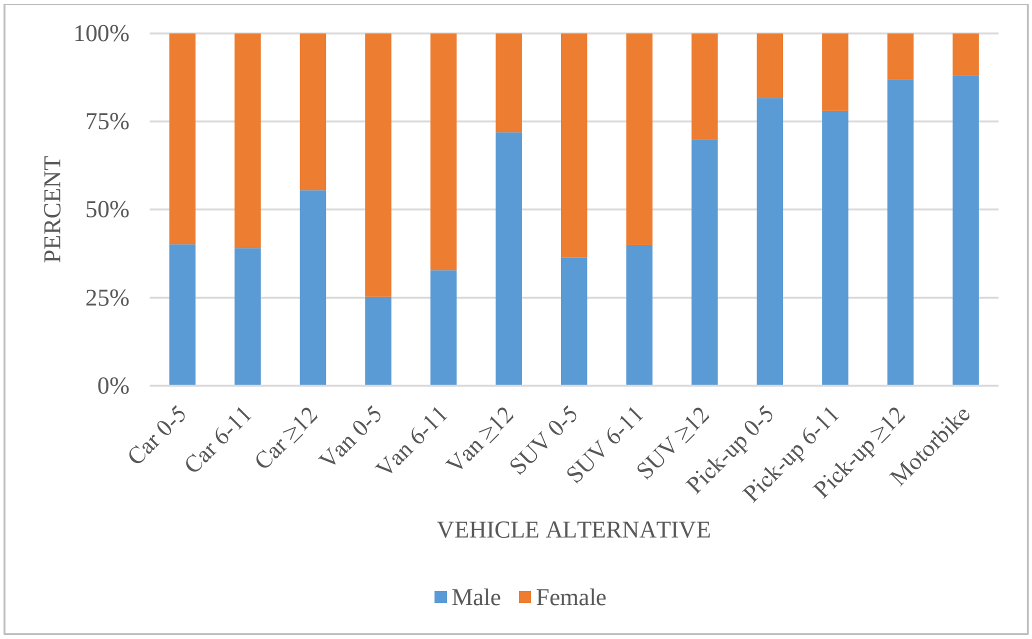

Figure 7.1. Vehicle type by primary driver's gender.

Figure 7.2 presents a similar comparison between workers and non-workers. It was observed that the primary driver profile for worker’s and non-workers is consistent with the person characteristics seen in the data. At the person level, it was found that about 60% of the respondents are workers and from the model estimation dataset also it was observed that 60% on vehicles are assigned to workers, while 40% of the vehicles are assigned to non-workers. From the figure it can be noticed that except for van body type, workers are assigned more as primary drivers to newer vehicle categories. This is understandable, as workers need to use their vehicle for daily commute and hence prefer a newer (reliable) vehicle over an older one. For van body type however, a higher proportion of vehicles have non-workers as primary drivers. This might point to stay-at-home moms/dads who utilize the van primarily to chauffer their kids. Motorbikes, which are usually considered as ‘recreational/hobby’ vehicles are also predominantly assigned to workers. Affordability might a factor driving this observation.

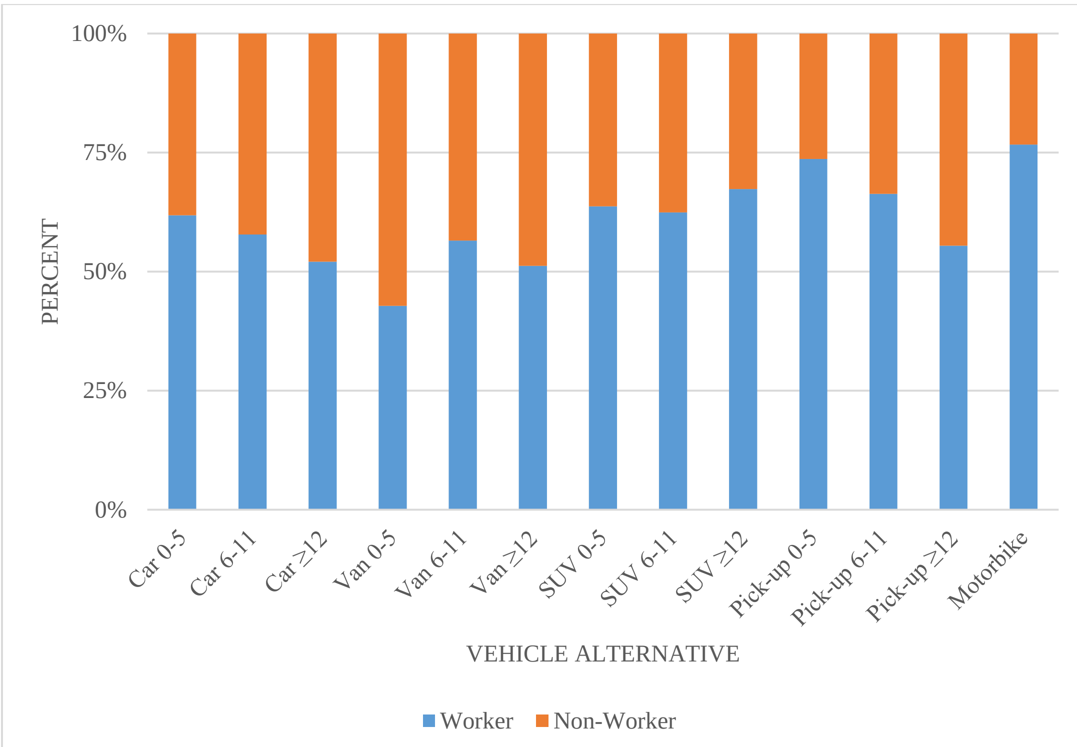

Figure 7.2. Vehicle type by primary driver's work status.

Figure 7.3 presents the results of a cross tabulation between vehicle type and education level of the primary driver. As observed from the descriptive statistics, majority of the primary drivers have an educational level of associates degree or higher. Educational attainment could act as a proxy variable for the occupation level of the individual. It is observed that pick-up truck body type has a majority of primary drivers with an associate’s degree or less, while other body types have an equal proportion of primary drivers with less than a bachelor's degree and primary drivers with a bachelor's degree or more. Individuals with a higher educational level (bachelor's degree or more) tended to be assigned to newer vintages than older ones across all body types. Following the analysis of data, an attempt was made to translate the empirical findings into a utility maximization model where a primary driver is maximizing his/her utility by choosing the vehicle that best fits his needs from a fleet of vehicles available in the household.

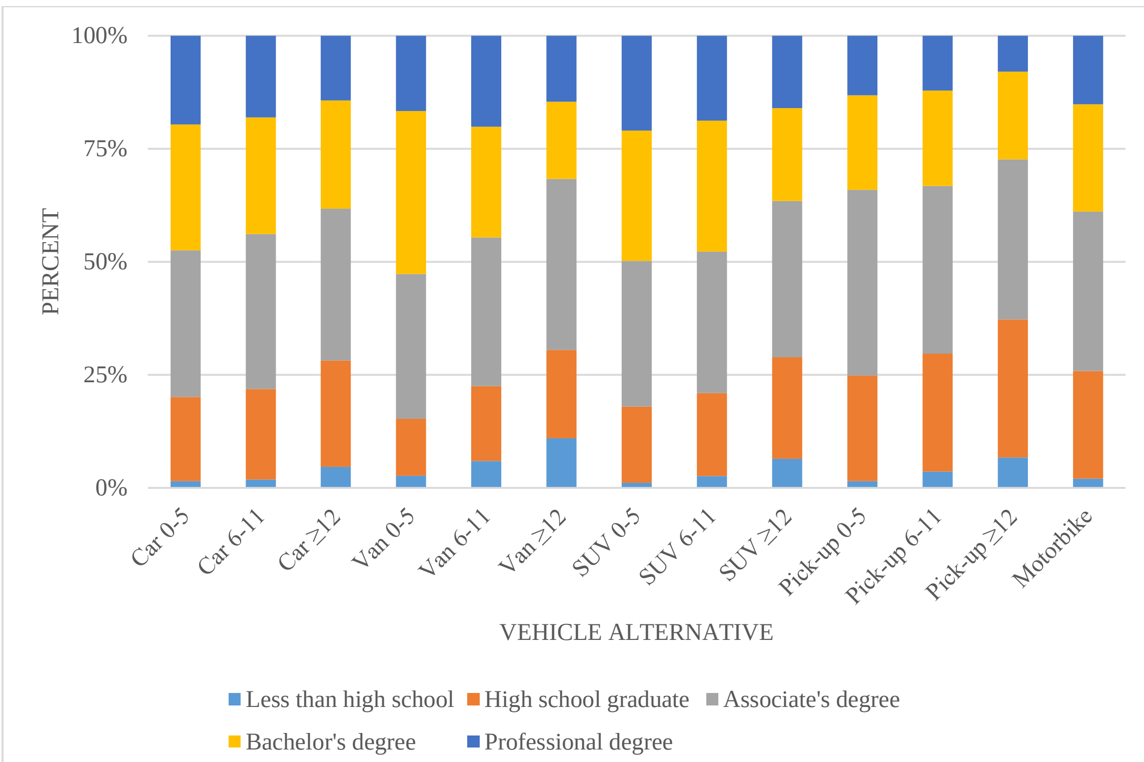

Figure 7.3. Vehicle type by primary driver's educational attainment.

Since the decision being modeled here is that of assigning a unique vehicle to a driver in the household, traditional single discrete choice models such as a multinomial logit or a nested logit model would perfectly fit the choice context. However, if an MNL model is chosen, there is the risk of modeling highly correlated alternatives together which violates the independence of irrelevant alternatives assumption observed by the MNL model. For example, in a household that owns 2 cars and 1 SUV, if one has to assign a vehicle to one of the drivers in the household, the choice context does not involve independent and irrelevant choices (choice between two cars). A straightforward way to handle this issue is considering a nested logit structure. A nested logit model retains the computational feasibility offered by an MNL model, and relaxes the independence assumption using a hierarchical structure.

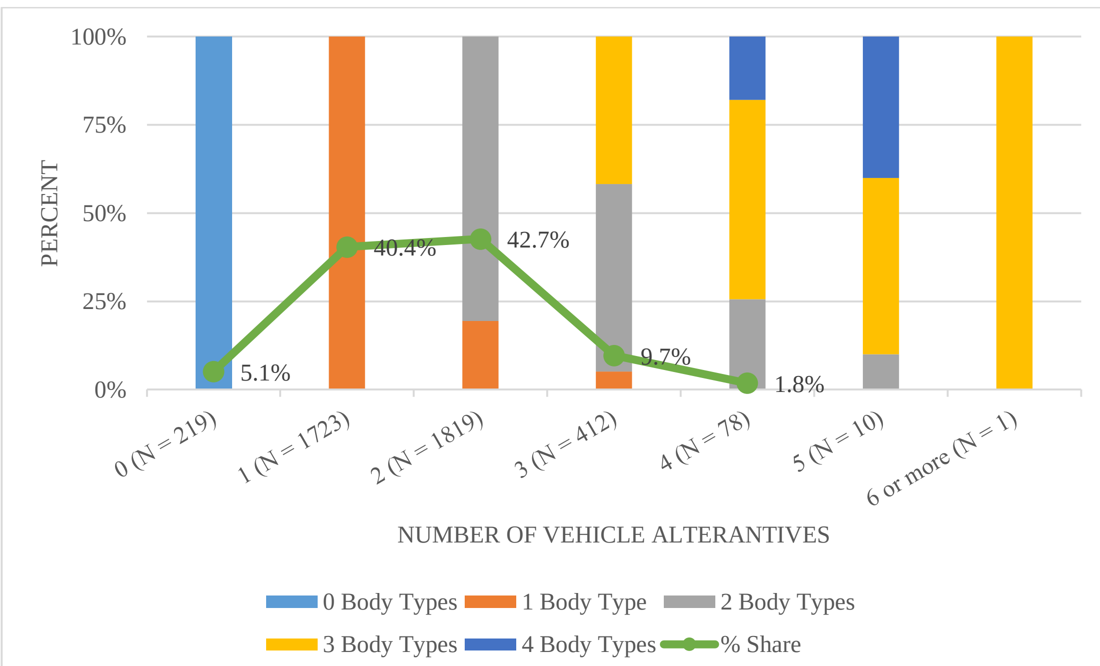

Figure 7.4. Cross classification between number of vehicles and number of body types owned by households.

Before deciding on a specific model structure, it was felt prudent to examine the data to find out how many households own multiple vehicles pertaining to the same body type (making the estimation of an MNL model infeasible). Figure 7.4 presents the results of the analysis of a cross classification between number of vehicle alternatives and the number of body types owned by households in the dataset. The green line in the figure identifies the market share of each vehicle ownership category. The number of vehicle alternatives in this figure points to the number of distinct motorized alternatives owned by each household from the 13 alternatives considered for vehicle fleet composition model system. From the figure, it can be observed that households owning zero and one vehicle alternatives own exactly zero and one vehicle body types, as expected. Amongst households owning 2 vehicles, 80% of the households own distinct body types and 20% (~360) households own both vehicles of the same body type. Amongst households who own three vehicles, only 5% of the households own all the three vehicles of the same body type, 53% of the households own vehicles belonging to 2 distinct body types and about 42% of the households own all three vehicles belonging to different body types. It should also be kept in mind that households owning 3 or more vehicle comprise only 11% of the dataset.

Further analysis was carried to observe the vehicle body type distribution in multivehicle households (2 or more vehicles). Results of the analysis are presented in Figure 7.5. The results depict what percent of such households own multiple vehicles of the same body type. From the figure, it can be observed that among multivehicle households, a majority (~23%) have multiple cars in their vehicle fleet. Number of households that own multiple SUVs and pick-up trucks are less than 5% and number of households that own multiple vans are almost negligible. Two important observations can be made from this analysis. One, the number of households that own multiple vehicles of the same body type are quite few in the data set. Two, even amongst the households that do own multiple vehicles, there is a dominance of households who own multiple cars compared to other body types.

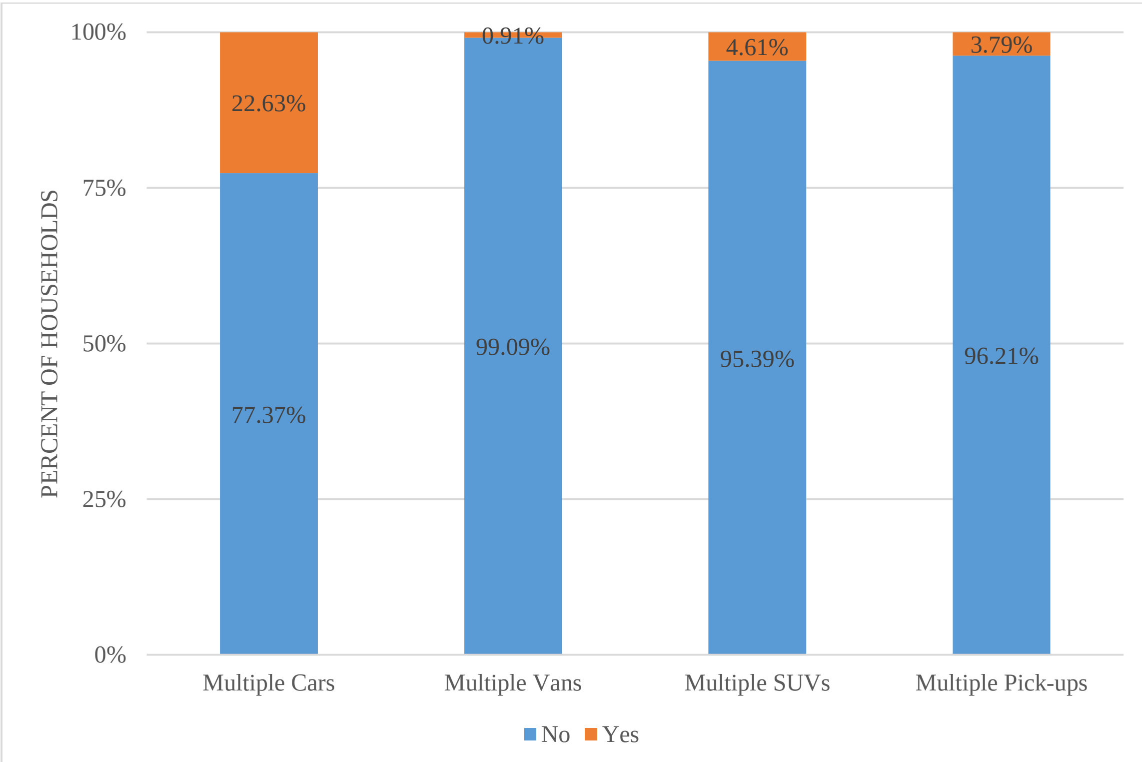

Figure 7.5. Vehicle body type distribution of multivehicle households.

From the findings in the data, it was felt prudent not to use a nested logit structure as the data at hand would not support the estimation of a robust model. An MNL model structure with a variable choice set was chosen for the primary driver allocation model. Future efforts should consider estimation of a nested logit model with body type classification as the upper nest and vintage classification as a subnest within each body type provided sufficient data is available. The estimation and validation results of primary driver allocation model are discussed in the next section.

Model estimation results. The primary driver allocation model is developed as an MNL model where an individual in the household is choosing a vehicle from the fleet owned by the household such that they maximize their utility. The individual is thus 'assigned' as the primary driver to the vehicle that he/she chooses. Previous studies successfully estimated primary driver assignment models in conjunction with a fleet composition component (Vyas et al., 2012), but such integrated models are difficult to implement in an application context. This research learns from the existing literature and develops a standalone primary driver assignment model for implementation in a tour level vehicle type choice framework. In application mode, the primary driver allocation model is run in an iterative fashion for every driver in the household starting with the head of the household. Once a person is assigned as primary driver to a vehicle, that vehicle is excluded from the choice set for primary driver assignment for rest of the individuals in the household. In cases where there are more household vehicles than drivers in the household, every driver is assigned as primary driver to a unique vehicle and the vehicles remaining after the allocation step are assumed to available to all the drivers in the household equally. An MNL model is estimated with motorbike as the base alternative. Estimation results of the primary driver allocation model are presented in Table 7.3.

Table 7.3

MNL Model of Primary Driver Allocation

Vehicle Type	Explanatory Variable	Coefficient	t-statistic
Car 0 -5 years old	Constant	-0.96	-7.82
	Female	1.82	14.88
	Age (> 16 and ≤ 24)	0.86	3.61
	Age (≥ 65)	-0.28	-2.17
	Respondent holds a graduate or professional degree	0.19	1.50
	Respondent is a high school graduate	-0.31	-2.42
	Worker from a highest income (≥ $100,000) household	0.15	1.35
Car 6-11 years old	Constant	-1.15	-8.61
	Female	1.74	14.03
	Age (> 16 and ≤ 24)	1.26	5.50
	Worker	0.19	1.72
Car 12 years or older	Constant	-0.25	-2.43
	Age (> 16 and ≤ 24)	1.15	4.06
Van 0-5 years old	Constant	-1.91	-8.81
	Female	3.60	13.71
	Respondent holds a graduate or professional degree	-0.45	-1.48
Van 6-11 years old	Constant	-1.32	-7.39
	Female	2.48	11.05
Van 12 years or older	Constant	-0.22	-1.15
SUV 0-5 years old	Constant	-1.16	-8.27
	Female	2.21	15.01
	Age (≥ 65)	-0.21	-1.16
	Worker from a highest income (≥ $100,000) household	0.22	1.57
SUV 6-11 years old	Constant	-1.03	-7.74
	Female	1.75	10.90
SUV 12 years or older	Constant	-0.33	-2.32
Pick-up 0-5 years old	Constant	0.16	1.30
	Female	-1.02	-5.37
Pick-up 6-11 years old	Constant	0.01	0.12
	Female	-0.67	-3.89
Pick-up 12 years or older	Constant	-0.28	-2.43
Goodness of fit statistics			
Sample Size (N)		6,842	
Likelihood ratio (df = 19)		23577.52	
ρ^2 Adjusted		0.78	

From the model results, it was observed that females tend to prefer newer vintages than older ones across car, van and SUV body types. Among these three categories females tended to prefer vans more than car and SUV body types. Females also had a negative proclivity to be allocated as primary drivers to pick-up trucks in the household. All these findings are in line with the observations from the descriptive statistics (see Figure 7.1). Younger individuals (>16 and ≤ 24 years old) had a greater preference for cars. Within the car body type, this cohort had a greater propensity to be allocated as primary drivers to older than newer vehicles. This finding is behaviorally consistent as majority of this category is college going students who might prefer cars over vans or SUVs which are purported as ‘family’ vehicles.

Older individuals (≥ 65 years old) are less likely to be assigned as primary drivers to newer vehicles in the household, a finding consistent with expectation. Individuals with educational attainment of a high school degree are less likely to be assigned as primary drivers to newer cars, while respondents with a graduate or professional degree are more likely to be assigned as primary drivers to newer cars. One reason for this might be the affordability factor where individuals with a greater educational attainment have a higher level of financial flexibility to own and use newer vehicles more than the individuals with lesser educational attainment. It was also observed that workers in general tend to choose cars over other body types. It is generally expected that work tours are undertaken as solo tours and hence workers tend to choose smaller (possibly more fuel efficient) vehicles to carry out their daily commute, leaving the larger vehicles (such as vans and SUVs) for household activities. Workers from highest income (≥ $100,000) households tended choose to newer cars and SUVs. It is intuitive that highest income households tend to own newer vehicles and hence workers in such households tend to get assigned as primary drivers to newer cars and SUVs. The likelihood ratio of the model is 23577.52 which is substantially greater than the critical χ^2 value at any reasonable level of significance.

Replication of observed patterns. In a traditional estimation/validation exercise, a model is estimated on 70% of the data and validated on the 30% hold out sample. In the current effort, 100% of the sample is used for estimation to ensure a reasonable sample size in each of 13 vehicle alternatives considered for model estimation. So, the model is applied on the entire estimation dataset to see how well it can replicate the observed primary driver allocation patterns.

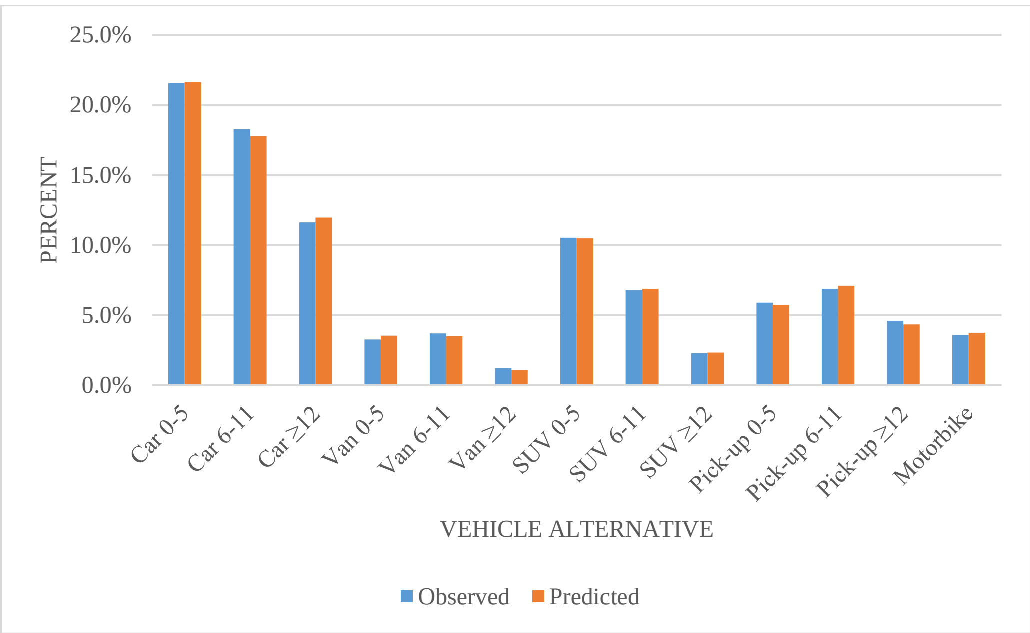

Figure 7.6. Observed vs. predicted vehicle allocation profile from primary driver allocation model.

Figure 7.6 shows the comparison between observed and predicted vehicle allocation profiles. The results presented are for an uncalibrated version of the primary driver allocation model. The model performance observed from this figure indicates that the overall primary driver allocation patterns were replicated accurately by the model. But the true litmus test to the primary driver allocation model is how well it can predict the vehicle allocation patterns of specific market segments in the data.

Figure 7.7. Observed vs. predicted vehicle allocation profiles for male and female respondents from primary driver allocation model.

Figure 7.7 shows the profiles for vehicles allocated to male and female market segments. Panel A presents the profile of vehicles chosen by male respondents in the dataset and Panel B depicts a similar comparison for female respondents. From the figure, it can be observed that male primary drivers have more or less an even vehicle allocation across different vintages in each body type while female primary drivers have a consistently higher representation in newer vintages than older ones across all body types except pick-up trucks. A comparison between male and female market segments also reveals that higher proportion of pick-up trucks are allocated to male primary drivers compared to females. Similarly, a higher proportion of vans are allocated to female primary drivers in comparison to males. The model is able to accurately depict the finer nuances observed in the data with respect to vehicle allocation characteristics between male and female respondents.

Figure 7.8 presents a comparison of vehicle allocation profile for workers and non-workers. From the figure it can be observed that for the most part the allocation profiles are similar for workers and non-workers. A higher proportion of newer vehicles (0-5 years) are allocated workers than non-workers in car and SUV body types. The behavioral reason behind this finding might be that workers prefer newer and more reliable vehicles for their daily commute. Non-workers have a slightly higher allocation of vans (all vintages) than workers. This might refer to stay-at-home moms/dads who tend to prefer larger vehicles to fulfill the chauffeuring needs of children in the family. The primary driver allocation model is able to replicate the vehicle allocation patterns at the aggregate level as well as for specific markets segments. With this information at hand, the next component in the framework, a tour level vehicle type choice model determines which vehicle amongst the household's fleet will be used to undertake a specific tour.

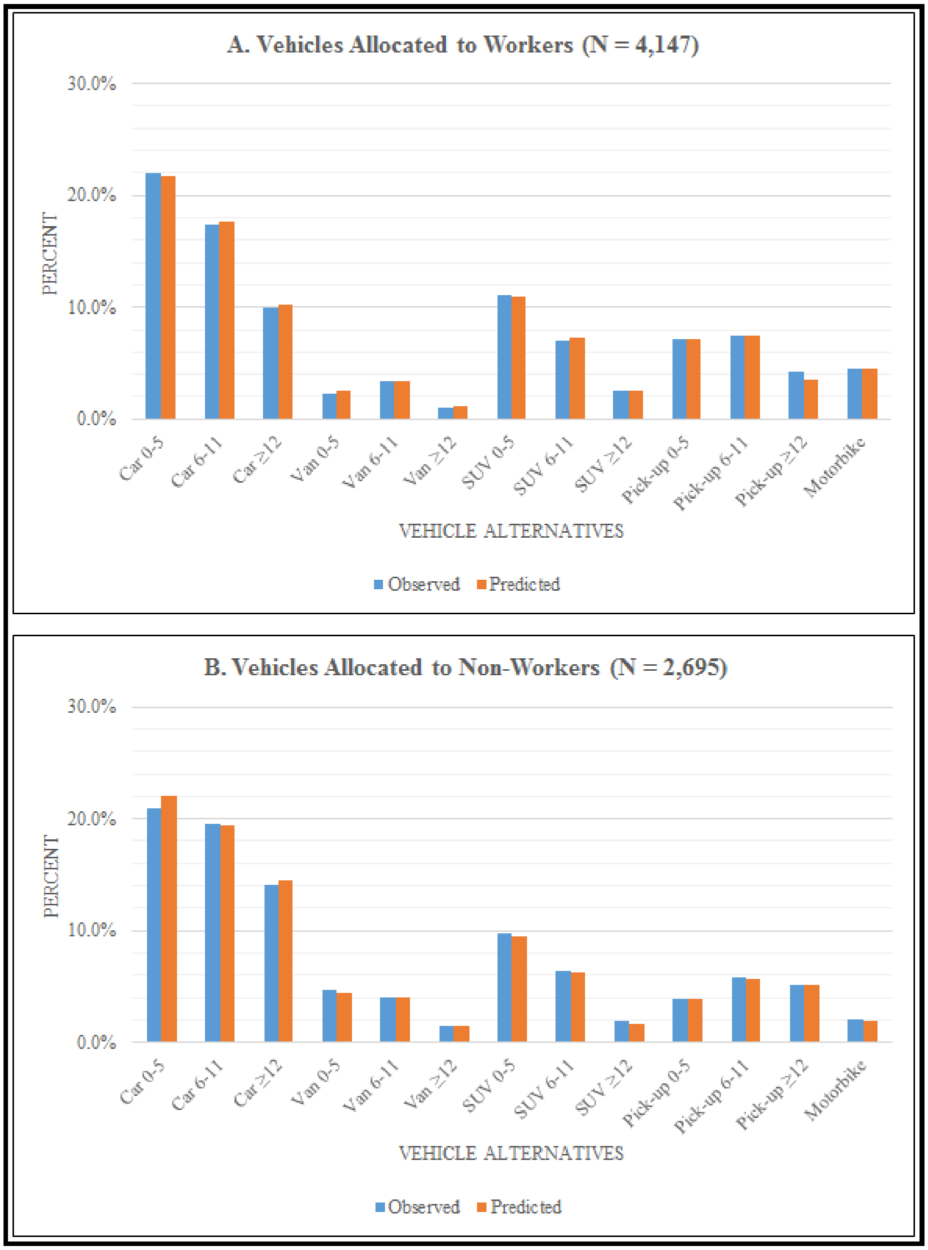

Figure 7.8. Observed vs. predicted vehicle allocation profiles for workers and non-workers from primary driver allocation model.

Tour Level Vehicle Type Choice Model

Data. The primary driver allocation model is a precursor to the tour level vehicle type choice model. In the primary driver allocation module, every vehicle in the household is assigned to a unique primary driver. This is a person level model where each person in the household chooses a vehicle for which he/she will be the primary driver. Once the assignment is done, vehicles assigned to respective primary drivers are assumed to be available to them throughout the day to travel to any activity. For non-primary drivers however (who do not have a vehicle assigned to them), a tour level vehicle type choice model determines which among the household vehicles will be utilized to undertake the tour under consideration. The idea behind the tour level vehicle type choice model is that individuals will choose a vehicle among the household's fleet to undertake a specific tour such that they maximize their utility. The tour level vehicle type choice model is an MNL model with variable choice-set. The choice-set is known in the estimation phase from the reported vehicle fleet characteristics of the household. In application mode, the vehicle fleet owned by each household will be predicted by the vehicle fleet composition module and provided as input to this component.

To estimate the tour level vehicle type choice model, trip level NHTS data is converted to tour level. In the NHTS trip level data, the respondent identifies which among the household vehicles was used to make a specific trip. Form the trip level data, the vehicle used for making each tour is imputed. First, all tours made by individuals greater than or equal to 16 years of age are selected. From these tours a sub-selection was made for home-based tours (tours that start and end at home). From the home-based tours, only tours made by auto modes (SOV or HOV) are selected for this analysis. Attributes such as vehicle used on the tour, primary purpose, tour accompaniment, travel party composition etc. are determined for the tour. From the activity schedules of different members in the household, each tour was also given a tag of joint/solo tour. Once the household vehicle used on each tour is identified, the next step is to generate a feasible choice set of vehicles available to make a specific tour. This is nothing but the vehicle fleet mix owned by the household. The vehicle fleet composition of the household is constructed from the vehicle file for the fleet composition modeling effort described in Chapter 5. The final data set used for model estimation consists of 5,165 home-based tours made by 3,385 persons from 2,365 households. Aggregate checks were made at household and person level to ensure that the dataset used for model estimation is representative of the regional level data (~4,700 households and ~10,400 persons). Table 7.4 presents the descriptive statistics of the data at household and person level.

Table 7.4

Tour Level Vehicle Type Choice Model: Data Description

Characteristic	Mean	Standard Deviation
Household (N = 2365)		
Number of vehicles in the household	2.08	1.013
Number of adults in the household	1.93	0.685
Number of drivers in the household	1.90	0.717
Number of workers in the household	1.09	0.881
Number of children in the household	0.57	1.007
% Households with income < $25,000	12.60%	0.490
% Households with income ≥ $75,000	39.79%	0.332
Person (N = 3385)		
Age of the respondents	53.35	16.054
% of Male respondents	46.26%	0.499
% of Respondents that are workers	62.90%	0.483
% of Respondents with an associate's degree	33.44%	0.472
% of Respondents with an bachelor's degree	25.58%	0.436
% of Respondents with a graduate or professional degree	17.90%	0.383

A comparison of household and person level characteristics between this data and the data prepared from primary driver allocation model (Table 7.1) reveals that both datasets have very similar characteristics assuring consistency between different model components in the framework. In addition to this, it was observed that the household level characteristics from both these tables are very similar to the characteristics of the dataset used in vehicle fleet composition modeling framework (Table 5.1). Table 7.5 presents the profile of vehicles used on home-based tours. An immediate observation from the table is that most of the home-based tours are made by car and SUV body types which is in line with the fleet composition patterns observed in the region (see Figure 5.1). It is however found that utilization of pick-up trucks (about 13% of all tours) is slightly less than their representation in the region's fleet (17%, from Figure 5.1) whereas the exact opposite pattern was observed for vans. Vans represent about 8% percent of the vehicle fleet in the region, but are used in about 10% of the home-based tours. These are intuitive observations in that pick-up trucks are not used as frequently for daily travel whereas households who own vans tend to use them to a higher degree (from results in Table 5.6 and entailing discussion).

Table 7.5

Tour Level Vehicle Type Choice Model: Vehicle Profile

	Body Type				
Age	**Car (%)**	**Van (%)**	**SUV (%)**	**Pick-Up (%)**	**Motorbike (%)**
0-5 years	46.9	46.2	55.0	40.4	
6-11 years	36.0	45.0	37.3	45.2	100.0
≥ 12 years	17.1	8.8	7.8	14.4	
Total No. of Tours	***2777***	***520***	***1146***	***688***	***34***

The intent of tour level vehicle type choice model is to advance the current practice in existing activity-based modeling systems to represent ‘vehicle type’ rather than just ‘mode’ (auto) at the tour level. The level of disaggregation at which current activity-based models represent an auto mode is an SOV or a HOV. When this data is sent to the network assignment module and subsequent emission computation models, an assumption has to be made regarding the emission profile of an ‘average’ SOV/HOV mode. To see if SOV and HOV modes are used for making specific types of tours, an exploratory analysis was carried out on the dataset and vehicle profile was constructed using a cross classification between the aggregate travel modes (SOV and HOV) and primary purpose of the tour. Figure 7.9 presents the aggregate vehicle profile by primary purpose of the tour.

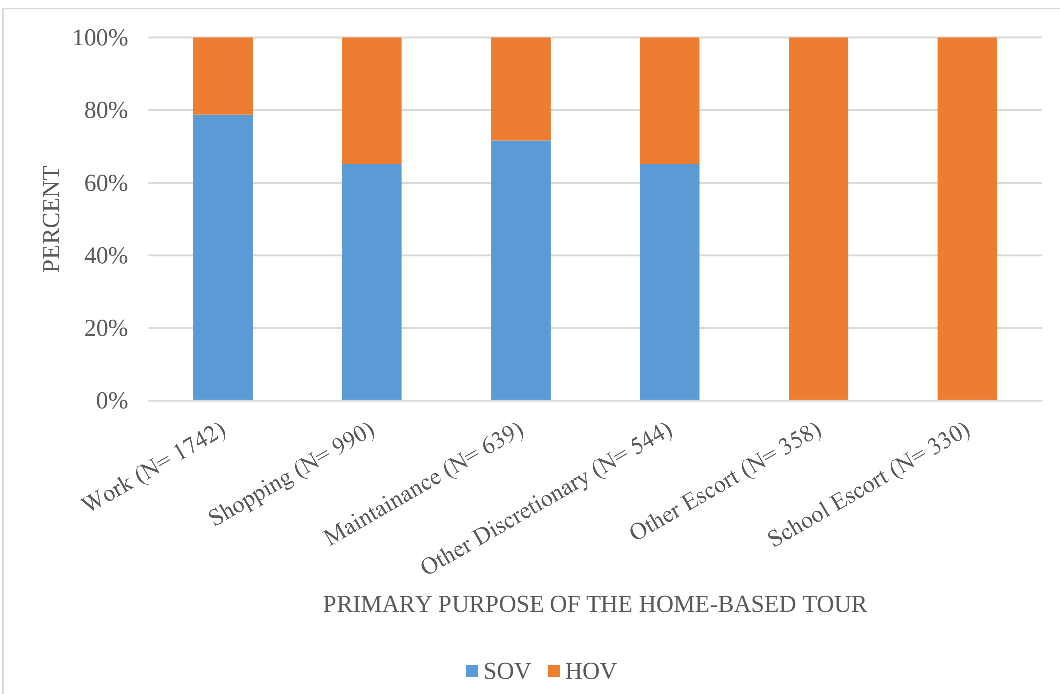

Figure 7.9. Aggregate vehicle profile by primary purpose of the tour.

From the figure, it can be observed that tours with work and maintenance activities as primary purpose tend to be undertaken predominantly using SOV mode while home-based shopping and home-based discretionary tours have a mix of SOV and HOV modes. Tours with some form of escorting are undertaken using HOV modes, which is intuitive. The question now is how the emission profile of an SOV/ HOV mode should be decided. Is there a representative body type that represents an SOV and/or a HOV mode? Is it safe to assume that SOV trips are made using smaller vehicles such as cars and HOV trips using larger body types such as vans? To answer these questions the same analysis as above was carried out except at the disaggregate level of body type of the vehicle. Results of the analysis are presented in Figure 7.10 and reveal some interesting findings. From the figure, it can be observed that there are noticeable differences in vehicle body type composition profiles for different types of tours. Home-based escort tours which were seen as only using HOV modes have a mix of all body types (small to large).Even within the home-based escort tours, the vehicle body type composition of school escort and other escort tours is significantly different. School escort tours which are mostly intended to drop-off/pick-up children below driving age at school tend to be undertaken more using ‘roomy’ vehicles (such as vans and SUVs) than the smaller cars. Other escort tours on the other hand have a significant proportion of tours made by cars.

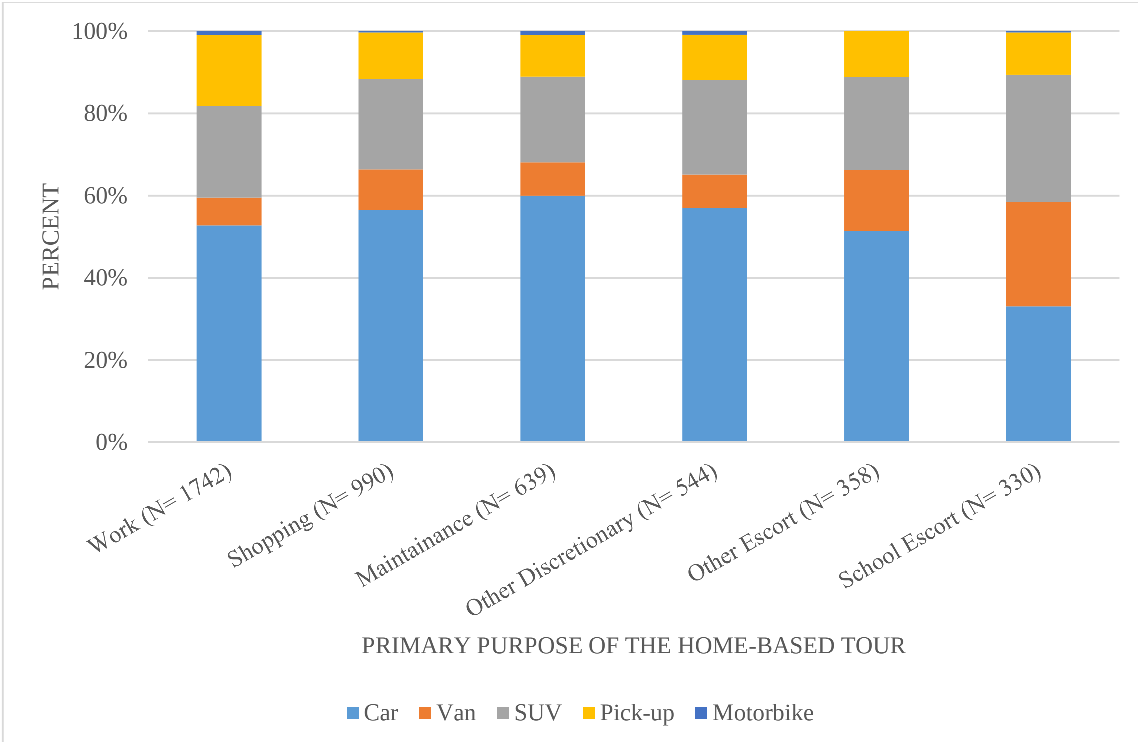

Figure 7.10. Vehicle body type profile by primary purpose of the tour.

Different types of tours call for usage of different types of vehicles (subject to their availability in the household) and hence it is necessary to model vehicle type choice at the tour level. The body type composition of vehicles used on different types of tours can be further disaggregated by the vintage classification of the body type. This information is made available by the vehicle fleet composition model discussed in Chapter 5. The tour level vehicle type choice model utilizes this information to model the specific household vehicle used on each tour. Table 7.6 provides the results of a similar analysis as Figure 7.10 but at the disaggregate level of body type and age. An MNL model structure was chosen owing to the data limitations discussed in the previous section.

Table 7.6

Vehicle Profile (Body-Type x *Age) by Primary Purpose of the Tour*

Vehicle Type	**Primary Purpose of the Tour**					
	Work (%/)	**Shopping (%)**	**Maintenance (%)**	**Other Discretionary (%)**	**Other Escort (%)**	**School Escort (%)**
Car 0-5 Years	25.3	25.9	26.8	24.8	25.1	19.4
Car 6-11 Years	19.5	19.1	20.0	23.9	18.2	8.5
Car ≥12 Years	7.9	11.5	13.1	8.3	8.1	5.2
Van 0-5 Years	2.1	4.7	3.3	3.9	8.7	13.3
Van 6-11 Years	3.7	4.0	3.9	3.7	5.6	10.6
Van ≥12 Years	0.9	1.1	0.9	0.6	0.6	1.5
SUV 0-5 Years	11.9	12.2	11.9	11.4	14.0	16.7
SUV 6-11 Years	8.4	8.1	7.8	10.5	6.1	11.5
SUV ≥12 Years	2.0	1.6	1.1	1.1	2.5	2.7
Pick-up 0-5 Years	7.3	4.2	3.8	3.3	6.1	6.1
Pick-up 6-11 Years	7.5	5.2	4.2	5.5	4.5	3.6
SUV ≥12 Years	2.4	2.0	2.2	2.2	0.6	0.6
Motorbike	1.0	0.3	0.9	0.9	0.0	0.3
Total Number of Tours	***1742***	***990***	***639***	***544***	***358***	***330***

Model estimation results. Following the analysis of the data, a tour level vehicle type choice model with variable choice-set (vehicle fleet owned by the household) is estimated. Various tour attributes are used as explanatory variables to define vehicle type chosen to undertake a specific type of tour. The vehicle type classification used in the model is the same as primary driver allocation model as well as the vehicle fleet composition module (4 body types x 3 vintage categories + motorbike). Results of the tour level vehicle type choice model are presented in Table 7.7.

Table 7.7

MNL Model of Tour Level Vehicle Type Choice

Vehicle Type	Explanatory Variable	Coefficient	t-statistic
Car 0-5 years old	Primary purpose of the tour is work	0.45	3.83
Car 6-11 years old	Constant	0.06	0.64
	Primary purpose of the tour is work	0.18	1.31
	Joint tour	-0.41	-2.15
Car 12 years or older	Constant	-0.19	-1.87
	Joint tour	-1.22	-4.28
Van 0-5 years old	Constant	0.91	4.75
	Primary purpose of the tour is work	-1.40	-4.64
	Primary purpose of the tour is discretionary	-0.76	-1.82
	Primary purpose of the tour is school escort	0.53	1.30
	Tour party consists only of adults	-1.48	-2.88
Van 6-11 years old	Constant	0.06	0.39
	Primary purpose of the tour is other escort	1.01	2.02
	Primary purpose of the tour is school escort	0.47	1.20
	Joint tour	1.48	3.12
	Tour party consists only of adults	-1.68	-2.82
Van 12 years or older	Constant	-0.09	-0.39
	Joint tour	-1.35	-1.62
SUV 0-5 years old	Constant	0.15	1.37
	Primary purpose of the tour is shopping	0.37	1.75
	Primary purpose of the tour is other escort	0.61	1.93
	Primary purpose of the tour is eat meal	0.49	1.28
	Primary purpose of the tour is school escort	0.48	1.74
	Number of escorting stops on the tour	0.23	2.04
	Joint tour	1.28	3.89
	Tour party consists only of adults	-1.01	-2.59
SUV 6-11 years old	Constant	0.08	0.73
	Primary purpose of the tour is maintenance	0.46	1.70
	Joint tour	0.64	2.06
	Tour party consists only of adults	-0.69	-1.79
SUV 12 years or older	Constant	-0.44	-2.62
	Joint tour	-1.10	-2.26
Pick-up 0-5 years old	Constant	-0.13	-1.03
	Primary purpose of the tour is work	0.62	3.20
	Primary purpose of the tour is discretionary	-0.54	-1.63
	Number of escorting stops on the tour	-0.34	-2.03
Pick-up 6-11 years old	Constant	-0.29	-2.46
	Primary purpose of the tour is work	0.39	2.18
	Primary purpose of the tour is school escort	-0.55	-1.42
Pick-up 12 years or older	Constant	-0.90	-4.69
	Primary purpose of the tour is work	0.40	1.46
	Primary purpose of the tour is other escort	-1.38	-1.82
	Primary purpose of the tour is school escort	-1.06	-1.36
	Joint tour	-1.92	-3.40

From the model results, it was observed that cars and pick-up trucks are most preferred vehicle types to undertake home-based work tours. While the use of cars to undertake work tours is intuitive (since most work tours tend to be solo tours, prompting the use of smaller household vehicles), the use of pick-up trucks for making work tours is intriguing and warrants further investigation. One possible reason might be that the occupation of individuals using pick-up trucks on their commute might require the use of such vehicle for their work purpose as well. Within car and pick-up truck body types, newer vintages are preferred to older ones, subject to their availability in the household. Van body type has a negative proclivity to be chosen for home-based work tours, which is consistent with expectation. Tours with an escorting (school/other) activity as the primary purpose tend to be undertaken using larger (and more comfortable) body types such as vans and SUVs. Within these body types, school escort tours had a greater proclivity to be undertaken by newer vintages. It is natural for parents to want the highest level of safety possible when chauffeuring their kids, which explains the choice of newer vintages for tours with school escort as the primary purpose.

Tours with meal and shopping activity as primary purpose had a greater probability of happening using newer SUVs, while tours with maintenance activity as primary purpose had a greater propensity to happen using slightly older SUVs. Car and pick-up truck body types had a lesser probability of being used for joint tours. This is finding is nicely complemented by an earlier finding that these are the preferred vehicle body types for home-based work tours which tend to be predominantly solo tours. SUVs and vans as expected had a greater propensity of getting utilized for undertaking joint tours. An interesting finding here is, while relatively newer vans/SUVs (< 12 years) tended to be chosen with a greater probability for making joint tours, older vans/SUVs had a negative propensity to be chosen for similar types of tours. This might translate to ‘cautionary’ behavior of individuals where travelers usually play it safe by not choosing unreliable vehicles for joint tours as any incident (such as a flat tire, engine troubles etc.) might impact the schedules of all the individuals involved in the joint travel.

Tours with greater number of escort stops tend to be undertaken more by newer SUVs, while pick-up trucks are less likely to be chosen to participate in tours with more escort stops. Tour composition also has an impact on the type of vehicle chosen to make a tour. Van and SUV body types tend to be chosen less to participate in tours where the travel party consisted only of adults. This finding is nicely complemented by the fact that the same body types are chosen more for tours with school escort as the primary purpose. All the findings from model estimation results are behaviorally intuitive and consistent with expectations.

Table 7.8

Tour Level Vehicle Type Choice: Goodness of Fit Measures

Statistic	**Value**
Sample size (N)	5,165 tours
L(0)	-13247.96
L(β)	-2808.76
Likelihood ratio	20878.41
$\chi^2_{33,0.001}$	63.87
$\rho^2 Adjusted$	0.79

Goodness of fit measures of the model are presented in Table 7.8. The likelihood ratio of the model is substantially greater than the critical χ^2value at any level of significance. Following the model estimation, an extensive replication exercise was carried out to see how well the model can predict vehicle type choices at the aggregate level as well as at the level of tours with different primary purposes.

Replication of observed patterns. The tour level vehicle type choice model was applied on the entire estimation dataset to test how well the model can replicate observed vehicle type choice patterns. It was necessary to use 100% of the data for model estimation to ensure reasonable sample sizes for all the body-type x age categories considered for model estimation, leaving no holdout sample for a true validation process. So, the process followed for testing the efficacy of the model is more of a replication than a validation process. In the presence of a larger dataset, this constraint can be overcome in a straightforward manner.

Figure 7.11 presents the comparison between observed and predicted vehicle type choice profile. The results presented are for uncalibrated version of the model. It can be observed that the model is able to replicate the aggregate vehicle type choice profile quite well. The model slightly under predicts the usage of newer cars (0-5 years) and over predicts the usage of newer SUVs. This can be handled with very minimal calibration of the model. The model also slightly over predicts the usage of motorbikes. This is because motorbike is used as the base alternative in model estimation and has no coefficient whatsoever to explain the preference for motorbikes to participate in specific type of tours. Model coefficient assertion for motorbike category might be warranted to handle this anomaly in a behaviorally consistent fashion. Though the model performs exceedingly well at the aggregate level, it does not ensure the effectiveness of the model to predict vehicle type choice at the level of individual tours. To examine this, comparisons are made between observed and predicted vehicle type choice profiles for specific type of tours. Figures 7.12-7.14 present such comparison for home-based tours with different primary purposes.

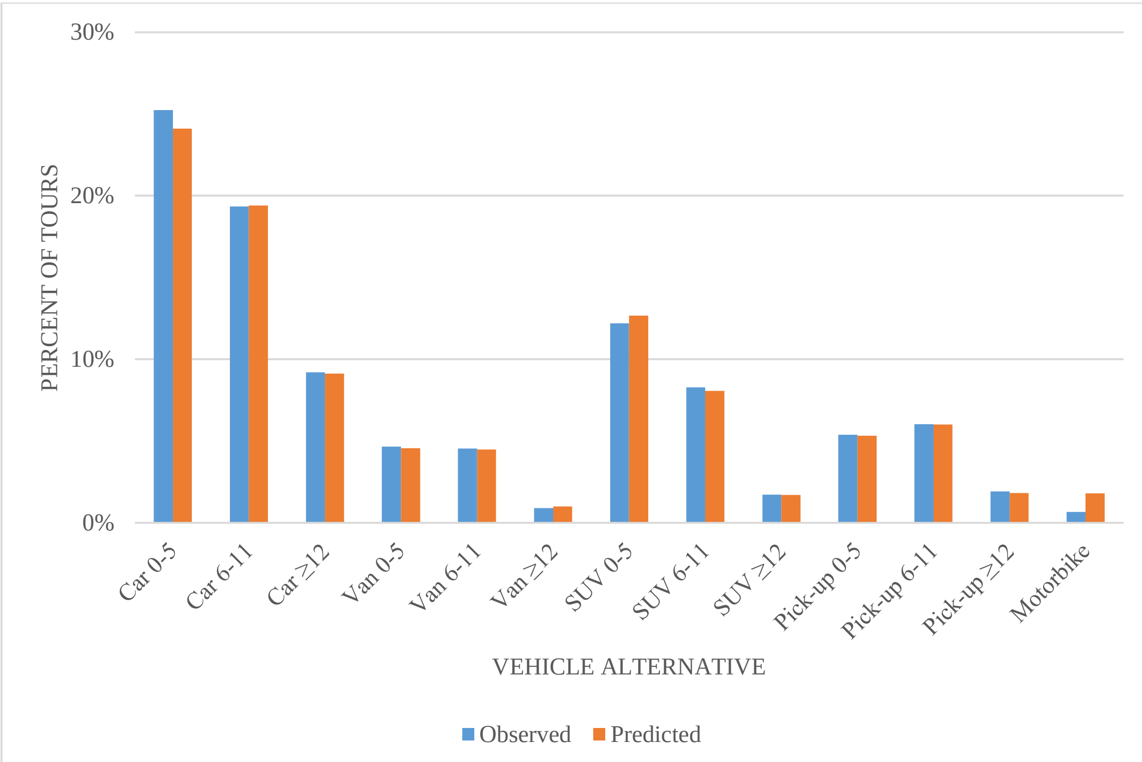

Figure 7.11. Observed vs. predicted vehicle type choice (aggregate comparison).

Figure 7.12 presents the vehicle type choice profile for home-based work and home-based shopping tours. Comparison between observed and predicted patterns for tours with work as primary purpose are depicted in Panel A of the figure.

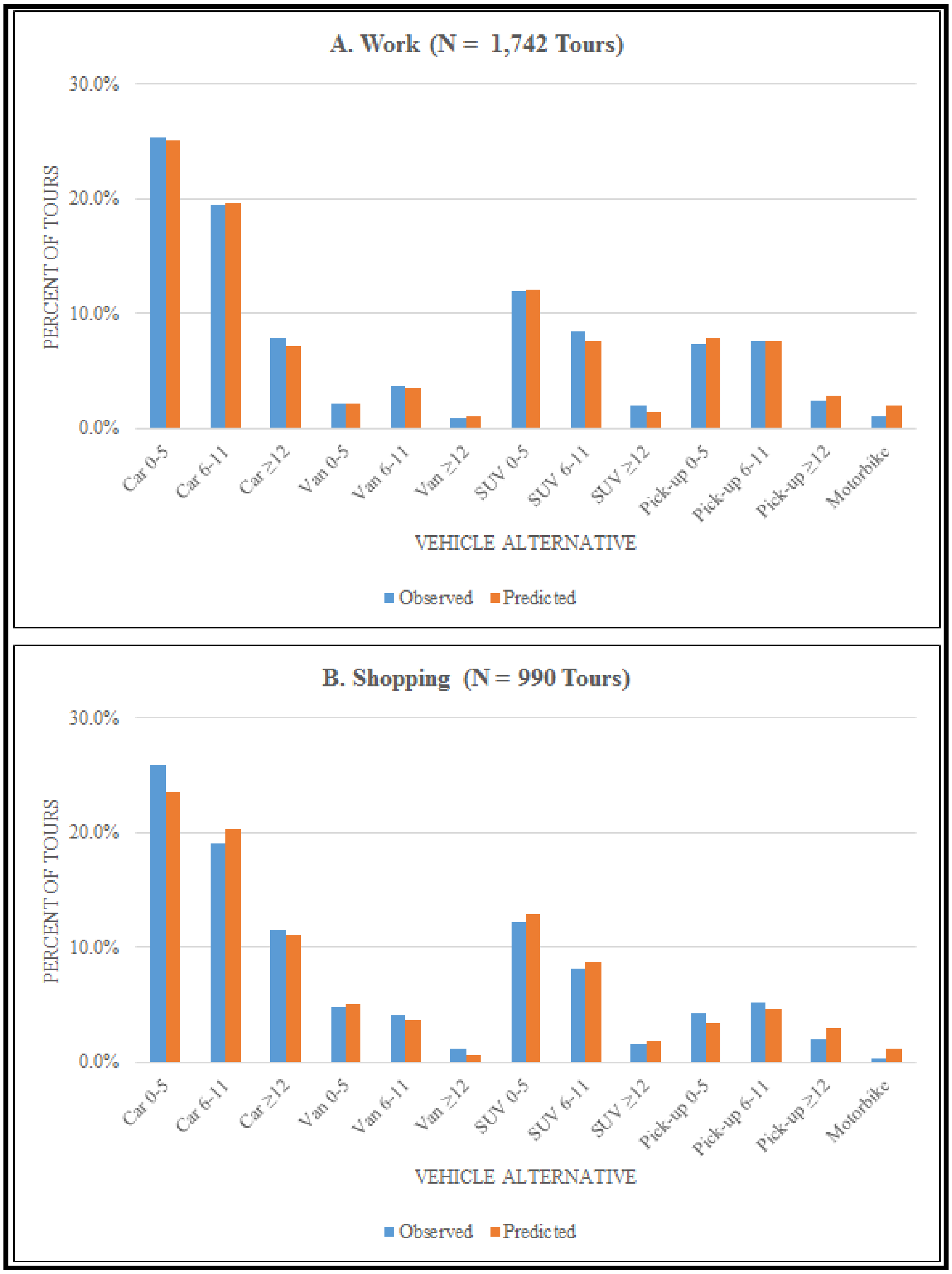

Figure 7.12. Vehicle type choice by primary purpose of the tour:
Panel A: Work, Panel B: Shopping.

The model is able to accurately predict the vehicle type choices observed in the data. Relative abundance of car and SUV body types is expected for work tours. It is interesting to see a significant representation of pick-up trucks in this category. Pick-up trucks are usually considered less fuel efficient and hence do not make an ideal commute vehicle. A probable reason for choosing pick-up trucks to go to work might be that the occupation of the individual might require such a vehicle for work purposes anyway. It is heartening to see that the model is able to predict this pattern very accurately. Vans are the least preferred vehicle to go to work, which is consistent with expectation. This is corroborated by negative sign in the utility equation for van in the model estimation result which directly translates to negative propensity of choosing vans for work purpose as observed in from the figure. Panel B presents a similar comparison for tours with shopping as primary purpose. SUVs have a slightly higher representation for this tour category, which is consistent with the observed data and findings from the model estimation results. Pick-up trucks on the other hand are not a popular choice to be used on tours with shopping as the primary purpose.

Figure 7.13 presents vehicle type choice profiles for tours with maintenance and other discretionary activities as primary purpose. The model is able to accurately depict the ‘pattern’ of vehicle type choices for both these primary purposes. Slight calibration might be warranted to exactly match the vehicle type choice profiles. Figure 7.14 presents the vehicle type choice profile for tours with escort activities as primary purpose. Panel A presents the comparison for home-based other escort tours and Panel B presents similar results for home-based school escort tours.

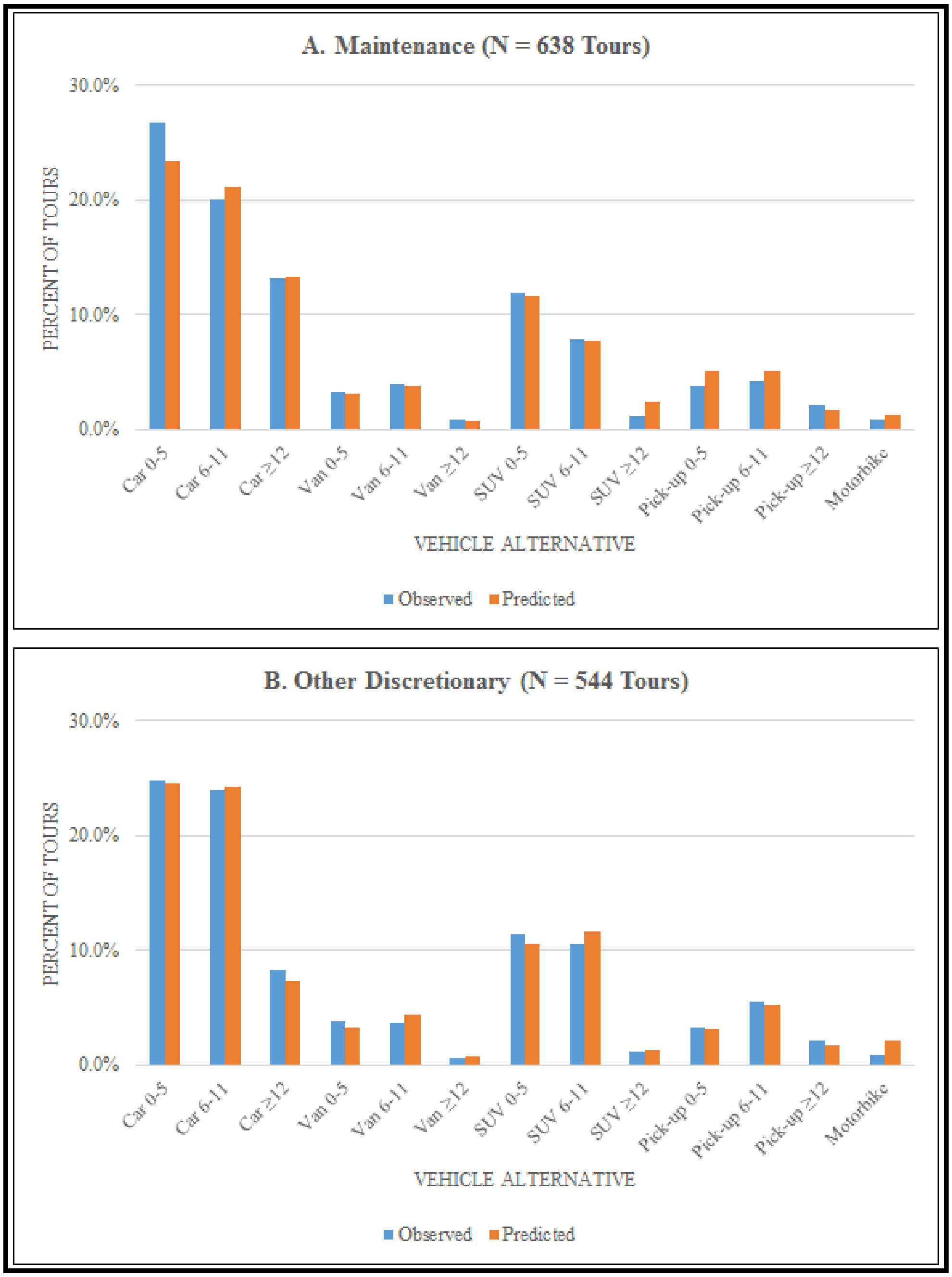

Figure 7.13. Vehicle type choice by primary purpose of the tour:
Panel A: Maintenance, Panel B: Other Discretionary.

An observation that is readily apparent from this figure is that the vehicle type choice profile of tours with escort activity as primary purpose is significantly different from other type of tours. A greater proportion of vans and SUVs are chosen for escort activities.

Within tours with escort activity as primary purpose, school escort tours have a greater proportion of vans chosen than that of other escort tours. This behavior is understood as vans are usually more convenient to chauffer children to school.

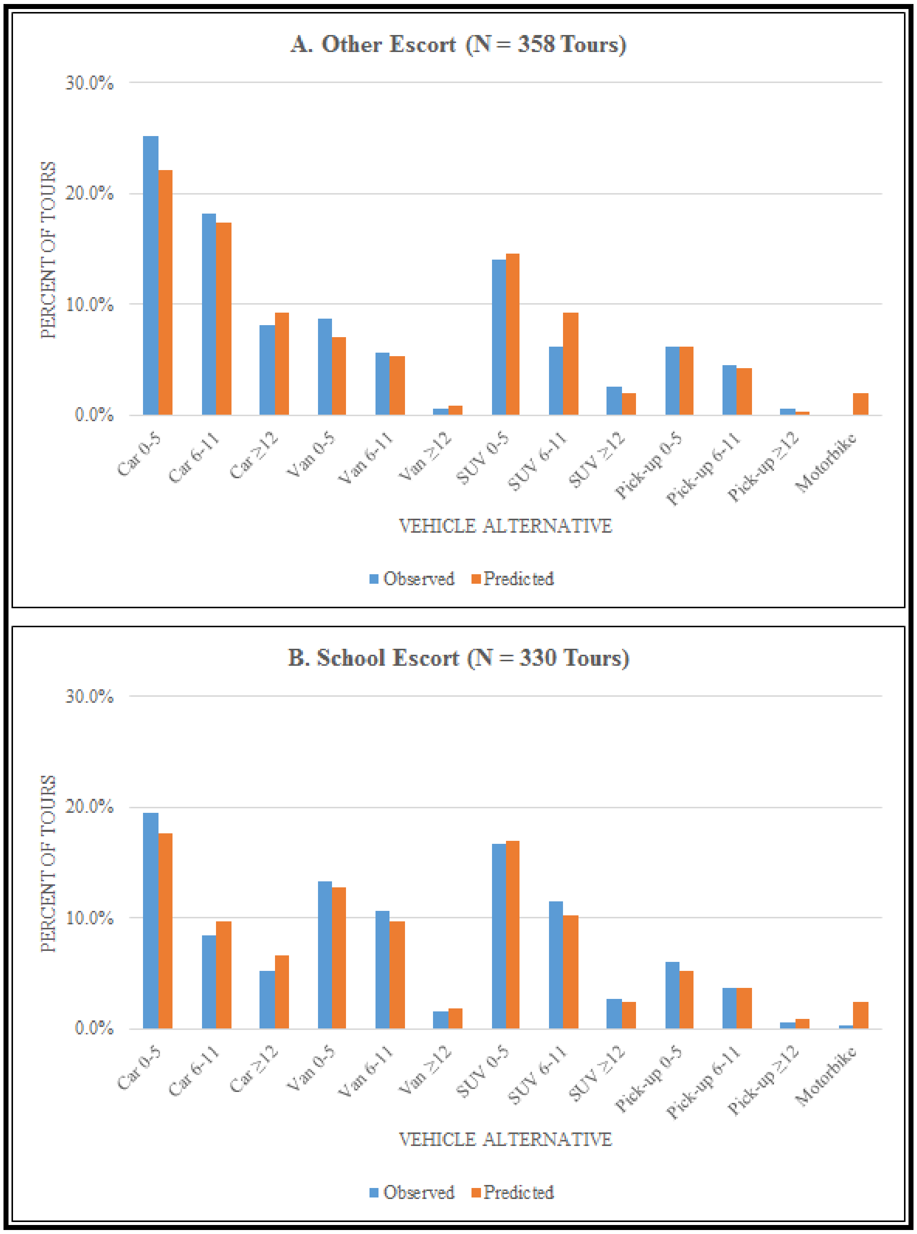

Figure 7.14. Vehicle type choice by primary purpose of the tour:
Panel A: Other escort, Panel B: School escort.

While the model does not exactly match the vehicle types chosen across all 13 vehicle alternatives, it can be observed that patterns in vehicle type choice are predicted quite well. The only anomaly from the model predictions is the choice of motorbikes to undertake escort tours. This can be handled with assertion of an alternative specific constant for motorbike category as discussed before. Overall, the model is able to predict the vehicle type choice patterns exceedingly well at the aggregate level as well as at the level of specific types of tours and shows great promise for implementation in the context of the overall framework of this research effort.

On any tour, if there are two primary drivers each with their own assigned vehicle, simple rule based heuristics determine the vehicle to be used on the tour. Between vehicles of different body types, the larger (and hence more conformable) body type is always chosen. Between vehicles of different vintage classifications, the newer (and hence more reliable) vintage is always chosen. Between vehicles of similar body type x age classification a random assignment is done. In case fuel efficiency of both vehicles is known (or predicted), the more fuel efficient (and hence more economical) vehicle is always chosen. While the rule based heuristics proposed are somewhat rudimentary in nature, they provide a decent starting point to handle the few instances in which there are two primary drivers engaging in a tour. An attempt was made to identify these rules based on trip level information collected in the NHTS, but the level of detail collected the survey data was not sufficient to form strong hypotheses for the rule based heuristics. Travel surveys should progress toward collecting information at the tour level to facilitate development of models that explain such finer nuances in tour characteristics at the household level. Information regarding the type of vehicle at the level of individual tours helps in accurate emission forecasts at the region level and provides planners with a powerful policy analysis tool. Such disaggregate level of information will also help in testing policies aimed as specific vehicle categories prior to their implementation so that informed decisions can be made by policy makers. Examples of such policies are restricting the entry of old pick-up trucks into specific set of zones or cash for clunkers programs for older vehicles etc.

Summary and Conclusions

The tour level vehicle type choice modeling framework developed as a part of this dissertation aims at enhancing the modeling methodology in current activity-based modeling methods to predict the type of vehicle (classified by 13 body-type x age combinations) used to undertake a specific tour. The current level at which most of the activity-based modeling systems represent personal travel is at the level of mode, there by aggregating all vehicle types into one ‘auto’ mode. The framework takes inputs from tour composition modules discussed in Chapter 4 regarding the primary purpose of the tour, tour composition etc. and from the fleet composition module discussed in Chapter 5 regarding the fleet of vehicle owned by the household. The framework allocates household resources (vehicles) to their (activity-travel) needs.

The framework starts with identifying a primary driver for each vehicle owned by the household. It is assumed that primary drivers use ‘only’ the vehicle assigned to them to travel to any activity in the day. This assumption is corroborated by the finding that 90% of tours used for estimating vehicle type choice model component are undertaken by individuals using the vehicle ‘assigned’ to them. The primary driver allocation model is a person level MNL model where an individual chooses a vehicle among the fleet owned by the household such that they maximize their utility. The choice set for every individual is constrained to include only the vehicles owned by the household. The vehicle type classification observed for this model is the same as the one used in the vehicle fleet composition modeling framework. This ensures seamless continuity between different model components. The model estimation results are fairly intuitive and the uncalibrated version of the model is found to replicate the observed primary driver allocation patterns quite effectively.

Next component in the model system is a tour level vehicle type choice model that determines which vehicle among the household’s fleet will be chosen to undertake a specific type of tour by an individual who has no vehicles assigned to him/her. This component is intended primarily to model the travel characteristics of auto-deficient households who have more number of drivers in the household than vehicles, but can be easily extended to auto-sufficient households as well. This is a tour level model where tour attributes influence the type of vehicles to be used. Household level socio-demographics are not used as explanatory variables in this model as such characteristics are assumed to proliferate into the models via fleet composition characteristics of the household. The model assumes that all of the household vehicles are ‘available’ to the non-primary driver to undertake a specific tour. The underlying assumption is that the non-primary driver will undertake the tour if and when the vehicle becomes available. The estimated model is able to replicate observed patterns exceedingly well at the aggregate level as well as the level of tours with specific primary purposes (work, shopping etc.). Both the primary driver allocation model and the tour level vehicle type choice model are developed as MNL models. One might argue that the alternatives in the MNL model structure for these models are highly correlated (different vintages across the same body type), but it should be kept in mind that choice set used for model estimation is constrained to include only the vehicles owned by the household, thereby alleviating this problem to some extent. Future efforts should explore nested logit structures for both these models, provided the data allows for estimation of a robust model.

The tour level vehicle type choice model is held down by the assumption that all household vehicles are available for non-primary drivers at all times in a day. This is not true in the real world and to be able to accurately depict real world behavior in a simulation environment, a real-time vehicle allocation and tracking framework is proposed in the next chapter. The framework is presented in the context of an integrated model system where there is continuous communication between an activity-based model system and a dynamic traffic assignment model. Future efforts in this domain should focus on including the make/model information of vehicle in addition to body-type x age classification to represent household's vehicle fleet.

CHAPTER 8

A CONCEPTUAL FRAMEWORK FOR REAL-TIME VEHICLE ALLOCATION AND TRACKING

This chapter presents a conceptual framework for real-time vehicle allocation and tracking framework that could potentially be implemented in an integrated model setting where an activity-based model and dynamic traffic assignment model are in close communication with each other. Previous chapters presented the frameworks as well as model estimation/validation results for

- *Tour characterization framework*: Determines all of the secondary stops made on different types of tours undertaken by persons in a household.
- *Vehicle fleet composition framework*: Predicts the vehicle fleet mix owned by a household, along with an estimate of the annual mileage consumed using each of the vehicles owned.
- *Tour level vehicle type choice*: Framework to allocate vehicles owned by the households to tours undertaken by them.

The tour level vehicle type choice framework utilizes the information from both tour characterization and fleet composition components. While this framework is a good starting point to introduce household vehicular constraints into vehicle allocation for various tours undertaken by members of the household, it is still held down by the assumption that all vehicles are available to household members at all times. In real world, in addition to auto-deficiency constraints, one often finds temporal constraints with regard to availability of household vehicles at different times of the day. For example, consider a (hypothetical) household with three licensed drivers and two cars. If two among the three drivers take the cars out to engage in different activities across the day, the third driver is constrained by unavailability of a car to undertake any activity. The third driver (and any other household member dependent on him for chauffeuring) will either have to wait until one of the household vehicles becomes available, choose an alternative mode of transportation or choose to engage in joint travel with one of the household members who is utilizing one of the household vehicles to begin with. This chapter proposes a real-time vehicle allocation and tracking framework, with an intent to translate this real world behavior into an integrated modeling framework.

Table 8.1

Mode Share of Work Trips in MAG Region

Mode	**Number of Trips in the Survey Data**	**Percent of Trips**
Auto	3356	93.95%
Bus	52	1.46%
Walk/Bike	120	3.36%
Other	36	1.01%
Airplane	8	0.22%
Total	**3572**	**100.00%**

The proposed framework assumes that all fixed activities (work trips/school trips) made by members of household are by auto modes. In case an auto mode is unavailable to make the work trip, it is assumed that individuals resort to private auto modes such as a taxi cab or shared ride facilities to reach work. Similarly, school trips for children are usually undertaken by a school bus or chauffeured by an adult. This assumption is made after a careful observation from the NHTS data collected for MAG region. Table 8.1

presents the share of different transportation modes used for work trips in MAG region. It can be observed from the table that number of work trips made using non-auto mode are very few and if walk trips are not considered, the percentage of non-auto mode share is less than 3%. This observation is consistent with the travel patterns observed in the region. Therefore the framework being proposed mainly concerns with non-mandatory activities and that too for households who have less number of vehicles than drivers in the household (auto-deficient households).

Prior to entering this framework, but after the fleet composition framework, a primary driver allocation module (discussed in the previous chapter) identifies and allocates each vehicle in the household to a 'primary driver'. This can be done by allocating vehicle fleet to individuals in a household such that they maximize their utility. Or this can be decided based on simple/complex heuristics involving one or more person level attributes (gender, income, age etc.). Once a vehicle is allocated to a primary driver, it is assumed that the vehicle is available to that driver throughout the day to travel to any activity. If the primary driver is not utilizing this vehicle in a given time period, it will become available to other members (licensed drivers) in the household.

Vehicle Tracking Framework

A real-time vehicle allocation and tracking framework can be realized only when the activity-based microsimulation model and dynamic traffic assignment (DTA) model are integrated with tight coupling and work in close conjunction with each other. Information exchange between both the systems should happen at a fine temporal resolution (every minute). Figure 8.1 provides a sample schematic for such a model system.

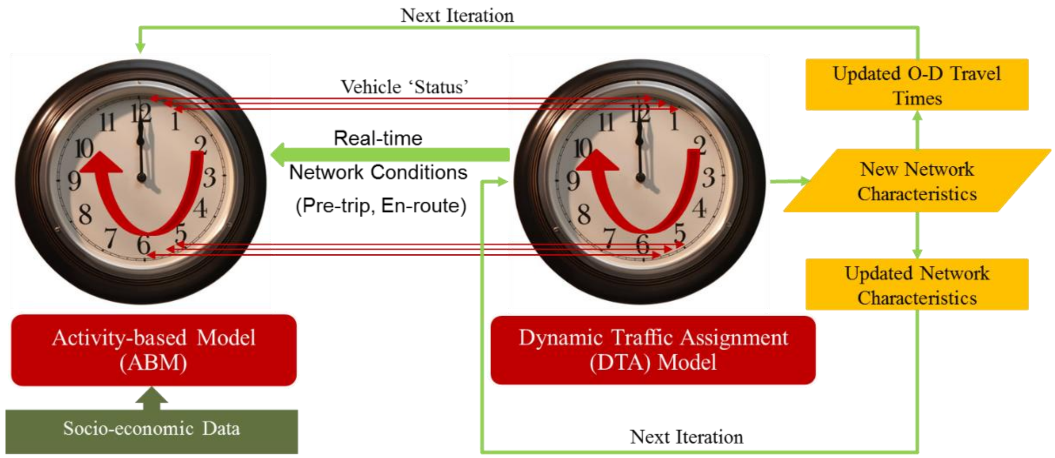

Figure 8.1. Integrated model design framework.

The activity-based model takes socio-economic data of a synthetic population as input and simulates the activity travel patterns of individuals in continuous-time. Every minute, the activity-based model sends all the trips departing in that minute to the dynamic traffic assignment model (each trip is tagged with the id of the vehicle being used on the trip). The dynamic traffic assignment model in return sends information regarding all vehicles on the network back to the activity-based model. The information sent back includes but is not limited to:

i. Travel time for trips that reached their destination.

ii. Current location for trips that are enroute.

iii. Network status for all vehicles being simulated on the network.

At the beginning of the day (minute zero), all the vehicles in the household have a network status of '1' meaning they are available for use by members of the household. Each vehicle is tagged with a home TAZ location. If any of the household members embark on a tour, the individual checks out a household vehicle that is assigned for the journey and vehicle is simulated on the network. Every minute, the dynamic traffic assignment model sends a status report to the activity-based model with a list of all the vehicles that are 'on the network' as well as the ones that reached their respective destinations. All the vehicles that are enroute (in other words have their current TAZ location ≠ their home TAZ location) to their final destination (home) will be assigned a value of '0'. Once the vehicle finishes a tour (reaches home TAZ location), the dynamic traffic assignment model sets the status of the vehicle back to '1' and sends this information to the activity-based model. At any given instance each and every vehicle in the fleet of the household either has a '0' (meaning that the vehicle is unavailable) or a '1' (meaning that the vehicle is available) value assigned to them. This way in any given minute, the sum of network status values of all vehicles in the household informs the activity-based model regarding the availability of household's vehicles to make a tour.

Vehicle Allocation Framework for Children's After School Activities

At the top level, the vehicle allocation framework separates children from all other individuals in the household. As mentioned before, school trips made by children are assigned an auto mode (randomly chosen from the household's fleet) by default. Framework for vehicle allocation and tracking for children's after school activities is presented in Figure 8.2.

For children, it is first checked if they have any non-mandatory activities scheduled after school. If so, it is checked whether this activity can be performed with another household member. The interpretation for such activities in the activity-based modeling jargon is a 'partially joint tour', where an adult in the household has free time available and can chauffer the child as a part of his journey. For example, a working adult picking up his/her kid after school and dropping him off at soccer practice is a partially joint tour. While the main purpose of this tour is work (which is a solo tour for the adult), it is still a partially joint tour for the portion where the adult tends to the chauffeuring needs of the child. The tour formation component (scheduling of fully/partially joint tours at the household level) is beyond the scope of this research effort and is not discussed here. The proposed framework considers this information as exogenously provided.

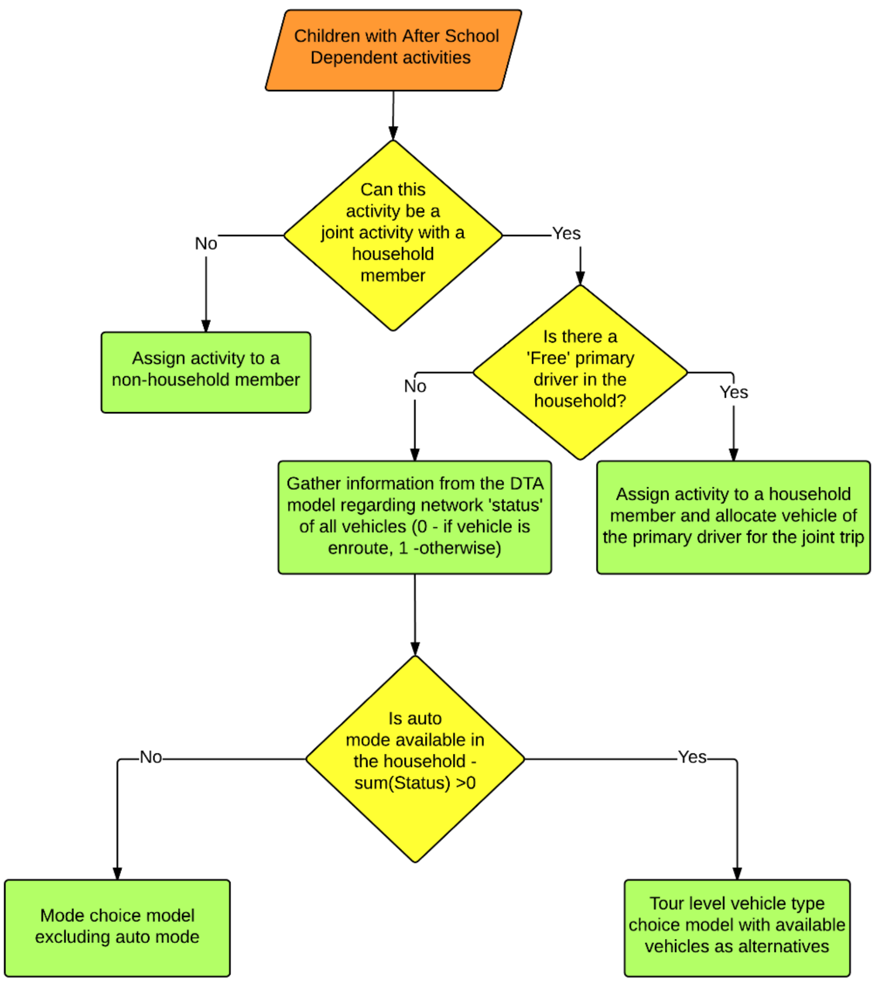

Figure 8.2. Vehicle allocation and tracking framework for children's afterschool activities.

Suppose a child has an after school activity scheduled, it is first checked if there are available adults in the household that are primary drivers. If yes, the child is assigned to that household member and the vehicle tagged to the primary driver is automatically assigned to the tour. Suppose a household member is available but he/she is not a primary driver, information gathered from the dynamic traffic assignment model regarding the network status of the vehicles is retrieved and summed across all the vehicles owned by the household. If sum of the status variable is greater than zero, it means that at least one of the household vehicles is available to be allocated to the non-primary driver in order to chauffer the child. A tour level vehicle type choice model with 'available vehicles' in the household determines the vehicle to be used on the tour. If not, the adult is still assigned the duty of chauffeuring the child to his after school activity, but a mode other than auto is assigned for the trip (determined using a mode choice model excluding the auto mode). Incase none of the household members are available in the time period in which the child needs to perform an after school activity, the child is assigned to a non-household member (synonymous to taking a cab to engage in the activity)

Vehicle Allocation Framework for Adult's Non-Mandatory Activities

Figure 8.3 depicts the vehicle allocation framework for non-mandatory activities carried out by adults. The framework starts with identifying whether the adult has free time to engage in a non-mandatory activity. To do this, it is checked whether the travel time to next fixed activity is less than the time available in the open time-space prism for the individual. If so, an activity type choice model determines the activity that will be pursued by the individual and a subsequent activity duration model determines the amount of time the person is going to spend at the chosen activity. The vehicle allocation module starts with checking whether the person is a primary driver or not. If the person is a primary driver, vehicle tagged to the person from a primary driver allocation module will be used to make this trip. Scope for performing this activity as a joint activity with another household member is checked (this information is provided exogenously to the framework). If yes, the activity is performed as a joint activity and based on whether there are one or more primary drivers on the trip, the vehicle to be used is determined based on the logic discussed in the vehicle type choice modeling framework. If not, then the person simply proceeds to the next fixed activity in his schedule.

On the other hand, if the person is not a primary driver, then vehicle availability is checked. To do this, vehicle status of all household vehicles at that instance is checked to see if there any 'idle' vehicles to carry out this trip. Information gathered from the DTA model regarding the network status of all the vehicles is retrieved and summed across all the vehicles owned by the household. If the sum of the status variable is greater than one, it means that at least one of the household vehicles is available to be allocated to the non-primary driver in order to carry out this trip. Destination choice for the trip is determined using auto travel times. Incase none of the household vehicles are available at the given instance to carry out a trip, auto mode will be excluded from the set of available modes in the mode choice step. Destinations are sampled based on the chosen mode (walk/bike/transit). After a viable destination is chosen for the trip, it is again checked if there is enough time to engage in the activity with the chosen mode. If the mode chosen for the trip is an auto mode, scope for joint activity participation is checked. For any other mode, the trip will be carried out as a solo activity.

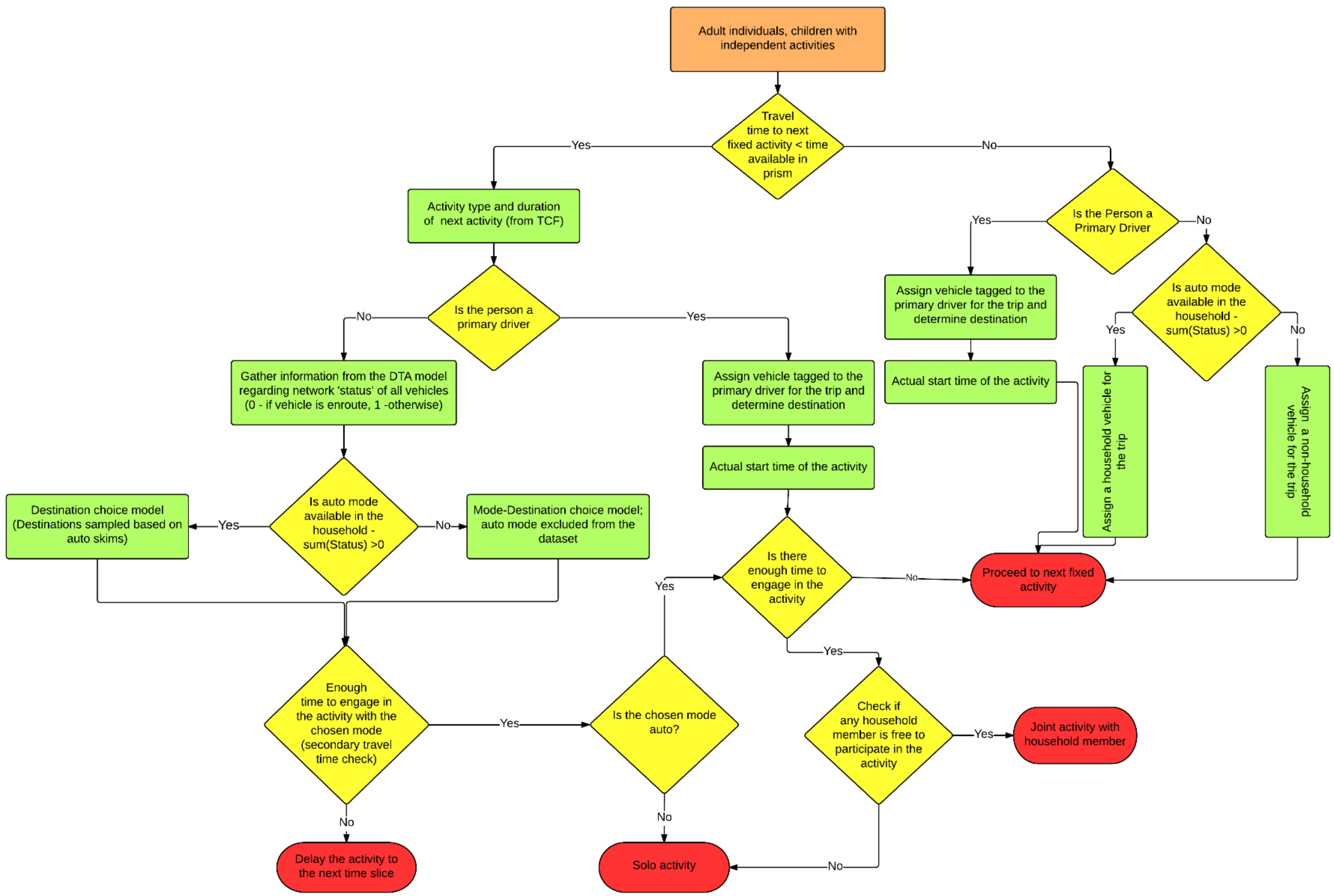

Figure 8.3. Vehicle allocation and tracking framework for adult's non-mandatory activities.

Once the destination is chosen for a trip, it is checked whether there is enough time to engage in the activity using the chosen mode. If yes, that activity is pursued by the individual. If not, then the individual can either delay the activity to the next time-slice and re-check for availability of vehicles or alter the activity and search for a new set of destinations based on the newly chosen activity. The models that need to be estimated in order to apply this framework are as listed below along with a proposed form of estimation.

i. *Model of primary driver allocation*: This is a model that allocates individuals in a given household to the vehicles such that the household's utility is maximized. In application step, this model will have a variable choice set. As soon as a driver in the household is allocated to a vehicle as the 'primary driver', he/she will be removed from the subsequent model applications so that each person is uniquely allocated to a vehicle in the household. This model is discussed in detail in the previous chapter.

ii. *Mode choice model with variable choice set*: The mode choice model being proposed in the framework is heavily influenced by the availability of auto mode. So two separate mode choice models need to be estimated, one with auto mode included in the choice set and another model without the auto mode. This model can be a simple MNL model with motorized and non-motorized modes as elemental alternatives or a nested logit model with modes grouped together by characteristic (auto vs non-auto etc.,)

iii. *Destination choice models*: The destination choice models being proposed are dependent on the mode chosen by a traveler. If the chosen mode is auto, destinations need to be sampled based on auto skims. If a non-auto mode is chosen destinations need to be sampled based on transit skims. In application mode, this model needs to be sensitive to the chosen mode and should be capable of doing intelligent sampling. Separate destination choice models will be estimated using transit and auto skims.

iv. Activity type and duration choice models: While the activity type choice is an MNL model, the duration choice model is a linear regression model.

Summary and Conclusions

The proposed vehicle allocation and tracking framework can be used as a starting point to implement vehicle accountancy in integrated model systems where an activity-based model and a dynamic traffic assignment model are in continuous communication with each other and exchange information regarding the network 'status' of all vehicles on a minute-by-minute basis. The proposed framework is comprehensive in that it covers all possible outcomes with respect to vehicle availability/unavailability in the household. The proposed frameworks 'mimics' the real world decision making process of auto-deficient households in which individuals are often constrained from engaging in activities due to unavailability of a vehicle. The concept of wedding a vehicle to a primary driver in the household is debatable as vehicle allocation might sometimes be dependent on the relative importance of the activity rather than which individual is assigned as a primary driver to the vehicle. In such a scenario, it is easy to rid the primary driver allocation module and do random assignment of household's vehicle to their travel needs and still use the same framework for vehicle accountancy. The merits/de-merits of including a primary driver allocation module and the efficacy of the model framework in application mode should be the focus of future work in this research area.

CHAPTER 9

SUMMARY AND FUTURE WORK

There has been tremendous advancement in the field of travel behavior research in the past few decades. Much progress has been made to enhance transportation planning process from the traditional four stage (trip-based) travel demand modeling to activity-based modeling methods. The intent of these advancements is to represent personal travel in a behaviorally realistic way. There have been parallel efforts on the fronts of research as well as practice in the profession to develop state-of-the art activity-based modeling systems that identify and model all the nuances in activity-travel patterns observed in day-to-day life. In this dissertation, frameworks are presented to enhance existing activity-based models on multiple fronts with the help of recent methodological progress made in the research arena. The primary objectives of this dissertation are threefold.

- Develop a tour characterization framework with an intent to enhance tour based activity model systems in practice to accommodate the continuous treatment of time.
- Develop a fleet composition modeling simulator capable of predicting the vehicle fleet owned by a household classified by body type and age with a goal better predict emission footprint of personal travel.
- Develop a tour level vehicle type choice modeling framework to advance the activity-based models currently in practice to model the exact type of vehicle used for specific types of tours rather than just modeling the mode (SOV, HOV etc.)

The National Household Travel Survey data from 2008, for the Greater Phoenix Metropolitan region was used to estimate and validate all the components of the proposed frameworks. A brief summary of each of these modeling frameworks is presented here along with possible avenues for future research.

Tour Characterization Framework

The motivation behind development of the tour characterization framework is to enhance the discrete time representation adopted in tour-based models currently in practice with an evolutionary continuous-time approach that is capable of leveraging both history of activity participation as well as anticipatory activity engagement details and determine the mix/sequence of stops undertaken on a tour. The reason for many tour-based model systems to resort to a discrete time representation is the computational burden involved in implementing a continuous-time approach. The proposed framework develops model components that are behaviorally intuitive yet computationally efficient.

The tour characterization framework consists of two components. The first component is an MDCEV model of activity type mix that predicts all of the secondary activities that an individual engages in, as a part of his/her tour. The second component is a stop sequencing model system that determines the sequence in which the activities are performed on multi-stop tours. An ‘epoch’ which is defined as the summation of travel time and activity duration is considered as the unit of analysis to facilitate continuous-time representation of activities. Model components of the tour characterization framework are presented for HBW tours and HBO tours made by non-workers. The model components are able to accurately predict the observed activity-travel patterns and show promise for implementation in a regional level activity-based travel demand model.

In the process of development and testing of the tour characterization framework, an important observation was made that in case of work tours, there is no constraint set in the modeling framework to limit the duration of activities predicted on the morning commute (journey to work). If a person starts at home at 8 am clock in the morning and reaches work at 9 am, the individual has an hour to engage in any/all activities and travel to work. Ideally the model system should predict activity epochs that sum to an hour minus the commute duration for the individual. But, this is not imposed as a hard constraint in either the MDCEV modeling methodology or the stop sequencing components of the framework. While the models developed should be able to take care of such nuances in a logically consistent fashion, it is felt prudent to handle this situation by incorporating a stop duration reconciliation module.

Future efforts should concentrate on developing a tour reconciliation module which checks the total epoch durations predicted by the model system for the outbound half tour of a HBW tour and open time available for the individual to engage in any activities. If the epoch durations exceed the available time, the activity (or activities) are moved to the inbound half tour schedule (work to home journey) instead of the outbound half tour. This is a simple yet behaviorally realistic way of handing subtlety in HBW tour scheduling. The model components developed should be validated against secondary data from regions that have similar activity travel patterns.

Vehicle Fleet Composition Framework

The vehicle fleet composition model system is developed with an intent to provide disaggregate level of information regarding the fleet composition of a region to activity-based model systems and subsequent emission modules so that the emission footprint can be accurately depicted at the regional level. Despite significant advancements in depicting activity-travel patterns of individuals, most of the activity-based model systems are still constrained to modeling the auto ownership (number of vehicles owned by the household), which is of limited use in calculating the emission footprint that results from the travel of all the household members. The proposed model system is developed in a parsimonious yet effective way to predict the vehicle fleet owned by a household (and thus the entire region) classified by vehicle body type and vintage. The model system also predicts the average annual mileage that the household puts on each of the vehicles owned. The model system is developed in open source coding platform ‘R’. The system takes household and built environment characteristics as input and predicts the vehicle fleet mix owned by the household. The components of the fleet composition modeling system are able to successfully replicate the vehicle ownership patterns classified by body type and age observed in the NHTS data for the Greater Phoenix Metropolitan Region. Future efforts should focus on validating the model components against secondary data sources (for example, data from the Department of Motor Vehicles)

Further, the fleet composition simulator presented should be enhanced to include a fleet evolution module. The evolution module should take the fleet predicted as input in the base year and evolve the vehicle acquisition, transaction and disposal of household vehicles over time. In the modeling framework, separate models are estimated and applied to predict the vehicle fleet mix and count of number of vehicles (of each type) owned by the household. There are might be common unobserved factors affecting both the consumption of alternative vehicle types and the number of vehicles owned within each vehicle type, which calls for modeling both these dimensions together. An integrated model of fleet composition and count is estimated and presented to this effect. Future research should concentrate on incorporating this integrated model into the fleet composition modeling framework.

Tour Level Vehicle Type Choice Modeling Framework

The tour level vehicle type choice framework is developed with an intent to advance the modeling methodology in activity-based modeling systems from predicting the 'mode' chosen to make a tour to predicting the exact type and age of the vehicle chosen. A limiting reason why almost none of the activity-based models in practice do not predict the type of vehicle used on a tour is due to the lack of information regarding the household's vehicle fleet to begin with. This problem is taken care of in an efficient way using the fleet composition model framework proposed in this dissertation. With this information at hand and information regarding the tour characteristics from the tour composition framework (coupled with some exogenous inputs), a novel methodology is proposed to allocate the vehicles owned by a household to the activity-travel needs of its members. The tour level vehicle type choice framework begins with a primary driver allocation model that assigns each vehicle owned by a household to a unique driver within the household. The model is run in an iterative fashion until the each driver in the household is assigned a unique vehicle (for auto sufficient households) or the household vehicles are exhausted (for auto deficient households).

The primary driver is assumed to 'only' utilize the vehicle that is assigned to him/her. This assumption resulted from a finding in the data that 90% of the tours used for analysis were made by individuals using his/her assigned vehicle. This observation is also not far from reality where individuals are usually 'wed' to their vehicles and use them for all their travel necessities. If a primary driver (an individual who has an assigned household vehicle) wants to embark on a tour, he/she simply use their assigned vehicle. If a non-primary driver wants to embark on a tour, a tour level vehicle type choice model (constrained by the household's vehicle fleet) determines the type of vehicle that will be used to engage in the trip-chain. The tour level vehicle type choice model is an MNL model with variable choice set (vehicle fleet owned by the household). The proposed model components performed exceptionally well at the aggregate as well as disaggregate levels.

While the tour level vehicle type choice framework is a good starting point to advance the activity-based models from modeling 'mode' to the specific 'vehicle type' owned by the household, it still assumes that all the vehicles are available to the household members at all times of day. This assumption can be overcome by incorporating a real-time vehicle allocation and tracking module. A framework is proposed to this effect in Chapter 8. The implementation of this framework requires an integrated modeling system where an activity-based model and a dynamic traffic assignment model communicate with each other in continuous time (every minute). Future efforts should aim at incorporating such a module alongside the tour characterization and vehicle fleet composition frameworks. The rule based heuristics adopted for the case when there is more than one primary driver on a tour need to be checked against observed data. This exercise could not be carried out as a part of the current effort as such disaggregate level of information is not readily available and extremely hard to construct from the trip level information in the NHTS data. Travel survey have the necessity to progress from trip level to tour level data collection as such information will facilitate development of robust models in the context of tour level activity-based models and provide a variety of options to validate the estimated models along multiple dimensions (particularly for joint tours). Both the primary driver allocation model and the tour level vehicle type choice model are developed as MNL models. Future efforts should explore nested logit structures for both these models so that the alternatives are not correlated across the choice dimension. It should be noted that the tour level vehicle type choice modeling framework proposed as a part of this dissertation is still in its incipient stages. Future research should focus on enhancing this framework from a methodological (model structures) as well as operational (real-time vehicle availability) standpoints.

In summary, this dissertation contributes to enhance existing knowledge in activity-based modeling techniques in research as well as practice. All of the proposed frameworks have been thoroughly tested for their efficiency in replicating the observed activity-travel patterns and vehicle ownership/utilization patterns. Future research should focus on implementing this framework (as a whole or in part) in a fully operational integrated urban model system at the regional scale to represent travel behavior as realistically as possible.

REFERENCES

Adler, T., and Ben-Akiva, M. (1979). A theoretical and empirical model of trip chaining behavior. *Transportation Research Part B: Methodological*, *13*(3), 243-257.

Algers, S., Daly, A., Kjellman, P., and Widlert, S. (1996). Stockholm model system (SIMS): Application. Volume 2: Modelling Transport Systems. In *World Transport Research. Proceedings of the 7th World Conference on Transport Research*.

Arentze, T. A., and Timmermans, H. J. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological*, *38*(7), 613-633.

Axhausen, K. (1998). Can we ever obtain the data we would like to have ?. In T. Gärling, T. Laitila and K. Westin (Eds.), *Theoretical foundations of travel choice modelling* (pp. 305-323), Oxford: Pergamon Press.

Axhausen, K. W., and Gärling, T. (1992). Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. *Transport Reviews*, *12*(4), 323-341.

Barrett, C. L., Beckman, R. J., Berkbigler, K. P., Bisset, K. R., Bush, B. W., Eubank, S., ... Williams, M. D. (1999). TRANSIMS (TRansportation ANalysis SIMulation System). Volume 0: Overview. *Los Alamos National Lab., Los Alamos, NM, Rep. LA-UR-99-1658*.

Berkovec, J., and Rust, J. (1985). A nested logit model of automobile holdings for one vehicle households. *Transportation Research Part B: Methodological*, *19*(4), 275-285.

Bhat, C. R. (1997). An econometric framework for modeling the daily activity-travel pattern of individuals. Presented at the Activity-based Analysis Workshop, National Institute of Statistical Sciences (NISS), NC.

Bhat, C. R. (1999). An analysis of evening commute stop-making behavior using repeated choice observations from a multi-day survey. *Transportation Research Part B: Methodological*, *33*(7), 495-510.

Bhat, C. R. (2005). A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, *39*(8), 679-707.

Bhat, C. R. (2008). The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological*, *42*(3), 274-303.

Bhat, C. R., and Guo, J. Y. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research Part B: Methodological*, *41*(5), 506-526.

Bhat, C. R., and Pulugurta, V. (1998). A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions. *Transportation Research Part B: Methodological*, *32*(1), 61-75.

Bhat, C. R., and Sen, S. (2006). Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. *Transportation Research Part B: Methodological*, *40*(1), 35-53.

Bhat, C. R., and Singh, S. K. (2000). A comprehensive daily activity-travel generation model system for workers. *Transportation Research Part A: Policy and Practice*, *34*(1), 1-22.

Bhat, C. R., Astroza, S., Sidharthan, R., and Bhat, P. C. (2014). A multivariate hurdle count data model with an endogenous multiple discrete–continuous selection system. *Transportation Research Part B: Methodological*, *63*, 77-97.

Bhat, C. R., Guo, J. Y., Srinivasan, S., and Sivakumar, A. (2004). Comprehensive econometric microsimulator for daily activity-travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, No. 1894,* Transportation Research Board of the National Academies, Washington, D.C., 57-66.

Bhat, C. R., Paleti, R., Pendyala, R. M., Lorenzini, K., and Konduri, K. C. (2013). Accommodating immigration status and self-selection effects in a joint model of household auto ownership and residential location choice. In *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2382*, Transportation Research Board of the National Academies, Washington, D.C., 142-150.

Bhat, C. R., Srinivasan, S., and Sen, S. (2006). A joint model for the perfect and imperfect substitute goods case: application to activity time-use decisions. *Transportation Research Part B: Methodological*, *40*(10), 827-850.

Bowman, J. L., and Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, *35*(1), 1-28.

Brownstone, D., Bunch, D. S., and Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B: Methodological*, *34*(5), 315-338.

Cao, X., Handy, S. L., and Mokhtarian, P. L. (2006). The influences of the built environment and residential self-selection on pedestrian behavior: evidence from Austin, TX. *Transportation*, *33*(1), 1-20.

Castro, M., Paleti, R., and Bhat, C. R. (2012). A latent variable representation of count data models to accommodate spatial and temporal dependence: application to predicting crash frequency at intersections. *Transportation Research Part B: Methodological*, *46*(1), 253-272.

Cervero, R., and Kockelman, K. (1997). Travel demand and the 3Ds: density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199-219.

Cervero, R., and Seskin, S. (1995). An evaluation of the relationships between transit and urban form. *Transit Cooperative Research Program*, *Transportation Research Board*, National Research Council, Washington, D.C.

Chapin, F.S. (1974). *Human activity patterns in the city: Things people do in time and space*. New York, NY: John Wiley and Sons.

Choo, S., and Mokhtarian, P. L. (2004). What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice. *Transportation Research Part A: Policy and Practice, 38*(3), 201-222.

de Jong, G. (1996). A disaggregate model system of vehicle holding duration, type choice and use. *Transportation Research Part B: Methodological, 30*(4), 263-276.

EIA, *Oil: Crude and Petroleum Products Explained*. (November 5, 2013). Retrieved August 15, 2014 from http://www.eia.gov/energyexplained/index.cfm?page=oil_home#tab2.

Eluru, N., Bhat, C. R., Pendyala, R. M., and Konduri, K. C. (2010). A joint flexible econometric model system of household residential location and vehicle fleet composition/usage choices. *Transportation*, *37*(4), 603-626.

EMFAC. *California Department of Transportation*. Retrieved April 10, 2014 from http://www.dot.ca.gov/hq/env /air/pages/emfac.htm.

EPA, *Sources of Greenhouse Gas Emissions*. (July 22, 2014). Retrieved August 16, 2014 from http://www.epa.gov/climatechange/ghgemissions/sources/transportation.html.

Ettema, D. F., Borgers, A. W. J., and Timmermans, H. J. P. (1993). A simulation model of activity scheduling behavior. In *Transportation Research Record, Journal of the Transportation Research Board, No. 1413*, Transportation Research Board of the National Academies, Washington, D.C., 1-11.

Ewing, R. H. (2008). Characteristics, causes, and effects of sprawl: a literature review. In *Urban Ecology* (519-535). Springer, US.

Ewing, R. H., and Cervero, R. (2001). Travel and the built environment: a synthesis. *Transportation Research Record: Journal of the Transportation Research Board, No. 1780*, Transportation Research Board of the National Academies, Washington, D.C., 87-114.

Frank, L. D., and Pivo, G. (1994). Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transportation*

Research Record: Journal of the Transportation Research Board, *No. 1466*, Transportation Research Board of the National Academies, Washington, D.C., 44-52.

Gärling, T., Brännäs, K., Garvill, J., Golledge, R. G., Gopal, S., Holm, E., and Lindberg E., (1989). Household activity scheduling, In *Transport Policy, Management and Technology Towards 2001, Vol. IV, Contemporary Developments in Transport Modeling, Western Periodicals* (pp. 235-248), Ventura, CA.

Gärling, T., Kwan, M. P., and Golledge, R. G. (1994). Computational-process modelling of household activity scheduling. *Transportation Research Part B: Methodological*, *28*(5), 355-364.

Golob, T. F., and Burns, L. D. (1978). Effects of transportation service on automobile ownership in an urban area. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 673*, Transportation Research Board of the National Academies, Washington, D.C., 137-145.

Golob, T. F., and Van Wissen, L. (1989). A joint household travel distance generation and car ownership model. *Transportation Research Part B: Methodological*, *23*(6), 471-491.

Golob, T. F., Bunch, D. S., and Brownstone, D. (1997). A vehicle use forecasting model based on revealed and stated vehicle type choice and utilisation data. *Journal of Transport Economics and Policy*, *31*(1), 69-92.

Golob, T. F., Kim, S., and Ren, W. (1996). How households use different types of vehicles: a structural driver allocation and usage model. *Transportation Research Part A: Policy and Practice*, *30*(2), 103-118.

Goulias, K., and Kitamura, R. (1989). Recursive model system for trip generation and trip chaining. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1236*, Transportation Research Board of the National Academies, Washington, D.C., 59-66.

Goulias, K., Pendyala, R. M., and Kitamura R. (1990). Practical method for the estimation of trip generation and trip chaining, *Transportation Research Record: Journal of the*

Transportation Research Board, *No. 1285*, Transportation Research Board of the National Academies, Washington, D.C., 47-56.

Hägerstraand, T. (1970). What about people in regional science? *Papers in regional science*, *24*(1), 7-24.

Hamed, M. M., and Mannering, F. L. (1993). Modeling travelers' postwork activity involvement: toward a new methodology. *Transportation Science*, *27*(4), 381-394.

Hensher, D. A. (1985). An econometric model of vehicle use in the household sector. *Transportation Research Part B: Methodological*, *19*(4), 303-313.

Hensher, D. A., and Le Plastrier, V. (1985). Towards a dynamic discrete-choice model of household automobile fleet size and composition. *Transportation Research Part B: Methodological*, *19*(6), 481-495.

Hocherman, I., Prashker, J. N., and Ben-Akiva, M. (1983). Estimation and use of dynamic transaction models of automobile ownership. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 944*, Transportation Research Board of the National Academies, Washington, D.C., 134-141.

Janssens, D., Wets, G., Brijs, T., Vanhoof, K., and Timmermans, H. J. P. (2003). Identifying behavioral principles underlying activity patterns by means of bayesian networks. Presented at the 82nd Annual Meeting of the Transportation Research Board, Washington, D.C.

Joh, C. H., Arentze, T. A., and Timmermans, H. J. P. (2001). A position-sensitive sequence-alignment method illustrated for space-time activity-diary data. *Environment and Planning A*, *33*(2), 313-338.

Jones, P. M., Koppelman, F. S., and Orfeuil, J. P. (1990). Activity analysis: State-of-the-art and future directions. *Developments in dynamic and activity-based approaches to travel analysis*, 34-55.

Kain, J. F., and Fauth, G. R. (1978). The impact of urban development on auto ownership and transit use. *Real Estate Economics*, *6*(3), 305-326.

Kitamura, R. (1988). An evaluation of activity-based travel analysis.*Transportation*, *15*(1-2), 9-34.

Kitamura, R., and Fujii, S. (1998). Two computational process models of activity-travel behavior. In T. Gärling, T. Laitala, and K. Westin (Eds.), *Theoretical foundations of travel choice modeling* (pp. 251-279). Oxford: Pergamon Press.

Kitamura, R., and Kermanshah, M. (1984). A sequential model of interdependent activity and destination choice. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 987*, Transportation Research Board of the National Academies, Washington, D.C., 81-89.

Kitamura, R., Fujii, S., and Otsuka, Y. (1996). An analysis of induced travel demand using a production model system of daily activity and travel which incorporates time-space constraints. Presented at the Fifth World Congress of the Regional Science Association International, Tokyo, Japan.

Kitamura, R., Golob, T. F., Yamamoto, T., and Wu, G. (2000). Accessibility and auto use in a motorized metropolis. Paper presented at the 79th Annual Meeting of the Transportation Research Board, Washington, D.C.

Kitamura, R., Lula, C. V., and Pas, E. I. (1993). AMOS: An activity-based, flexible and truly behavioral tool for evaluation of TDM measures. Proceedings of the 21st Summer Annual Meeting: Transportation Planning Methods (pp. 283-294). London: PTRC Education and Research Services, Ltd.

Konduri, K. C., Ye, X., Sana, B., and Pendyala, R. M. (2011). Joint model of vehicle type choice and tour length. *Transportation Research Record: Journal of the Transportation Research Board, No. 2255*, Transportation Research Board of the National Academies, Washington, D.C, 28-37.

Krizek, K. J. (2003). Neighborhood services, trip purpose, and tour-based travel. *Transportation*, *30*(4), 387-410.

Kwan, M. P. (1997). GISICAS: An activity-based travel decision support system using a GIS-interfaced computational process model. In: D. F. Ettema, and H. J. P Timmermans (Eds.), *Activity-based approaches to activity analysis* (pp. 263-282). Oxford: Pergamon Press.

Lave, C. A., and Train, K. (1979). A disaggregate model of auto-type choice. *Transportation Research Part A: General*, *13*(1), 1-9.

Lerman, S., and Ben-Akiva, M. (1976). *A behavioral analysis of automobile ownership and mode of travel, Vol. 2. Technical Summary*. Washington, D.C.: U.S. Department of Transportation.

Mannering, F. (1983). An econometric analysis of vehicle use in multivehicle households. *Transportation Research Part A: General*, *17*(3), 183-189.

Mannering, F., and Winston, C. (1985). A dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics*, *16*(2), 215-236.

Manski, C. F., and Sherman, L. (1980). An empirical analysis of household choice among motor vehicles. *Transportation Research Part A: General*, *14*(5), 349-366.

McGuckin, N., Zmud, J., and Nakamoto, Y. (2005). Trip-chaining trends in the United States: understanding travel behavior for policy making. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1917*, Transportation Research Board of the National Academies, Washington, D.C., 199-204.

McNally, M. G. (1995). An activity-based microsimulation model for travel demand forecasting. Presented at the Conference on Activity-Based Approaches, Eindhoven University of Technology, Eindhoven, Netherlands.

Mensah, J. (1995). Journey to work and job search characteristics of the urban poor. *Transportation*, *22*(1), 1-19.

Miller, E. J., and Roorda, M. J. (2003). A prototype model of 24-hour household activity scheduling for the Toronto area. In *Transportation Research Record*, *Journal of the Transportation Research Board*, *No. 1831*, Transportation Research Board of the National Academies, Washington, D.C., 114-121.

Miller, E. J., Roorda, M. J., and Carrasco, J. A. (2005). A tour-based model of travel mode choice. *Transportation*, *32*(4), 399-422.

Mohammadian, A., and Miller, E. J. (2003). Empirical investigation of household vehicle type choice decisions. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1854*, Transportation Research Board of the National Academies, Washington, D.C., 99-106.

MOVES, *Motor Vehicle Emission Simulator*. Retrieved April 10, 2014 from http://www.epa.gov/otaq/models/moves/.

Musti, S., and Kockelman, K. M. (2011). Evolution of the household vehicle fleet: anticipating fleet composition, PHEV adoption and GHG emissions in Austin, Texas. *Transportation Research Part A: Policy and Practice*, *45*(8), 707-720.

My Car Stats. *Manufacturer Car Warranty Expiration Chart*. (October, 2010). Retrieved August 18, 2014 from http://www.mycarstats.com/content/car_warrantylimits.asp.

National Household Travel Survey. *Derived Total HH Income*. (2009). Retrieved August 2, 2014 from http://nhts.ornl.gov/tables09/CodebookPage.aspx?id=949.

NHTS User's Guide, U.S. Department of Transportation (2009). Retrieved August 31, 2014 from http://nhts.ornl.gov/2009/pub/UsersGuideV2.pdf.

NHTS, Codebook Browser. Retrieved September 27, 2014 from: http://nhts.ornl.gov/tables09/CodebookBrowser.aspx

Paleti, R., Bhat, C. R., Pendyala, R. M., and Goulias, K. G. (2013). Modeling of Household Vehicle Type Choice Accommodating Spatial Dependence Effects. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2343*, Transportation Research Board of the National Academies, Washington, D.C., 86-94.

Paleti, R., Eluru, N., Bhat, C. R., Pendyala, R. M., Adler, T. J., and Goulias, K. G. (2011a). Design of comprehensive microsimulator of household vehicle fleet composition, utilization, and evolution. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2254*, Transportation Research Board of the National Academies, Washington, D.C., 44-57.

Paleti, R., Pendyala, R. M., Bhat, C. R., and Konduri, K. C. (2011b). A joint tour-based model of tour complexity, passenger accompaniment, vehicle type choice, and tour length. Paper presented at the 90th Transportation Research Board Annual Meeting, Washington, D.C.

Parsons Brinckerhoff. *MAG CT-RAMP Activity Based Model Phase I, Prepared for Maricopa Association of Governments*. (2010). Retrieved August 24, 2014 from https://www.azmag.gov/Documents/TRANS_2010-12-06_MAG-CT-RAMP-Activity-Based-Model-PhaseI-Model-Estimation-Results.pdf.

Pendyala, R. M., Bhat, C. R., Goulias, K. G., Paleti, R., Konduri, K. C., Sidharthan, R., Hu, H.-H., Huang, G., and Christian, K. P. (2012a). Application of socioeconomic model system for activity-based modeling. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2303*, Transportation Research Board of the National Academies, Washington, D.C., 71-80.

Pendyala, R. M., Kitamura, R., and Prasuna Reddy, D. V. G. (1998). Application of an activity-based travel-demand model incorporating a rule-based algorithm. *Environment and Planning B*, *25*(5), 753-772.

Pendyala, R. M., Konduri, K. C., Chiu, Y. C., Hickman, M., Noh, H., Waddell, P., You, D., and Gardner, B. (2012b). Integrated land use-transport model system with dynamic time-dependent activity-travel microsimulation. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2303*, Transportation Research Board of the National Academies, Washington, D.C., 19-27.

Pinjari, A. R., and Bhat, C. R. (2011). An efficient forecasting procedure for kuhn-tucker consumer demand model systems: application to residential energy consumption analysis. Technical Paper, Department of Civil & Environmental Engineering, University of South Florida.

Pinjari, A. R., Eluru, N., Bhat, C. R., Pendyala, R. M., and Spissu, E. (2008a). Joint model of choice of residential neighborhood and bicycle ownership: accounting for self-selection and unobserved heterogeneity. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2082*, Transportation Research Board of the National Academies, Washington, D.C., 17-26.

Pinjari, A. R., Eluru, N., Srinivasan, S., Guo, J. Y., Copperman, R. B., Sener, I. N., and Bhat C. R., (2008b). CEMDAP: Modeling and microsimulation frameworks, software development, and verification. Proceedings of the *87th* Annual Meeting of the Transportation Research Board, Washington D.C.

Recker, W. W., McNally, M. G., and Root, G. S. (1986a). A model of complex travel behavior: Part I—Theoretical development. *Transportation Research Part A: General*, *20*(4), 307-318.

Recker, W. W., McNally, M. G., and Root, G. S. (1986b). A model of complex travel behavior: Part II—an operational model. *Transportation Research Part A: General*, *20*(4), 319-330.

Roorda, M. J., Miller, E. J., and Habib, K. (2008). Validation of TASHA: A 24-h activity scheduling microsimulation model. *Transportation Research Part A: Policy and Practice*, *42*(2), 360-375.

Shiftan, Y. (1998). Practical approach to model trip chaining. *Transportation Research Record: Journal of the Transportation Research Board, No. 1645*, Transportation Research Board of the National Academies, Washington, D.C., 17-23.

Train, K. (1986). *Qualitative choice analysis: Theory, econometrics, and an application to automobile demand* (Vol. 10). Cambridge, MA: MIT press.

Vovsha, P., and Bradley, M. (2006). Advanced activity-based models in context of planning decisions. *Transportation Research Record: Journal of the Transportation Research Board, No. 1981*, Transportation Research Board of the National Academies, Washington, D.C., 34-41.

Vovsha, P., Freedman, J., Livshits, V., and Sun, W. (2001). Design features of activity-based models in practice: Coordinated travel-regional activity modeling platform. *Transportation Research Record: Journal of the Transportation Research Board, No. 2254*, Transportation Research Board of the National Academies, Washington, D.C., 19-27.

Vyas, G., Paleti, R., Bhat, C. R., Goulias, K. G., Pendyala, R. M., Hu, H. H., Adler, T. J., and Bahreinian, A. (2012). Joint vehicle holdings, by type and vintage, and primary driver assignment model with application for California. *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2302*, Transportation Research Board of the National Academies, Washington, D.C., 74-83.

Wilson, W. C. (1998). Activity pattern analysis by means of sequence-alignment methods. *Environment and Planning A*, *30*(6), 1017-1038.

World Research Institute, *CATI 2.0: WRIs Climate Data Explorer*. (July 2014). Retrieved September 17, 2014 from http://cait2.wri.org/wri/Country%20GHG%20Emissions?indicator[]=Total%20GHG%20Emissions%20Excluding%20Land-Use%20Change%20and%20Forestryandindicator[]=Total%20GHG%20Emissions%20Including%20Land-Use%20Change%20and%20Forestryandyear[]=2011andchartType=geo.

BIOGRAPHICAL SKETCH

Venu Madhav Garikapati was born on 6th February 1986 to Subrahmanyam and Kamala in Gudivada, India. He received high school and college education in his native state, Andhra Pradesh and pursued his undergraduate studies in Civil Engineering at Bapatla Engineering College, Bapatla, India. Towards the end of his bachelor's program, he started developing an interest towards transportation engineering. Upon completing his undergraduate degree, he secured a seat in Indian Institute of Technology, Bombay to pursue his graduate studies where his interest in transportation engineering further got streamlined to travel behavior modeling. After his master's studies, he worked as a transportation modeler for one and a half years. He started his doctoral study at Arizona State University in Spring 2011 under the guidance of Dr. Ram M. Pendyala. In addition to academics, he was actively involved in the ASU student chapter of the Institute of Transportation Engineers and served two terms on the graduate student assembly representing the School of Sustainable Engineering and the Built Environment. He was awarded the Graduate Student Service award for his services to the graduate student community. He won the first place in Civil engineering graduate student research symposium at ASU for the year 2013-14. His research interests include activity-based modeling, transportation planning, policy analysis and travel demand modeling.