MICROSIMULATION OF ACTIVITY PARTICIPATION, TOUR COMPLEXITY, AND MODE CHOICE WITHIN AN ACTIVITY-BASED TRAVEL DEMAND MODEL SYSTEM

by

Naznin Sultana Daisy

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
March 2018

"I dedicate this dissertation to my late father,
To a father, who showed me what it means to be a daughter"

Table of Contents

List of Tables

List of Figures

Abstract

Over the past few decades, trip-based travel demand approaches have been replaced by activity-based microsimulation travel demand techniques, which are able to capture the latent demand for activity participation, interdependency among trips, and household interactions. Activity-based models consider trips as a derived demand which arise from activity engagement behavior. This research aims to depict the daily activity-travel behavior of travelers as a result of choice decision making processes through the development of the Scheduler for Activities, Locations, and Travel (SALT) microsimulation travel demand model. The SALT model is comprised of five main components: population synthesis, time-use activity pattern recognition, tour mode choice, activity destination choice, and activity/trip scheduling. A series of advanced econometric micro-behavioral modules are developed to model behavioral mechanisms of different population groups in the region.

An under-recognized issue in most of the econometric activity-based models is that they treat all out-of-home travelers, whether workers or non-workers, as undifferentiated groups, decreasing the ability to predict activity-travel decisions. To this end, an advanced disaggregated modeling framework is developed that can derive separate utility functions for both in-home and out-of-home activities for travelers with heterogeneous daily-activity patterns, along with simulation of correlation matrices. Additionally, a cluster-based technique is developed to model trip chaining, tour complexity, and tour mode choice of worker and non-worker clusters. These models capture associations between socio-demographics characteristics, trip attributes, and land use patterns in order to predict travel tour incidence and type, and mode choice. For empirical analysis of activity-travel behavior this study employs data from the large Halifax Space Time Activity Research (STAR) household time-use and travel survey, which consists of GPS-verified data for 2,778 person-days.

This study also contributes by designing and conducting the first Canadian university-based travel-diary survey (EnACT), to better understand activity-travel patterns and trip making frequencies of university commuters. In addition, a synthetic pseudo-panel modeling framework is developed to explore the longitudinal activity-travel behavior of urbanities. In summary, the disaggregated modeling framework presented in this study is useful for deeper understanding of individuals' activity-travel decisions, and may be operationalized to examine sensitive policy issues such as transportation control measures and congestion-pricing.

List of Abbreviations and Symbols Used

ADAPTS	Agent-based Dynamic Activity Planning and Travel Scheduling
AIC	Akaike Information Criteria
ALBATROSS	A Learning-Based Transportation-Oriented Simulation System
AMOS	Activity MObility Simulator
AT	Active Transportation
BIC	Bayesian Information Criterion
CARLA	Combinatorial Algorithm for Rescheduling Lists of Activities
CART	Classification and Regression Trees
CATI	Computer-Assisted Telephone Interviewing
CDF	Cumulative Density Function
CEMDAP	Comprehensive Econometric Microsimulator for Activity-Travel Patterns
CMA	Census Metropolitan Area
C-MVP	Cluster-based Multivariate Probit Model
CO	Combinatorial Optimization
CPM	Computational Process Models
DAP	Daily Activity Pattern
EnACT	Environmentally Aware Travel Diary Survey
FAMOS	Florida Activity Mobility Simulator
FBS	Fitness Based Synthesis
GHK	Geweke-Hajivassiliou-Keane
GPS	Global Positioning System
GSS	General Social Survey
HAPP	Household Activity Pattern Problem

HEH	Home-Entertainment-Home
HGH	Home-Organizational/Hobbies-Home
HOT	High Occupancy Toll
HOV	High Occupancy Vehicle
HPH	Home-Shopping-Home
HRM	Halifax Regional Municipality
HSH	Home-School-Home
HTH	Home-Sports-Home
H-W-P-H	Home-Work & Shopping-Home
H-W-S-H	Home-Work & School-Home
H-W-G-H	Home-Work & Organizational-Home
H-W-G-H	Home-Work & Entertainment-Home
H-S-P-H	Home-School & Shopping-Home
H-S-G-H	Home-School & Organizational-Home
H-P-G-H	Home-Shopping & Organizational-Home
H-P-E-H	Home-Shopping & Entertainment-Home
H-P-T-H	Home-Shopping & Sports-Home
H-E-G-H	Home-Entertainment & Organizational-Home
H-S-E-H	Home-Entertainment & School-Home
H-T-G-H	Home-Sports & Entertainment-Home
H-T-E-H	Home-Sports & Organizational-Home
ICT	Information and Communication Technologies
IID	Identically and Independently Distributed
IPF	Iterative Proportional Fitting
IPU	Iterative Proportional Updating

LCAH	Latent Class Accelerated Hazard
MCMC	Markov Chain Monte Carlo
MNL	Multinomial Logit
MVN	Multivariate Normal
NB	Negative Binomial
NHTS	National Household Travel Survey
PCATS	Prism-Constrained Activity Travel Simulator
PDF	Probability Density Function
SALT	Scheduler for Activities, Locations, and Travel
SFCTA	San Francisco County Transportation Authority
SMASH	Simulation Model of Activity Scheduling Heuristics
STAR	Space-Time Activity Research
STARCHILD	Simulation of Travel/Activity Responses to Complex Household Interactive Logistic Decisions
TASHA	Travel Activity Scheduler for Household Agents
TAZ	Traffic Analysis Zone
TDM	Travel Demand Management
TESS	Transportation and Environmental Simulation Studies
TOD	Time-Of-Day
TRANSIMS	Transportation Analysis and Simulation System
TURP	Time Use Research Program
W/S	Work/School Activity
ZINB	Zero-Inflated Negative Binomial
ZIP	Zero-Inflated Poisson
a	Selected Individual
b	Iteration Number

i	Observations
k	Largest Possible Count Value
k	Latent Class
n	Panel Units
s	Index Representing the Various Cells in the Count Table
t	Time Period
u	Index Representing Both Count and Control Tables
Ω	Correlation Matrix
α	Vector of Regression Parameters
θ	Dispersion Parameter
λ	Scale Parameter
$\Pr(Y_n)$	Probability of Y_n Tours Performed by the N Th Individual
$\Gamma(.)$	Gamma Function
Φ	Density Function
$E[y_k]$	Expected Number of Trips by Mode
G	Vuong Statistic
P	Shape Parameter
$P(y_k)$	Probability of Numbers of Trips
R	Sample Size
U	Total Number of Both Count and Control Tables
$Var[y_k]$	Variance of The Number of Trips
X	Vector of Explanatory Variables
Z	Parametric Hazard Model
$\overline{W}_{ct}$	Vector of Time Variant Variables

$\overline{X}_{ct}$	Mean Value for All Individuals Classified into Cohort c at Time Period t
$\overline{Z}_c$	Vector of Time Invariant Variables
$[l]_{mn}$	Lower Triangular Matrix
$\ln L$	Log Likelihood
F_U^{ab}	Fitness Value for Control Table U
JK_{us}^a	Contribution of the a^{th} Individual in the Seed Data to the s^{th} Cell in Control table u
L_i	Log-Likelihood Function
M_{us}	Amount of Cell s in Control Table u
PM_{us}^{b-1}	Value of Cell s in the Count Table u
$S_{(t)}$	Survival Function
S_q	Standard Deviation,
T_{us}^{b-1}	Difference Value Between Control and Count Tables for Cell s in Control table u
U_{in}	Systematic Utility
X_n	Vector of Explanatory Variables
Y_n	Number of Home-Based Tours
Y_n^*	Latent and Continuous Variable
Z_{in}	Vector of Observed Attributes
$f_{(t)}$	Density Function of the Weibull Distribution
$f_1(.)$	Distribution Function of the ZINB Distribution
$f_2(.)$	Distribution Function of the Parent NB Distribution,
f_{in}	Choice Indicator
g^u	Selected Individual Type According to the Fitness Value

km_b	Set of Selected Individuals for Adding into the Count Tables and Synthesized population list
lm_b^{us}	Selected Individual for the Cell s in the Count Table u in the Iteration b
u_{nt}	Composite Error Term
y_m^*	Binary Dependent Variables
α_m	Conformable Parameter Vectors
β_{in}	Corresponding Parameter of Vector Z_{in}
β_{nt}	Fixed Individual Effects
ε_n	Error Term
ε_{r1}	Random Number Generator
θ_k	Cutoff Points
$\lambda_{(t)}$	Weibull Distribution
λ_k	Poisson Parameter for Observation k
μ_n	Expected Value of Y_n
φ_k	Unknown Parameter Vector for Each Class k
ϵ_m	Random Errors

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Lei Liu. A true and caring mentor, who created such an excellent research environment that stimulated my PhD to be one of the best and most productive periods of my life. Dr. Liu's kind words helped me to strive for continuous progress in everything I do, both professionally and personally. Dr. Liu, thank you for allowing me to work and contribute to the development of the teaching materials design and teaching in both Optimization Methods course and Environmental Assessment and Management course.

I would also like to thank Professor Hugh Millward for his extreme patience, support, and time towards my research works. This research work wouldn't be possible without the continuous support and guidance of Dr. Millward. Dr. Millward, thank you for spending many hours with me to improve this work, and for providing me with the STAR activity travel diary data. I would also like to thank Professor Brian Baetz and Professor Mysore Satish for providing review comments to improve the work presented here.

I want to thank Professor Marty Leonard from the bottom of my heart for showing me the meaning of being compassionate and passionate for the things we dreamed for. I also would like to thank all the board members of DAGS, DSU from 2014-2016 academic years. This journey of four and half years wouldn't be possible without my friends in Halifax. My heartfelt thanks to my friends, Amy, Hannah, Mohammad Hesam, Levi, Siobhan, Cara, Niki, Bria, Lauren, Marc, Kym, and Sean for being there when I needed them the most. I want to thank my colleague and friend, Dr. Mohammad Hesam, for all the research ideas we have generated and executed together over the past few years.

I would like to thank Nova Scotia Entrance Scholarship (NSGS), Natural Sciences and Engineering Research Council (NSERC), Faculty of Graduate Studies (FGS), Nova Scotia Research and Innovation Graduate Scholarship, and Faculty of Engineering for funding this research. I would like to acknowledge the Transportation and Environmental Simulation Studies (TESS) research group for providing the feedback at different stages of my research. I also would like to acknowledge that the data for this research were provided by the Halifax STAR Project, supported through the Atlantic Innovation Fund from the Atlantic Canada Opportunities Agency, Project No.181930.

Last but not the least, I want to thank my family and relatives. I want to thank my eldest sister Nilufar and elder sister Rabea for always being there for me like a guardian angel. I want to thank my younger sister Nigar and my youngest brother Fahim to allow me to be their high school and college mentor, and for spurring my educator self to me. I want to thank my five years old niece for always entertaining me with her questions and comments! I want to thank my hundred-crossed grandfather for telling me again and again to enjoy the life. I want to thank my grandma for feeding me throughout my life especially when I used to study. I want to thank my mother for being the most loyal listener of my random talks, for being my strength and truest friend over this four and half years. My loudest thanks to the biggest inspiration of my life, my late father who always wanted to see me as a PhD, who taught me to do right and to stand for right, who will always be my guiding star, a father who was my wings, a father who showed me the meaning of being a daughter... Father, together we made it!

Chapter 1 Introduction

1.1 Background

Rapid urbanization, population growth, and technological progress have increased the level of complexity in travel behavior. Expansion of the transportation network with more diverse alternative transport modes and accessibility to perform activities remotely have changed the activity participation pattern of household members. Suburban development patterns, flexible work hours, and increasing participation in out-of-home activities are making the travel patterns of individuals more complex than ever. This demand for activity participation arises from the need/desire to complete spatially separated activities, for instance, work, study, shopping, recreation, and social interactions (Chapin 1974; Becker 1976). These realities demand a disaggregated travel demand modeling system that can more realistically mimic and predict the travel demands of different population segments in the region, both at the individual and household levels (Ortuzar and Willumsen 1994). This research aims to depict the daily activity-travel behavior of travelers as a result of choice decision making processes through the development of a new comprehensive cluster-based disaggregated modeling framework that is based on the activity-based travel demand modeling concepts and theories. A series of advanced econometric micro-behavioral modules are developed to model and micro-simulate short-term behavioral mechanisms of population groups with homogeneous activity patterns in the region, including activity participation, tour complexity, and mode choice.

Over the past few decades, transportation professionals have witnessed a paradigm shift from the aggregated travel demand models to disaggregated activity-based models for travel demand management and planning. A disaggregated analysis of travel demand can lead to better plans, policies, and demand management systems that can improve the overall transportation system as a more efficient and sustainable system, both environmentally and economically (Ben-Akiva and Lerman 1985).

Dissatisfaction with the traditional travel demand models (i.e. four-stage and trip-based models), along with the increasing availability of more detailed data and model estimation methods and tools (advanced econometric and machine learning techniques), have triggered the shift from aggregated travel demand models to disaggregated activity-based model (Goulias and Kim 2005). The fundamental idea of the activity-based model is that travel is a derived demand that results from participation in different spatially separated out-of-home activities (Manheim 1979). Generally, activity-based travel demand models are expected to provide better behavioral realism, as they incorporate the interdependence of travel decisions undertaken by individuals and households along with resource allocation, time-use pattern, share of household responsibilities, and joint activity participation (Jones, Koppelman and Orfeuil 1990; Arentze and Timmermans 2000). The activity-based approach postulates activity-travel patterns as an outcome of a sequence of inter-dependent decisions made by individuals and households constrained by space and time dimensions (Hagerstrand 1970). In comparison to the single-facet travel behavior models (e.g. departure time, mode choice, or route choice), transport modelers have developed more disaggregated and comprehensive activity-based models that can simulate multi-faceted daily activity-travel patterns. Some features of activity-based models, including activity-travel patterns, time allocation, scheduling decisions, and household interactions, have received more attention. These features are more relevant for applications in the transportation planning and policy. Conversely, the transportation management system deals more with short-term (typically daily) features of activity-travel patterns. In the short-term, activity-travel patterns are uncertain due to vagaries in the transportation system and urban environment, most importantly, uncertain individual travel decisions.

Transportation planners and engineers have a significant role in shaping the way people travel, through designing and planning a better transportation system that can accommodate current and future travel demands. Moreover, transport modelers are trying to better understand the peoples' travel behavior to improve the estimation accuracies of modeling travel demands to better predict travel decisions. To lessen these uncertainties related to short-term travel decisions of travelers, accurate and detailed travel information is one of the key requirements.

Thus, disaggregated activity-based modeling technique that could model and micro-simulate activity-travel behavior mechanism of distinct population groups, rather than considering them in aggregation, would be more desirable for forecasting travel demand and comprehensive evaluation of the related sensitive policy issues, such as analyzing transportation control measures and congestion-pricing.

1.2 The Activity-Based Approach

The activity-based approach is the latest generation of travel demand models. The first generation started with modeling at the aggregated level (also known as four-stage models), which forecast travel demand based on the observed relations for groups of travelers or, on average, relations at a zonal level (Ortuzar and Willumsen 1994). Some of the major limitations of the aggregated modeling approach are the lack of integrity, which refers to consistencies and congruence among various sub-models (i.e. destination sub-model and trip assignment sub-models), lack of interdependencies (i.e. dependency between trips within a trip chain, or between trips undertaken on a given day), higher temporal (i.e. considers only peak and off-peak periods) and spatial aggregation (i.e. trip origins and destinations are considered as a single entity in space), and lack of behavioral realism (i.e. do not consider constraints on individuals or households behavior or choice mechanism) (Rasouli and Timmermans 2013). Therefore, there was a gradual transition from modeling travel demand as an aggregated entity to a disaggregated one. The second generation of travel demand models, known as disaggregated trip-based models, utilized individuals or households, unlike the traffic analysis zones used in the aggregated approach, as the unit of modeling. The disaggregated approach examined the movements of individuals in time and space, as they participated in different activities over a 24-hour day or longer periods. In the late 1960s, nearly 28000 time-use diaries were collected from 12 countries to better understand the activity-travel behavior of individuals (Szalai 1972). Travel diaries collected the detailed timing, duration, sequencing, and purposes of trips. The data collected from travel diaries allowed transportation modelers to analyze the movement patterns of individuals or households within zones, whereas aggregate models considered only trips between the zones. The practice of determining trip generation by zones was replaced by trip rates for different household types, and gravity models based on zonal attractiveness were replaced by destination choice logit models.

The latest generation of travel demand models is known as activity-based models. Modeling travel demand from an activity-based perspective originated from the principle that the need for travel is derived from the need for activity participation (Hagerstrand 1970; Ellegard and Syedin 2012). The last few decades have witnessed the evolution of activity-based models from conceptual ones to operational ones and the paradigm shift from the traditional travel demand models to the activity-based approach. Currently, many cities and regions in North America (mainly in the United States), Europe, and Japan have developed activity-based travel demand models and use them in practice. Figure 1.1 illustrates a typical activity-based model structure for modeling short-term and long-term mobility decisions.

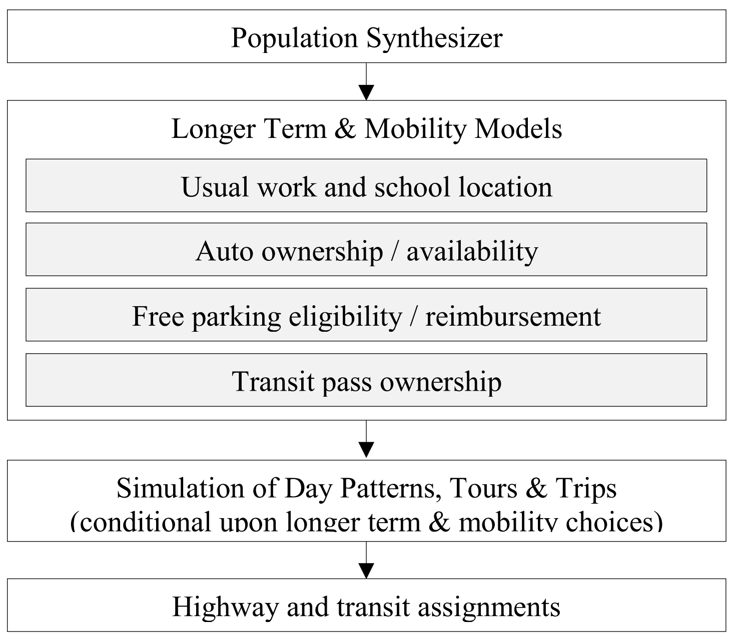

Figure 1.1 Short-term and longer-term and mobility choice models in an activity-based model (Castiglione et al. 2015, p.103)

The common components of most activity-based travel demand models are population synthesizer, activity generator, activity scheduler, tour and trip destination choice, tour and trip time of day, tour and trip mode choice, and network assignment. In comparison to earlier travel demand models, an activity-based model can provide a reasonable behavioral basis to assess the potential travel responses of travelers to sensitive policy actions by examining the process of modification of activity participation (Bhat and Koppelman 1999; Pendyala and Goulias 2002; Arentze and Timmermans 2004).

In general, activity-based models can be classified based on their modeling approach into some major categories: constraint-based models, rule-based models, hybrid models, and econometric models. Constraint-based prototypes were developed based on the seminal research of activity-participation under time-space constraints. An example of constraint-based prototypes is CARLA (Jones et al. 1983). The rule-based approach, often referred to as Computational Process Models (CPM), uses heuristic decision rules (if A, then B) to mimic the underlying decision-making process and to make decisions about activity participation and travel (Jones et al. 1983). STARCHILD (Recker 1986) was the first rule-based modeling framework and was later extended as a mathematical Household Activity Programming problem (HAPP). SCHEDULER (Garling et al. 1989), SMASH (Ettema et al. 1993), AMOS (Kitamura et al. 1993; Pendyala et al. 1998), and ALBATROSS (Arentze et al. 1999; Arentze and Timmermans 2004) are some other examples of rule-based activity-based models. Hybrid activity-based models are those that use more than one concept for the computational process, production rules, and utility maximization to handle the different behavioral and temporal-spatial constraints in modeling different facets of activity-travel patterns. TASHA (Miller and Roorda 2003) and ADAPTS (Auld and Mohammadian 2009) are two well-known examples of hybrid models.

To date, activity-based travel demand modeling frameworks that adopt the econometric approach are among the most comprehensive and well-developed ones. Generally, there are two streams of econometric models: 1) with individual daily activity patterns and 2) with coordinated daily activity patterns. The Portland Metro Model (Bowman 1998), San Francisco County Transportation Authority (SFCTA) Model (Bradley et al. 2001), SACOG (Bradley et al. 2007), Denver (DRCOG) Model (Sabina and Rossi 2006), Seattle (PSRC) Model (Nichols et al. 2014), Jacksonville (NFTPO) Model (Lawe 2010) and Houston (HGAC) Model (Rossi et al. 2013) are some examples of econometric models with individual daily activity patterns. These models treat individuals as the unit for activity-travel decision-makings. CEMDAP is another well-known framework characterized by its unique activity generation-allocation-scheduling process and continuous representation of time through hazard-based duration (Bhat et al. 2004).

1.3 Motivations and Context of this Study

This study presents the development of the Scheduler for Activities, Locations, and Travel (SALT) travel demand microsimulation model. The work specifically relates to the development of the advanced econometric micro-behavioral modules within the SALT modeling framework. The SALT model system, which is currently under development at the Department of Civil and Resource Engineering of Dalhousie University, adopts activity-based travel demand concepts and theories. It has five core modules: population synthesis, time-use activity pattern recognition, tour mode choice, activity destination choice, and activity/trip scheduling. A combined series of advanced machine learning and econometric models are articulated in the SALT modeling framework. Initially, the SALT model system derives population clusters with homogeneous time-use activity patterns using advanced machine learning technique. To this end, this research depicts an advanced econometric cluster-based modeling framework for activity participation, trip chaining, and tour mode choice modules within the SALT system. These modules can analyze the activity-travel decision-making process of the identified worker and non-worker clusters at the finest disaggregated analysis unit. In addition, this study develops mode specific trip frequency models for university population segments and also presents the development of a cohort-based pseudo panel modeling framework for longitudinal travel behavior analysis.

1.4 Objectives and Scope

The main objective of this dissertation is to develop the Scheduler for Activities, Locations, and Travel (SALT) travel demand model. Particularly, this study addresses the decision-making process of activity-travel behavior, such as why and how an individual shall participate to a particular activity in a given day through the development of a series of advanced econometric based micro-behavioral models. It is also hypothesized that clustering daily activity patterns based on the homogeneous time-use patterns can be an efficient tool to capture uncertainty related to short-term activity-travel choices, which are latent inputs to activity-based travel demand models. Furthermore, the choice of activity participation, trip chaining, number of tours, and mode choices strongly differ between travelers based on their socio-demographic, socio-economic, and land use attributes. To achieve these goals, the following main objectives were carried out during the course of development of the SALT model system and advancement of its econometric micro-behavioral modules:

- To understand the recurrent pattern of daily activity-travel behavior, the activity participation of worker and non-worker clusters are analyzed using a new developed Cluster-based Multivariate Probit Model (C-MVP) models;
- To understand the tour complexity including trip chaining behavior and tour frequency per day of worker and non-worker clusters;
- To develop the tour mode choice of worker and non-worker groups by modeling their relationships to socio-demographics, trip attributes, and land use patterns;
- To develop mode specific trip frequency model for the university population that can be utilized to better understand the activity-travel behavior of a university community at the disaggregated level; and
- To develop a cohort based synthetic pseudo panel modeling framework to analyze longitudinal changes in activity participation and duration over longer time periods.

1.5 Thesis Structure

To examine the behavioral mechanisms of SALT's model system, this study aims to develop a series of advanced econometric micro-behavioral modules. This dissertation is organized into eleven chapters, where chapter two and chapter four to nine are prepared as independent journal paper format. A brief description of the content of each chapter is summarized below:

Chapter 1 provides the background and motivation for this research and articulates the goals and methodological considerations. Chapter 2 reviews the mainstream research works in the area of activity-based modeling and econometric models. Chapter 3 presents the data sources and methods used in this study. Chapter 4 reports the investigation of the activity-travel behavior of five out-of-home worker clusters. Chapter 5 examines the activity-travel behavior of five non-worker clusters within the SALT model. Chapter 6 investigates how socio-demographics, trip attributes, and land use patterns shape and/or predict tour complexity and mode choices for five out-of-home worker clusters. Chapter 7 presents a cluster-based approach to model trip chaining, tour complexity, and tour mode choice of five non-worker clusters within the SALT model. Chapter 8 describes the daily activity-travel behavior of undergraduate students, graduate students, faculty, and staff at a large Canadian university. Chapter 9 depicts repeated cross-sectional data in a pseudo panel data approach to investigate an individual's daily participation in out-of-home discretionary activities. Chapter 10 presents a novel approach of pseudo-panel based duration modeling of discretionary activities. Finally, chapter 11 concludes the thesis. Major conclusions of this research are drawn, limitations are discussed, and possible avenues for future research are identified.

Chapter 2 Individuals' Activity-Travel Behavior in Travel Demand Models: A Review of Recent Progress[1]

2.1 Introduction

Activity-based travel demand analysis has received considerable progress over the past decades because of its ability to examine the complexity and variability of activities an individual participates in during a given time window. In general, it examines the underlying behavioral mechanisms for individual decisions to engage in various activities at different times and in different geographical locations. More specifically, activity-based analysis attempts to understand and model the behavioral basis of why, when, where, how, and with whom an individual participates in an activity. These behavioral aspects of travel are associated to factors such as needs, preferences, habits, travel service characteristics, socio-demographic characteristics, built environment and urban form characteristics, and so on.

Interest in analyzing travel demand management policies has led to a paradigm shift, and the major focus of travel demand models has changed from aggregated-level trip-based modeling to disaggregated-level activity-based modeling. The trip-based approach fails to analyze the time-use context to predict trip-related decisions. Most trip-based methods miss the broader context within which an individual makes their travel decisions (for more details, see Kitamura 1988; Jones 1979; Axhausen et al. 2002). In contrast, the activity-

[1] A version of this chapter has been accepted:
Daisy, N. S., Millward, H., and L. Liu. Individuals' activity-travel behavior in travel demand models: A review of recent progress. Peer reviewed ASCE proceedings of the 18th COTA Conference International Conference of Transportation Professionals (CICTP). Shanghai, China., 2018.

based approach assumes travel as a derived demand, stemming from the need to participate in different activities at different times and in different geographic locations (Recker 1995). This approach focuses on the patterns of activity-travel behavior, utilizes the 24-hour day as the unit of analysis, and recognizes the complex connections between activity and travel behavior. The activity-based approach can address behavioral sensitivities to short term transport policies such as staggered working hours, congestion pricing, and ridesharing incentives.

The activity-based approach originated from Hagerstrand's (1970) time-geography theory and Chapin's (1974) study of the activity patterns of urban populations. This approach is fundamentally different from the trip-based approach (Pas 1996). For instance, the trip-based approach considers 'time' as the 'cost' of a trip whereas the activity-based approach contemplates time as a continuous component within which an individual participates in activities or in travel to participate in activities (Kurani and Lee-Gosselin 1996). The activity-based approach assumes that the time-use pattern of an individual determines the activity-travel patterns of that individual. Over the past few decades, the activity-based travel demand approach has received considerable attention, as explained in the remainder of this study. Although 'behavioral' mechanisms are better mimicked in activity-based models compared to trip-based models, however, a significant improvement in behavioral realism is still required in travel demand models. The next section examines the differences between trip-based and activity-based approaches in more detail.

2.2 Trip Based Versus Activity-Based Model

In trip-based models, travel is considered as a combination of ‘trips’, and trips are viewed as independent from each other. The approach typically ignores the associations between the choice attributes of mode, travel time, and destinations, which leads to error in trip chain prediction and in predicting the impact of policy actions on travel behavior. In contrast, in activity-based models (ABMs) ‘travel’ is viewed as a derived demand (Jones 1979; Bhat et al. 2004; and Davidson et al. 2007). This approach emphasizes ‘activity participation behavior’. Trip chaining is explicitly accommodated by using ‘tours’ in the modeling stage to predict travel patterns. Tour-based ABM models capture the interdependencies among the trips within a tour, and can also identify the linkages among other tours completed in the same day. ABMs also incorporate patterns of activity sequences in the modeling system, which can address many travel demand management (TDM) issues by predicting how individuals will change their activity participation behavior in response to certain policy actions.

Another fundamental difference between trip-based and activity-based approaches is how they consider ‘time’. Trip-based models consider ‘time’ as the cost of the trip and a day is considered as a combination of peak and off-peak hours. In contrast, the activity-based approach considers ‘time’ as a continuous domain and individuals’ time-use decisions and activity-travel decisions as being intertwined, and reflected in their activity-travel patterns. Individuals’ activity-travel decisions are affected by their socio-demographics, and by spatial and temporal aspects of the transportation system and the built environment.

Trip-based and activity-based approaches also differ considerably in the degree of spatial aggregation. Trip-based models predict trips for aggregated spatial zones called traffic analysis zones (TAZs), and predict trip totals and mode breakdown between pairs of zones (Rasouli and Timmermans 2013). They either ignore the impact of socio-demographic attributes, or consider them in very limited fashion, which in turn limits the ability of trip-based models to assess changes in transport demand due to socio-demographic changes. In contrast, activity-based models analyze activity and travel behavior for individuals or homogeneous groups of individuals. They can accommodate any number of socio-demographic factors, and thus can be useful in predicting the effects of socio-demographic, land use, or policy changes. In a nutshell, the activity-based approach offers a disaggregated modeling environment which focuses on individual behavioral responses.

2.3 Roots of Activity-Based Model

Chapin (1974) provided a framework of social constraints and individual motivations that shape the daily activity pattern. He argued that individuals participate in activities to fulfill their essential needs and/or wishes. However, Chapin didn't consider the geography or spatial context of activity participation and travel. During the same time, Hagerstrand (1970) explicitly discussed the relationships between the time and location of activities, which are modeled by three types of constraints: authoritative constraints (e.g. shops' business hours), capability constraints (such as sleeping, eating, and personal care), and coupling constraints (such as availability of an individual who interacts with another individual). Though Hagerstrand explained the relationship between time and space, Jones (1979) extended his work by explicitly addressing the relationships between time, space, activities, and travel. He defined travel as a derived demand stemming from the need or wish to participate in various activities at different locations at different periods of time. Simultaneously, utility maximization microeconomic theories of the allocation of time to activities were being developed in the regional science field (Becker 1965). Subsequently, random utility maximization discrete choice theory was proposed by McFadden (1973), and is still one of the most commonly used and popular applied modeling approaches in activity-travel behavior analysis. As mentioned earlier, the shortcomings of trip-based models motivated the paradigm shift to activity-based models. For instance, one of the limitations of trip-based models was their limited behavioral realism (Jones 1979; Axhausen et al. 2002). Activity-based travel demand models view travel as a derived demand necessitated from the need/desire to participate in different activities, and thus provide a clear and explicit rationale for activity participation behavior (Jones 1979; Bhat et al. 2004; and Davidson et al. 2007). Additionally, and unlike trip-based models, ABMs can address ridesharing, or strategies like high occupancy vehicle (HOV) or high occupancy toll (HOT).

The central idea of the activity-based approach is to analyze how an individual decides to use the 24 hours in a day among different activities and travel. One stream of researchers has studied this decision-making process as activity time allocation studies and another stream have studied it as activity episode analysis. Based on the individual/household characteristics, activity time allocation studies categorize activities into various target types and then investigate the time-use allocated to those activity categories. Many time allocation studies employ sociological and economic approaches which emphasize 'resource theory' and study the effects of attitudes towards the role of gender on time allocation of individuals living together. For instance, Geerken and Gove (1983) postulated an integrated socio-economic theory of 'imperfect' utility maximization and their focus was on time expenditure on household work by individuals living in the same space as a function of time expenditure at work. However, they didn't formulate the time allocation process. In planning, Chapin (1974) depicted a framework in which the propensity of activity participation has been viewed as the outcome of the interactions between constraints imposed by society and inherent motivations. Subsequently, Reichmann (1976) categorized activities into three types: subsistence activities, or work related; maintenance activities, or procurement and consumption of goods; and leisure activities, or social, recreational, and other discretionary activities. This classification, or categorizations similar to this, has been adopted in most of the empirical time allocation researches. For instance, Aas (1982) used the following four categories of activities: free time, necessary (personal) time, contracted (paid) work time, and committed (unpaid) work time.

The other strand of research is activity episode analysis, which developed based on Hagerstrand's (1970) concept of the space-time prism. Cullen and Phelps (1975) and Heideman (1981) considered information related to individuals' mental maps or perceptions of their activity-spaces as the determinants of their activity episode patterns. Ellegard and Vilhelmson (2004) suggested that home is the hub of activities ranging from watching TV and reading a magazine to preparing dinner or engaging with friends and other household members living in the home. An individual performs activities in a specific order for a particular project. The locations of activities are linked together through daily travel (Vilhelmson 1999). Heideman (1981) considered information related to individuals' perception of their mental capabilities and activity-space as the determinants of their activity episode patterns. The following section presents an overview of the existing activity-based travel demand models.

2.4 Activity-Based Travel Demand Models

Broadly, activity-based models can be categorized into two modeling approaches, which are utility maximization-based econometric models and rule-based computational process models. Other modeling systems include the time-space prism and constraints, the agent-based approach, and mathematical programming models. However, these modeling techniques are not exclusive or exhaustive.

2.4.1 Rule-Based Computational Process Models

Rule-based computational models employ a set of rules in the form of conditional actions that stand for the process of solving the task (Garling et al. 1994). One stream of rule-based modelers builds activity schedule-building models and another stream builds switching models (Jovicic 2001). In activity schedule-building models, the activity schedule is built from scratch, whereas switching models update the previous schedule as an outcome of hypothesized changes. Rule-based modelers assume that complex activity-travel behavior cannot be explained fully by utility maximization (Rasouli and Timmermans 2013). Rule-based models assume that individuals use context-dependent choices heuristically for participating in activities and travel.

A number of rule-based models have been developed so far. CARLA (Clarke 1986) was the earliest scheduling model developed based on the combinatorial algorithm. During the same time, STARCHILD (Recker et al. 1986) was developed in two phases. STARCHILD added a logit model to expand the feasible daily activity pattern generation framework of CARLA. Subsequently, SCHEDULER (Garling et al. 1994) was developed, where a small set of activity episodes with higher priority were selected as the list for short-term execution. Those short-term activities were then sequenced and their locations were identified based on a ‘distance-minimization’ approach. In 1996, the mobility simulator AMOS (Kitamura et al. 1996) was developed that takes the observed daily activity-pattern (DAP) as the input and identifies the constraints based on heuristic rules. It then synthesizes the probable changes in DAP if the activity-travel environment changes. SMASH (Ettema et al. 1993) was also developed at this time, and postulates the activity scheduling procedure as a step-wise and sequential process which starts with a null activity schedule at each step.

One of the most advanced and comprehensive rule-based models, ALBATROSS (Arentze and Timmermans 2005), was developed in the Netherlands. ALBATROSS takes individual activity diary, a list of constraints, socio-demographic attributes, zonal data, and transportation system attributes as the inputs into the modelling process. ALBATROSS used observed data to determine the decision-making process and it also incorporated machine learning techniques in the development phase of the decision-making rules (Garling et al. 1994; Arentze and Timmermans 2005; and Rasouli and Timmermans 2013).

However, rule-based models still have some unresolved issues. For instance, it is difficult to determine the statistical significance of the factors affecting the individual’s decisions relating to their activity schedule. Another aspect is that rule-based models consider the activity episode generation process as an exogenous event and focus instead on the scheduling or sequencing of the activities. Consequently, it is difficult to compute the decision-rules for the activity-travel scheduling process of individuals.

2.4.2 Utility Maximization-Based Econometric Model Systems

The utility maximization-based econometric models were developed from the consumer choice theory (Becker 1965) of economics. Consumer choice theory assumes that an individual makes their activity-travel choices to maximize their utility from the selected choice. The utility maximization-based econometric modelling system consists of a series of discrete choice models that are utilized to forecast the attributes related to individuals' activity-travel decisions. These models employ utility maximization-based equations to identify the associations between an individual's socio-demographics and their activity-travel characteristics. There are several criticisms about this modelling approach. One of them is that individuals are not rational utility maximisers (Timmermans et al. 2002), and another is that these models do not explicitly consider the underlying decision-making process of activity-travel behavior. Some of these models were developed by planning agencies, as for example San Francisco SFCTA (Bradley et al. 2001), New York NYMTC (Vovsha et al. 2002), Columbus MORPC (PB Consult 2005), and Sacramento SACOG (Bowman and Bradley 2005-2006). Other models were developed by the research community, such as CEMDAP and FAMOS.

TRansportation ANalysis SIMulation System (TRANSIMS)

TRANSIMS was developed by the Los Alamos National Laboratories in Portland, Oregon (USDOT 1997). An exemplary state-of-art feature in econometric modelling was the inclusion of a daily activity scheduler developed by Bowman and Ben-Akiva (2001) into the TRANSIMS system. This model system is the first true operational activity-based modelling approach developed for regional travel demand modelling. The front phase of this model is the activity-based model of Bowman and Ben-Akiva (2001). It is linked with a population synthesizer and a microsimulation model of activity-travel behavior. The generation of activity patterns for synthetic populations based on skeletal base patterns is comparable to the method proposed by McNally (1995). TRANSIMS reflects the need for realistic behavioral abstraction in travel demand forecasting, but it is dependent on extensive data definition. This system is based on multinomial logit and nested logit models with an activity hierarchy. The application of this modelling system can also be found in the documentation of travel demand models developed for the Columbus and Atlanta regions (PB Consult 2005).

SACRAMENTO

The SACRAMENTO model (Bowman and Bradley 2005-2006) has an activity-travel forecasting system called DaySim. It predicts the 24-hour activity and travel schedule for each individual, which can be labeled as the 'full individual day pattern'. DaySim has three hierarchical tiers, which are daily activity pattern (DAP) choice models, tour choice models, and trip/stop choice models. In this system, all the activity-travel choices are derived using a nested logit model or a multinomial logit model. The DAP model has two segments: the daily activity pattern model and the tour frequency model. The output of these models includes the number and occurrence of home-based tours, and occurrence of any additional trip/stop for each activity type. In contrast, the tour-level models forecast the primary activity destination, mode of travel, time-of-day (TOD) of travel, and information related to additional stops/trips. They also include work-based tour prediction and stop-level models. The output of the activity generation phase includes an activities list, tours, and trips per person per day.

Comprehensive Econometric Microsimulator for Activity-Travel Patterns (CEMDAP)

CEMDAP is an activity-travel forecasting system developed by Bhat et al. 2004; and Pinjari et al. 2006. It is a continuous time prediction system based on a variety of econometric models, including regression models, discrete choice models, and hazard-based duration models. It encompasses hierarchical daily-activity pattern characteristics, tour-level characteristics, and stop-level characteristics. This model is different from earlier models as it represents a continuous time-based DAP inside the space-time constraints imposed by mandatory activities (i.e. school and work). It develops a separate framework for workers and non-workers. The workers segment includes before-work pattern, home-work commute pattern, work-based pattern, work-home commute pattern, and post home arrival pattern. Tour level models include mode of travel, number of stops, and home-stay duration, along with a sequence of tours. Stop-level models include the type of activity, travel time from previous stop, activity duration, location of stop, and sequence of stops within a tour. Activity-patterns of non-workers are modelled as a set of out-of-home activity stops spaced between in-home activity stays (Bhat et al. 2004). Pattern-level characteristics consist of the occurrence of stops in the day, number of stops for each activity category, and the sequence of activities per day. Tour-level attributes comprise just the travel mode choice. Attributes considered in the stop-level model include activity duration, travel time from previous episode, and stop location. All these models can be categorized into two categories: the activity generation-allocation model and the scheduling model (Pinjari et al. 2006). CEMDEP is one of the main models of the travel demand model for Southern California, the SIMAGENT model (Goulias et al. 1996).

FAMOS (Florida Activity Mobility Simulator)

FAMOS (Pendyala et al. 2005) is similar to CEMDAP in several ways, including the explicit adoption of space-time constraints, and the use of a continuous time model for forecasting. It uses Hagerstrand's (1970) space-time prisms to model the temporal and spatial constraints related to activities and trips. A Prism-Constrained Activity Travel Simulator (PCATS) is used to simulate trips and activities completed by an individual, along with mode of travel, duration, travel time, location of the activity, and sequence of activities. The limits (or boundaries) of the space-time prisms are derived through stochastic frontier models (Pendyala et al. 2005). Individual's activity-travel patterns are assumed to take place within the boundaries (or frontiers), and then DAPs are simulated within the frontiers.

2.5 Activity-Travel Behavior Dimensions and Implications

This section provides a synopsis of the various dimensions of activity-travel behavior, and considers those aspects that have gained significant emphasis in the activity-based paradigm.

2.5.1 In-Home and Out-of-Home Activity Substitution

A number of works have focused on the trade-offs between in-home and out-of-home activity, since these can play an important role in trip generation. Kitamura et al. (1996) studied the time allocation behavior between two discretionary activities by using a utility-maximization based discrete-continuous model. This study suggested that individuals who worked are less likely to engage in out-of-home discretionary activities. However, individuals who work more hours in the week or who had longer commute time, spent more time in out-of-home discretionary activities. Other socio-demographic attributes, including child care and household size, were also found to have significant effects on in-home versus out-of-home trade-offs. Lawson (1998) studied the decision-making process of in-home and out-of-home activity participation and its determinants. His study suggested that household composition, age, life-style choices, and work attributes of individuals affect the decisions.

2.5.2 Intra-Household Interactions

There has been increasing attention to the role of intra-household interactions in relation to activity-travel behavior. Activity-travel behavior decisions can be affected by, for example, joint activity participation and travel, allocation of household maintenance responsibilities, allocation of household resources (vehicles), and activity participation derived from the mobility-dependency of other household members (for instance, chauffeuring children, the elderly and others). Srinivasan and Bhat (2005) studied the time allocation and participation in maintenance activities of household members. Their study found that household socio-demographic attributes such as age, gender, household responsibilities, income level, vehicle ownership, presence of children, and employment have significant effects on joint and solo participation in maintenance activities. A study by Kato et al. (2009) examined the joint activity time allocation for household members. The results of the study suggest that households with a higher number of children allocate more time to joint leisure activity, and households where husbands have more non-working days place less emphasis on their out-of-home leisure activities. This is also essential to understand the mobility needs of children and other mobility-dependent individuals on a traveler's activity-travel pattern.

2.5.3 Daily Activity-Travel Patterns

After the seminal work of Bowman (1998) many researchers studied and extended the empirical modelling of daily activity-travel patterns. Wen (1998) developed an operational activity generation model. This model includes tour and stop generation, assignment of stops to tours, and location and mode for each tour. Lee and McNally (2006) utilized doubly-censored Tobit models to examine time-use behavior in the household context. They defined five types of households: single non-worker, single worker, couple non-worker, couple one-worker, and couple two-worker households, and studied their trip chaining behavior. They found that type of household, household structure, and intra-household interactions all have significant influence on trip chaining propensity.

2.5.4 Time-Frame of Activity-Travel Analysis

In early work, most of the studies examined the activity-travel behavior on a single day (24-hour) travel diary survey. Analyses based on a single day exhibit an implied assumption of behavioral consistency in activity decisions and processes from the observed day to the unobserved ones. It was later recognized that there might be day-to-day dependence and variations in activity-travel patterns (Axhausen et al. 2002; Bhat et al. 2004; and Spissu et al. 2009). More research is needed, therefore, to determine periodic weekly, monthly, and annual fluctuations in the incidence and duration of activities, particularly for time-flexible and discretionary activities.

2.5.5 Space-Time Interactions

Since the 1950s, traffic analysis zones (TAZs) have been used for trip-based travel demand modeling and forecasting, as inputs to transportation planning. However, in the shift from trip-based to activity-based modelling the TAZ spatial unit was found to be too coarse. The space-time prism focusses on localized spatial-temporal constraints (Hagerstrand 1970), and to employ this concept operationally in activity-based modelling both allows and requires a finer representation of spatial units (i.e. parcels). The TAZ based system is too spatially coarse to accurately represent network attributes such as transit stops. Finer resolutions are needed both for long-term location choices (i.e. work, home) and for short-term activity-travel decisions, such as discretionary activity locations. Earlier researches also concluded that individual's perceptions of space and 'neighborhood' affect their activity-travel decisions. Their knowledge of built environment attributes and the land-use mix of the local residential neighborhood may constrain the activity-travel decisions because of accessibility to potential activity destinations.

2.6 Conclusion and Recommendations

Understanding activity-travel behavior is important to improve our knowledge of transportation planning issues, such as multimodal choices, transit/pedestrian oriented development, HOV/HOT lanes, telecommunications, flexible work schedules, social exclusion, environmental justice etc. Over the past decades, the activity-based approach has seen significant research attention and progress. This review has detailed the major conceptual and technical advances for study and modelling of activity-travel behavior. There has been substantial progress in modelling frameworks and techniques, but there is much still to understand regarding how households or individuals make decisions that drive their activity-travel behavior and patterns. Several useful research directions may be suggested:

- In modelling, it is important to consider the inter-dependency between activities completed in the 24-hour day or over successive days. Most studies to date fail to do this, and treat activity episodes as unrelated. However, in reality activity choices are affected by the choices an individual makes for previous activities.
- It is also important to better understand the activity-travel behavior of non-workers. Traditionally more focus has been given on workers, as they contribute strongly to peak hours traffic. However, there is an increasing trend towards off-peak hours congestion levels, which warrants more attention to non-worker activity-travel behavior and non-work trips.
- It has been confirmed in the literature that trips should not be analyzed in isolation, and that tour complexity influences mode choice decisions. However, tour frequency, trip chaining, and tour mode choices exhibit heterogeneity among individuals with different socio-demographic attributes or activity patterns. For instance, individuals who work from 9 a.m. to 5 p.m. might have more trip chaining propensity than those workers with shorter work-days. We therefore require a modelling approach that can capture this heterogeneity.
- Intra-household interactions can greatly constrain an individual's activity-travel behavior. In particular, the constraints are more onerous in larger household and those with young children. More attention should be paid to these constraints in modelling.

- Both the built environment and land-use of the neighborhood within which an individual operates affect the activity-travel decisions of that individual and/or other household members.
- It is important to study activity-travel behavior decisions longitudinally through time to better understand responses to technological and societal change.
- Recent studies have revealed the need to develop sub-models for major sub-populations or special trip generators, such as large universities or large port areas. Therefore, one potential avenue of future work is to develop sub-models for such special trip generators in regional travel demand models.

Chapter 3 Data and Methods

3.1 Scheduler for Activities, Locations, and Travel (SALT)

The primary goal of this dissertation is to develop the Scheduler for Activities, Locations, and Travel (SALT) travel demand model. This study utilized disaggregated travel demand modeling approaches, in particular, the activity-based approach. Activity-based models consider individual or household as the unit of analysis, which can capture the behavioral realism at the finest level. To this end, this study developed a series of advanced econometric micro-behavioral modules within the SALT modeling framework to better understand the activity-travel behavior mechanism of the population segments. As shown in Figure 3.1, the SALT model comprises of five core components:

- Population synthesizer: This module expands the sample households with respect to the control table derived from the census data on individual and household attributes.
- Time-use activity pattern recognition: The core of the SALT model is the pattern recognition module. Individuals with homogeneous daily activity patterns are grouped together into clusters based on their time-use activity patterns. Information related to activity type, timing, duration probability distribution, and sequential arrangement of activities can be derived for each identified cluster.
- Tour mode choice: Tour type choice, tour frequency, number of intermediate stops, and mode choices for all of the identified clusters in the SALT model system are estimated. Socio-economic and socio-demographic characteristics, trip and travel characteristics, and land use attributes are incorporated into the model estimation.

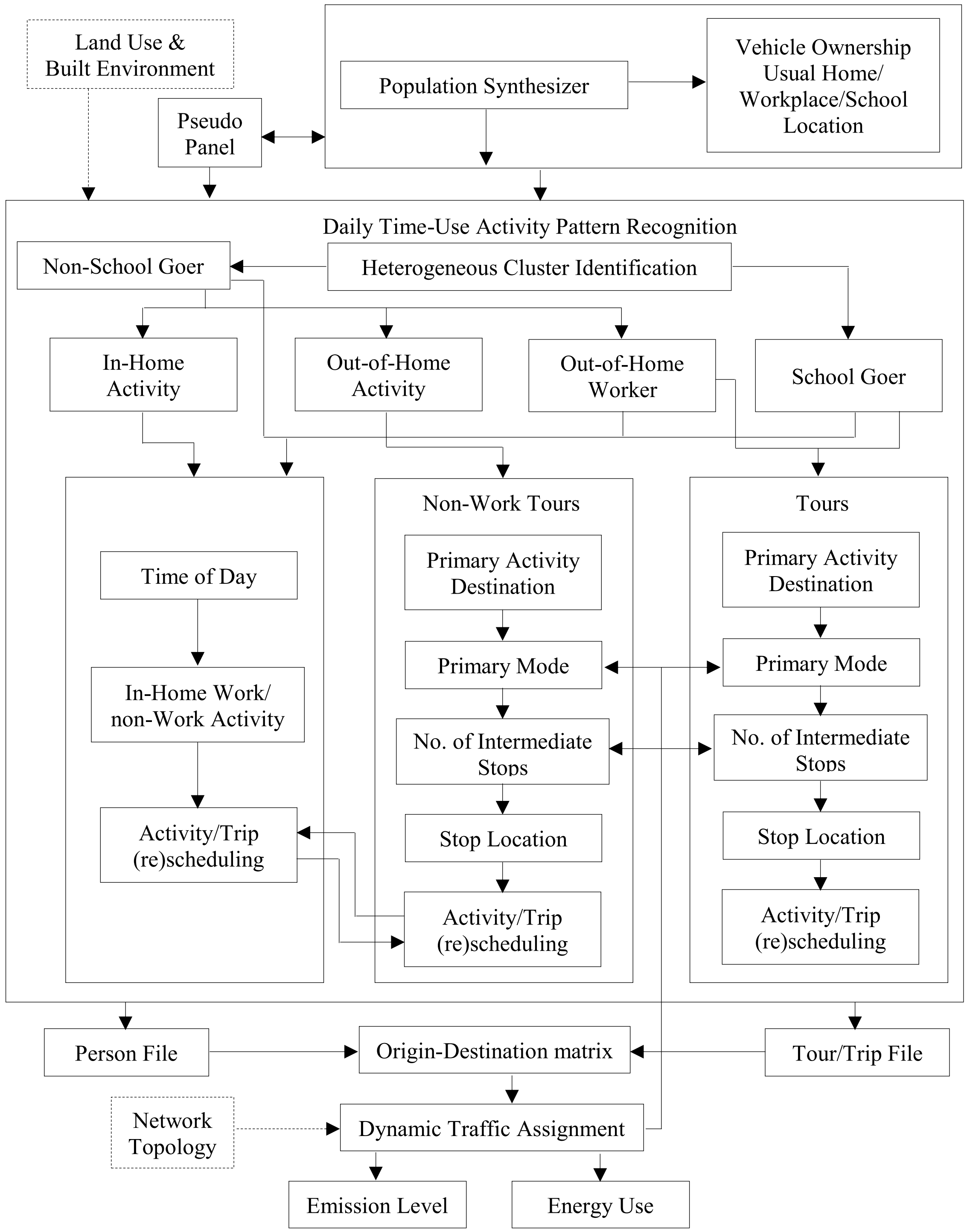

Figure 3.1 A conceptual framework of the Scheduler for Activities, Locations, and Travel (SALT)

- Activity destination choice: The daily activity agenda formation is modeled with the determination of the type of activities required/desired, their frequency, and their sequential arrangements.
- Activity/trip scheduling: The temporal attributes for every activity type in the agenda is estimated, and the 24-hour activity schedule is formed through a rule-based decision algorithm.

Specifically, this study concentrated on the development of the advanced econometric cluster-based modules for modeling activity participation, trip chaining, and tour mode choice. In addition, this study presents mode specific trip frequency models for university population commuters. Furthermore, the development of a cohort based pseudo panel modeling framework for longitudinal travel behavior analysis is presented.

3.2 Data

The main data sources used in this study are drawn from the Space-Time Activity Research (STAR) time-use and travel survey, the Environmentally Aware Travel Diary Survey (EnACT), and the General Social Survey (GSS) data of Statistics Canada. Table 3.2 presents a summary of data sources used for building the micro-behavioral modules within the SALT modeling system.

3.2.1 Halifax Space Time Activity Research Survey (STAR)

The activity participation, tour complexity and mode choice analysis are based on the STAR (Space-Time Activity Research) time-use and travel survey, conducted in Halifax Regional Municipality from 2007 to 2008. This survey collected information from 1,971 randomly selected households in the urban, suburban, and exurban areas of the municipality. Primary respondents aged 15 or over were randomly selected in each household; they maintained a time diary over two consecutive days (2880 minutes), and wore a GPS tracking device for all out-of-home activity. The diary data were validated through GPS-assisted prompted recall computer assisted telephone interviews. This translates into 3,919 diary person-days of information, comprised of 108,529 episodes of time diary information. For each of these minutes the data collector retrieved: (i) what was being done, (ii) what else was being done at the same time, (iii) where it was done, (iv) how long it was done for, (v) who it was done with, and (vi) purpose/for whom it was done. After data preprocessing, the final cleaned data is comprised of activity patterns of 2,778 person-days.

The STAR data include socio-demographic information, household size, accommodation type, motor vehicles and modes of transportation, parking availability and type, household energy usage, residential locations, education status, employment statistics (e.g. number of working adults in the household, occupation type, work hours, location, etc.), commitment (family, work, etc.), travel behavior (purpose of trip, duration etc.), spatial information on activities (latitude, longitude, address, municipality information, frequency of visit, etc.), routing information, distance of trip, and trip accompaniment. Full descriptions of the survey design and the socio-demographic characteristics of respondents can be found in TURP (2008) and Millward and Spinney (2011). Hafezi et al. (2017a, b) derived distinct worker and non-worker population clusters from the STAR dataset which include individuals with homogeneous time-use activity patterns as illustrated in Figure

3.2. This study utilizes the identified clusters for micro-behavioral model development as described in Chapter 4 to Chapter 7.

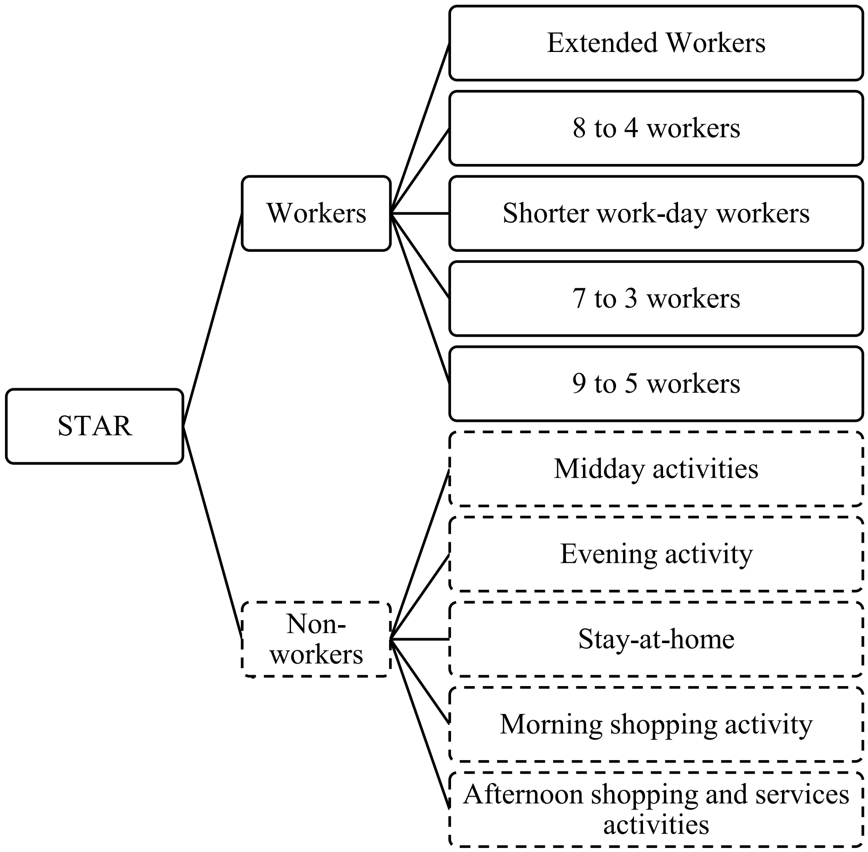

Figure 3.2 Worker and non-worker clusters identified from the STAR dataset

3.2.2 Environmentally Aware Travel Diary Survey (EnACT)

A unique web-based Environmentally Aware Travel Diary Survey (EnACT) is designed, implemented and collected in Spring 2016 among Dalhousie university commuters. The purpose of the EnACT survey was to better understand and model activity-travel behavior of different university segments as special trip generators in travel demand models. A survey link was sent to all Dalhousie community segments (undergraduate students, graduate students, faculty members, and staff) via e-mail and social media (Facebook and

Twitter). The EnACT survey included six sections: (1) household information, (2) individual information, (3) environmental attitudes and behavior, (4) attitudes toward transportation, (5) information and communication technologies (ICT) related information, and (6) 24-hour travel log. Dalhousie University has four campuses, three of them located in the city of Halifax (Halifax Regional Municipality (HRM)) and another one in the town of Truro. The land use and neighborhood characteristics of the Truro campus are different from those of the other three campuses. All these three campuses are located in the inner-city of Halifax on flat terrain, providing a friendly environment for Active Transportation (AT) users living near to the campuses.

Table 3.1 demonstrates the comparisons between the surveyed sample and the university population using Dalhousie Analytics Data (provided by the university). As shown in Table 3.1, samples by gender, age and employment status were almost evenly distributed among different university population segments. However, the sample size obtained for faculty members is slightly under-represented. Also, female staff are over-represented in the sample compared to male staff, and the percentages of part-time staff and part-time undergraduate students are higher than those in the university population. From analyzing the average ages of each of the four groups, it is found that no significant differences exist between those of the population groups and those of the sample groups. Additionally, home postal code and work (destination) postal code of all the respondents were geocoded in ArcGIS 10.2.2, and network commuting distances were calculated using the network analyst tool. This dataset is used in Chapter 8 for mode specific trip frequency modeling. A comprehensive descriptive analysis of all six sections of the EnACT survey can be found in Liu et al. (2016) and Daisy et al. (2018a).

Table 3.1 Comparison of the EnACT sample and total Dalhousie university population

		Undergraduate students		Graduate students		Faculty		Staff	
		Population	**Sample**	**Population**	**Sample**	**Population**	**Sample**	**Population**	**Sample**
Total		14132	129	3194	126	1531	24	1807	67
Gender (%)	Male	45.19	44.19	45.02	41.27	53.95	33.33	36.41	19.40
	Female	54.81	55.81	54.98	58.73	46.05	66.67	63.59	80.60
Age (years, avg.)		22.2	23.2	29.8	28.3	49.0	49.5	46.7	44.2
Employment status (%)	Full time	89.17	79.84	78.37	83.33	73.35	87.50	94.74	86.57
	Part time	10.83	20.16	21.63	21.63	26.65	12.50	5.26	13.43

3.2.3 General Social Survey (GSS)

The public use micro data of the General Social Survey (GSS) of 1986, 1992, 1998, 2005, and 2010 are utilized for pseudo model development. This random stratified survey series started in 1982 and each wave includes a survey of individual household information, personal characteristics, socio-demographic information, and a 24-hour time-diary episode file for each participant. The dataset utilized in this study is based on pooling data from all these five waves. The GSS activity episode file has 188 sub-categories of activities. However, the GSS public use micro data does not offer the locational information for public access. The data therefore lack locational attributes of the built environment, which are very important to analyze the correlates/determinants of mandatory activities and activity spaces. Recently, the 2015 GSS database is released which can be used for further extensions of the proposed in Chapter 9 and Chapter 10.

Table 3.2 Data sources used in the development of econometric micro-behavioral modules within the SALT model system

Data Objects	Data Sources	Data Descriptions	Unit	SALT's Micro-Module(s)
Representative time-use microdata sample at the household level	STAR[1]	Time-diary and GPS geo-coordinate microdata sample data Land use and built environment data	Three dimensions (temporal, socio-demographic and spatial) data with five minutes intervals Parcel level	Pattern recognition Ensemble learning Activity participation Trip Chaining Tour mode choice
Representative time-use microdata sample of university population (undergraduate student, graduate student, staff and faculty)	EnACT[2]	Time-diary microdata sample	Three dimensions (temporal, socio-demographic and spatial)	Travel behavior characteristics of university community Transport-related GHG emissions
Time-use microdata sample at the individual level	GSS[3]	Time-diary microdata sample	Activity duration Episode duration	Synthetic pseudo panel
Marginal population data at the DA and regional level	CCS[4]	Distribution of the socio-demographic characteristics of the marginal population data	Dissemination area (DA) level Regional level	Population synthesis
Road network	NRN[5]	National road network layer in the ArcGIS platform	Street level Highway level	Network building
Transport service location and road network	HRM[6]	Transit stop locations, Transit and road networks	Street level	Network building

[1]Space-Time Activity Research, [2]Environmentally Aware Travel Diary Survey, [3]General Social Survey, [4]Canada's Census, [5]National Road Network and Environment Canada archive, [6]Halifax Regional Municipality Geodatabase

Figure 3.3 shows the advanced econometric micro-behavioral modules incorporated in the SALT model system. In the following section, major sub-models developed in this dissertation are outlined.

3.3 Activity Participation

This phase presents an innovative modeling framework to determine workers' and non-workers' decisions relating to activity participation. To better understand the activity-travel behavior of workers and non-workers of Halifax Regional Municipality (HRM), this study utilizes a cluster-based Multivariate Probit (C-MVP) model along with estimation of transition matrices of activity episodes. Household travel-diary data from the Halifax Space-Time Activity Research (STAR) was utilized to investigate activity-travel behavior of workers. From the STAR dataset, five worker clusters with homogeneous daily activity patterns are identified, which were extended workers, shorter-work day workers, 8 to 4 workers, 7 to 3 workers, and 9 to 5 workers; and a series of C-MVP models are estimated for each worker cluster. Non-workers (e.g., homemakers, retirees, and unemployed individuals) represent a significant portion of the urban population, and their activity-travel behavior is still an under-researched area in the literature. From the STAR dataset, based on the daily activity patterns, five non-worker clusters were identified, which are non-worker midday activities, non-worker evening activity, stay-at-home non-workers, non-worker morning shopping activity, and non-worker afternoon shopping and services activities. More details on the methods and results relevant to activity participation of worker and non-worker clusters can be found in Chapter 4 and Chapter 5, respectively.

3.4 Tour Frequency, Trip Chaining, and Tour Mode Choice

Given the estimated set of activity types in the agenda, the next step is to identify the tour frequency, trip chaining, and mode choice related to individual's daily travel agenda. In this phase, a cluster-based disaggregated approach is introduced to model trip chaining, tour complexity, and tour mode choice, to better understand their relationships to socio-demographics, trip attributes, and land use patterns. The identified five workers and five non-workers clusters drawn from the Space-Time Activity Research (STAR) survey for Halifax, Canada, are utilized for tour formation and empirical models, and home-based tours are utilized as the unit of analysis. Number of tours made in a day, number of stops made in each tour, activity purpose at each stop, and mode choices for each tour are identified for each worker and non-worker cluster. The home-to-home journey, for which origin and destination is home without any occurrence of intermediate home stops, is defined as a home-based simple tour. Tours with more than one stop are identified and named as complex tours. Tours starting or ending outside the participant's home were eliminated due to interpretability problems. Thus, a total of nineteen categories of tours are identified for each cluster.

After the identification of the tour type, the number of stops per tour is estimated for each individual. Then the travel mode for each tour is selected based on the longest in-mode travel time. Multimode tours are also identified through data mining techniques. A total of ten categories of mode/multimode tours are identified for further empirical models. The number of tours per day for all clusters are modeled using a Poisson regression model. Trip chaining is then modeled using an Ordered Probit model. Finally, tour mode choice is modeled using a Multinomial Logit (MNL) model. More details on the empirical models and results can be found in Chapter 6 and Chapter 7.

3.5 Mode Specific Trip Frequency Model

Recent literature concluded that university populations can be considered as a special trip generator and should be modeled separately in the regional travel demand model. Halifax Regional Municipality (HRM) is the base for several universities, and the universities generate a significant amount of travel demand in the transportation system. However, activity-travel behavior student group is not well recognized in both the GSS and STAR surveys. To this end, this dissertation contributes in analyzing activity-travel behavior of the largest university population of the Canadian Maritime Provinces, Dalhousie University. The data were derived from the first university-based activity travel diary survey (EnACT). The daily activity-travel behavior and activity-travel demands for auto, active transportation (AT), and transit trips of undergraduate students, graduate students, faculty, and staff of university commuters are modeled. Number of trips generated by individuals for each of these three modes are estimated using advanced statistical techniques. Finally, a series of Zero-Inflated Negative Binomial (ZINB) models are estimated for auto, AT, and transit trip frequencies. More details on methods and findings can be found in Chapter 8.

3.6 Synthetic Pseudo Panel

The objective this phase is to utilize repeated cross-sectional data in a pseudo panel data approach to investigate an individual's daily participation in out-of-home discretionary activities. The pseudo panel data methodology uses repeated cross-sectional data of the General Social Survey (GSS) conducted by Statistics Canada over the period 1992-2010 to investigate longitudinal activity participation in out-of-home discretionary activities. In this study, out-of-home discretionary activities are defined as an individual's daily participation in shopping, grocery, social, recreational, entertainment, and organizational activities in each surveyed year. Based on gender and the birth year of the respondent, cohorts are defined, and their activity behavior is traced over time in each of the cross-sectional data sets. A random coefficient model is developed on the basis of the cohort data. The estimation of the random coefficient model of the pseudo panel data suggests that personal and household socio-demographic characteristics have long run effects on out-of-home discretionary activity participation.

Public-use micro-data provide only small samples for each surveyed year. Hence, expansion of the sample is necessary to develop reasonable pseudo panel cohorts for empirical model estimation. Therefore, this research uses a population synthesis technique to generate synthetic populations for pseudo panel modeling. The sample sizes of 645, 672, 1105 and 962 are expanded to 5% of the total population of Nova Scotia, Canada, for each of the surveyed years. After the population synthesis, based on birth year, gender, and education level, pseudo cohorts are formed. These cohorts are then utilized for empirical models. In order to capture panel effects, a Latent Class Accelerated Hazard (LCAH) model is used. The study demonstrates that the pseudo panel methodology can be utilized for exploring longitudinal dynamics of activity duration where panel travel survey is absent. More detailed information on the methods and results can be found in Chapter 9 and Chapter 10.

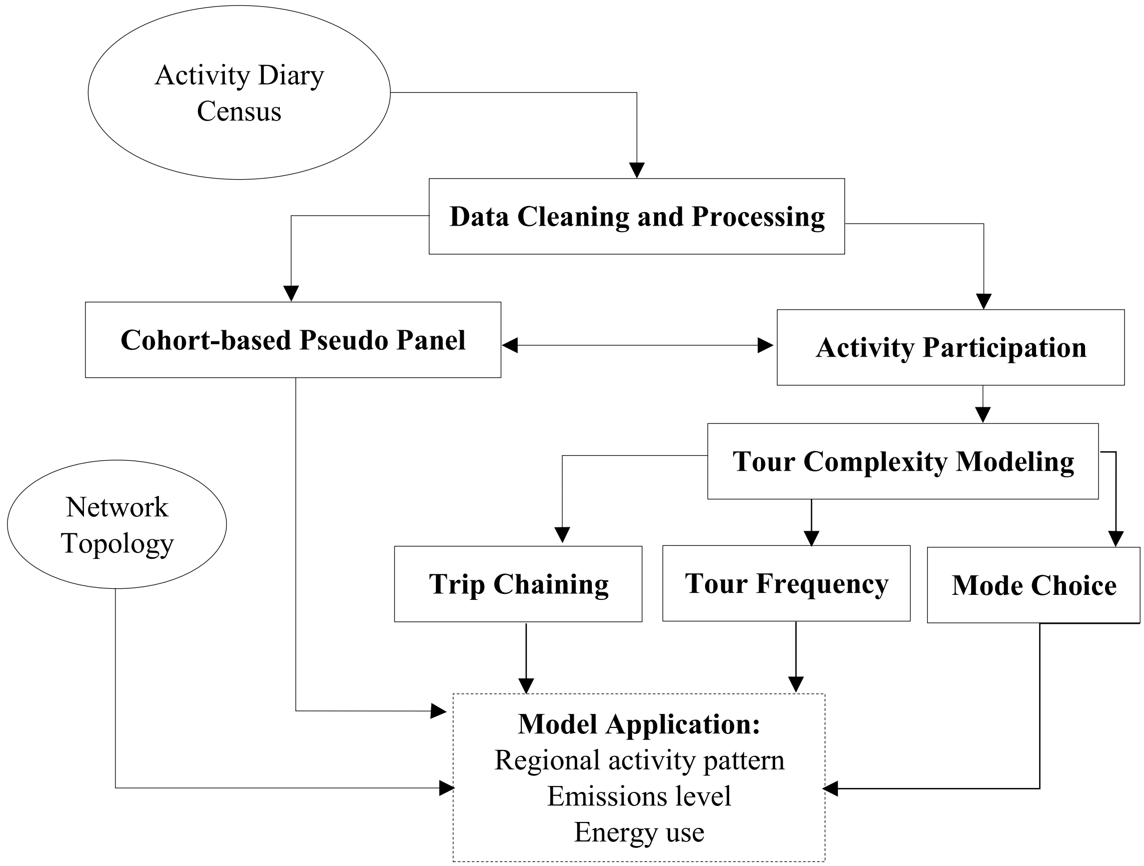

Figure 3.3 Econometric micro-behavioral modules incorporated in the SALT model system

Chapter 4 Modeling Activity-Travel Behavior of Out-of-Home Workers with Homogeneous Activity Patterns[2]

4.1 Introduction

Analyzing the activity-travel behavior of individuals has become a major concern in activity based travel demand models (Arentze and Timmermans 2000; Auld and Mohammadain 2012). The largest population segment in urban areas is out-of-home workers. Transportation professionals have traditionally focused on work-related travel and commute trips to manage peak hour congestion. However, non-work travel demand exhibits greater flexibility and variability across the worker population segments, as well as among non-workers. In developed nations, historically the demand for non-work travel is increasing (Toole-Holt et al. 2005; Daisy et al. 2018b). This study presents an innovative modeling framework to determine workers' decisions relating to activity participation, with emphasis on non-work trips. All activities taken inside the home are classified together as in-home activities, while those undertaken outside the home are classified as work, school, shopping and services, organizational/hobbies, entertainment, and sports. Since non-work/non-school activities are flexible in time and location in comparison to work and school activities, we kept all the four discretionary categories to capture the variability and determinants related to each.

In this study, initially we identify representative daily activity patterns of working individuals (Hafezi et al. 2017a, b). These clusters have significantly different activity

[2] An earlier version of this chapter has been presented:
Daisy, N. S., Millward, H. and Liu, L. Analyzing time windows and time allocation to in-home and out-of-home activities in workers' activity patterns. Proceedings of the 53rd Canadian Transportation Research Forum (CTRF). Ottawa, Canada., 2018.

pattern along with socio-demographic differences in comparison to non-workers. Then, we model activity type choices, both in-home and out-of-home, with Multivariate Probit models (C-MVP). The reason for utilizing the C-MVP model structure is that the activities in each day are correlated with each other. Most earlier studies assumed the activity participation as an independent phenomenon in multivariate cases, resulting in either logit or mixed logit models. However, activity participation in one activity is associated with both the previous and next activity. This interdependency between activities can be captured with correlation matrix of C-MVP model.

In this study, we estimate full correlation matrices for six activity categories: in-home activity, out-of-home mandatory, out-of-home shopping, out-of-home organizational, out-of-home entertainment, and out-of-home sports. We also provide a transition matrix analysis with broader set of activity categories (9 types) to better understand the trip chain behavior of each cluster. The results of this study are expected to be incorporated into the Scheduler for Activity, Location and Travel (SALT) for Halifax. The next section will provide a brief introduction to current literature relevant to workers' activity participation.

4.2 Literature Review

Activity participation is a core component in activity-based travel demand models (Arentze and Timmermans 2000; Daisy et al. 2017a). It is important to understand how individuals choose their activities in a day, why they wish to pursue them, with whom they participate, at which frequency, and the inter-dependency between activities. Activity type choices have been modeled explicitly by Hamed and Mannering (1993) within the post-work activity patterns framework. Bhat and Singh (2000) and Rajagopalan et al. (2009)

analyzed activity participation for weekends and weekdays. Other approaches predicted activity frequencies by utilizing Poisson related methods (Ma and Goulias 1999), structural equations (Lu and Pas 1999), etc. Several studies focused on participation in discretionary activities (Meloni et al. 2007; Pinjari et al. 2009) or on maintenance activities (Srinivasan and Bhat 2005; Vovsha et al. 2004). Another stream of research analyzed the trade-offs between in-home and out-of-home activity engagement (Srinivasan and Bhat 2005; Kuppam and Pendyala 2001). The activity type choice models often analyzed participation in an activity with the assumption of independence from other activities (Chu 2005), or modelled activity purposes as a single category, such as maintenance purposes (Pendyala and Bhat 2004), or discretionary activity purposes (Yamamoto et al. 2000). Some studies analyzed activity engagement based on the commute time-of-the day (Rajagopalan et al. 2009; Castro et al. 2011; Chu 2017).

Although significant contributions have been made on activity type choice and activity timing, these studies are limited in various ways: (a) some studies did not differentiate between activities by purposes (Rajagopalan et al. 2009); (b) some studies focus more on a specific activity category, for instance maintenance or discretionary categories; and (c) some studies focus only on a specific time-of-the day, for instance, post-home arrival of workers (Bhat at al. 2004).

Researchers have analyzed activity participation behavior by employing random utility models (Adler and Ben-Akiva 1979; Bhat at al. 2004), rule-based and need-based models (Arentze and Timmermans 2009), complex simulation models (Pendyala et al. 2005), and mathematical programming models (Recker 2001). Allahviranloo and Recker (2014) utilized the C-MVP model with Gibs Sampling and Data augmentation to investigate out-of-home activities. However, earlier studies identified the difficulties of choice set generation for the time block selection. Individuals may not participate in all the selected time blocks, as they have different start and end times for work and non-work activities. Most of the previous studies assumed that the choice set is constant across the population. This necessitates the inclusion of a daily activity pattern clustering framework before modeling the activity participation.

This study contributes to the literature on the analysis of activity participation and activity timing by developing a comprehensive modeling framework that classifies individuals based on their daily activity pattern and then models their activity-travel behavior. Specifically, the Multivariate Probit (C-MVP) model is utilized to analyze activity involvement, since it can explicitly capture the correlation between activities. To our best knowledge, the C-MVP model with the Geweke-Hajivassiliou-Keane (GHK) estimator has not previously been utilized for modeling activity selection behavior of workers.

4.3 Data

This work described here will form one component of an operational activity-based model for Halifax, a mid-sized Canadian city. The Halifax Regional Municipality (HRM) is the largest municipality in Atlantic Canada as well as the capital of Nova Scotia. It is a mid-sized metropolitan area (c. 400,000), has a diverse and developing economy, and registers 0.5% per year population growth. We employed data from the Halifax Space-Time Activity Research (STAR) project conducted between April 2007 and May 2008. STAR was the world′s first large survey to use global positioning system (GPS) tracking for verification of household activity-travel diary data (Bricka 2008). The sample size consists of 1,971 randomly selected households in HRM, which represents one household in 78. Activity-travel diary and questionnaire data were collected over 373 days, with a total participation rate of 21% (Millward and Spinney 2011).

The primary respondent in each sample household was selected randomly, and had to be more than 15 years age. These respondents completed a detailed time-diary for two consecutive days. The time diary coding and questionnaire on household characteristics were based on Statistics Canada′s (2006) General Social Survey (GSS), Cycle 19. Primary respondents also carried a GPS-device (Hewlett Packard iPAQ hw6955) for all out-of-home activity, programmed to collect GPS data every 2 s. The GPS data provided precise start and end times for all "stops" on travel routes with more than 2 minutes stopping duration. These GPS data were used with CATI software in day-after interviews with respondents, to verify and enhance the time-diary data.

4.3.1 Description of Clusters

A pattern recognition model was applied to the Halifax STAR household activity data (Hafezi et al. 2017a, b). A subtractive clustering algorithm was utilized to initialize the total cluster number and cluster centroids. Identification of individuals with homogeneous activity patterns was accomplished using a fuzzy c-means clustering algorithm, and sets of representative activity patterns were identified using a multiple sequence alignment method. Advanced decision tree models were used to explore inter-dependencies interdependencies in each identified cluster, and characterization of cluster memberships through their socio-demographic attributes was achieved by use of the CART algorithm.

Figure 4.1 shows the representative daily activity patterns of individual activities for the five identified worker clusters. These clusters are as follows: extended day workers, 8:00 am to 4:00 pm workers, shorter work-day workers, 7:00 am to 3:00 pm workers, and 9:00 am to 5:00 pm workers.

Figure 4.1 Representative activity patterns of identified five worker clusters

NNNNNNNNNNNNNNHHHHHHHHWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWH
HHHHHHHHHHHHHHHHHHHHHHHHHLLLLLLLLLLLLLLLLLLNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Cluster#1: Extended work-day workers

NNNNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHHHWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNN

Cluster#2: 8-4 workers

NNNNNNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHL
LLLLLLLLLLLLLLLLLLLLLLLLLLLLLNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNN

Cluster#3: Shorter work-day workers

NNNNNNNNNNNNNNHHHHHHHHHHHHWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
LLLLLLLLLLLLLLLLLLLLLLLLNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNN

Cluster#4: 7-3 workers

NNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW
WWWWWWWWWWWHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNN

Cluster#5: 9-5 workers

**Representative pattern starts from 4:00 a.m. to 3:59 a.m. on the next day (288 five minutes intervals); In-home (H/L/N), Workplace/School (W/S), Shopping (P), Organizational/hobbies (G), Entertainment (E), Sports (T)*

We used six activity categories for C-MVP model estimation: in-home, out-of-home mandatory (work/school), out-of-home shopping and services, out-of-home organizational and hobbies, out-of-home entertainment, out-of-home sports for each cluster. Table 4.2 shows the details of the activity categories. Table 4.1 and Table 4.2 show the socio-demographic attributes and activity time-use of all the worker's clusters, respectively. The details of each cluster presented in Table 4.1 and Table 4.2 are as follows:

Cluster#1 is the extended day workers group. The individuals belonging to this cluster typically participate in work activity for a longer duration, starting from 8:00 am to 8:00 pm. This cluster predominantly comprises middle-aged females aged between 36 and 55 years old (67.0%). Almost 76.0% of them are high-school graduates, and 73.0% are full-time workers. Individuals from this group are mostly middle income (60.0%), and the majority of the workers (55.0%) had no flexibility in their work schedule.

Cluster#2 is the 8:00 am to 4:00 pm worker cluster. This cluster mostly comprises middle-aged males with high-school graduation or better. More than 92% of the workers in this cluster work full-time, and their income level is middle-income. Workers of this cluster participate in discretionary activities typically in the evening.

Cluster#3 is the shorter work-day workers, who work less than 5 hours a day and who typically finish their work in the early afternoon before 2:00 pm. The majority of this cluster are middle-aged females between 36 to 55 years old (71.0%). Additionally, 85.0% of individuals in this group are high school graduates, and 56.0 % had some flexibility in their work schedule.

Cluster#4 consists mostly of 7:00 am to 3:00 pm workers. The majority of individuals from this group are middle-aged males between 36 to 55 years old, and 47.0% have middle-income. Nearly all individuals in this cluster are full-time workers (93.0%), and 63.0% of them had no flexibility in work schedule.

Cluster#5 mostly comprises individuals who work from 9:00 am to 5:00 pm. Unlike, cluster 4, individuals from cluster 5 usually travel to and from work during the morning and evening peak hours. A large proportion of individuals from this cluster are middle-aged females between 36 to 55 years old (53.0%) with middle-income, and most are high school graduates. The majority of the workers (60.0%) indicate that they have some flexibility in their work schedule.

Table 4.1 Analysis of worker clusters data: Share of socio-demographic variables

Social demographic variables		**Sample mean (%)**	**Mean of cluster (%)**				
			#1	#2	#3	#4	#5
Gender	Female	0.50	0.53	0.44	0.52	0.47	0.53
Age	Young adults (ages 15-35 years)	0.09	0.12	0.10	0.10	0.05	0.09
	Middle-aged adults (ages 36-55 years)	0.69	0.67	0.66	0.71	0.72	0.70
	Older adults (aged older than 55 years)	0.22	0.20	0.24	0.19	0.23	0.22
Education Level	Bachelor degree and above	0.44	0.45	0.44	0.56	0.22	0.55
Occupation	Regular shift	0.87	0.73	0.93	0.87	0.93	0.89
	Irregular schedule	0.10	0.22	0.03	0.09	0.07	0.08
	Student	0.01	0.01	0.00	0.01	0.00	0.01
	Retired	0.01	0.02	0.02	0.01	0.00	0.00
	Work at home	0.20	0.23	0.13	0.30	0.06	0.26
Flexible schedule	Have no flexibility in a work schedule	0.51	0.55	0.54	0.44	0.63	0.40
Job number	Have more than one job	0.07	0.09	0.04	0.05	0.07	0.08
Income	Low-income (<= $ 40,000)	0.27	0.28	0.22	0.32	0.29	0.26
	Middle-income ($ 40,000 - $ 100,000)	0.61	0.60	0.68	0.55	0.64	0.59
	High-income (> $ 100,000)	0.12	0.12	0.10	0.13	0.08	0.15
Total cluster membership			137	401	171	229	348
Percentage in total (number of person-days)			4.93	14.43	6.16	8.24	12.53

Table 4.2 Analysis of Cluster Data: Share of activity time-use of all worker clusters

Activity categories	Descriptions	Share of daily activity engagement (%)				
		#1	#2	#3	#4	#5
In-home (*H/L/N*)	Home Chores (H): Working at home, eating/meal preparation, indoor or outdoor cleaning, interior or exterior home maintenance, child care or other in-home activities.	12.87	15.54	23.29	17.89	18.21
	Home Leisure (L): Watching TV/listening to radio, reading books/newspapers, etc.	6.17	9.26	11.18	10.53	9.46
	Night sleep (N)	32.26	31.09	34.47	30.95	34.11
Workplace/School (W/S)	Work (W): Work/job, all other activities at work, work related (conferences, meetings, etc.).	43.47	36.44	24.70	36.29	33.08
	School/college related (S): Class participation, all other activities at school.	-	0.06	0.13	0.12	0.06
Shopping (P)	Shopping for goods and services, routine shopping.	0.78	1.39	1.73	0.98	1.14
Organizational/ hobbies (G)	Organizational, voluntary, religious activities. Hobbies done mainly for pleasure, cards, board games, all other hobbies activities.	1.14	1.43	1.59	0.72	1.00
Entertainment (E)	Eat meal outside of home, all other entertainment activities.	2.27	3.05	2.01	1.52	1.64
Sports (T)	Walking, jogging, bicycling, all sports related activities.	1.03	1.74	0.91	1.02	1.30
Total (%)		*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

4.4 Methods

Probit models have not been applied widely in activity-based modeling, and their application is limited to univariate binary models, or mixed models (Allahviranloo and Recker 2014). For instance, Bhat and Srinivasan (2005) utilized an orde1-red binary probit model to investigate the stop frequency modeling for different activities. Lamondia et al. (2010) employed binary probit models to analyze activity participation behavior of individuals. Ruiz and Roorda (2008) estimated a multivariate probit model using weighted least-squares with mean and variance correction estimator to examine the decision-making process of activity companionship, planning, scheduling and execution. Allahviranloo and

Recker (2014) estimated a multivariate probit model with full correlation matrix with markov chain monte carlo (MCMC) to study the dependency between daily activity type choices of individuals and their socio-demographic characteristics along with correlation among activity choices.

This study utilizes the C-MVP model with GHK estimator for activity type choice of individuals for six categories: in-home, work/school, shopping, organizational/hobbies, entertainment, and sports. This problem can be analyzed empirically in two ways, by either multinomial or multivariate probit models. Multinomial models assume that random error terms of choice equations are independent (Greene 2003) and a traveler may choose only one alternative (Katchova 2013). In the case of activity type choice of workers, the six choices are non-independent alternatives and variables might have association with each other and the choice set is not purely independent. Thus, confirming a multivariate case of analysis. For our current study, the choices of activity types are not mutually exclusive, and therefore the error terms of the activity type choices may be mutually inclusive and correlated. Consequently, we chose to use a cluster-based multivariate probit (C-MVP) model, which allows for the possible correlation in the activity choices simultaneously.

The C-MVP model consists of a set of b binary dependent variables y_b^* (observation subscript $i = 1,2, \ldots\ldots, n$ has been suppressed), where the B-equation multivariate probit model framework:

$$y_1^* = x_1'\beta_1 + \varepsilon_1, \quad y_1 = 1 \quad if\ y_1^* > 0 \tag{1}$$

$$y_2^* = x_2'\beta_2 + \varepsilon_2, \quad y_2 = 1 \quad if\ y_2^* > 0 \tag{2}$$

$\vdots$

$$y_b^* = x_b'\beta_b + \varepsilon_b, \quad y_b = 1 \;\; if \; y_b^* > 0 \tag{3}$$

$$E[\varepsilon_b | x_1, \dots \dots, x_B] = 0 \tag{4}$$

$$Var[\varepsilon_b | x_1, \dots \dots, x_B] = 1 \tag{5}$$

$$Cov[\varepsilon_1, \varepsilon_b | x_1, \dots \dots, x_B] = \rho_{ib} = \sum \tag{6}$$

$$(\varepsilon_1, \dots \dots, \varepsilon_B) \sim Multivariate\ normal\ (MVN)[0, \Omega] \tag{7}$$

Where, b is the activity types, i.e. in-home, shopping, mandatory, sports, organizational, entertainment. X is a vector of explanatory variables, $\beta_1, \beta_2, \dots \dots \beta_b$ are conformable parameter vectors, and $\epsilon_b, B = 1 \dots . B$ are random errors distributed as multivariate normal distribution with a mean of zero, unitary variance and a correlation matrix, which is 6×6 for this study, $Q = [q_b]$. The density function of the equation will be $\Phi(\varepsilon_1, \varepsilon_2, \dots \dots \varepsilon_b; Q)$. Ω is the correlation matrix, and $\sum$ is the covariance matrix. Each individuals equation is a standard probit model.

In GHK simulation, the approximation is computed based on averaging R draws from a certain multivariate normal distribution, for each observation (Greene 2003; Hajivassiliou and Ruud 1994*)*. It assumes that a multivariate normal distribution function can be articulated as the product of sequentially conditioned univariate normal distribution functions.

The joint probabilities of the observed events, $[y_{i1}, y_{i2}, \dots \dots, y_{iB} | x_{i1}, x_{i2}, \dots \dots, x_{iB}], i = 1, \dots \dots, n$ that form the basis for the log-likelihood function are the M-variate normal probabilities,

$$L_i = \sum_{i=1}^{N} \Phi_B(q_{i1} x'_{i1} \beta_1, \dots \dots, q_{iB} x'_{iB} \beta_B | \Omega^*) \tag{8}$$

Where:

$q_{ib} = 2y_{ib} - 1,$

$\Omega^*_{bn} = q_{ib} q_{in} \rho_{bn}$

The practical obstacle to this extension is the evaluation of the B-variate normal integrals and their derivatives. Among the simulation methods examined in Greene (2003), the GHK smooth recursive simulator appears to the most accurate. The general approach uses,

$$P = pr[a_1 < \varepsilon_1 < c_1, \dots \dots, a_b < \varepsilon_b < c_b] \approx \frac{1}{R} \sum_{r=1}^{R} \prod_{b=1}^{b} Q_{rb} \tag{9}$$

Where, Q_{rb} are easily computed univariate probabilities. The probabilities Q_{rb} are computed according to the following recursion: we first factor $\sum = \rho_{jb}$ using the Cholesky factorization $\sum = LL'$.

We get:

$$\begin{bmatrix} \varepsilon_1 \\ \cdots \\ \varepsilon_B \end{bmatrix} = N_B \begin{bmatrix} 0 \\ \dots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \dots & \rho_{1B} \\ \cdots & \dots. & \dots \\ \rho_{1B} & \dots & 1 \end{bmatrix} \tag{10}$$

Where L is a lower triangular matrix. The elements of L are $L = [l]_{bn}$, a lower triangular matrix. where $l_{bn} = 0 \;\; if \; n > b$. The recursive computation of probability, P, starts with,

$$Q_{r1} = \Phi(c_1 \,/\, l_{11}) - \Phi(a_1 \,/\, l_{11}) \tag{11}$$

Where, $\Phi(\mathrm{q})$ is the standard normal CDF evaluated at (q). Using the random number generator, ε_{r1} is a random draw from the standard normal distribution truncated in the range, $A_{r1} = a_1/l_{11}$ to $C_{r1} = c_1 \,/\, l_{11}$. The draw from the distribution is obtained using Gweke's method. It follows for steps $b = 2, \dots \dots, B$, and compute,

$$A_{rb} = \left[a_b - \textstyle\sum_{q=1}^{b-1} l_{bq}\varepsilon_{rb}\right] / \, l_{bb} \tag{12}$$

$$C_{rb} = \left[c_b - \textstyle\sum_{q=1}^{b-1} l_{bq}\varepsilon_{rb}\right] / \, l_{bb} \tag{13}$$

$$Q_{rb} = \Phi(C_{rb}) - \Phi(A_{rb}) \tag{14}$$

Finally, in preparation for the next step in the recursion, we generate random draws from the truncated standard normal distribution in the range A_{rb} to C_{rb}. This process is replicated R times, and the estimated probability is the sample average of the simulated probabilities. The GHK simulator has been found to be impressively fast and accurate for fairly moderate numbers of replications (Hajivassiliou and Ruud 1994*)*. Its main usage has been in computing functions and derivatives for maximum likelihood estimation of models that involve multivariate normal integrals.

4.5 Discussion of Results

The parameter estimates from C-MVP models of five worker clusters are depicted in Table 4.3 to Table 4.9. In these tables, night sleep, in-home chores, and in-home leisure activities are combined as in-home activities, while work and school activities are combined as a single work/school activity category. Some of the explanatory variables exhibit statistical significance within the 95% confidence interval (t-statistic greater than 1.96). Other variables with t-statistic less than the threshold value have been retained in the final model specification, with an assumption that if a larger data set were available, these parameters might show statistical significance. For the sake of brevity, we discuss only the more significant variables in this section.

4.5.1 Results of C-MVP Correlation Matrices

According to the simple correlation matrix, in Table 4.3, in-home activities have a negative correlation with all other out-of-home activities. For extended workers, work/school activities have a negative correlation with in-home activities, out-of-home entertainment, and sports activities, and positive correlation with shopping and organizational/hobbies activities. It implies that extended workers would choose to trip chain in home-work-home tour for shopping and organizational/hobbies activities. For 8 to 4 workers, work/school activity has a negative correlation with in-home and entertainment activities and positive correlation with shopping, organizational/hobbies, and sports activities. Individuals from 8 to 4 worker cluster appeared to be inclined to trip chain in their home-work-home journey for shopping, organizational/hobbies and sports activities. on the other hand, for shorter work-day workers, 7 to 3 workers and 9 to 5 workers clusters, the correlation between work and non-work activities have positive coefficients indicating that workers from all these three clusters are more likely to trip chain non-work activities with home-work-home trip chain. Consequently, for all the workers clusters, shopping, organizational/hobbies, entertainment, and sports activities are all positively correlated. It implies that it is more likely that workers would trip chain for non-work activities.

Table 4.3 Correlation matrix between different activity types for all workers clusters

Correlation Matrix for Cluster #1: Extended work-day workers						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-1.49*	-0.47*	-0.39*	-0.22*	-0.21*
Work/school		1	0.18*	0.12*	-0.05*	-0.01*
Out-of-Home Shopping			1	0.24*	0.28*	0.28*
Out-of-Home Organizational/hobbies				1	0.23*	0.14*
Out-of-Home Entertainment					1	0.20*
Out-of-Home Sports						1
Correlation Matrix for Cluster #2: 8-4 workers						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-0.89*	-0.46*	-0.43*	-0.40*	-0.36*
Work/school		1	0.04*	0.10*	-0.02*	0.04*
Out-of-Home Shopping			1	0.10*	0.21*	0.22*
Out-of-Home Organizational/hobbies				1	0.15*	0.10*
Out-of-Home Entertainment					1	0.13*
Out-of-Home Sports						1
Correlation Matrix for Cluster #3: Shorter work-day workers						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-1.34*	-0.69*	-0.68*	-0.56*	-0.56*
Work/school		1	0.30*	0.35*	0.25*	0.34*
Out-of-Home Shopping			1	0.31*	0.41*	0.30*
Out-of-Home Organizational/hobbies				1	0.30*	0.32*
Out-of-Home Entertainment					1	0.37*
Out-of-Home Sports						1
Correlation Matrix for Cluster #4: 7-3 workers						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-1.65*	-0.56*	-0.54*	-0.46*	-0.40*
Work/school		1	0.32*	0.33*	0.23*	0.18*
Out-of-Home Shopping			1	0.24*	0.27*	0.34*
Out-of-Home Organizational/hobbies				1	0.28*	0.25*
Out-of-Home Entertainment					1	0.29*
Out-of-Home Sports						1
Correlation Matrix for Cluster #5: 9-5 workers						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-1.60*	-0.61*	-0.60*	-0.51*	-0.55*
Work/school		1	0.33*	0.37*	0.27*	0.34*
Out-of-Home Shopping			1	0.22*	0.47*	0.30*
Out-of-Home Organizational/hobbies				1	0.22*	0.42*
Out-of-Home Entertainment					1	0.20*
Out-of-Home Sports						1

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

4.5.2 Results of the C-MVP Parameter Estimation

4.5.2.1 In-Home Activity Participation

According to the results presented in Table 4.4, individuals from all the population clusters are less likely to participate in in-home activities jointly. This may reflect differences of activity schedules with other household members. The size of the household is positively related to in-home activity participation for all the worker clusters except extended workers. Male individuals from all the clusters except 7 to 3 workers are less likely to engage in in-home activities. This is consistent with many other studies (Bhat et al. 2004). Presumably it reflects the trend of women to take a major part in in-home household maintenance responsibilities. On the other hand, perhaps male members of the 7 to 3 worker cluster share in-home household responsibilities more in the later part of the day as they finish work activity earlier than other clusters. With the increase in age the probability to engage in in-home activities for extended and 9 to 5 worker clusters decreases whereas for 8 to 4 workers it increases. Among the work-related attributes, being a paid worker, daytime work schedule, and more than one job are also found to be significant, but with mixed positive and negative signs among the five clusters. On the other hand, hours worked at main job, and flexible work schedule are found to have the same sign for all the clusters. With the increase in hours worked at main job, the propensity to engage in in-home activities decreases, due to the limited time budget.

Individuals living in duplex housing partake more in-home activities, though the reasons are readily apparent. An increase in mean commute time is significantly related to decreased in-home activity participation for extended workers and 9 to 5 workers.

Presumably this reflects that extended workers have limited time available for all other activities, and that 9 to 5 workers travel during peak hours, making their commute times longer and limiting time availability for other activities. Living in the urban core area, population density in the home neighborhood, and land use mix in the home neighborhood are also found to be significant for some clusters, but with both positive and negative effects.

Table 4.4 Output of C-MVP parameter estimates for in-home activity participation

Explanatory variables	**Extended**	**8 to 4**	**Shorter**	**9 to 5**	**7 to 3**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	-0.80*	-0.30*	-0.18*	-0.13*	-0.37*
Size of the Household	-0.02*	0.01*	0.01*	0.06*	0.03
Male (1, if the gender of the individual is male, 0 otherwise)	-0.19*	-0.10*	-0.12*	-0.17*	0.07*
Married (1, if the individual is married, 0 otherwise)	0.05	0.12*	-0.08	0.36*	0.14*
Age of the individual	-0.01*	0.02*	0.04	-0.01*	-0.01
Paid worker (1, if the individual is a paid worker, 0 otherwise)	-0.31*	0.31*	0.26*	-0.17*	-0.07
Working Day time (1, if the individual works in the day time, 0 otherwise)	-0.01	0.27*	-0.22*	0.13**	
Hours worked at main job	-0.03*	-0.03**	-0.01	-0.01	-0.06*
More than one job (1, if the individual works in more than one, 0 otherwise)	-0.14	0.11	0.01**	0.04	-0.01
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	0.04**		0.04**	0.01*	
Low income level (1, if the individual belongs to low income level, 0 otherwise)	0.01	0.06	-0.08	-0.12**	-0.02
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)		0.10**	0.01	0.18	0.13**
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)		-0.31*			0.12**
Mean commute time	-0.02*	0.01		-0.04*	0.02
Urban Core (1, if the individual is living in urban core, 0 otherwise)	0.01	-0.03*	-0.05*	0.02	0.01*
Retail floor area ratio in the home neighborhood			0.03*		0.01
Population density of the home neighborhood		0.01*	0.03	-0.01	0.01*
Land use mix in the home neighborhood		0.03		-0.02	-0.03*
Constant	1.24*	-0.72*	0.66*	1.07*	0.56*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

4.5.2.2 Out-of-Home Work/School Activity Participation

The parameter estimates of out-of-home work/school activity for worker clusters are presented in Table 4.5. For all the clusters it is less likely that they will conduct the work/school activity jointly. Male individuals from extended worker, 8 to 4 worker, and 7 to 3 worker clusters are more likely to engage in work activity. On the other hand, male individuals from the shorter work-day cluster are less likely to engage in work activity compared to female counterparts. This can be because the shorter work-day cluster has 71.0% female individuals with a flexible work schedule. Individuals from the extended worker cluster are less likely to engage in work/school activity as age increases. Possession of a valid driver license or bus pass, marital status, work related attributes, residential location in the urban core, and land use characteristics are found to be significant for some clusters, but with varying signs.

For extended workers, being male and paid worker show the highest coefficient values, both positive, whereas for shorter work-day workers the same variables are strongly negative to work activity. For 8 to 4 workers the two most significant factor affecting work activity participation are day time work schedule and having a low-income level, again both positive. For 7 to 3 workers marital status and multi-unit housing are important (both negative), and for 9 to 5 workers being a paid worker and possession of a valid driver license are most important (bot positive). The varying strengths and signs of the associations with workers' work/school activity participation confirms the heterogeneity among workers in their daily activity patterns.

Table 4.5 Output of C-MVP parameter estimates for out-of-home work/school activity participation

Explanatory variables	Extended	8 to 4	Shorter	7 to 3	9 to 5
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	-0.07*	-0.09*	-0.03	-0.10*	-0.01
Duration of the activity episode	0.01*			0.03*	0.02*
Size of the Household	0.01	-0.01	-0.03	-0.04	0.05
Male (1, if the gender of the individual is male, 0 otherwise)	0.23*	0.17*	-0.26*	0.11*	0.07
Married (1, if the individual is married, 0 otherwise)	-0.13	0.11*	0.08	-0.18*	-0.25**
Age of the individual	-0.01*	0.01**	0.06	0.01	0.01*
Driver License (1, if the individual has a valid driver license, 0 otherwise)	-0.15	-0.04		-0.03	0.26*
Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)	0.07	-0.12*			
Paid worker (1, if the individual is a paid worker, 0 otherwise)	0.21**	0.11*	-0.36*	0.09*	0.28*
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.01	0.31*	0.21*		0.06
Hours worked at main job	0.06*	0.04*	0.06**	0.01**	0.08*
More than one job (1, if the individual works in more than one, 0 otherwise)	0.19*	-0.01	-0.01*	0.01	-0.06**
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)		0.04*	-0.03	0.03	-0.08*
Low income level (1, if the individual belongs to low income level, 0 otherwise)	0.04	0.21*	0.16*		
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	-0.11	-0.08	0.02	-0.05*	0.22
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	0.10	-0.20		-0.10	0.09**
Mean commute time		-0.07**		-0.04*	-0.05
Population density of the home neighborhood	0.01		-0.01	-0.06	0.01*
Urban Core (1, if the individual is living in urban core, 0 otherwise)	0.20*	0.05**	0.01	0.12*	-0.05*
Land use mix in the home neighborhood		-0.02**		0.01	0.03**
Constant	-1.36*	-0.54*	-0.91*	-0.78*	-1.89*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

4.5.2.3 Out-of-home Shopping and Services Activity Participation

Table 4.6 presents the parameter estimates for shopping and services activities for all worker clusters. Across all the clusters, it is less likely that workers will participate in shopping and services activities jointly. This is because workers available time window may not match with others to conduct shopping activity jointly. Also, with the increase in the duration of shopping activity episode, workers from all the clusters are less likely to participate in shopping activities. This represents the law of diminishing marginal utility for shopping activities. Shorter work-day workers and 9 to 5 workers are more likely to engage in shopping activities if the household size increases. This is because larger household needs more good and supplies that may motivate these two clusters to participate in shopping and services activities. On the other hand, 8 to 4 workers are less likely to engage in shopping activities if the household size increases. This is relatable with the cluster membership as this cluster is mostly comprises middle-aged males with 92% having a full-time work. As expected being male is negatively associated to shopping and services activity participation for 8 to 4 workers. Individuals from 8 to 4 workers, 7 to 3 workers and 9 to 5 workers, with a valid driver license are more likely to participate in more shopping activities compared to others. This is because possession of a valid driver license provides greater convenience and opportunity to carry shopped goods via auto. Among the work-related attributes, being paid worker, working day time, hours worked at main job, having more than one job significantly affect different clusters in various magnitude. However, working in more than one job is significantly associated with more shopping activity participation for 8 to 4 workers and shorter work-day workers. Presumably, working in more than one job offers more affordability among workers from this cluster which motivates them to engage in more shopping and services activities. On the other hand, extended workers are less likely to engage in shopping activities if they work in more than one job. This is obviously related to the time constraints that affect the shopping activity engagement for extended workers.

Table 4.6 Output of C-MVP parameter estimates for out-of-home shopping and services activity participation

Explanatory variables	Extended	8 to 4	Shorter	7 to 3	9 to 5
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	-0.19	-0.03	-0.21*	-0.07**	-0.36*
Duration of the activity episode	-0.01*	-0.01*	-0.01*	-0.01*	-0.02*
Size of the Household	0.02	-0.05*	0.03**	-0.02	0.08*
Male (1, if the gender of the individual is male, 0 otherwise)		-0.07*	-0.03		0.05
Married (1, if the individual is married, 0 otherwise)	0.14*	0.10**	0.20*	-0.02	-0.66*
Age of the individual	0.09*	-0.03	-0.01**	-0.09*	
Driver License (1, if the individual has a valid driver license, 0 otherwise)	-0.17	0.08**		0.07*	0.23**
Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)	-0.36	-0.09**		0.13**	
Paid worker (1, if the individual is a paid worker, 0 otherwise)	0.25	-0.24*	0.05	0.22*	0.39*
Working Day time (1, if the individual works in the day time, 0 otherwise)	-0.02		0.15		0.14*
Hours worked at main job	0.01	-0.04	0.01	0.01*	0.01
More than one job (1, if the individual works in more than one, 0 otherwise)	-0.15**	0.01**	0.03*	0.05	
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	0.01*		-0.02	0.01	-0.01**
Low income level (1, if the individual belongs to low income level, 0 otherwise)	0.07	-0.04	0.09	-0.07*	-0.28*
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	-0.32	0.02	0.05	0.10**	0.10
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	-0.45	-0.12**		-0.28**	
Mean commute time	-0.03**	-0.08*		-0.01	-0.03*
Population density of the home neighborhood	-0.02	0.06	-0.03*	0.03	0.04
Urban Core (1, if the individual is living in urban core, 0 otherwise)	0.02	0.03**	0.01	0.04*	0.01
Land use mix in the home neighborhood	-0.04**	0.01**			0.03*
Constant	-2.31*	-1.30*	-1.44*	-0.96*	-2.71*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Individuals living in the urban core from shorter work-day and 7 to 3 worker clusters are more likely to engage in shopping activities. This may be because core living offers them greater accessibility, and also because these groups have more flexible time in a given day compared to other workers, which might positively affect their shopping activity participation. Finally, population density and land use mix are found to be significant for several clusters, with both positive and negative coefficient values.

4.5.2.4 Out-of-home Organizational and Hobbies Activity Participation

Table 4.7 shows the C-MVP parameter estimates for out-of-home organizational/hobbies activities. Unlike shopping activity, it is more likely that individuals of all the clusters will engage in out-of- organizational/hobbies jointly. This may reflect the social nature of these activities, which are often seen as family events. Participation in organizational/hobbies activity decreases for all worker clusters if the duration of the activity increases. This could be because workers employ their flexible time budget on these kinds of activities, and are more able to limit time expenditures for them, in comparison with mandatory activities. Married individuals from extended worker and 9 to 5 worker clusters are less likely to engage in organizational/hobbies activities compared to others. Presumably, married workers have shared responsibilities of other household related activities to perform in their leisure time, compared to those who are single. Individuals with flexible work schedule from the 8 to 4 worker and shorter work-day worker groups are more likely to engage in organizational and hobbies activities compared to others. Obviously, a flexible work schedule provides more flexible time in a given day, which motivates workers to perform these non-work activities.

Having a low annual income level is significantly related to participation in organizational and hobbies activities for extended workers, 8 to 4 workers, and 9 to 5 workers. This may reflect their lack of financial resources, since many such activities carry money costs.

Population density in the home neighborhood has a negative coefficient value for extended workers in the case of participation to out-of-home organizational/hobbies activities. Residential location in the urban core and land use mix in the home neighborhood have both positive and negative impacts on participation in organizational/hobbies activities.

Table 4.7 Output of C-MVP parameter estimates for out-of-home organizational and hobbies activity participation

Explanatory variables	Extended	8 to 4	Shorter	7 to 3	9 to 5
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	0.28*	0.22*	0.39*	0.38*	0.47*
Duration of the activity episode	-0.01*	-0.04*	-0.06*	-0.02*	-0.01*
Size of the Household	0.05	0.02**	0.11*	-0.03**	0.06
Male (1, if the gender of the individual is male, 0 otherwise)	-0.26*	0.07	-0.07**	0.11	
Married (1, if the individual is married, 0 otherwise)	-0.05*	-0.06	-0.04	-0.06	-0.34**
Age of the individual	-0.01	0.02*	-0.01	-0.05	-0.05**
Driver License (1, if the individual has a valid driver license, 0 otherwise)	0.27*	0.29*		0.31*	-0.26
Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)					
Paid worker (1, if the individual is a paid worker, 0 otherwise)	0.22**	-0.20*	0.20*	-0.07	
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.23**	0.05			-0.12
Hours worked at main job	0.04	0.04	-0.11**	-0.03**	-0.07
More than one job (1, if the individual works in more than one, 0 otherwise)	0.14**		-0.07	0.05	
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	0.02	0.04**	0.01*		0.01
Low income level (1, if the individual belongs to low income level, 0 otherwise)	-0.23*	-0.15**	0.03		-0.11**
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	0.22	0.02	-0.20*	0.11**	0.20**
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	-0.96	-0.18**	-0.01	-0.32**	-0.39
Mean commute time	-0.06*	-0.05*			0.08
Population density of the home neighborhood	-0.04*		-0.01		
Urban Core (1, if the individual is living in urban core, 0 otherwise)	-0.02*	-0.01**	0.03	0.05**	0.02*
Land use mix in the home neighborhood	0.07*	0.04		-0.02**	-0.01**
Constant	-2.48*	-2.25*	-1.64*	-0.99*	-0.31*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

4.5.2.5 Out-of-home Entertainment Activity Participation

Table 4.8 presents the C-MVP parameter estimates for out-of-home entertainment activities for worker clusters. Across the clusters, individuals are more likely to participate in entertainment activities jointly and less likely to participate if the duration of the entertainment activity increases. The probability to participate increases with household size for extended workers and 9 to 5 workers, but decreases for 8 to 4 workers. Increased household size provides greater availability of companionship for engaging in entertainment activities. It is interesting to note that being male is positively associated with entertainment activity participation for all worker clusters compared to their female counterparts. Individuals from extended worker and 7 to 3 worker clusters with a valid driver license are more likely to participate in entertainment activities compared to those without licenses. The reason is obvious as possession of a valid driver license provides the travelers with the ability to drive to entertainment locations and makes them far more accessible, and thus more likely to be visited. Also, as found earlier, workers are more likely to engage in entertainment activities jointly which might be easier to travel by an auto than other modes of travel. As expected having a valid bus pass is negatively associated with participation in entertainment activity for extended workers. Those reliant on bus travel have less free time for entertainment, owing to longer commutes, and also are less able to access entertainment locations.

Work-related variables are also found to significantly affect entertainment activity participation. For instance, hours worked at main job has a significant negative association to entertainment activities. Presumably, this is because extra work hours reduce free time available for entertainment activities. Household structure and low annual income both have mixed impacts on entertainment activities across all the clusters.

Table 4. 8 Output of C-MVP parameter estimates for out-of-home entertainment activity participation

Explanatory variables	**Extended**	**8 to 4**	**Shorter**	**7 to 3**	**9 to 5**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	1.11*	1.24*	0.95*	1.20*	0.88*
Duration of the activity episode	-0.02*	0.02*	-0.01**	-0.03*	-0.02*
Size of the Household	0.07*	-0.05**	-0.04	0.06	0.15*
Male (1, if the gender of the individual is male, 0 otherwise)	0.04	0.15*	0.13	0.06**	0.71*
Married (1, if the individual is married, 0 otherwise)	0.43**	-0.13	0.14**	0.20	0.30**
Age of the individual	0.07	0.02**	-0.01	-0.03	-0.07**
Driver License (1, if the individual has a valid driver license, 0 otherwise)	0.41*	-0.17		0.46*	
Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)	-0.29**	0.07		0.09	
Paid worker (1, if the individual is a paid worker, 0 otherwise)	0.51	-0.05**	0.02		
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.37*	-0.12**	-0.29*		-0.44*
Hours worked at main job	-0.06**	-0.06	-0.04**	-0.01*	
More than one job (1, if the individual works in more than one, 0 otherwise)	-0.48**	-0.09	-0.04**	0.06**	0.08
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	0.02	0.06*	0.05**	0.01	0.05
Low income level (1, if the individual belongs to low income level, 0 otherwise)	-0.18*	-0.17*	-0.13	0.06**	0.30
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	0.52*		0.19*	0.18	
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	-0.78	-0.41		-0.18	0.18**
Mean commute time	0.01			0.04	-0.07*
Population density of the home neighborhood	-0.03*	-0.05	0.05*		-0.05*
Urban Core (1, if the individual is living in urban core, 0 otherwise)	-0.02	0.04**	0.01	0.02**	0.03**
Land use mix in the home neighborhood		0.03*		0.01	0.04**
Constant	-4.32*	-1.56*	-2.33*	-3.00*	-3.01*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Individuals from 8 to 4 workers, shorter work-day workers, and 7 to 3 workers living in the urban core are more likely to engage in entertainment activities. Obviously, this is related to the locational advantage of higher accessibility to entertainment opportunities, compared to those who live far from the downtown core. Higher neighborhood land use mix is positively related to participation in entertainment activities for 8 to 4 workers and 9 to 5 workers. Higher land use mix often brings more alternative entertainment activity destinations within convenient proximity, increasing the probability that they will be visited.

4.5.2.6 Out-of-home Sports Activity Participation

Table 4.9 presents the C-MVP parameter estimates for out-of-home sports activity participation for all five worker clusters. Similar to entertainment activity, out-of-home sports activities are also more likely to be undertaken jointly for all the worker clusters. Also, with an increase in activity duration, the participation probability decreases, owing to diminishing marginal utility. With an increase in household size, the probability of participating in sports activity decreases significantly for workers from 8 to 4 and shorter work-day clusters. This may relate to the household resource availability needed for sports activity participation or increased time commitment needed for other household related activities. Male individuals from extended worker and 8 to 4 worker clusters are more likely to engage in sports activities than female counterparts, whereas males from shorter work-day and 7 to 3 clusters are less likely to engage in sports activities outside the home. However, married workers from 8 to 4 and 7 to 3 worker clusters are more likely to engage in out-of-home sports activities than those not married. Age is negatively associated with sports activity participation for 9 to 5 workers. This cluster comprises many older females, who are less sporty due to both physical constraints and their higher share in in-home activities. Possession of a valid driver license has significant negative impact on workers from 8 to 4 and 9 to 5 clusters. Presumably, driving provides them greater convenience and opportunity to engage in other discretionary activities, which affects their engagement in sports activities.

Work related attributes affect the out-of-home sports activity participation significantly with various magnitude for all the clusters. For instance, extended workers and 8 to 4 workers with flexible work schedule are less likely to engage in out-of-home sports activity. Workers with low income level are less likely to engage in sports activity, presumably because they have more resource constraints or limitation of time availability. Individuals from 8 to 4, shorter work-day, and 7 to 3 worker groups living in the urban core are less likely to engage in out-of-home sports activity than those living outside the core. This may reflect the age and health composition of these groups in the core, which tend to be older. However, activity participation in sports activity increases if the population density of the home neighborhood increases, for all the worker clusters. Neighborhood with higher population density are likely to have more provision of sports activity centers, which motivates workers to engage in out-of-home sports activities in their leisure time. On the other hand, land use mix has both positive and negative effects on sports activity participation among worker clusters.

Table 4.9 Output of C-MVP parameter estimates for out-of-home sports activity participation

Explanatory variables	Extended	8 to 4	Shorter	7 to 3	9 to 5
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	0.62*	0.34*	0.59*	0.42*	0.47*
Duration of the activity episode	-0.02**	0.01*	-0.03*	-0.01**	-0.04*
Size of the Household	-0.08	-0.06*	-0.11*	-0.04	
Male (1, if the gender of the individual is male, 0 otherwise)	0.16**	0.05**	-0.23**	-0.28*	
Married (1, if the individual is married, 0 otherwise)	0.04	0.12**	-0.05**	0.27*	-0.42
Age of the individual	0.05**	-0.04	-0.02	0.01*	-0.01**
Driver License (1, if the individual has a valid driver license, 0 otherwise)	0.12	-0.24*		-0.13	-0.03**
Paid worker (1, if the individual is a paid worker, 0 otherwise)	-0.13	0.20	0.38*		0.38
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.02	0.25	0.09		-0.36
Hours worked at main job	-0.01	0.09*	-0.07*	-0.01	-0.02
More than one job (1, if the individual works in more than one, 0 otherwise)	-0.26**		0.03	-0.08	0.07
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	-0.07*	-0.09*	0.04		-0.06
Low income level (1, if the individual belongs to low income level, 0 otherwise)	0.01	-0.07	-0.06	-0.07*	-0.12**
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	0.30	-0.04	-0.35*	0.37*	0.39
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	-0.52	0.25		0.37	0.11
Mean commute time	0.02	-0.07		0.04	-0.02*
Population density of the home neighborhood	0.04*	0.05*	0.01*	0.01*	0.04*
Urban Core (1, if the individual is living in urban core, 0 otherwise)	-0.02	-0.03**	-0.01*	-0.03**	0.01
Land use mix in the home neighborhood	-0.02*	0.01		0.02*	-0.01
Constant	-0.74*	-3.64*	-2.51*	-2.35*	-1.79*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

4.5.3 Transition Matrices

Activity episode sequences of each worker cluster were analyzed to produce transition probability matrices. Each transition matrix shows the likelihood of the occurrence of a certain activity category as a successive episode. The rows in Table 4.10 represent the activity category of the current episode, and the columns represent the category of the subsequent activity episode.

For C*luster#1*, in-home leisure, in-home household chores, and night sleep activities are typically followed or preceded by other in-home activities. When individuals from this cluster follow in-home activities by an out-of-home activity, it is most likely to be work. Out of the home, work is followed most often by entertainment, followed by home chores, shopping, and organizational/hobbies. Both organizational/hobbies and entertainment are frequently followed by work, whereas sports are often followed by home chores. It is clear from the transition matrix that all the out-of-home activities tend to have more than one stop in the trip chain.

In the case of *Cluster#2*, in-home activities are mostly preceded and followed by one or more other in-home activity. The activity episode subsequent to work is often entertainment or home chores. Shopping activities are most often followed by work activity, and to a lesser extent by entertainment activity and organizational/hobbies activity. The immediate subsequent activity of organizational/hobbies activities is either work or in-home chores. Entertainment activity is typically followed by a work activity, whereas sports activity is followed by in-home chores or work activity. Shopping, organizational/hobbies, and entertainment activities typically have more than one out-of-home stop in the trip chain.

Cluster#3 is the cluster of shorter work-day workers. For this cluster, the transitions from current activity are similar to those for cluster 2, except that organizational/hobbies and sports are far more likely to be followed by home chores.

The transitions from current to next activity for *Cluster#4* are similar to those for *Cluster#3,* except individuals from this cluster are more likely to undertake discretionary activities before work activity, or produce tours for discretionary activities. This is because, except work, all the discretionary activities are either followed by work activity or in-home household chores. On the other hand, work activity often has more than one stop in the trip chain, for engaging in entertainment and sports activities.

Table 4.10 Activity episode transitions (in percentage) matrix

Cluster#1: Extended work-day workers

	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.48	0.31	0.10	-	-	0.05	-	0.06
L	0.59	-	0.37	0.03	-	-	-	0.01	0.01
N	0.93	0.04	-	-	-	-	-	-	0.02
W	0.27	0.03	-	-	0.16	-	0.16	0.31	0.08
P	0.11	-	0.05	0.63	-	-	0.11	0.11	-
S	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00	0.00
G	0.28	0.11	-	0.44	0.08	-	-	0.06	0.03
E	-	-	0.03	0.82	0.05	-	0.08	-	0.03
T	0.74	0.07	-	0.15	-	-	0.04	-	-

Cluster#2: 8-4 workers

	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.59	0.28	0.05	0.01	-	0.02	0.00	0.05
L	0.64	-	0.30	0.02	0.00	-	0.01	0.01	0.02
N	0.96	0.03	-	0.00	-	-	0.00	-	0.00
W	0.21	0.03	0.00	-	0.11	-	0.16	0.34	0.14
P	0.02	0.07	-	0.49	-	0.05	0.14	0.16	0.07
S	0.00	0.00	0.00	1.00	0.00	-	0.00	0.00	0.00
G	0.21	0.06	0.04	0.62	0.02	-	-	0.05	-
E	0.04	0.01	0.02	0.81	0.01	-	0.01	-	0.10
T	0.51	0.10	-	0.27	0.05	-	0.01	0.06	-

Cluster#3: Shorter work-day workers

	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.59	0.23	0.07	0.01	-	0.03	0.01	0.05
L	0.72	-	0.23	0.03	0.00	-	0.00	0.00	0.01
N	0.94	0.04	-	-	-	-	-	-	0.02
W	0.47	0.05	-	-	0.14	-	0.05	0.24	0.05
P	0.15	0.12	-	0.42	-	-	0.19	0.04	0.08
S	0.00	0.00	0.00	0.00	0.00	-	1.00	0.00	0.00
G	0.47	0.06	0.03	0.24	0.03	0.03	-	0.12	0.03
E	0.14	0.03	-	0.76	0.07	-	-	-	-
T	0.77	0.08	-	0.13	0.03	-	-	-	-

Cluster#4: 7-3 workers

	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.64	0.24	0.06	0.00	-	0.02	-	0.03
L	0.70	-	0.24	0.03	0.00	-	0.00	0.00	0.02
N	0.97	0.01	-	0.01	-	-	0.01	-	0.00
W	0.26	0.10	0.01	-	0.08	0.01	0.09	0.34	0.11
P	0.38	-	-	0.38	-	-	0.08	0.08	0.08
S	-	-	-	1.00	-	-	-	-	-
G	0.34	0.09	-	0.44	0.06	-	-	0.03	0.03
E	-	0.02	-	0.94	-	-	0.03	-	0.02
T	0.40	0.13	-	0.38	0.04	-	0.02	0.02	-

Cluster#5: 9-5 workers

	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.69	0.17	0.03	0.01	-	0.03	0.00	0.06
L	0.74	-	0.20	0.01	0.00	-	0.02	0.00	0.02
N	0.86	0.13	-	0.00	-	-	0.01	-	-
W	0.70	0.09	-	-	0.09	-	0.06	0.03	0.03
P	0.19	0.04	-	0.19	-	0.02	0.29	0.15	0.13
S	0.00	0.00	0.00	1.00	0.00	-	0.00	0.00	0.00
G	0.61	0.11	0.02	0.03	0.11	-	-	0.05	0.08
E	0.06	0.06	-	0.06	0.28	-	0.33	-	0.22
T	0.61	0.22	0.01	0.01	0.03	-	0.07	0.04	-

H = Home chores
L = Home leisure
N = Night sleep
W = Workplace
P = Shopping
S = School/college
G = Organizational/hobbies
E = Entertainment
T = Sports
Horizontal axis = preceding activity
Vertical axis = succeeding activity

In the case of *Cluster#5*, the activity subsequent to work is in-home chores. The activity subsequent to shopping is organizational/hobbies, shopping, or work. Similarly, entertainment activity has trip chains of shopping, organizational/hobbies, and sports activity episodes. Sports activity is typically followed by in-home activities. Unlike other clusters, individuals from this cluster produce discretionary trip chains which are not necessarily preceded or followed by any work activity.

4.6 Conclusions

The aim of this study was to investigate the activity-travel behavior of workers through an innovative cluster-based Multivariate Probit Modeling (C-MVP) framework. Five worker clusters had previously been identified using a daily activity pattern recognition method: Cluster#1: extended work-day workers, Cluster#2: 8 to 4 workers, Cluster#3: shorter work-day workers, Cluster#4: 7 to 3 workers, and Cluster#5: 9 to 5 workers.

The correlation matrices for all the clusters show that in-home activities have a negative correlation with all the out-of-home activities, with varying magnitude for each cluster. For shorter work-day workers, 7 to 3 workers and 9 to 5 workers, all the out-of-home activities have positive mutual dependence. However, work/school activity has a negative correlation with entertainment and sports activity for extended workers, whereas for 8 to 4 workers work/school activity has a negative correlation with entertainment activity. This implies that extended-day workers are less likely to trip chain entertainment and sports activities within the home-work-home tour, and 8 to 4 workers are less likely to trip chain entertainment activity. However, shorter work-day workers, 7 to 3 workers, and 9 to 5 workers are more likely to trip chain the non-work activities within the home-work-home tour. As well as for all the clusters, workers are more likely to make complex non-work tours.

This study uses these identified worker clusters for C-MVP model estimation, assuming a non-zero correlation between the types of activities in which individuals participate in a given day. The explanatory variables include both individual and household characteristics, and characteristics of the neighborhood of residence. Dependent variables are six activity categories: in-home activities, work/school, shopping, organizational/hobbies, entertainment, and sports. Based on the results, we conclude that activity participation of workers is significantly associated with their socio-demographic characteristics, individual characteristics, household structure, accompaniment arrangement, commute time, and neighborhood land use attributes. The model coefficients vary considerably between clusters in magnitude, sign, and significance, showing the value of segmenting the population into homogeneous clusters.

Across all the clusters, it is more likely that in-home activities will be conducted in solo. For out-of-home activities, shopping activities of extended workers, shorter work-day workers, and 9 to 5 workers, and work/school activities of 8 to 4 worker, are less likely to be jointly conducted. Duration of each activity also has a significant impact on the participation propensity for that specific activity. As household size increases, extended-day workers are more likely to engage in out-of-home entertainment activities and less likely to engage in in-home activities. With an increase in population density in the home neighborhood, participation in out-of-home sports activities increases. Also, with an increase in land use mix, the propensity to participate in out-of-home organizational/hobbies activities increases.

To better understand the activity dependence, we estimated the transition probability matrices for each cluster. Transition matrices show the most frequent subsequent activity after the current activity episode. In-home activities are most frequently followed by work, for most clusters. Extended-day workers and 7 to 3 workers have similar activity transitions and simpler out-of-home trip chains than other clusters. In the case of shorter work-day workers and 7 to 3 workers, shopping, work/school, and organizational/hobbies activities have at least one out-of-home subsequent activity other than work activity. The transition matrix for 9 to 5 workers shows that the subsequent activity of a current discretionary activity will be another discretionary activity (for instance, shopping will be followed by organizational/hobbies activity) or individuals will make dedicated discretionary tours rather than chain the maintenance and discretionary activities with work activity.

Workers' population groups contribute the largest share to total urban traffic. Moreover, they have a regular and relatively inflexible activity schedule in comparison to non-workers. Thus, this study contributes through offering a cluster-based modeling approach to better analyze the activity-travel behavior of workers. The application of C-MVP with correlation matrices estimated by GHK is also a contribution that can capture the dependency among activities. Future study includes comparing activity-travel behavior of workers with other population groups, including non-workers and students. The results and modeling framework are expected to be incorporated into the Scheduler for Activity, Location and Travel (SALT) model for Halifax.

Chapter 5 Out-of-Home Activity Choices and Activity Transitions for Non-Worker Population Groups[3]

5.1 Introduction

Analyzing activity patterns and travel behavior has become a major concern in transportation research due to the shift toward activity-based models (Arentze and Timmermans 2000; Auld and Mohammadain 2012). In activity-based modeling, it is important to understand why an individual participates in an activity, with whom, for how long, and how frequently. Activities of individuals are divided into mandatory and non-mandatory (discretionary) activities. Most studies focus on modeling the mandatory and non-mandatory activities of working groups of individuals. For instance, Damm (1980) examined the timing of non-work activity. Stopher et al. (1996) developed an activity-based model of travel for mandatory, flexible, and optional activity categories. Arentze and Timmermans (2000) developed a computational process model with fixed and flexible activities. In 2004, Timmermans et al. (2004) developed a hybrid model for activity-travel patterns of leisure and vacation travel. The same year, Bhat et al. (2004) developed a comprehensive Econometric Micro-simulator for simulating the daily activity-travel behavior of workers as well as non-workers. However, the activity-travel behavior of non-workers can be very different from that for workers (Misra 1999), as they don't have any fixed activities at fixed locations, e.g. work or school activities.

---

[3] An earlier version of this chapter has been presented:
Daisy, N. S., Millward, H. and Liu, L. Modeling activity-travel behavior of non-workers grouped by their daily activity patterns. Proceedings of the 15th International Conference on Travel Behavior Research, Santa Barbara, California, USA., 2018.

In reality, due to their flexible activity-schedule, non-workers' participation in out-of-home activities in a typical day is more difficult to model and analyze (Daisy et al. 2017b). Non-workers' activity patterns are more flexible and discretionary in nature. Non-workers mostly include retired individuals, homemakers, job seekers, and individuals without a job, most of whom have more time and flexibility to choose a daily activity-travel schedule in comparison to workers. The difference between workers' and non-workers' activity-travel behavior has a significant impact on the transportation system, as well as on transportation-related policies and planning (Daisy et al. 2017c). For instance, Bricka (2008) studied the trip chaining decisions of workers and non-workers and found that policies need separate measures for individuals who work and individuals who don't. However, the empirical investigation of non-workers' activity-travel behavior is very limited compared to that for workers (Bricka, 2008; Misra and Bhat 2000; Manoj and Verma 2015; Habib et al., 2016; Hafezi et al. 2018a, b). Therefore, a comprehensive econometric investigation with clustered activity patterns as input can reveal insights to better understand non-workers' activity-travel behavior. This study contributes to this research area by employing such an investigation.

This study employs the Halifax Space-Time Activity Research (STAR) data for empirical models. The first step was to cluster non-workers based on their daily activity patterns. Hafezi et al. (2017a, b) identified five non-worker clusters from the STAR data. These clusters have significantly different activity patterns along with heterogenous socio-demographic attributes. Thus, it is important to identify the relationships between socio-demographic characteristics and activity-travel behavior of individuals, particularly for planning and policy implications. We utilize a multivariate probit modeling (C-MVP) approach to determine the significant factors related to activity patterns for each non-worker cluster. The use of an C-MVP model is regarded as an important contribution in this research area, since activities in each day are time-correlated phenomena rather than independent choices.

5.2 Literature Review

To achieve the ultimate benefits from people-friendly transportation planning, it is important to analyze the activity-travel behavior of segments of travelers, including workers and non-workers. Recent research emphasizes the need to comprehend the activity-travel behavior of non-workers. For instance, Azari et al. (2013) found that it is more likely that non-workers will be more sensitive to policies like cordon pricing or parking pricing, but they are less sensitive to increased travel time in comparison to workers. Another study by Shiftan and Golani (2005) found that workers are more likely to change their mode of travel, whereas non-workers are more likely to change the trip destination or to abandon the trips. Susilo and Kitamura (2005) examined long-term activity-travel behavior and observed that non-workers were more likely to spread their activities over the study period whereas workers are less likely to do so. Ye et al. (2008) found significant differences among individuals who work and those who don't with regard to time-allocation to different activities. Bricka (2008) found that non-workers highly chain their trips compared to worker counterparts. Dharmowijoyo et al. (2014) found that non-workers make a lower number of trips, and have less dependency on auto travel compared to workers. A study by Liu et al. (2014) revealed that weather condition has a significant impact on non-workers' activity and travel engagements in comparison to workers. In summary, the studies mentioned above show that activity-travel behavior and factors influencing it are significantly different between workers and non-workers.

Non-workers comprise a growing portion of the adult population, so it is important to study their activity-travel behavior for effective transportation planning and policies. Misra and Bhat (2000) observed that non-workers' activity travel behavior needs more empirical investigation. They showed that non-workers significantly contribute to household-related activities and they schedule those activities earlier in the day. Bhat and Misra (2001) developed a comprehensive econometric model to analyze the activity-travel behavior of non-workers. A discrete choice model (multinomial logit) was utilized to study the daily-activity-pattern of non-workers, and it was found that they arranged services to other household members earlier in the day and these activities were not linked with other activities.

Yamamoto et al. (2000) studied the responsiveness of non-workers regarding policies like congestion pricing. They found that there were significant differences between workers and non-workers in utility perception and values related to transportation, and different travel behavior. Manoj and Verma (2015) studied the activity-travel behavior of non-workers from various socio-economic settings for individuals in Bangalore city, India. By utilizing univariate regression, the authors estimated the effects of socio-demographic attributes, land use characteristics, and travel features on non-workers' activity engagement, time-of-day choice, trip chaining, and mode choice. Habib et al. (2016) studied non-workers' activity-travel schedule behavior with a comprehensive random utility maximizing travel options model. The study found that socio-demographic characteristics and the presence of children have a significant impact on non-workers' activity schedules.

However, non-workers' activity-travel behavior is still an under-researched area (Misra and Bhat 2000; Manoj and Verma 2015; Habib et al. 2016). The present study contributes to the literature by shedding light on the activity-travel behavior of non-workers from Halifax city. Recent work by Hafezi et al. (2017a) identified five representative activity-travel patterns of non-workers based on their daily activity patterns, and noted significantly different socio-demographics between the five population groups. Here we extend the empirical study of these five identified clusters regarding their activity participation behavior.

5.3 Data

This study utilizes data from the Halifax Space-Time Activity Research (STAR) project conducted between April 2007 and May 2008 (TURP, 2008). STAR was the world's first large survey to use global positioning system (GPS) tracking for verification of household activity-travel diary data (Bricka 2008). HRM is the largest municipality in Atlantic Canada as well as the capital of Nova Scotia. It is a mid-sized metropolitan area (c.400,000 population) with a diverse and developing economy, with 0.5% per year population growth. The sample size consists of 1,971 randomly selected households in Halifax Regional Municipality (HRM), which represents almost one household in 78. Activity-travel diary and questionnaire data of almost were collected through 373 days of the year, with a total participation rate of 21% (Millward and Spinney 2011).

The primary respondent in each sample household was selected randomly, and had to be more than 15 years age. These respondents completed a detailed time-diary for two consecutive days. The time diary coding and questionnaire on household characteristics were based on the General Social Survey (GSS) Cycle 19 Survey of Statistics Canada′s (2006). Primary respondents also carried a GPS-device (Hewlett Packard iPAQ hw6955) for all out-of-home activity, programmed to collect GPS data every 2 s. The GPS data provided travel routes and precise start and end times for all "stops" with more than 2 minutes stopping duration. These GPS data were used with CATI software in day-after interviews with respondents, to verify and enhance the time-diary data.

5.3.1 Description of Clusters

Initially, a pattern recognition model was applied to the Halifax STAR household activity data (Hafezi et al. 2017a, b). A subtractive clustering algorithm was utilized to initialize the total cluster number and cluster centroids. Identification of individuals with homogeneous activity patterns was accomplished using a fuzzy c-means clustering algorithm, and sets of representative activity patterns were identified using a multiple sequence alignment method. Advanced decision tree models were used to explore inter-dependencies in each identified cluster, and characterization of cluster memberships through their socio-demographic attributes was achieved by use of the CART algorithm (see for more details Hafezi et al., 2017c). Figure 5.1 shows the representative daily activity patterns of individual activities for the five identified non-worker clusters.

Figure 5.1 Representative activity patterns of identified five non-worker clusters

NNNNNNNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHPPPPPPPPPPPPGGGGGGGGGGGGGGGGGGHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHLLLLLLLLLLLLLLLLLLLLLLLN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Cluster#1: Non-worker, midday activities

NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEELLLLLLLLLLLLLLLLLLLLNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Cluster#2: Non-worker, evening activity

NNNNNNNNNNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHLLLLL
LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Cluster#3: Stay-at-homes

NNNNNNNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHHHTTTTTTTPPPPPPPP
PPPPPPPHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Cluster#4: Non-worker, morning shopping

NNNNNNNNNNNNNNNNNNNNNNNNNNNNNHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHPPPPPPPPPPPPPHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Cluster#5: Non-worker, afternoon shopping

**Representative pattern starts from 4:00 a.m. to 3:59 a.m. on the next day (288 five minutes intervals); In-home (H/L/N), Workplace/School (W/S), Shopping (P), Organizational/hobbies (G), Entertainment (E), Sports (T)*

Table 5.1 and Table 5.2 shows the socio-demographic attributes and activity time-use of all the non-worker clusters, respectively. We used six activity categories for model estimation: in-home, out-of-home mandatory (work/school), out-of-home shopping and services, out-of-home organizational and hobbies, out-of-home entertainment, and out-of-home sports. Table 5.2 shows the details of the activity participation in these categories, in terms of average time allocation, for the five population clusters. Percentage values in bold show important activities for each group. Our group names reflect both the time allocations and the activity sequences for each group. The details of each cluster are as follows:

Cluster#1 is the non-worker midday activities cluster. The individuals belonging to this cluster participated in entertainment and organizational/hobbies activities predominantly in the midday. This cluster mostly comprises individuals more than 55 years age (66.0%), retired (52.0%), and female (53.0%). Only 35% members have bachelor or above degree and belong to low or middle income level. The predominant out-of-home activities of this cluster are organizational/hobbies, entertainment, shopping and services, and sports activities.

Cluster#2 is the non-worker evening activity cluster. These non-workers participate in out-of-home organizational/hobbies and entertainment activities, mostly in the evening. Similar to the previous cluster, a large proportion of individuals are female (59.0%), and older than 55 years (53.0%). Among all the individuals 37% have bachelor or above degree, and mostly retired (39.0%) or work at home (15.0%). These individuals have low to middle income, and those who work have a flexible work schedule (54.0%). Among the out-of-home activities, this group predominantly participate in organizational/hobbies and entertainment activities.

Cluster#3 comprises non-workers with a stay-at-home representative activity pattern. A large proportion of this group is old-aged females. Individuals in this group mostly belong to the low-income partition. This cluster has the largest membership in comparison to other clusters.

Cluster#4 is the non-worker morning shopping activity cluster. Similar to earlier clusters, this cluster also largely comprises females more than 55 years age. Many individuals in this cluster do not have a high-school diploma and most have a low-income level (53.0%). They are mostly retired. The majority of those who work have flexible working hours (52.0%), and some portion work at-home (11.0%). Shopping and services are the most predominant out-of-home activity, along with sports and organizational/hobbies activities.

Cluster#5 comprises individuals who do shopping and services activities in the afternoon. Similar to other clusters, the members of this group are mostly older females. A large proportion have low-income, but their education level is fairly high.

Table 5.1 Analysis of non-worker cluster data: Share of socio-demographic variables

Social demographic variables		**Sample mean (%)**	**Mean of cluster (%)**				
			#1	#2	#3	#4	#5
Gender	Female	0.56	0.53	0.59	0.56	0.54	0.59
Age	Young adults (ages 15-35 years)	0.08	0.05	0.10	0.11	0.09	0.07
	Middle-aged adults (ages 36-55 years)	0.32	0.29	0.38	0.32	0.29	0.32
	Older adults (aged older than 55 years)	0.60	0.66	0.53	0.57	0.63	0.61
Education Level	Bachelor degree and above	0.34	0.35	0.37	0.35	0.30	0.35
Occupation	Regular shift	0.23	0.22	0.26	0.24	0.19	0.24
	Irregular schedule	0.09	0.10	0.10	0.11	0.07	0.07
	Student	0.02	0.00	0.04	0.01	0.03	0.02
	Retired	0.45	0.52	0.39	0.41	0.53	0.41
	Work at home	0.12	0.10	0.15	0.16	0.11	0.09
Flexible schedule	Have no flexibility in a work schedule	0.47	0.48	0.46	0.43	0.48	0.51
Job number	Have more than one job	0.09	0.04	0.16	0.08	0.11	0.05
Income	Low-income (<= $ 40,000)	0.48	0.44	0.48	0.49	0.53	0.48
	Middle-income ($ 40,000 - $ 100,000)	0.45	0.46	0.45	0.45	0.42	0.46
	High-income (> $ 100,000)	0.07	0.10	0.07	0.07	0.05	0.06
Total cluster membership			225	238	419	247	262
Percentage in total (number of person-days)			8.10	8.57	15.08	8.89	9.43

Table 5.2 Analysis of cluster data: Share of activity time-use of all non-worker clusters

Activity categories	Descriptions	Share of daily activity engagement (%)				
		#1	#2	#3	#4	#5
In-home (*H/L/N*)	Home Chores (H): Working at home, eating/meal preparation, indoor or outdoor cleaning, interior or exterior home maintenance, child care or other in-home activities.	28.36	32.75	41.10	35.94	35.75
	Home Leisure (L): Watching TV/listening to radio, reading books/newspapers, etc.	16.54	11.41	18.28	17.70	17.09
	Night sleep (N)	36.76	34.96	35.98	35.02	36.26
Total in-home (%)		*81.66*	*79.11*	*95.36*	*88.65*	*89.10*
Workplace/School (W/S)	Work (W): Work/job, all other activities at work, work related (conferences, meetings, etc.).	0.87	1.34	0.38	1.20	1.43
	School/college related (S): Class participation, all other activities at school.	0.18	0.25	0.02	0.36	0.14
Shopping & services (P)	Shopping for goods and services, routine shopping.	5.00	3.20	1.40	**3.66**	**4.17**
Organizational/ hobbies (G)	Organizational, voluntary, religious activities. Hobbies are done mainly for pleasure, cards, board games, all other hobbies activities.	**5.42**	**6.53**	0.92	2.40	2.39
Entertainment (E)	Eat meal outside of home, all other entertainment activities.	3.21	**7.05**	0.48	0.99	1.27
Sports (T)	Walking, jogging, bicycling, all sports related activities.	3.65	2.52	**1.44**	2.75	1.50
Total (%)		*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

5.4 Methods

A wide variety of methods have been utilized to study the propensity to select, execute, and schedule activities through perceived utility functions of individuals. Among these are Random Utility Models (Adler and Ben-Akiva 1979; Bhat et al. 2004), mathematical programming models (Recker 2001), rule-based and need-based models (Arentze and Timmermans 2009), and complex simulation models (Pendyala et al. 2005). However, the application of probit models in activity-based modeling is mostly limited to univariate binary models, or mixed models (Allahviranloo and Recker 2014). For instance, Bhat and

Srinivasan (2005) utilized an ordered binary probit model to investigate the stop frequency modeling for different activities. Lamondia et al. (2010) employed binary probit models to analyze activity participation behavior of individuals. Ruiz and Roorda (2008) estimated a multivariate probit model using weighted least-squares with mean and variance correction estimator to examine the decision-making process of activity companionship, planning, scheduling, and execution. Allahviranloo and Recker (2014) estimated a multivariate probit model with full correlation matrix with Markov Chain Monte Carlo (MCMC) to study the dependency between daily activity type choices of individuals and their socio-demographic characteristics along with correlation among activity choices.

Most earlier studies assumed the activity participation as an independent phenomenon in multivariate cases, resulting in either logit or mixed logit models. However, activity participation in one activity is associated with both the previous and next activity, and varying sequences have greater likelihoods. These interdependencies between activities can be captured with the correlation matrix of the C-MVP model. These interdependencies between activities can be captured with the correlation matrix of the C-MVP model. The empirical specification of activity choice over the six categories (in-home, work/school, shopping & services, organizational/hobbies, entertainment, and sports) can be analyzed empirically in two ways, by either multinomial or multivariate probit models. Multinomial models assume that random error terms of choice equations are independent (Greene, 2003). For our current study, the choices of activity types are not mutually exclusive, and therefore the error terms of the activity type choices may be mutually inclusive and correlated. Consequently, we chose to use a multivariate probit (C-MVP) model, which allows for the possible correlation in the activity choices simultaneously.

The C-MVP model consists of a set of m binary dependent variables y_m^* (observation subscript $i = 1,2, \ldots\ldots, n$ has been suppressed), where the M-equation multivariate probit model framework:

$$y_1^* = x_1'\alpha_1 + \varepsilon_1, \quad y_1 = 1 \quad if\ y_1^* > 0 \tag{1}$$

$$y_2^* = x_2'\alpha_2 + \varepsilon_2, \quad y_2 = 1 \quad if\ y_2^* > 0 \tag{2}$$

$$\vdots$$

$$y_m^* = x_m'\alpha_m + \varepsilon_m, \quad y_m = 1 \quad if\ y_m^* > 0 \tag{3}$$

$$E[\varepsilon_m | x_1, \ldots \ldots, x_M] = 0 \tag{4}$$

$$Var[\varepsilon_m | x_1, \ldots \ldots, x_M] = 1 \tag{5}$$

$$Cov[\varepsilon_1, \varepsilon_m | x_1, \ldots \ldots, x_M] = \rho_{im} = \sum \tag{6}$$

$$(\varepsilon_1, \ldots \ldots, \varepsilon_M) \sim Multivariate\ normal\ (MVN)[0, \Omega] \tag{7}$$

Where, m is the activity types, i.e. in-home, shopping, mandatory, sports, organizational, entertainment. X is a vector of explanatory variables, $\alpha_1, \alpha_2, \ldots\ldots \alpha_m$ are conformable parameter vectors, and $\epsilon_m, M = 1 \ldots . M$ are random errors distributed as multivariate normal distribution with a mean of zero, unitary variance and a correlation matrix, which is 6×6 for this study, $Q = [q_m]$. The density function of the equation will be $\Phi(\varepsilon_1, \varepsilon_2, \ldots \ldots \varepsilon_m; Q)$. Ω is the correlation matrix, and $\sum$ is the covariance matrix. Each individuals equation is a standard probit model.

In Geweke-Hajivassiliou-Keane (GHK) simulation, the approximation is computed based on averaging R draws from a certain multivariate normal distribution, for each observation (Greene 2003; Hajivassiliou and Ruud 1994). It assumes that a multivariate normal distribution function can be articulated as the product of sequentially conditioned univariate normal distribution functions.

The joint probabilities of the observed events, $[y_{i1}, y_{i2}, \dots \dots, y_{iM} | x_{i1}, x_{i2}, \dots \dots, x_{iM}], i = 1, \dots \dots, n$ that form the basis for the log-likelihood function are the M-variate normal probabilities,

$$L_i = \sum_{i=1}^{N} \Phi_M (q_{i1} x_{i1}' \alpha_1, \dots \dots, q_{iM} x_{iM}' \alpha_M | \Omega^*) \tag{8}$$

Where:

$q_{im} = 2y_{im} - 1,$

$\Omega_{mn}^* = q_{im} q_{in} \rho_{mn}$

The practical obstacle to this extension is the evaluation of the M-variate normal integrals and their derivatives. However, given the speed of modern computers, simulation-based integration using the GHK simulator or simulated likelihood methods allows for estimation of relatively large models. Among the simulation methods examined by Greene (2003), the GHK smooth recursive simulator appears to the most accurate. The general approach uses,

$$P = pr[a_1 < \varepsilon_1 < b_1, \dots \dots, a_m < \varepsilon_m < b_m] \approx \frac{1}{R} \sum_{r=1}^{R} \prod_{m=1}^{m} Q_{rm} \tag{9}$$

Where, Q_{rm} are easily computed univariate probabilities. The probabilities Q_{rm} are computed according to the following recursion: we first factor $\sum = \rho_{jm}$ using the Cholesky factorization $\sum = LL'$. We get:

$$\begin{bmatrix} \varepsilon_1 \\ \cdots \\ \varepsilon_M \end{bmatrix} = N_M \begin{bmatrix} 0 \\ \ldots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \ldots & \rho_{1M} \\ \cdots & \ldots. & \ldots \\ \rho_{1M} & \ldots & 1 \end{bmatrix} \tag{10}$$

Where L is a lower triangular matrix. The elements of L are $L = [l]_{mn}$, a lower triangular matrix. where $l_{mn} = 0 \ if \ n > m$. The recursive computation of probability, P, starts with,

$$Q_{r1} = \Phi(b_1 \,/\, l_{11}) - \Phi(a_1 \,/\, l_{11}) \tag{11}$$

Where, $\Phi(\mathrm{q})$ is the standard normal CDF evaluated at (q). Using the random number generator, ε_{r1} is a random draw from the standard normal distribution truncated in the range, $A_{r1} = a_1/l_{11} \ to \ B_{r1} = b_1 \,/\, l_{11}$. The draw from the distribution is obtained using Gweke's method. It follows for steps $m = 2, \ldots \ldots, M$, and compute,

$$A_{rm} = \left[a_m - \textstyle\sum_{q=1}^{m-1} l_{mq}\varepsilon_{rm}\right] / \, l_{mm} \tag{12}$$

$$B_{rm} = \left[b_m - \textstyle\sum_{q=1}^{m-1} l_{mq}\varepsilon_{rm}\right] / \, l_{mm} \tag{13}$$

$$Q_{rm} = \Phi(B_{rm}) - \Phi(A_{rm}) \tag{14}$$

Finally, in preparation for the next step in the recursion, we generate random draws from the truncated standard normal distribution in the range A_{rm} to B_{rm}. This process is replicated R times, and the estimated probability is the sample average of the simulated probabilities.

The GHK simulator has been found to be impressively fast and accurate for fairly moderate numbers of replications (Greene 2003). Its main usage has been in computing functions and derivatives for maximum likelihood estimation of models that involve multivariate normal integrals.

5.5 Discussion of Results

The parameter estimates of the C-MVP model for different non-worker clusters are depicted in Table 5.3 to Table 5.9. In these tables, night sleep, in-home chores, and in-home leisure activities are combined as in-home activities, and work and school activities are combined as a single work/school activity category.

5.5.1 Results of C-MVP Correlation Matrices

The correlation matrices for all the non-worker clusters estimated through MVP models have been presented in Table 5.3, all values over 0.15 are significant at p=0.05. In-home activities have a negative correlation with all out-of-home activities, whereas all five out-of-home activities (work/school, shopping, entertainment, organizational/hobbies, sports) have positive mutual dependence. For non-workers with midday activities, it is more likely that they will trip chain shopping and services activities with work or with other discretionary activities, whereas entertainment activities are more likely to be chained with shopping and services activities. For non-workers with evening activity, it is more likely that they will trip chain entertainment activities with shopping and services activities. For morning shoppers and afternoon shoppers, it is more likely that they will trip chain entertainment and sports activities in the same tour.

Table 5.3 Correlation matrix between different activity types

	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
Correlation Matrix for Cluster#1 non-worker midday activities						
In-Home	1	-0.659*	-1.248*	-0.804*	-0.644*	-0.725*
Work/school		1	0.45*	0.391*	0.267*	0.385*
Out-of-Home Shopping			1	0.379*	0.37*	0.374*
Out-of-Home Organizational/hobbies				1	0.282*	0.338*
Out-of-Home Entertainment					1	0.342*
Out-of-Home Sports						1
Correlation Matrix for Cluster#2 non-worker evening activity						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-0.539*	-0.751*	-0.623*	-0.557*	-0.529*
Work/school		1	0.282*	0.244*	0.133*	0.268*
Out-of-Home Shopping			1	0.202*	0.286*	0.133*
Out-of-Home Organizational/hobbies				1	0.041*	0.14*
Out-of-Home Entertainment					1	0.105*
Out-of-Home Sports						1
Correlation Matrix for Cluster#3 non-workers who stay-at-home						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-0.846*	-0.918*	-0.722*	-0.646*	-0.858*
Work/school		1	0.445*	0.428*	0.5*	0.502*
Out-of-Home Shopping			1	0.312*	0.313*	0.386*
Out-of-Home Organizational/hobbies				1	0.317*	0.367*
Out-of-Home Entertainment					1	0.351*
Out-of-Home Sports						1
Correlation Matrix for Cluster#4 non-worker morning shopping activity						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-0.779*	-0.867*	-0.677*	-0.57*	-0.73*
Work/school		1	0.461*	0.327*	0.406*	0.375*
Out-of-Home Shopping			1	0.241*	0.187*	0.292*
Out-of-Home Organizational/hobbies				1	0.148*	0.257*
Out-of-Home Entertainment					1	0.357*
Out-of-Home Sports						1
Correlation Matrix for Cluster#5 Afternoon shopping and services activities						
	In-Home	Work/school	Shopping	Org./Hobbies	Entertainment	Sports
In-Home	1	-0.638*	-0.958*	-0.688*	-0.537*	-0.585*
Work/school		1	0.298*	0.341*	0.326*	0.239*
Out-of-Home Shopping			1	0.224*	0.225*	0.25*
Out-of-Home Organizational/hobbies				1	0.181*	0.291*
Out-of-Home Entertainment					1	0.294*
Out-of-Home Sports						1

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

5.5.2 Results of C-MVP Parameter Estimation

Some of the explanatory variables exhibit statistical significance within the 95% confidence interval (t-statistic greater than 1.96). Other variables with t-statistic less than the threshold value have been retained in the final model specification, with an assumption that if a larger data set were available, these parameters might show statistical significance. For the sake of the brevity, we discuss only the more significant variables in this section.

5.5.2.1 In-Home Activity Participation

Table 5.4 presents the covariate matrix of in-home activities obtained from the C-MVP model for all the non-worker clusters. According to the results, it is less likely that an in-home activity will be conducted jointly compared to solo, for all clusters. Many respondents live alone, and besides, home chores are typically solo activities. With an increase in household size, participation in in-home activities increases for all the clusters, perhaps because larger household size demands greater household responsibilities from non-working members. Male individuals are less likely to engage in in-home activities in comparison to female individuals, which is consistent with earlier studies that females shoulder a higher share in household responsibilities (Castro et al. 2011). Age has both positive and negative effects on in-home activity participation. Married individuals with midday activities are less likely to engage in in-home activities, whereas other married non-worker clusters are more likely to engage in in-home activities. Those who work in the daytime have both positive and negative association with in-home activity participation, and individuals with flexible work hours are more likely to participate in in-home activities.

Individuals living in the urban core are less likely to engage in in-home activities than others, except for individuals with a stay-at-home activity pattern. This presumably reflects that individuals living in the core have more opportunities to engage in out-of-home activities compared to those living elsewhere. For similar reasons, if the population density in the home neighborhood is higher, or land use mix is greater, non-workers with evening activities and afternoon shopping activities are less likely to engage in in-home activities.

Table 5.4 Output of C-MVP parameter estimates for in-home activity participation

Explanatory Variables	**Cluster #1**	**Cluster #2**	**Cluster #3**	**Cluster #4**	**Cluster #5**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	-0.72*	-0.74*	-0.61*	-0.50*	-0.02
Size of the Household	0.06*	-0.02	0.04*	0.01	
Male (1, if the gender of the individual is male, 0 otherwise)	-0.08	-0.19*	-0.12*	-0.26*	-0.08*
Married (1, if the individual is married, 0 otherwise)	-0.02*	0.12*	0.10*	0.06	0.02
Age of the individual	0.01	-0.01*	0.03**	0.04*	0.01**
Paid worker (1, if the individual is a paid worker, 0 otherwise)	-0.02	-0.02	0.06	0.07	0.09
Working Day time (1, if the individual works in the day time, 0 otherwise)	-0.15**	0.17*	-0.02	0.21*	
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	-0.02	0.03	0.01*	0.02	0.02**
Low income level (1, if the individual belongs to low income level, 0 otherwise)	-0.01	-0.02	-0.03	-0.17*	-0.02
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	-0.08**	0.11*	0.01	-0.05**	0.03
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	-0.25*	-0.24*	0.11		-0.02
Urban Core (1, if the individual is living in urban core, 0 otherwise)	-0.03*	-0.11*	0.05	-0.04*	-0.08*
Mean commute time	0.02	-0.02	0.03	0.01	-0.02
Population density of the home neighborhood	0.01	-0.01*	-0.01	0.02	0.01**
Land use mix the home neighborhood	-0.02	-0.01*	-0.02		-0.03*
Constant	1.09*	0.93*	0.95*	0.14*	0.05*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

5.5.2.2 Out-of-Home Work/School Activity Participation

Table 5.5 shows the parameter estimates for out-of-home work/school activity participation. Non-workers with midday activities, evening activities, and stay-at-home clusters are more likely than not to engage in work/school activities jointly with another individual. With an increase in household size, participation in work/school activities is lower for all the clusters except stay-at-home non-workers. This is perhaps because non-workers in larger households often have care-giving activities related to other household members. Male individuals of all the clusters are more likely to engage in work/school activities. Age of the individuals has both positive and negative effects on work/school activity participation. Possession of a driving license increases the probability of participation in work activity for the evening activities and afternoon shoppers clusters, but decreases the probability for morning shoppers. The work-related attributes of flexible work hours, daytime working hours, and paid work have both positive and negative effects on the non-worker clusters. Housing structure type are also found to be significant for some clusters.

Members of the evening activities and afternoon shoppers clusters living in the urban core are more likely to engage in work activity than those living elsewhere, whereas morning shoppers living in the core are less likely to engage in work activities. The mean commute time is also significantly related to work activity participation for non-worker clusters, but the signs are both positive and negative. Population density and land use mix in the home neighborhood are also found to have significant mixed effects on non-workers for work activity participation.

Table 5.5 Output of C-MVP parameter estimates for out-of-home work/school activity participation

Explanatory Variables	Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	0.22*	0.21*	0.16**	0.12	-0.02
Duration of the activity episode	-0.01	0.01*	0.01	0.02*	0.02*
Size of the Household	-0.07	-0.10**	0.09**	-0.04**	-0.06
Male (1, if the gender of the individual is male, 0 otherwise)	0.25**	0.39*	0.06	0.61*	0.38*
Married (1, if the individual is married, 0 otherwise)	0.31	-0.01	-0.32*	0.31	-0.15
Age of the individual	0.02**	-0.02*	0.01*	-0.03	-0.01*
Driver License (1, if the individual has a valid driver license, 0 otherwise)	2.63	0.77**	-0.65	-0.16**	0.22**
Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)	-2.89	0.22**	-2.90	-0.71*	
Paid worker (1, if the individual is a paid worker, 0 otherwise)	0.09	0.48*	-0.08**	0.37*	-0.32*
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.75*	0.31*	-0.21**	0.27*	-0.06
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	0.11*	0.01*	-0.01	-0.01**	-0.01*
Low income level (1, if the individual belongs to low income level, 0 otherwise)	-0.14		-0.08	0.04	0.15**
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	0.43*	-0.09	-0.76*	-0.10	-0.29**
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	0.87*	0.27	-0.08	0.16**	-0.10
Urban Core (1, if the individual is living in urban core, 0 otherwise)	0.02	0.04*	0.01	-0.02**	0.06*
Mean commute time	-0.02*	0.01*	0.02*	-0.03	-0.01
Population density of the home neighborhood	-0.01	0.02**	0.01		0.02
Land use mix the home neighborhood	0.02	-0.02*	-0.02**	0.02	0.04
Constant	-6.06*	-2.34*	-1.85*	-2.19*	-1.51*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

5.5.2.3 Out-of-Home Shopping and Services Activity Participation

Table 5.6 presents the parameter estimates for out-of-home shopping and services activity participation. Non-workers with evening activity, stay-at-home, and afternoon shopping activity patterns are less likely to engage in shopping activities jointly compared to solo, whereas non-workers with midday activities, and stay-at-home activity pattern are more likely to engage in shopping activities jointly. Reasons for these relationships probably relate to the socio-demographic profiles of these clusters. Non-workers spending more time in a shopping activity episode are less likely to participate in more shopping related activities, particularly those in clusters 4 and 5. This is because for these clusters the shopping activity duration is longer than for other clusters. Spending additional time in shopping thus reflects a diminishing marginal utility.

An increase in household size increases the propensity to participate in shopping activities for non-workers with midday activities, but decreases the probability of participation for non-workers with stay-at-home activity pattern and morning shopping activities. It is possible that greater household related responsibilities generate greater need to participation in shopping and services activities for individuals in these clusters. Male individuals in the afternoon shopping cluster are less likely to participate in shopping activities whereas male individuals from all other four clusters are more likely to engage in them. Individuals with a valid driver license participate in more shopping activities compared to others, presumably because the ability to drive provides greater convenience and opportunity to carry shopped goods. Those with a daytime work schedule are less likely to participate in out-of-home shopping activities. It is possible that work hours conflict with shopping activity locations' closing hours, or they rely on other household members to shop for them.

Individuals living in the urban core are more likely to engage in shopping activities, presumably because they live in closer proximity to many shopping destinations. Similarly, greater land use mix in the home neighborhood increases the propensity to participate in shopping activities for all the non-worker clusters. Mixed land uses usually offer greater accessibility to shopping destinations.

Table 5.6 Output of C-MVP parameter estimates for out-of-home shopping and services activity participation

Explanatory variables	**Cluster #1**	**Cluster #2**	**Cluster #3**	**Cluster #4**	**Cluster #5**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	0.23*	-0.05**	0.16*	-0.06**	-0.15*
Duration of the activity episode	-0.01*	-0.02*	-0.02*	-0.03**	-0.01**
Size of the Household	-0.05**	-0.01	0.01**	0.02**	-0.03
Male (1, if the gender of the individual is male, 0 otherwise)	0.05	0.11*	0.04**	0.07	-0.03**
Married (1, if the individual is married, 0 otherwise)	0.01	-0.04	-0.06**	-0.21**	-0.08**
Age of the individual	-0.04	0.03	0.02	0.01	-0.04*
Driver License (1, if the individual has a valid driver license, 0 otherwise)	0.21**	0.65*	0.11**	0.11**	0.09*
Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)	-0.29	0.19	0.12	-0.02	0.01
Paid worker (1, if the individual is a paid worker, 0 otherwise)	0.05	0.15	-0.11		-0.03
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.11	-0.07**	-0.01	-0.30*	0.04*
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	-0.01*	-0.01	0.01	0.03**	-0.01
Low income level (1, if the individual belongs to low income level, 0 otherwise)	0.03	0.03	0.09**	0.10*	-0.02
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	-0.23*	0.04	0.07	0.15*	0.01*
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	0.21*	0.03	-0.08	-0.05	-0.04*
Urban Core (1, if the individual is living in urban core, 0 otherwise)	0.11*	0.08*	0.04	0.14	0.10*
Mean commute time	0.01*	-0.02	-0.01**	-0.01	-0.01*
Population density of the home neighborhood	-0.02**	0.02*	0.02*	-0.04**	0.01*
Land use mix the home neighborhood	0.04*	0.01*	0.02*	0.03*	0.01*
Constant	-1.27*	-2.07*	-1.99*	-1.42*	-0.99*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

5.5.2.4 Out-of-Home Organizational and Hobbies Activity Participation

Table 5.7 shows the C-MVP parameter estimates for out-of-home organizational/hobbies activities. As with shopping activity, it is more likely that individuals of all the clusters will engage in out-of- organizational/hobbies jointly, rather than solo. Individuals with evening activities, morning shopping, and afternoon shopping activity patterns are more likely to engage in more organizational/hobbies activities if the duration of the activity episode increases. Married individuals from all clusters show a negative coefficient, indicating that they are less likely to engage in organizational/hobbies activities, compared to those who are single. This may reflect married individuals often have less free time than singles, owing to shared household responsibilities. With an increase in age, the propensity to participate in out-of-home organizational/hobbies decreases for non-workers with midday activities. This may relate to the membership in this cluster: its age profile is the youngest of the five clusters.

Individuals living in the urban core from midday activity and evening activity clusters are more likely to engage in organizational/hobbies activities, whereas individuals from morning shopping activity and afternoon shopping activity clusters are less likely to engage in organizational/hobbies activities. Population density, commute time, and land use mix in the home neighborhood have both positive and negative relationships with participation in organizational/hobbies activities, but are not significant for all clusters.

Table 5.7 Output of C-MVP parameter estimates for out-of-home organizational and hobbies activity participation

Explanatory variables	Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	0.71*	0.62*	0.54*	0.64*	0.21*
Duration of the activity episode	0.03	0.03*	-0.01	0.01**	0.01*
Size of the Household	0.01	-0.02	0.07*	0.02	0.10*
Male (1, if the gender of the individual is male, 0 otherwise)	0.01*	0.03*	0.01	-0.07*	-0.02*
Married (1, if the individual is married, 0 otherwise)	-0.13	-0.03	-0.05*	-0.35*	-0.02
Age of the individual	-0.01*	0.02*	0.03	-0.04	-0.01
Driver License (1, if the individual has a valid driver license, 0 otherwise)	0.84*	-0.09**	0.15**	0.30**	0.14
Paid worker (1, if the individual is a paid worker, 0 otherwise)	-0.11	0.03	0.03	0.17	0.02
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.03	0.12	-0.17**	-0.10**	-0.05
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	0.01	-0.02	-0.01**	-0.04**	-0.03
Low income level (1, if the individual belongs to low income level, 0 otherwise)	-0.03	-0.06	-0.15*	0.15*	-0.12*
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	0.16	-0.08	-0.08**	0.03	0.02
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	0.09*	0.46*	0.32	-0.11	
Urban Core (1, if the individual is living in urban core, 0 otherwise)	0.03*	0.02*	-0.01	-0.04*	-0.02*
Mean commute time	0.04	0.01*	-0.02	-0.03**	0.01*
Population density of the home neighborhood	0.01*	-0.01	0.01	-0.04*	0.02
Land use mix the home neighborhood	-0.01	0.02**	0.01	0.02	-0.03*
Constant	-2.62*	-2.76*	-2.88*	-2.26	-2.12*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

5.5.2.5 Out-of-Home Entertainment Activity Participation

Table 5.8 presents the C-MVP parameter estimates for out-of-home entertainment activities. Across the clusters, individuals are more likely to participate in entertainment activities jointly and also if the duration of the entertainment activity increases. Except for non-workers with evening activity pattern, with an increase in household size the probability of participation in entertainment activities decreases for all clusters. This may reflect an increase in shared household related activities that reduce the time available for out-of-home entertainment activities. Being male is positively associated with entertainment activity participation, for all the clusters. Individuals with midday activities, evening activities, and morning shopping activities with a valid driver license are likely to participate in more entertainment activities, compared to others. Conversely, having a valid bus pass is negatively associated with participation in entertainment activity. These findings reflect that entertainment locations tend to be far from residential areas, and often in locations poorly served by bus transit. In addition, residential areas more than 15 km away from downtown Halifax are not served by transit. Household structure and low annual income both have mixed impacts on entertainment activities across the clusters.

Individuals living in the urban core with midday activities, evening activities, and morning shopping are more likely to engage in entertainment activities than those living elsewhere. Greater land use mix in the home neighborhood also increases participation in entertainment activities for non-workers: more mix brings more alternative entertainment activity destinations within convenient proximity.

Table 5.8 Output of C-MVP parameter estimates for out-of-home entertainment activity participation

Explanatory variables	**Cluster #1**	**Cluster #2**	**Cluster #3**	**Cluster #4**	**Cluster #5**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	1.42*	1.42*	1.00*	1.10*	1.00*
Duration of the activity episode	0.01	0.01*	0.02*	0.02*	0.03*
Size of the Household	-0.19*	0.03**	-0.08	-0.06	-0.14*
Male (1, if the gender of the individual is male, 0 otherwise)	0.06**	0.14*	0.10	0.24*	0.17**
Married (1, if the individual is married, 0 otherwise)	0.01	-0.28*	-0.20	-0.19	
Age of the individual	-0.02	0.01*	-0.04**	-0.01	
Driver License (1, if the individual has a valid driver license, 0 otherwise)	0.35*	0.06*	-0.22**	0.17*	-0.24
Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)	-0.06	-0.47*	-3.12	-0.01	-0.08
Paid worker (1, if the individual is a paid worker, 0 otherwise)	-0.07	-0.06	-0.23	-0.10**	-0.08
Working Day time (1, if the individual works in the day time, 0 otherwise)	-0.12**		0.33**		-0.10
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	-0.01	0.03	-0.04	-0.01**	-0.02
Low income level (1, if the individual belongs to low income level, 0 otherwise)	0.01	-0.04**	-0.09**	0.15	-0.02
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	0.11**	-0.02	-0.37**	-0.14**	0.31*
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	-0.12**	0.11	0.13	-0.37	-0.04
Urban Core (1, if the individual is living in urban core, 0 otherwise)	0.05*	0.03*	-0.04*	0.01**	0.06
Mean commute time	-0.02	0.01	0.01	-0.01	0.01
Population density of the home neighborhood	0.01**	0.02	-0.03	-0.03	-0.02*
Land use mix the home neighborhood	0.02*	0.01	0.02*	0.01**	0.11
Constant	-2.35*	-3.02*	-1.68*	-3.33*	-2.12*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

5.5.2.6 Out-of-Home Sports Activity Participation

Table 5.9 presents the C-MVP parameter estimates for out-of-home sports activity participation. As with organizational/hobbies activity and entertainment activity, out-of-home sports activities are also more likely to be undertaken jointly. With an increase in household size, the participation probability for sports activity decreases for non-workers with midday activities, evening activities, and stay-at-home clusters. This may relate to the resource constraints needed for sports activity participation, or to shared household responsibilities. Married individuals with midday activity patterns are more likely to engage in out-of-home sports activities.

Table 5.9 Output of C-MVP parameter estimates for out-of-home sports activity participation

Explanatory variables	**Cluster #1**	**Cluster #2**	**Cluster #3**	**Cluster #4**	**Cluster #5**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Joint (1, if the activity is performed jointly with another individual, 0 otherwise)	0.62*	0.61*	0.56*	0.61*	0.37*
Duration of the activity episode	0.02*	0.03*	0.02*	0.04*	0.02*
Size of the Household	-0.04	-0.04**	-0.13*	0.01	-0.02
Male (1, if the gender of the individual is male, 0 otherwise)	-0.02	0.06**	-0.05**	0.07**	-0.04
Married (1, if the individual is married, 0 otherwise)	0.24**	0.20	0.01		
Age of the individual	-0.03*	-0.01**	-0.04	-0.01	-0.06*
Driver License (1, if the individual has a valid driver license, 0 otherwise)	-0.27	0.45	-0.11**	0.30	-0.61*
Paid worker (1, if the individual is a paid worker, 0 otherwise)	-0.10**	-0.08**	-0.08**	0.07	-0.16
Working Day time (1, if the individual works in the day time, 0 otherwise)	0.09**	-0.04	0.41*	-0.06	0.09
Flexible work Schedule (1, if the individual has a Flexible work Schedule, 0 otherwise)	0.02		0.04	-0.04**	-0.12
Low income level (1, if the individual belongs to low income level, 0 otherwise)	-0.09	-0.03	0.05**		
Duplex Housing (1, if the individual is living in duplex house, 0 otherwise)	0.06	0.07	0.04	0.15	-0.23
Multiunit Housing (1, if the individual is living in Multiunit house, 0 otherwise)	-0.27	-0.17	-3.71	0.32*	-0.32
Urban Core (1, if the individual is living in urban core, 0 otherwise)	-0.02**	-0.01	-0.05*	-0.03	-0.02**
Mean commute time	-0.04	0.02**	0.02	0.02**	-0.02
Population density of the home neighborhood	-0.01*	0.02	-0.03	0.03*	0.15*
Land use mix the home neighborhood	-0.02*	0.02	-0.01	0.05*	0.04**
Constant	-2.03*	-3.34*	-2.14*	-3.25*	-3.18*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Age is negatively associated with sports activity participation for all the clusters. The median age for all the clusters is more than 55 years, and participation in sports activities is often constrained by age. Possession of a driving license has significant negative impact on sports activity for non-workers with stay-at-home and afternoon activity patterns.

Residential location in the urban core is significantly negatively associated with sports activity participation for non-workers with midday activities, stay-at-home, and afternoon activity patterns. Individuals living in the core have more opportunities to engage in shopping and entertainment activities, which constrains the time budget for active sports activity participation. Increased population density and land use mix in the home neighborhood have mixed impacts on out-of-home sports activity participation.

5.5.3 Transition Matrix

The activity episode sequence of each non-worker cluster may be analyzed as a transition matrix. Each transition matrix in Table 5.10 shows the likelihood of a certain activity category succeeding a preceding activity type. The rows represent the activity category of the preceding episode, and the columns represent the category of the subsequent episode. For out-of-home activity sequences, these matrices are useful in revealing trip-chaining activity. In general, across all non-worker clusters, the most frequent transitions are from in-home chores to in-home leisure, or the reverse (leisure to chores). The transition matrices also show that in-home activities are seldom followed by out-of-home activities, for all the non-worker clusters. This is because non-worker individuals spend their time mostly at home.

Table 5.10 Activity episode transitions (in percentage) matrix

Cluster#1: Non-worker, midday activities									
	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.647	0.183	0.015	0.015	0.000	0.068	0.012	0.061
L	0.749	-	0.208	0.002	0.002	0.000	0.014	0.004	0.020
N	0.896	0.095	-	0.000	0.000	0.000	0.004	0.000	0.004
W	0.667	0.000	0.000	-	0.000	0.000	0.222	0.111	0.000
P	0.206	0.088	0.000	0.176	-	0.029	0.206	0.265	0.029
S	0.000	0.000	0.000	0.000	1.000	-	0.000	0.000	0.000
G	0.514	0.149	0.014	0.014	0.122	0.000	-	0.041	0.149
E	0.074	0.000	0.000	0.074	0.370	0.000	0.148	-	0.333
T	0.570	0.152	0.000	0.000	0.051	0.000	0.089	0.139	-

Cluster#2: Non-worker, evening activity									
	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.566	0.234	0.043	0.014	0.002	0.058	0.010	0.074
L	0.754	-	0.206	0.007	0.002	0.000	0.018	0.002	0.011
N	0.898	0.090	-	0.000	0.000	0.000	0.004	0.000	0.008
W	0.706	0.176	0.000	-	0.059	0.000	0.029	0.000	0.029
P	0.345	0.138	0.000	0.207	-	0.000	0.172	0.103	0.034
S	1.000	0.000	0.000	0.000	0.000	-	0.000	0.000	0.000
G	0.594	0.116	0.000	0.029	0.072	0.014	-	0.116	0.058
E	0.152	0.000	0.061	0.091	0.212	0.000	0.182	-	0.303
T	0.644	0.137	0.014	0.000	0.014	0.000	0.041	0.151	-

Cluster#3: Stay-at-homes									
	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.695	0.172	0.017	0.013	0.000	0.027	0.001	0.075
L	0.783	-	0.169	0.003	0.004	0.000	0.008	0.004	0.029
N	0.878	0.102	-	0.011	0.000	0.000	0.002	0.000	0.007
W	0.676	0.081	0.000	-	0.135	0.000	0.054	0.027	0.000
P	0.438	0.031	0.000	0.188	-	0.000	0.094	0.156	0.094
S	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
G	0.600	0.182	0.073	0.036	0.073	0.000	-	0.000	0.036
E	0.308	0.000	0.000	0.154	0.385	0.000	0.000	-	0.154
T	0.767	0.173	0.007	0.000	0.013	0.000	0.027	0.013	-

Cluster#4: Non-worker, morning shopping									
	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.676	0.181	0.014	0.011	0.000	0.041	0.004	0.073
L	0.789	-	0.161	0.006	0.002	0.003	0.006	0.002	0.029
N	0.902	0.086	-	0.000	0.000	0.000	0.004	0.000	0.008
W	0.474	0.158	0.000	-	0.158	0.000	0.158	0.053	0.000
P	0.217	0.043	0.000	0.217	-	0.000	0.217	0.217	0.087
S	0.333	0.000	0.000	0.000	0.333	-	0.000	0.333	0.000
G	0.593	0.148	0.000	0.074	0.111	0.019	-	0.019	0.037
E	0.143	0.071	0.000	0.214	0.214	0.000	0.143	-	0.214
T	0.726	0.189	0.000	0.000	0.032	0.011	0.032	0.011	-

Cluster#5: Non-worker, afternoon shopping									
	H	**L**	**N**	**W**	**P**	**S**	**G**	**E**	**T**
H	-	0.693	0.172	0.031	0.011	0.000	0.029	0.005	0.059
L	0.744	-	0.203	0.009	0.005	0.000	0.015	0.005	0.020
N	0.855	0.134	-	0.004	0.000	0.000	0.007	0.000	0.000
W	0.697	0.091	0.000	-	0.091	0.000	0.061	0.030	0.030
P	0.188	0.042	0.000	0.188	-	0.021	0.292	0.146	0.125
S	0.000	0.000	0.000	0.000	0.000	-	0.000	0.000	0.000
G	0.609	0.109	0.016	0.031	0.109	0.000	-	0.047	0.078
E	0.056	0.056	0.000	0.056	0.278	0.000	0.333	-	0.222
T	0.611	0.222	0.014	0.014	0.028	0.000	0.069	0.042	-

H = Home chores
L = Home leisure
N = Night sleep
W = Workplace
P = Shopping & services
S = School/college
G = Organizational/hobbies
E = Entertainment
T = Sports
Horizontal axis = preceding activity
Vertical axis = succeeding activity

For Cluster#1, individuals participate in entertainment activities after shopping and services activities. The subsequent activity episode after school/college is shopping and services activities. Among all discretionary out-of-home activities, entertainment is the only activity that usually has subsequent discretionary activities, of shopping/services and sports. Another important insight for this cluster is that the trip chains of work, shopping/services, organizational/hobbies, and sports are usually simple (i.e. home-shopping-home) and confined to one out-of-home stop only, whereas school/college trips and entertainment activities typically have more than one out-of-home stop and are complex in nature. In the case of Cluster#2, except for entertainment activities, all other in-home and out-of-home activities are usually followed by an in-home household chores activity episode. Entertainment activity often has a complex trip chain that is chained with shopping & services and sports activities. Cluster#3 is the stay-at-home cluster, with the most likely reason for leaving the home being sports activity. The subsequent activity episode of nearly all their activities takes place in the home, so that complex tours are highly unlikely. In the case of Cluster#4, the subsequent activities of work, shopping & services, school/college, and entertainment are often other, so complex trip chains occur quite frequently. The transition probabilities from current to next activity for Cluster#5 are similar to those for Cluster#4, except Cluster#5 has no school/college activity.

5.6 Conclusions

The main objective of this study was to investigate the activity-travel behavior of non-workers though an innovative cluster-based Multivariate Probit Modeling (C-MVP) framework. We employed five non-worker clusters, which were previously identified from the Halifax Space-Time Activity (STAR) dataset using a daily activity pattern recognition method (Hafezi et al. 2017a, b). These were labelled as: Cluster#1: non-workers Midday activities, Cluster#2: non-workers evening activities, Cluster#3: non-workers stay-at-home, Cluster#4: non-workers morning shopping activity, Cluster#5: non-workers afternoon shopping activity. The present study uses these non-worker clusters for C-MVP model estimation, along with the estimation of inter-correlations between activity participation, and activity episode transition probability matrices.

The empirical models clearly reveal that each cluster has a separate utility function for each activity participation. The explanatory variables include individual and household characteristics such as age, gender, income, job flexibility, commute time, driving license ownership, household size, household structure, accompaniment arrangement, and land use characteristics (population density in the home neighborhood, and land use mix). The dependent variables are six activity categories: in-home activities, work/school, shopping & services, organizational/hobbies, entertainment, and sports. Based on the results we conclude that activity participation of non-workers is significantly associated with their socio-demographic characteristics, individual characteristics, household structure, accompaniment arrangement, commute time, and land use attributes.

Across all the clusters, it is more likely that in-home activities will be conducted solo rather than jointly, whereas out-of-home activities are more likely to be conducted jointly. For out-of-home activities, all the independent variables have some significant effects, but typically these effects are specific to certain clusters, and to particular activities. Often the coefficient signs differ across the clusters, even for the same activity types. The Male variable, however, has fairly consistent effects: negative effects for participation in in-home activities, and positive ones for work/school and entertainment. Age has consistent effects negative effects for entertainment and sports, while Married has consistent negative relationships with shopping and organizational/hobbies activity. Housing type and location, and neighborhood characteristics, have highly varied effects across the clusters and activity types.

The correlation matrices for all the clusters show that in-home activities have negative correlations with all the out-of-home activities, with varying magnitude for each cluster.

To better understand the activity dependence, we estimated the activity transition matrices for each cluster. For all non-worker clusters, in-home chores, leisure, and night sleep have high transition probabilities with each other, whereas these three activity types seldom transition to out-of-home activities. Among all the activities for all the clusters, entertainment activity is more likely to be combined with one or more out-of-home activities in a discretionary trip chain.

The non-worker population contributes a significant share to total urban traffic. Moreover, they have a flexible activity schedule in comparison to workers. Thus, this study contributes through offering a cluster-based modeling approach to better analyze the activity-travel behavior of non-workers. The application of C-MVP with correlation matrices estimated by GHK is also a contribution that can capture the dependency among activities. This study offers a straightforward and practical approach for incorporating the behavior of population clusters based on activity patterns into activity-travel model systems. Results from this activity-travel modeling may be employed to generate activity and trip sequences as inputs to trip generation and traffic assignment model components. Future study includes comparing the activity-travel behavior of non-workers with other population groups, including worker and student clusters. The results and modeling framework are expected to be incorporated into the Scheduler for Activity, Location and Travel (SALT) model for Halifax.

Chapter 6 Trip Chaining and Tour Mode Choice of Non-Workers Grouped by Daily Activity Patterns[4]

6.1 Introduction

In recent years, travel behavior analysis has received more attention from transportation researchers and planners. Travel demands are growing, and increased vehicles miles traveled are associated with greenhouse effects, energy consumption, decreased environmental quality, traffic congestion, and traffic accidents (Akar et al. 2016). Thus, transportation planners and policy makers try to regulate and manage activity-travel patterns and travel distances (Van Acker and Witlox 2011). Travel distances and travel demand are connected with activity-travel patterns of individuals through trip chaining and number of trip tours (Duncan 2015; Bricka 2008). However, non-worker segments of the population have flexible activity-travel patterns, which are more difficult to model, predict, and analyze. Moreover, earlier studies show that workers and non-workers responses to transportation planning and policies are significantly different. Policies therefore require separate modeling and measures for individuals who work and individuals who do not. However, empirical investigation of non-workers' activity-travel behavior is very limited in comparison to that for workers (Bricka 2008; Manoj and Verma 2015; Habib et al. 2016). Therefore, this study aims to investigate non-workers' tour frequency, trip chaining, and tour mode choice behavior through an activity-pattern based clustered framework.

---

[4] A version of this chapter has been conditionally accepted:
Daisy, N. S., Millward, H., and L. Liu. Trip chaining and tour mode choice of non-workers grouped by daily activity patterns. 2018.

Tours are usually defined as home-to-home loops. Tours with two trips are called simple tours (e.g., home to work and then work to home); and tours with more than two trips are called complex tours (Paleti et al. 2011, Krizek 2003). It has been confirmed in the literature that trips should not be analyzed in isolation, and that tour complexity influences mode choice decisions (Pendyala and Ye 2005). Undoubtedly, for individuals who choose complex tours, the preferred mode is auto due to its greater flexibility and convenience in trip chaining. Also, for short distance tours, the preferred mode is walk. Strathman and Dueker (1990) found that complex tours tend to be more car-oriented than simple tours. Trip chaining is widely defined as a home-based tour that connects multiple out-of-home activities (Primerano et al. 2008). Better knowledge on trip chaining expands the understanding of associations between activity participation and travel (Ho and Mulley 2013). Therefore, studying trip chaining and tour mode choice for non-workers is important to promote policies such as switching to public transit (2000). To model daily travel of individuals, most activity-based models (ABMs) initially produce activities, then tours, and then add tour frequency and intermediate stops involving a series of tour frequency and stop-generation models (Bradley and Bowman 2009; Goulias et al. 2010).

This study contributes by investigating the tour characteristics of non-worker groups through a cluster-based modeling framework, and by employing group socio-demographics and urban form at tour origins as exogenous variables to better predict trip chaining, tour complexity, and tour mode choices. We utilize five non-worker population clusters drawn from the STAR (Space-Time Activity Research) household travel survey conducted in Halifax Regional Municipality, Canada. Distinct population groups with similar distributions of start time, activity type, and frequency, mostly comprising women and/or individuals aged more than 55 years old were previously identified by Hafezi et al. (2017a, b). Each cluster produces vital information such as activity type, start time, end time, duration probability distribution, and sequential arrangement of activities. Though identified by their activity patterns, members of these groups also vary in their personal characteristics, such as age, gender, income, and occupation, and these variables can be employed as predictors of activities and hence travel. Since activity generation modules directly affect the overall model prediction accuracy, it is important that individuals with similar characteristics are grouped into distinct homogeneous clusters. Generating more accurate activity patterns and prediction of tour frequency, intermediate stops, and mode choice is a significant step in moving current activity-based models closer to replication of reality. Overall, this study therefore provides a new approach to activity-based research, which should improve modeling of activity-travel behavior.

6.2 Literature Review

Much research has examined both trip chaining and travel distances, and investigated their causal determinants using socio-demographic and urban-form variables. Work by McGuckin and Murakami (1999) provided an early focus on trip chaining, and more recently Primerano et al. (2008) defined trip chaining as the combination of one or more secondary activities with a primary activity, via trips which start and end at home. Some studies have utilized trip chaining to measure tour complexity in terms of stop frequency within the tour (Liu 2012; Wang 2014). These studies found that gender, age of the individuals, and employment status play significant roles in trip chaining (McGuckin and Murakami 1999; Liu 2012; Kitamura and Susilo 2005; Susilo and Kitamura 2008). These studies also concluded that women, older adults, and workers tend to chain more trip segments. Among household characteristics, household income level and the presence of children in the household have positive influences on trip chaining (Van Acker and Witlox 2011; Krizek 2003; Liu 2012; Wang 2014; Kitamura and Susilo 2005; Susilo and Kitamura 2008; Bhat et al. 1999). However, increases in household size and number of vehicles decrease the propensity for trip chaining (Van Acker and Witlox 2011; Bricka 2008; Kitamura and Susilo 2005; Susilo and Kitamura 2008).

Other studies have showed that the built environment characteristics of home locations have significant effects on trip chaining. For example, people who live in low population density neighborhoods are more prone to trip chaining, and also tend to make more complex tours (Bricka 2008). However, another study suggests that non-workers living in suburban areas have lower tendency to trip chaining (Kitamura and Susilo 2005).

Tour frequency has been found to be affected by employment-related variables, household structure variables, accessibility, location variables, and mobility-related variables (Bhat et al. 1999). Among household socio-demographic characteristics, number of adults, number of employees, number of vehicles, household income, home-to-work distance, work neighborhood accessibility, and work residential accessibility were found to be significant determinants of tour frequency (Krizek 2003). However, these studies combined tours for workers and non-workers in the same framework. Presumably, non-workers would not have work-related variables to influence their travel behavior, and we expect significant differences in travel behavior between workers and non-workers.

The investigation of interconnections between trip chaining and mode choice has a long history. Several studies have confirmed that in general complexity of the tour, measured by stop frequency, positively influences the choices of both auto and walk modes (the latter only if the trip is short and a sidewalk is available), but negatively influences the transit mode choice (Bhat et al. 1999; Yun et al. 2014; Harding et al. 2015). However, number of stops is not the only determinant of mode choice. Other factors, for instance, location patterns also affect the choice of mode for a complex tour (Ho and Mulley 2013). Since the non-worker portion of the population has the most flexible time-use pattern, a separate analysis of trip chaining, tour complexity, and mode choice for non-worker groups is needed, and has not previously been undertaken, to our best knowledge. Therefore, we aim in this study to provide a comprehensive econometric investigation for non-worker groups, to better understand their tour complexity and mode choices. The results of this study are expected to be incorporated into the Scheduler for Activity, Location and Travel (SALT) for Halifax.

6.3 Data

6.3.1 Data Source

Data used in this study were obtained from the Halifax Space-Time Activity Research (STAR) project conducted between April 2007 and May 2008. It was the world′s first large survey to use global positioning system (GPS) technology to augment and verify household activity-travel diary data (Bricka, 2008). Halifax is the capital of the Canadian province of Nova Scotia, and Halifax Regional Municipality (HRM) is the largest municipality in Atlantic Canada. It is a mid-sized metropolitan city (c.400,000 population) with a diverse and growing economy, and population growth of 0.5% per year. The STAR project collected travel and household information of 1,971 randomly selected households in HRM, which represents almost one household in 78. The total participation rate was 21%. STAR collected activity-travel diary data distributed through seven days of the week and 12 months of the year (TURP 2008). In STAR, the entire HRM was sub-divided into the following four zones (based on Millward and Spinney 2011):

- Urban Core: areas within 5km walking range from the downtown. These are the older (pre-1960) high-density areas of Halifax and Dartmouth.
- Suburbs: urbanized areas near to the urban core, with central water and sewerage systems and medium density.
- Urban Fringe (Inner Commuter): transitional areas beyond the suburbs but within 25 km distance from the downtown. These commuter areas are impacted by low-density large-lot subdivision development and related land uses.
- Rural: areas beyond 25 km from downtown, less affected by exurban development, and largely dependent on resource activities.

The primary respondent in each sample household was selected randomly, and had to be more than 15 years age. These respondents completed a detailed time-diary for two consecutive days. The time diary coding and questionnaire on household characteristics were based on the General Social Survey (GSS) Cycle 19 Survey of Statistics Canada (2006). Primary respondents also carried a GPS-device (Hewlett Packard iPAQ hw6955) for all out-of-home activity, programmed to collect GPS data every 2 s. The GPS data provided precise start and end times, and travel routes for all "stops" with more than 2 minutes stopping duration. These GPS data were used with CATI software in day-after interviews with respondents, to verify and enhance the time-diary data. Land use characteristics and built environment variables used in this study were obtained from work by Neatt et al. 2017.

6.3.2 Cluster Description

Initially, a pattern recognition modeling framework was applied to the Halifax STAR household activity-travel data (Hafezi 2017a, b). A subtractive clustering algorithm was utilized to initialize the total cluster number and cluster centroids. Individuals with homogeneous activity patterns were identified using a fuzzy c-means clustering algorithm, and a set of representative activity patterns were recognized by utilizing a multiple sequence alignment method. The characterization of cluster memberships based on individuals' socio-demographic attributes was attained by using the CART algorithm. Activities were categorized in nine types: in-home chores, in-home leisure, in-home night sleep, out-of-home work, out-of-home shopping & services, out-of-home school, out-of-home organizational/hobbies, out-of-home entertainment, and out-of-home sports.

Table 6.1 shows the personal attributes and time allocations to in-home and out-of-home activities for the five identified non-worker clusters in the dataset. Note that a few individuals in these clusters have out-of-home work or school activity, but their activity profiles are closer to those of non-workers than workers. The details of each cluster are as follows:

Cluster#1 is the non-worker midday activities cluster. The individuals belonging to this cluster participated in entertainment and organizational/hobbies activities predominantly in the midday. This cluster predominantly comprises individuals more than 55 years age

(66.0%), retired (52.0%), and female (53.0%). Many members of this group are not university graduates and belong to the low or middle income level.

Cluster#2 is the non-worker evening activity cluster. These non-workers participate in out-of-home organizational/hobbies and entertainment activities, mostly in the evening. Similar to the previous cluster, a large proportion of individuals are female (59.0%), and older than 55 years (53.0%). Many individuals in this cluster have bachelor degree or above, and many are retired (39.0%). These individuals have low to middle income, and those who work have a flexible work schedule (54.0%).

Cluster#3 comprises non-workers with a stay-at-home activity pattern. A large proportion of this group is older-aged females. Individuals in this group mostly belong to the low-income partition. This cluster has the largest membership in comparison to other clusters. On an average, individuals from this cluster spend only 4.6% of their total time in out-of-home recreational activities.

Cluster#4 is the non-worker morning shopping activity cluster. Similar to earlier clusters, this cluster also largely comprises females more than 55 years age. Many individuals in this cluster do not have a bachelor degree or above and most have a low income level (53.0%). They are mostly retired. The majority of those who work have flexible working hours (52.0%), and some portion work at-home (11.0%).

Cluster#5 comprises individuals who do shopping and services activities in the afternoon. Similar to other clusters, the members of this group are mostly older females. A large proportion belongs to the low-income category, and their education level is similar to other non-worker clusters.

Table 6.1 Details of cluster characteristics

Socio-Demographic variables		Sample mean (propn.)	Mean of cluster (propn.) #1	#2	#3	#4	#5
Gender	Female	0.56	0.53	0.59	0.56	0.54	0.59
Age	Young adults (ages 15-35 years)	0.08	0.05	0.10	0.11	0.09	0.07
	Middle-aged adults (ages 36-55 years)	0.32	0.29	0.38	0.32	0.29	0.32
	Older adults (aged older than 55 years)	0.60	0.66	0.53	0.57	0.63	0.61
Education Level	Bachelor degree and above	0.34	0.35	0.37	0.35	0.30	0.35
Occupation	Regular shift	0.23	0.22	0.26	0.24	0.19	0.24
	Irregular schedule	0.09	0.10	0.10	0.11	0.07	0.07
	Student	0.02	0.00	0.04	0.01	0.03	0.02
	Retired	0.45	0.52	0.39	0.41	0.53	0.41
	Work at home	0.12	0.10	0.15	0.16	0.11	0.09
Flexible schedule	Have no flexibility in a work schedule	0.47	0.48	0.46	0.43	0.48	0.51
Job number	Have more than one job	0.09	0.04	0.16	0.08	0.11	0.05
Income	Low-income (<= $ 40,000)	0.48	0.44	0.48	0.49	0.53	0.48
	Middle-income ($ 40,000 - $ 100,000)	0.45	0.46	0.45	0.45	0.42	0.46
	High-income (> $ 100,000)	0.07	0.10	0.07	0.07	0.05	0.06
Total cluster membership			225	238	419	247	262
Average Number of tours per person per day			1.73	2.18	1.05	1.75	1.65
Percentage in total (number of person-days)			8.10	8.57	15.08	8.89	9.43
Activity categories	**Descriptions**		**Share of daily activity engagement (%)**				
Home chores (H)	Working at home, eating/meal preparation, indoor or outdoor cleaning, interior or exterior home maintenance, child care, other in-home activities.		34.73	41.39	43.10	40.54	40.12
Home leisure (L)	Watching TV/listening to radio, reading books/newspapers, etc.		20.26	14.42	19.17	19.96	19.18
Night sleep (N)	Night sleep		45.01	44.19	37.73	39.50	40.70
Total in-home (%)			*100.0*	*100.0*	*100.0*	*100.0*	*100.0*
Workplace (W)	Work/job, all other activities at work, work related (conferences, meetings, etc.).		4.76	6.41	8.16	10.53	13.12
Shopping & services (P)	Shopping for goods and services, routine shopping.		27.27	15.31	30.19	**32.26**	**38.27**
School/college (S)	Class participation, all other activities at school.		1.00	1.20	0.42	3.16	1.28
Organizational/ hobbies (G)	Organizational, voluntary, religious activities. Hobbies are done mainly for pleasure, cards, board games, all other hobbies activities.		**29.57**	**31.28**	19.90	21.11	21.94
Entertainment (E)	Eat meal outside of home, all other entertainment activities.		17.49	**33.73**	10.35	8.75	11.66
Sports (T)	Walking, jogging, bicycling, all sports related activities.		19.90	12.06	**30.98**	24.20	13.74
Total out-of-home (%)			*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

6.3.3 An Average Weekday

In this study, weekday personal activity diaries were utilized for cluster identification and analysis of typical travel activity patterns of each non-worker cluster. Although the current study focuses on modeling trip tours and trip modes, we also employed the STAR survey's GPS tracking to map the location of cluster members throughout Halifax for an average weekday. The four panels on Figure 6.1 show the spatial distribution of out-of-home activity for non-workers at 10:00 AM, 1:00 PM, 4:00 PM, and 7:00 PM, with these times being selected as ones with considerable out-of-home activity. The maps clearly show that the five non-worker clusters display distinct patterns of activity. At 10:00 AM, the majority of the morning shopper group are outside their home compared to others, and clustered in commercial areas. At 1:00 PM, a large percentage of the midday activities group and some stay-at-homes are doing organizational and hobbies and entertainment activities. In the 4:00 PM map, afternoon shoppers are mostly out for shopping, whereas at 7:00 PM non-workers with evening activities predominate, and are outside the home for organizational and hobbies, and entertainment activities.

6.3.4 Tour Formation Behavior

Following Shiftan (1998), we defined a home-based tour as a home-to-home journey comprising a sequence of out-of-home trips, for which origin and destination is home without any occurrence of intermediate home stops. Then the number of stops in each home-based tour was measured, to reflect the concept of trip chaining. The cutoff point for number of stops was seven trips, as very few individuals partake in more than seven trips in a tour. Then the number of tours made in a day, number of stops made in each tour, and mode choices for each tour were noted.

After identifying the tours, the tour purpose for work and school tours was assigned based on the priority order by Stopher et al. (1996), which awards the highest priority to work, followed by education. However, very few individuals from the five non-worker clusters participate in work/school (W/S) activity. Work and school tours comprise less than 5% of total tours across all the five clusters. The activity purpose for non-work tours was identified based on the longest activity duration. The explanatory variables used in the empirical model has been reported in Table 6.2. Simple tours were classified into six categories: home-work-home (HWH), home-school-home (HSH), home-shopping-home (HPH), home-organizational/hobbies-home (HGH), home-entertainment-home (HEH), and home-sports-home (HTH). Complex tours (with two non-home stops) were categorized as shown in Table 6.3. The travel mode/modes for each tour were selected based on the longest in-mode travel time. The choice set generates ten choices, which are car drive, car passenger, transit, bike, walk, car drive & walk, car passenger & walk, transit & walk, bike & walk, and car & transit. The number of stops and average duration of time spent at each stop was calculated through data mining methods.

6.3.5 Descriptive Statistics

The descriptive statistics in Table 6.1 show that on average travelers across all the clusters make between one and two home-based tours per day. However, individuals from Cluster 3 makes only one tour per day, whereas Cluster 2 makes more than two tours per day. The tours are categorized in 19 categories. From Figure 6.2, across all the clusters, more than

25% of tours are simple home-shopping-home (HPH) tours. The second highest tour category is home-organizational/hobbies-home (HGH). Among all the clusters, work and school related tours comprise less than 5% of the total tours.

From Figure 6.3(a), we can see that almost 99% of individuals make at least one home-based tour per day, except in Cluster 3. Almost 35% of individuals in Cluster 3 make no home-based tour. Non-workers across all the clusters engage in trip chaining and complex tours. From Figure 6.3(b), we can see that the highest five modes/multi-modes for tours are car drive, car passenger, walk, car drive & walk, and car passenger & walk. Presumably, this is because individuals are making complex tours and they find car and walk to be the most convenient mode of choice. These five mode/multi-mode categories were used for further empirical modeling, while transit and bike were excluded from the choice set. Figure 6.4 illustrates the modal shares for all tour categories and displays considerable variation among the clusters. However, the use of car drive is dominant across all the clusters for all activity categories.

Table 6.2 Details of exploratory variables

Abbreviation	***Variable Description***
HHSIZE	Household Size
DUTRVL	Total Travel Time per tour
HPH	Home-Shopping-Home tour (1, if the tour is H-P-H tour, 0 otherwise)
HGH	Home-Organizational/hobbies-Home tour (1, if the tour is H-G-H tour, 0 otherwise)
HEH	Home-Entertainment-Home tour (1, if the tour is H-E-H tour, 0 otherwise)
HTH	Home-Sport-Home tour (1, if the tour is H-T-H tour, 0 otherwise)
HPGH	Home-Shopping-Organizational-Home tour (1, if the tour is H-P-G-H tour, 0 otherwise)
HPEH	Home-Shopping-Entertainment-Home tour (1, if the tour is H-P-E-H tour, 0 otherwise)
HEGH	Home-Entertainment-Organizational-Home tour (1, if the tour is H-E-G-H tour, 0 otherwise)
TOURN	Number of tours per day
STOPS	Number of stops in a tour
MALE	Male (1, if the gender of the individual is male, 0 otherwise)
MARRID	Married (1, if the individual is married, 0 otherwise)
AGE	Age of the individual
DIPLOMA	Diploma or university certificate (1, if the individual has obtained bachelor degree or higher, 0 otherwise)
HISCHL	High school graduate (1, if the individual has obtained diploma or certificate; some university; some community college, trade, technical or business college; high school secondary; 0 otherwise)
DRIVRLIC	Driver License (1, if the individual has a valid driver license, 0 otherwise)
BUSPASS	Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)
LOW	Low income level (1, if the individual belongs to low income level, 0 otherwise)
HIGH	High Income (1, if the gender of the individual has high income level, 0 otherwise)
DWLOWN	Home Ownership (1, if the individual owns a home, 0 otherwise)
HV1	Number of car in the household
URBCR	Urban Core (1, if the individual is living in urban core, 0 otherwise)
SUBRB	Suburban (1, if the individual is living in suburban area, 0 otherwise)
URBFR	Urban Fringe (Commuter shed) (1, if the individual is living in Urban Fringe area, 0 otherwise)
RURAL	Rural (1, if the individual is living in rural, 0 otherwise)
POPDEN	Population density of the home neighbourhood
LNDUS	Land use mix in the home neighbourhood
RESID	Residential area density in the home neighbourhood
COMM	Commercial area density in the home neighbourhood
SIDEWALK	Sidewalk density in the home neighbourhood
INTERSC	Intersection density in the home neighbourhood
RETAIL	Retail floor area ratio in the home neighbourhood

Table 6.3 Tour typology

Type	Description
H-W-H	Home-Work-Home
H-P-H	Home-Shopping & Services-Home
H-S-H	Home-School/College-Home
H-G-H	Home-Organizational/Hobbies-Home
H-E-H	Home-Entertainment-Home
H-T-H	Home-Sports-Home
H-W-P-H	Home-Work & Shopping-Home
H-W-S-H	Home-Work & School-Home
H-W-G-H	Home-Work & Organizational-Home
H-W-G-H	Home-Work & Entertainment-Home
H-S-P-H	Home-School & Shopping-Home
H-S-G-H	Home-School & Organizational-Home
H-P-G-H	Home-Shopping & Organizational-Home
H-P-E-H	Home-Shopping & Entertainment-Home
H-P-T-H	Home-Shopping & Sports-Home
H-E-G-H	Home-Entertainment & Organizational-Home
H-S-E-H	Home-Entertainment & School-Home
H-T-G-H	Home-Sports & Entertainment-Home
H-T-E-H	Home-Sports & Organizational-Home

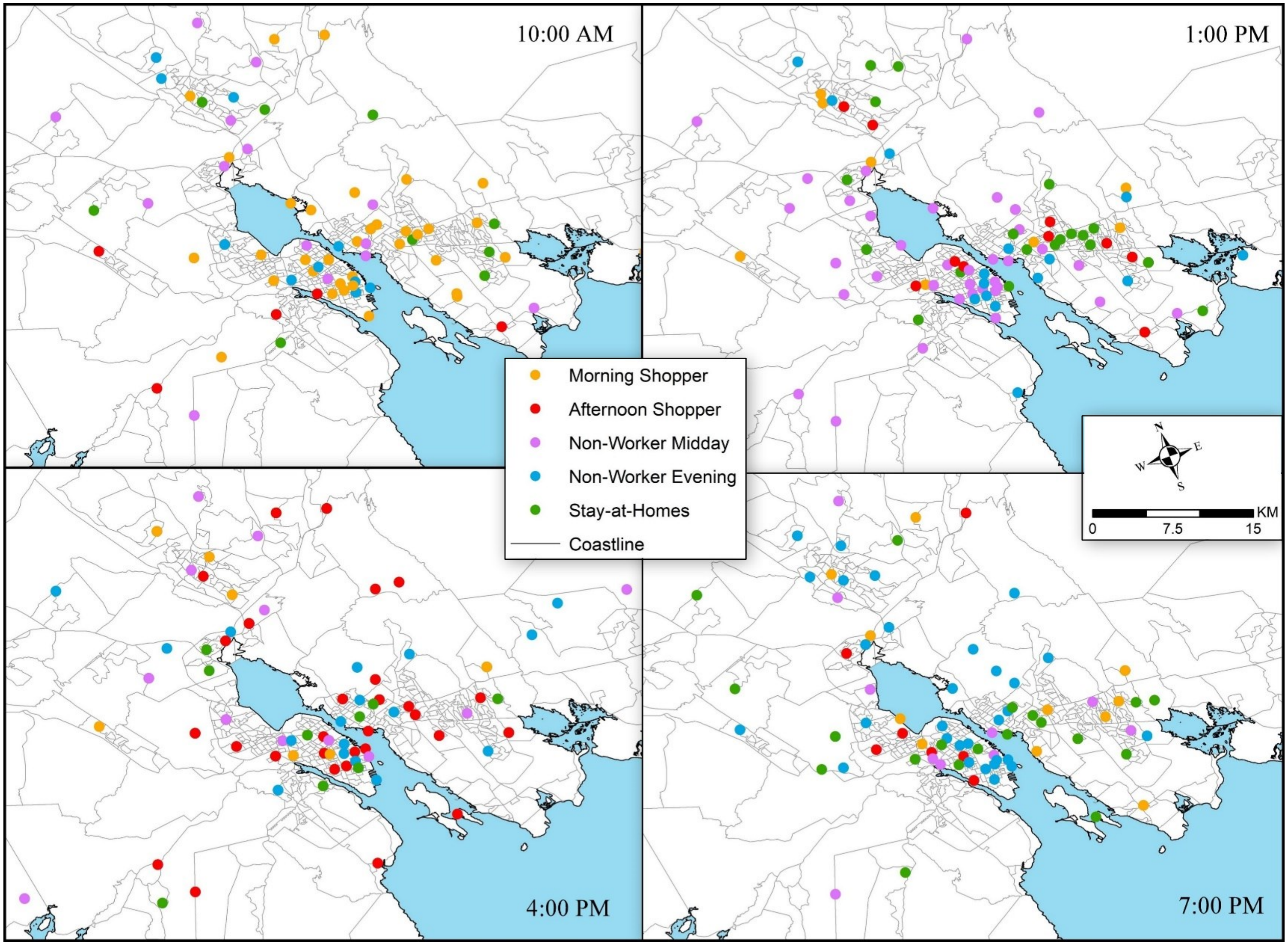

Figure 6.1 Spatial distribution of out-of-home activities of non-workers clusters at different times-of-day on a weekday

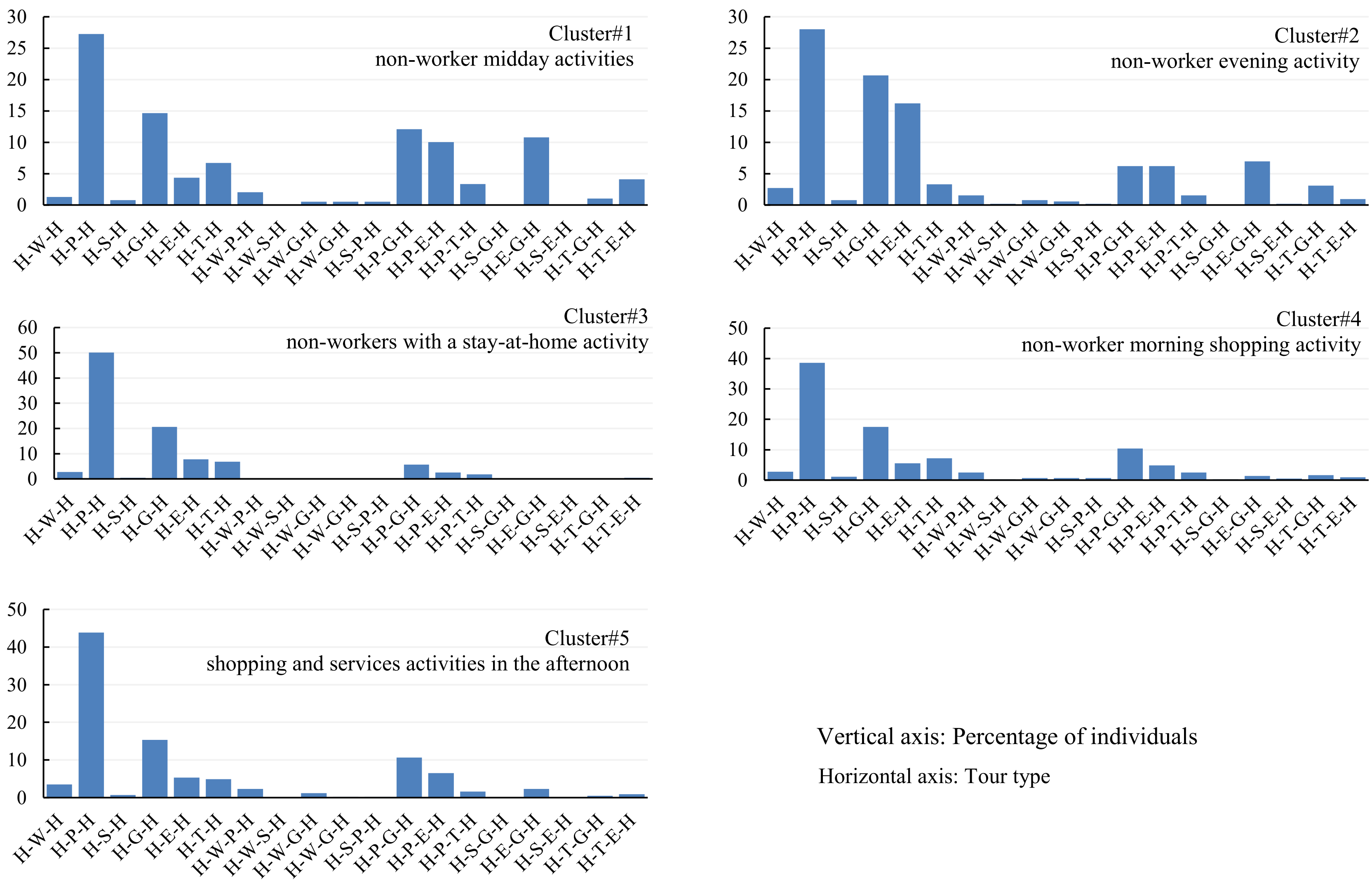

Figure 6.2 Share of Different Tour Types of Individuals, by Cluster

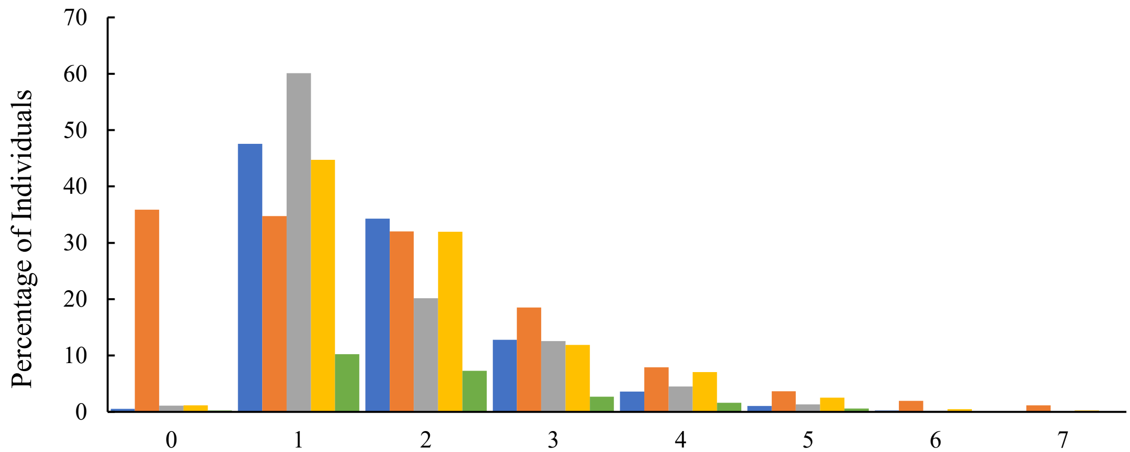

Figure 6.3 (a) Number of tours per day

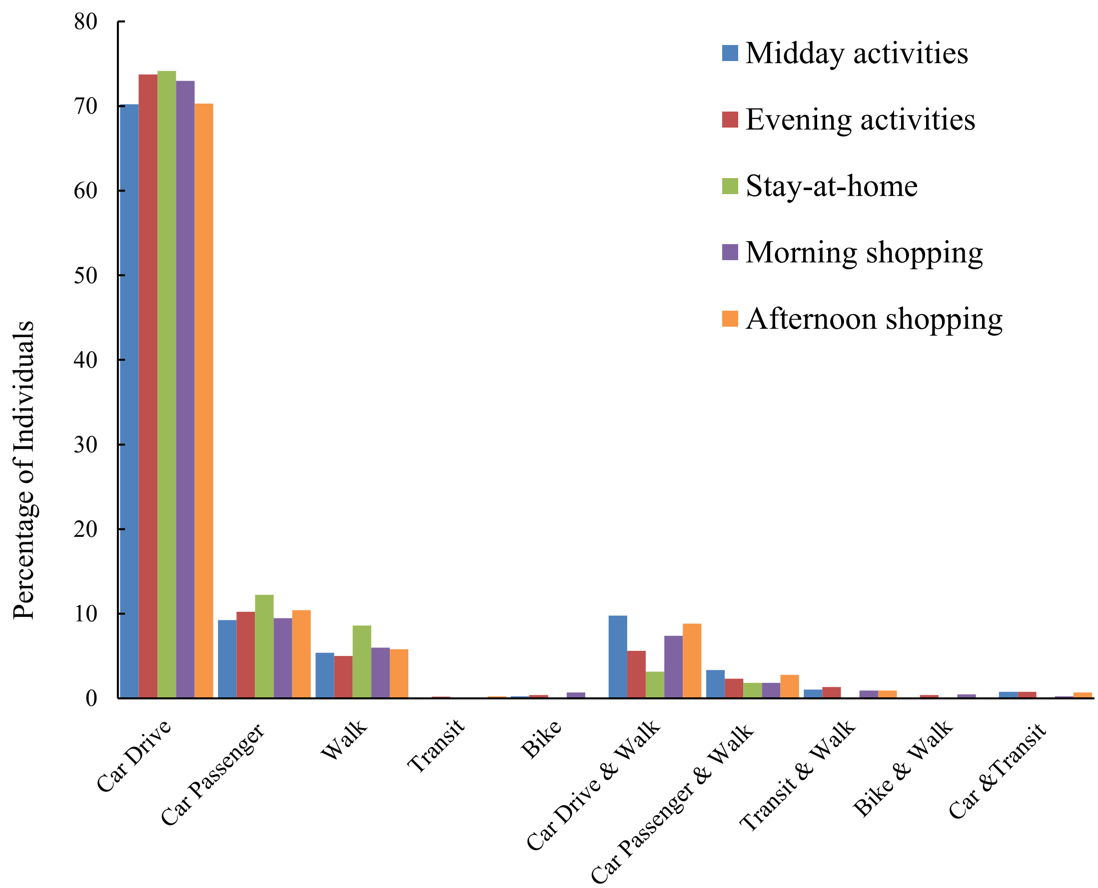

Figure 6.3 (b) Distribution of Modal Share

Figure 6.3 Distribution of Modal Share of Tours and Number of Tours Per Day, by Cluster

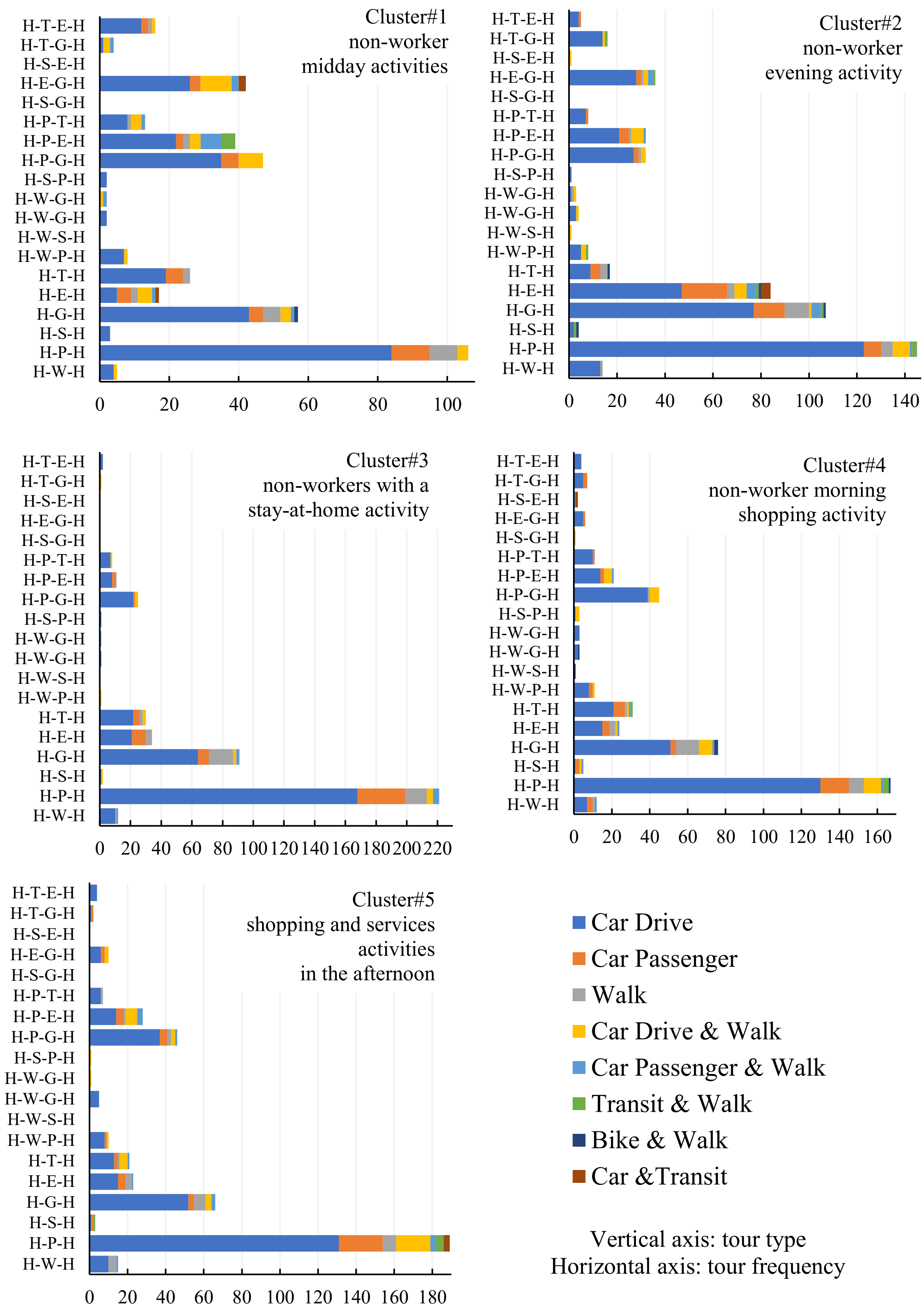

Figure 6.4 Distribution of Modal Share of Tours by Different Tour Types, by Cluster

6.4 Methods

6.4.1 Poisson Regression Model

The frequency of tours per day was modeled by utilizing Poisson regression. Earlier trip and tour studies utilized the Poisson regression model mostly in an aggregated manner, whereas this study models non-workers clustered by their homogeneous daily time-use patterns. It is hypothesized that tour frequency per day for individuals with homogeneous patterns is more likely to follow a Poisson distribution. Moreover, compared to other alternative econometric models, the Poisson regression model can treat tours as countable and number of tours as a non-negative integer. To demonstrate the modeling approach taken in this study, let Y_n be the number of home-based tours completed in a given day by an individual n. Then, a Poisson model can be written as:

$$\Pr(Y_n) = \frac{e^{-\mu_n}\mu_n^{Y_n}}{Y_n}, \mathrm{n} = 1,2,3 \ldots \ldots \mathrm{N} \quad (1)$$

where, $\Pr(Y_n)$ is the probability of Y_n tours performed by the N th individual, and $\mu_n$is the expected value of Y_n, which can be represented as:

$$E(Y_n) = \mu_n = \alpha X_n \quad (2)$$

Here, α represents a vector of regression parameters to be estimated, and X_n is a vector of variables describing individual's personal and household characteristics, land use attributes, and tour attributes. The partial effects in this non-linear regression are obtained from $\mu_n\alpha$ (Greene 2006).

6.4.2 Ordered Probit Model

An increase in number of trips per chain increases tour complexity. The observed number of stops per tour made by individuals is ordered. Therefore, it is assumed that there exists some level of utility associated with the different levels of trip chaining. Consequently, an ordered probit modeling structure was used to investigate the predictors that affect the trip chaining behavior of non-workers for home-based tours. Unlike previous applications of ordered probit modeling, this study applies it to population clusters with distinct activity patterns. Let us assume Y_n^* is the latent and continuous measure of tour stop frequency for non-work tours for individual N. Assuming ε_n as an error term and α as a vector of parameters associated with explanatory variables X_n, the ordered probit model can be written as follows:

$$Y_n^* = \alpha X_n + \varepsilon_n; n = 1,2 \dots \dots N \tag{3}$$

Where X_n is a vector of observed explanatory variables including personal characteristics, household socio-demographic characteristics, land use characteristics, and tour attributes.

The ordered probit model estimates $k - 1$ threshold values that horizontally divide the underlying continuous variable to forecast the observed count values, where k is the largest possible count value. Therefore, trip chain Y_n takes on values starting from 0 through R to generate an ordered segment of the latent trip chaining propensity into the observed stop frequencies. To convert Y_n^* to Y_n the cut points θ are introduced as written:

$$Y_n^* \begin{cases} 0 \text{ if } Y_n \leq \theta_0 \\ 1 \text{ if } \theta_0 \leq Y_n \leq \theta_1 \\ 2 \text{ if } \theta_1 \leq Y_n \leq \theta_2 \\ \vdots \\ K \text{ if } \theta_{k-1} \leq Y_n \leq \theta_k \end{cases} \tag{4}$$

Where θ_k are the cutoff points for the possible outcomes. The threshold frequency for this study ranges from 0 to 6. The estimation process of the ordered probit model is straightforward. Assuming that θ and ε are independently distributed, the probability of the trip chaining can be obtained as a mixture of a designated distribution at zero and assumed distribution of the response variable Y_n. The parameters of the model are estimated by using the maximum likelihood method. This model is an extension of a probit model for a binary outcome where the predicted probability of observing a particular ordinal outcome can be generated as follows:

$$\Pr[Y_n = k|X_n] = \Pr[\varepsilon_n \leq \theta_k - \alpha X_n] - \Pr[\theta_{k-1} - \alpha X_n] \tag{5}$$

Assuming an indicator variable ω that equals 1 if the traveler makes k stops in a tour, and 0 otherwise, the log likelihood (see Pratt 1981; Greene 2007a; Greene 2008a) can be written as:

$$\ln L = \sum_{n=1}^{N} \sum_{k=0}^{K} \omega \ln[F(\theta_k - \alpha X_n) - F(\theta_{k-1} - \alpha X_n)] \tag{6}$$

Where, $F(.)$ represents the cumulative density function (CDF) for ε_n.The log-likelihood was estimated by using the maximum likelihood method. Finally, Akaike Information Criteria (AIC) is used to evaluate the goodness of fit of the estimated models.

6.4.3 Multinomial Logit Model (MNL)

A choice set with unimodal and multimodal choices was generated through machine learning and data mining techniques. In the case of more than two modes in a tour, the modes with the two highest in-mode time were selected. It is assumed that the choices are independent from irrelevant alternatives (IIA), and a random utility-based Multinomial Logit (MNL) model was applied to investigate the mode choices. Five modes, including two multi-modes, were considered in the choice set: auto drive, auto passenger, walk, car drive and walk, and car passenger and walk. Transit and bike were excluded from the choice set as their proportions were very small. The random utility theory (McFadden 1974; Ortuzar and Willumsen 2011) postulates that utilities can be expressed as a sum of measured attractiveness and a random term as follows:

$$U_{in} = \beta_{in} Z_{in} + \epsilon_{in} \tag{7}$$

Where, $U_{in}$is the systematic utility, Z_{in} is a vector of observed attributes of the alternative mode i and individual N and ϵ_{in} is the random error. If the Pr is the probability of an individual N choosing mode i in our given choice set, then the MNL model can be written as follows (McFadden, 1974):

$$Pr_{in} = \frac{e^{U_{in}}}{\sum_{i=1}^{m} e^{U_{in}}} = \frac{e^{\beta_{in} Z_{in}}}{\sum_{i=1}^{m} e^{\beta_{in} Z_{in}}} \tag{8}$$

Where, β_{in} is the corresponding parameter of vector Z_{in}. The number of alternative choices is expressed as m in the above equation where m = 5 for this study.

If f_{in} is the choice indicator (=1 if i chosen by individual N and 0, otherwise) and Pr_{in} is the probability that individual N chooses alternative i, then the log-likelihood function can be written as follows:

$$LL(\beta) = \sum_i \sum_N d_{in} \ln(Pr_{in}(\beta)) \tag{9}$$

Parameters, β_{in} are estimated by using the maximum-likelihood estimation. Parameter values were obtained by maximizing the likelihood function obtained by equating the first derivative of the likelihood function to zero.

6.5 Discussion of Results

6.5.1 Poisson Regression Models for Tour Frequency

Table 6.4 presents the parameter estimates of tour frequency models for all five clusters. Among the personal characteristics, gender and age have positive signs, which implies that males generate more tours per day in comparison to female counterparts. Age is found to be significant for the non-worker with midday activities cluster. As expected, household size is statistically significant and positive for tour generation (Krizek 2003). Also, married individuals are more likely to make more tours per day than others. The number of cars in the household has positive association with tour frequency, presumably because more cars in the household offer greater opportunity to participate in more tours. Among the tour attributes, as expected, trip chaining has significantly negative association with tour frequency. This is because trip chaining allows an individual to combine more activities in a single trip. Our model also sheds light on the tour types associated with tour frequency: HPH, HGH, HEH, and HTH are positively associated with tour frequency. Thus, simple tours are positively associated with greater number of tours. As expected, complex tours are negatively associated with tour frequency: they combine or bundle more activities that an individual wants to complete in the day, obviating the need to return home and generate new tours.

Among the land use characteristics, the results are consistent with the existing literature. Residential location has significant impact on tour frequency per day. Individuals from urban core and suburban areas make more tours compared to those from urban fringe and rural areas. Individuals living in these areas have greater opportunity to travel to activity destinations in 15-20 mins compared to those who live in rural and urban fringe areas. It is interesting to note that land use mix has negative impact on tour frequency. If the land use mix in the home neighborhood is higher, it offers individuals greater accessibility to activity destinations, which will motivate individuals to trip chain rather than making new tours (Chen and Akar 2017). Intersection density, retail density, and residential density have both positive and negative impacts across the clusters.

Table 6.4 Estimation results of Poisson regression models for tour frequency

Explanatory Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Constant	0.58	-0.33	0.25	0.45	0.52
HHSIZE	0.06	-0.12*	0.10*	-0.03**	
DUTRVL	-0.02**			-0.03	-0.01**
HPH	0.24*	-0.06	0.20	0.36*	0.12
HGH	0.08	0.21*	0.19	0.25**	0.13
HEH	0.24	0.30*	0.25	0.54*	
HTH			0.13	0.05	0.09
HPGH		0.07	0.34	-0.20	
HPEH	-0.12**	0.05			
HEGH		-0.11*			
CHAIN	-0.03	-0.03*	-0.04**	-0.04	-0.02**
MALE	0.04	0.08	0.05	0.06**	
AGE		0.01*		-0.01	0.02
MARRID			0.21**	0.09	
DRIVRLIC	0.15	0.36	0.08	0.30	0.12
LOW	0.05	-0.09		0.06	0.10
DWLOWN	0.12	-0.22			
HV1		0.05**	0.08	0.05*	0.06*
URBCR	-0.10	0.01	0.04*		0.19
SUBRB	0.65	0.47*	0.49	-0.27	
URBFR					-0.01*
RURAL	-0.02*	0.16	-0.01*	0.05	0.31
POPDEN	0.53	0.08	0.13	0.38*	-0.25
LNDUS	-0.83	-0.25	-0.48**	-0.45**	-0.07
COMM	-0.01	-0.03	-0.05		
RESID		0.04**			-0.06
INTERSC	-0.02	0.03*	-0.02	-0.02*	1.19
RETAIL	-1.08	1.32			
Log likelihood (constant only)	-549.53	-849.28	-622.66	-653.01	-618.79
Log likelihood (full Model)	-531.49	-809.64	-610.05	-624.24	-608.34
AIC	2.84	3.22	2.85	2.98	2.90

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

6.5.2 Ordered Probit Model for Trip Chaining

Table 6.5 presents the parameter estimates of trip chaining behavior of individuals. Variation among the clusters indicates the latent heterogeneity among non-workers that could not be captured if they were modeled as a single group.

Among the personal characteristics, male individuals of the midday activities cluster and evening activities cluster are negatively associated with trip chaining, whereas being male is positively associated with trip chaining for other non-worker clusters. That is, men with stay-at-home, morning shopping, and afternoon shopping patterns make longer tours compared to females, which is consistent with previous findings (Chen and Akar 2017). Age is found to be positively associated with trip chaining for the non-worker with evening activity cluster. Household size is positively associated with trip chaining for non-workers. Similar to the findings of Chen and Akar (2017) and Susilo and Kitamura (2008), individuals from larger households are more likely to make simpler and shorter tours. Arguably, this may be because individuals from larger households have higher numbers of destination in different locations, which lessens the propensity of trip chaining.

The possession of a driver license is found to be negatively associated with trip chaining but positively associated with tour frequency. Perhaps the option to drive provides non-workers with greater ability to make multiple tours instead of engaging in multiple trips in the same tour. Number of cars in the household has negative association with trip chaining. The availability of more vehicles offers greater opportunity and convenience to engage in more tours rather than burdening one tour with more trips. Low-income level is negatively associated with trip chaining for non-workers in the midday activities cluster, but positively associated with trip chaining for non-workers with morning shopping activity cluster.

Among the tour attributes, as expected, simple tours tend to be negatively associated with trip chaining, whereas complex tours are positively associated. Among the complex tours, HPGH, HPEH, HEGH are more likely to have more trip segments than other tour types. Tour number per day is negatively associated with trip chaining. This is likely because trip chaining reduces the need for a higher number of tours.

Table 6.5 Estimation results for ordered Probit model for trip chaining

Explanatory Variables	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Constant	2.24*	2.55*	1.39*	1.88*	1.54*
HHSIZE	0.03	0.04	0.04	0.02	-0.07
DUTRVL	0.02*	0.02*	0.04*	0.03*	0.02*
HPH	-0.31	0.11		-0.44*	-0.06
HGH	-0.41*	-0.54*	-0.64*	-1.12*	-0.86*
HEH	-0.60*	-0.43*	-0.58*	-1.54*	-0.91*
HTH			-0.32	-1.04*	-0.67*
HPGH	0.76*	0.82*	0.95*	0.37*	0.56*
HPEH	1.05*	0.65*	0.61**	0.08	1.13*
HEGH	1.01*	0.75*			
TURN	-0.09	-0.09*	-0.06	-0.10**	-0.04
MALE	-0.26*	-0.03	0.03	0.13*	0.23**
MARRID	0.42		-0.08	-0.21	0.25
AGE		0.01			
DRIVRLIC	-0.08	-0.14	-0.27*	-0.27**	-0.14*
BUSPASS	-1.07	0.28**			
LOW	-0.26**	-0.13	0.15	0.33*	-0.02
DWLOWN	-0.30	-0.37	0.35	-0.16	0.62
HV1	0.01	-0.26*	-0.09	0.03	-0.07
URBCR	0.07	-0.58	-0.39	0.19	-0.19
SUBRB	0.10	-0.19	-0.56**	-0.10	-0.24
RURAL	1.20	0.71**		1.39*	
POPDEN	0.74*	-0.24	0.48*	-1.69**	
LNDUS	-0.25	0.19*	0.64*	0.26*	-0.52
RESID	0.02	0.03**	0.04**	0.03	0.03
COMM		-0.04	0.05	0.05*	0.05
INTERSC	-0.06	0.03	0.01	0.10*	-0.05
RETAIL		-1.10	-1.76		-1.59
Threshold Parameter (θ)					
θ (01)	1.80*	2.20*	1.89*	2.01*	2.07*
θ (02)	2.17*	2.54*	2.43*	2.44*	2.40*
θ (03)	2.75*	3.22*	3.19*	3.11*	2.99*
θ (04)	3.03*	3.56*	3.54*	3.46*	3.40*
θ (05)	3.65*	3.89*	4.14*	3.81*	3.77*
Log likelihood (constant only)	-657.75	-882.88	-747.24	-756.12	-745.89
Log likelihood (full Model)	-495.29	-722.40	-600.05	-592.49	-599.38
AIC	2.70	2.91	2.85	2.87	2.91

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Residential location affects trip chaining significantly. As expected, individuals living in the urban core and suburbs of Halifax are less likely to make complex tours: residence in these areas reduces travel distances by offering greater accessibility and land use mix. Population density in the home neighbourhood, residential density, and land use mix are positively associated with trip chaining, meaning that tours undertaken from and to compact and densified residential areas tend to be short and complex due to the nearness of activity destinations, which is consistent with Kitamura and Susilo (2005). Similarly, intersection density and commercial density have positive associations with trip chaining. Presumably, intersection density offers a better transportation network as well as travel options which may motivate people to make complex tours. Retail floor area ratio is found to be negatively associated with trip chaining, which is consistent with the findings of Chen and Akar (2017).

6.5.3 Tour Mode Choice Results of Multinomial Logit (MNL) Model

Before proceeding to tour mode choice modeling, we conducted robust analysis to generate the choice set. At the beginning, ten modes/multi-modes were considered, which were car drive, car passenger, walk, transit, bike, car drive & walk, car passenger & walk, transit and walk, bike & walk, and car & transit. However, modal shares for transit, bike, transit & walk, bike & walk, and car & transit were all less than 2%, and these were therefore excluded. Thus, only car drive, car passenger, walk, car drive & walk, and car passenger & walk were modeled, utilizing a series of MNL models.

Table 6.6, Table 6.7, Table 6.8, Table 6.9 and Table 6.10 describes the results from the mode choice models for all five non-worker clusters. Consistent with results for previous tour mode choice models, house size, number of autos in the household, gender, age, marital status, educational level, and income level, were found to be significant in the model. Male individuals tend to drive a car rather than to be a car passenger in a tour. Males are also less likely to choose walk mode for a home-based tour. Age has positive association with the choice of driving a car for tour mode choice, likely because a car provides independence, convenience, and comfort for older aged travelers. Age has both positive and negative association among clusters for car passenger, walk, and car drive & walk mode choices, indicating variation between clusters in terms of mode choice decisions. Larger household size increases the propensity of choosing car passenger, walk, and car drive & walk modes for home-based tours. Presumably, larger households offer greater opportunity for joint travel. It is interesting to note that individuals from cluster 5 who obtained bachelor degree or above (Table 6.2) are more likely to choose driving car and car passenger modes for home-based tours.

Among the tour attributes, the sign of travel time is negative and significant as expected. Number of tours per day has positive association with the car drive and walk mode choices. This is presumably because car drive and walk modes both offer greater freedom and convenience to undertake a simple and short tour at any time of the day. In contrast, car passenger is dependent on the schedule of other household or non-household members. Trip chaining has positive association with driving car mode choice, which is consistent with earlier studies which found that complex tours are more auto-oriented in comparison to simple tours (Strathman and Dueker 1990; Ye et al. 2007). In contrast, trip chaining has negative association with the walk mode choice.

Type of tour as well as purpose of the tour are also found to be significant in the model. Individuals are more likely to drive to accomplish simple shopping and entertainment tours, probably because choosing car increases the convenience of carrying shopping goods, and perhaps individuals travel longer for entertainment activities. On the other hand, walk, and car drive & walk are found to be negatively associated with shopping and entertainment tours, implying the opposite reason for driving a car. Complex tours of type HPEH have positive signs for car drive and car passenger and negative signs for walk, and car drive & walk modes. Presumably, individuals travel longer distances for entertainment activity and for goods shopping, which make car a preferred mode in comparison to other modes.

Residential location significantly affects the mode choice for tours. Individuals living in suburban areas are more likely to choose car mode, and car & walk mode for home-based tours, and less likely to choose walk mode. This may be because individuals in the suburbs need to travel longer distances for non-work trips. Land use mix, population density, sidewalk density, retail density, and intersection density are found to be significant for tour mode choice. Higher land use mix and sidewalk density have positive effects on walk mode choice and negative effects on car drive, car passenger, and car drive & walk mode choices. This is presumably because greater land use mix offers a higher number of activity destinations in closer proximity, and greater sidewalk density offers greater perceived safety to choose the walk mode.

Table 6.6 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Car drive

Explanatory Variables	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Constant	3.91*	1.79*	1.10	-0.37	2.33**
HHSIZE1	-0.28**	-0.32	-0.67	-0.39	-0.02
DUTRVL1	-0.01	-0.02**	0.01	-0.04*	-0.01
TOURN1	0.44*	0.81*	-0.20	1.03	0.16*
HPH1		0.78*	-0.43	0.08	0.63**
HGH1	-0.61	2.00**	-1.02**	0.26	-0.85**
HEH1	3.30*	2.90*	0.67*	-2.01	-2.32
HTH1				-1.36	-1.10
HPEH1	2.22*				
CHAIN1	0.45*	0.31	0.21*	0.55*	0.38*
MALE1	0.18	2.55**		0.87	1.57*
AGE1				0.06*	0.06*
MARRID1	0.24	0.20		0.68**	-1.38
DIPLOMA1					1.58*
HISCHL1	0.83				
HIGH1		1.48		0.64	0.49
LOW1		0.22			
HV11	0.33	1.13	1.07	2.93*	1.43*
URBCR1	-0.78		1.38	0.51	0.05
SUBRB1	0.12	3.89*	0.74*		
POPDEN1	-2.74	3.57	-1.27*	-0.51**	-4.13
LNDUS1		-1.28*	1.26		
SIDEWALK1		-3.45*	-1.09**	1.90	-0.65
INTERSC1	0.34		1.33*	-0.50	-0.06

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 6.7 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Car passenger

Explanatory Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Constant	4.23*	-0.81	-0.78	-3.23	5.40*
HHSIZE2	-0.78**	-0.47	0.84*	-0.58	-0.94*
DUTRVL2	0.01**	-0.01	0.02**	-0.03*	-0.01
TOURN2	0.30	0.79	-0.24	1.03	0.14
HPH2		0.20	-0.37	-0.07	-0.94
HGH2	-0.98	-1.59	-1.35	-0.54	-2.19**
HEH2	-0.38	-1.54	0.66*	-0.96	-1.59
HTH2				-0.67	-1.48**
HPEH2	2.60*				
CHAIN2		-0.44**			
MALE2	-1.97*	0.69		-0.91	0.13
AGE2				-0.08*	0.02**
MARRID2		0.57		1.46	1.58
DIPLOMA2	2.41*				1.41
HISCHL2	0.56				
HIGH2		2.08		0.57	0.61**
LOW2		0.57			
HV12	-0.85	0.72**	1.22	2.87*	0.48**
URBCR2	-0.77		0.08	0.17	0.31
SUBRB2	1.06	-3.42**	0.74		-1.00
POPDEN2	-10.85	3.95	-1.40*	0.03	
LNDUS2		1.50*	1.27		
SIDEWALK2		-2.94**	-0.46	0.87	-0.74**
INTERSC2	0.57		1.48*	-0.72**	-0.21

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 6.8 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Walk

Explanatory Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Constant	1.33	0.34	2.54	-1.78	1.33
HHSIZE3	-0.55	-0.08	-0.70	0.05**	0.50*
DUTRVL3	-0.05	-0.01**	0.03	-0.04*	-0.02
TOURN3	0.82	-0.12	-0.80**	1.27*	-0.05
HPH3		0.09	-0.42	-0.05	0.94**
HGH3	-0.18	-1.81	0.05	1.67	-0.26
HEH3	-1.56	-2.94*		-0.44	0.02
HTH3				-1.58**	-0.50
HPEH3	-1.92				
CHAIN3		-0.91*			
MALE3	-0.85**	2.56**		-0.33	2.61**
AGE3				0.06**	0.02
MARRID3	1.27	-1.27		0.77	1.52*
DIPLOMA3					-2.26
HISCHL3	1.00				
HIGH3		1.76		-0.31	-2.01*
LOW3		2.63*			
HV13	-0.15	-0.01	-0.29	1.74	0.27
URBCR3	0.34		1.79	2.06	2.30
SUBRB3	1.13**	-3.96*	-2.17*		-1.10
POPDEN3	3.67**	8.41	-0.60*	-0.34	
LNDUS3		1.56*	1.05		
SIDEWALK3		3.02	0.27	2.43*	-0.21
INTERSC3	0.16		0.75	-0.51	-0.19

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 6.9 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Car drive and walk

Explanatory Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coeff.	**Coeff.**	**Coeff.**	**Coeff.**	**Coeff.**
Constant	-2.45	-2.42**	-3.20**	-7.27*	-1.19
HHSIZE4	0.36**	-0.36	-0.78	-0.45	0.34**
DUTRVL4	-0.03	-0.04	-0.02**	-0.03*	-0.02
TOURN4	0.18	0.50**	0.22	1.31*	0.39*
HPH4		0.58	-1.75	-0.10	-0.04
HGH4	-0.02	-2.29**	-1.52	1.96	-0.38
HEH4	-0.17	-1.60		-1.02	-10.36*
HTH4				-0.92	1.35**
HPEH4	-1.91*				
CHAIN4		0.11			
MALE4	0.26	3.34*		0.87	1.47**
AGE4				0.06*	0.02
MARRID4	0.13	-0.08		0.61	1.16*
DIPLOMA4					-2.19
HISCHL4	0.59**				
HIGH4		1.90		0.86**	-0.13*
LOW4		1.02			
HV14	0.35**	0.73**	1.15	2.86*	1.70
URBCR4	0.74		3.69	0.36	-0.14
SUBRB4	1.58	3.81*	1.97		
POPDEN4	-0.37	3.34	-2.19*	-0.29	5.04
LNDUS4		1.25*	0.65**		
SIDEWALK4		-2.48	-0.54	2.16	-0.33*
INTERSC4	0.07		1.30	-0.53	-0.76

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 6.10 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Fitness parameters

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Log likelihood (constant only)	-387.81	-468.86	-384.74	-407.79	-432.64
Log likelihood (full Model)	-293.87	-327.19	-311.72	-297.95	-333.03
AIC	2.69	1.54	1.67	1.71	1.89

6.6 Conclusions

Land use and transportation planners have been increasingly interested to better understand relationships between land use characteristics and household travel behavior.

It is now well understood that the form, density, and land-use mix of the built environment shapes travel behavior. Most studies have focused primarily on their impacts on worker travel, but urban form also shapes the travel behavior of those who do not work outside the home. The non-worker population contributes a significant share to total urban traffic. Moreover, non-workers have a flexible activity schedule in comparison to workers. Modeling non-workers' tour formation in combination with that of workers will increase error for both groups. Therefore, this study presented tour modeling for non-worker population clusters, including tour formation, tour typology, tour frequency, trip chaining, and tour mode choice, to better understand their travel behavior.

Five non-worker clusters were drawn from the Halifax Space-Time Activity (STAR) dataset using a daily activity pattern recognition method: Cluster#1 non-workers Midday activities, Cluster#2 non-workers evening activities, Cluster#3 non-workers stay-at-home, Cluster#4 non-workers morning shopping activity, Cluster#5 non-workers afternoon shopping activity. Fully 95% of the home-based tours made by non-workers are non-work tours. On an average, a non-worker produces at least one home-based tour per day.

Tour frequency of five non-worker clusters was modeled by the Poisson regression model. Along with socio-demographic variables, urban form attributes are found to be significant in the model. Individuals from urban core and suburban areas make more tours compared to those from urban fringe and rural areas. It is worth noting that greater land use mix has a negative impact on tour frequency; the presence of more nearby destinations may encourage trip-chaining and thereby reduce tour frequency.

A series of ordered probit model was utilized to study the trip chaining behavior of travelers. Household size is negatively associated with trip chaining. Possession of driver license was also found to be negatively associated with trip chaining, whereas it was positively associated with tour frequency. Number of cars in the household has negative association with trip chaining. Among the complex tour sequences, HPGH, HPEH, HEGH are more likely to have more trip segments than other tour types.

A series of MNL models was estimated to understand non-workers' tour mode choices. The results of the models suggest that with better street connectivity, retail density, land use mix, and sidewalk density, individuals are more likely to choose the walk mode. Moreover, travel time has a negative association with all the tour mode choices. These findings have implications for land use planning policy, vindicating current policies promoting more efficient public transport, greater land use mix, and densification.

While the results of this research offer important insights into the role of the land use pattern in shaping travel patterns and tour complexity, there are some limitations to the current study. This can be further improved by analyzing the household interaction, accompaniment pattern, and joint travel of individuals in non-worker clusters. It would also be interesting to do the same study among worker clusters and compare the results. It is evident from this study that segmenting non-workers into clusters based on their activity-pattern can be a useful component to be incorporated in tour analysis of individuals and households.

Chapter 7 Trip Chaining Propensity and Tour Mode Choice of Workers: Evidence from a Mid-Sized Canadian City[5]

7.1 Introduction

Over the past years, activity-travel behavior has received increasing attention from transportation planners and professionals. This is because the combined effects of suburban development patterns, flexible work hours, and increasing participation in out-of-home activities all contribute to growth in vehicles miles traveled, which is directly associated with increased energy consumption, greenhouse effects, traffic congestion, and accidents (Akar et al. 2016), and decreased environmental quality. Therefore, it is important for transportation planners and policy makers to reduce drive-alone travel, as well as to regulate and manage individual's activity-travel demand and vehicles miles traveled (Acker and Witlox 2011). Previous studies showed that vehicles miles traveled and travel demand are associated with the activity-travel patterns of individuals (Duncan 2015; Bricka 2008). For the population segment working outside the home, the time-of-day for work activity is not consistent for all the individuals, and it is, therefore, useful to cluster workers by their time-use behavior. Thus, this study contributes to the literature by analyzing trip chaining behavior, tour complexity, and mode choices of worker groups clustered by their time-use activity patterns.

Most paid workers live and work in two geographically separate locations, which initiates the need for commuting trips. Other activity locations may then be visited on the way to

---

[5] A version of this chapter has been conditionally accepted:
Daisy, N. S., Millward, H., and L. Liu. Trip chaining propensity and tour mode choice of workers: Evidence from a mid-sized Canadian city. 2018.

or from work, thus producing trip chaining tours. Most activity-based models produce the tours first and then add stops to these tours to model daily travel of individuals (Bradley et al. 2009, Goulias et al. 2010). Most often in the literature, tours are defined as home-to-home loops. Tours with one stop are called simple tours (e.g., home to work and then work to home); and tours with more than one stop are called complex tours (Paleti et al. 2003; Krizek 2003; Primerano et al. 2008). Better knowledge of trip chaining is important to better understand the relationship between out-of-home activity engagement and travel; for instance, how non-work activities and trips are linked to work trips (Ho and Mulley 2013). Kitamura and Susilo (2005), Susilo and Kitamura (2005, 2008) and Dharmowijoyo et al. (2016) argued that trip chaining behavior of individuals is governed by their mode choice behavior or their usage of the motorized mode. Strathman and Ducker (1994) found that auto commuters are more likely to chain non-work trips to their work trip. Presumably, the auto mode offers greater flexibility and convenience to trip chain compared to other modes. Hence, studying trip chaining and tour mode choice behavior for out-of-home workers is important to promote policies such as switching to public transit (Hensher and Reyes 2000).

This study examines trip chaining tours using Halifax, Canada, as a case study. Like other mid-sized cities, Halifax, Canada, is experiencing traffic congestion caused by growing drive-alone travel as well as due to the increased number of commuters. The number of workers in Halifax Census Metropolitan Area (CMA) increased from 174,700 to 189,300 between 2001 and 2006, and from 189,300 to 224,595 between 2006 to 2011, increases of 8.4% and 18.6%, respectively. This increasing trend of workers in Halifax Regional Municipality (HRM) contributes to road congestion and decreased commuting speeds. A recent report shows that on average a Halifax commuter spent 23.7 min on the journey to work, and 76.6% of commuters' mode of transportation was car, truck, or van (National Household Survey 2011). To ameliorate these growing transportation issues, transportation planners and policy makers have adopted an 'integrated mobility plan' for HRM which seeks to integrate transportation and land use planning policies. This plan recognizes that the activity-travel patterns of individuals can be altered by changing neighborhood density and design characteristics, and the locations of employment and activity destinations. Illustrative of such policies that combine transport and land use are, in the USA, the New Urbanism movement (Handy 2005); and in Europe, the Compact City Policy (Maat et al. 2005). Numerous studies confirm that there exist strong associations between land use and travel behavior (Crane 2001; Bagley and Mokhtarian 2002; Frank et al. 2007; Frank et al. 2008; Ewing and Cervero 2010; Koohsari et al. 2017). Planners focus particularly on the land-use and neighborhood design features which promote the use of walking and transit modes. Many empirical studies have been conducted on these issues, and several meta-studies are now available (Saelens and Handy 2008; Sallis et al. 2009; Renalds et al. 2010; Koohsari et al. 2015).

Many studies have employed a trip-based approach (Guo et al. 2007), or an aggregated approach to travel behavior (Giuliano and Dargay 2006). In contrast, activity-based approaches consider that individuals schedule their daily activities in a daily activity agenda, and then plan tours rather than simple trips (Bhat and Koppelman 1999; Primerano et al. 2008). Thus, it is important to understand tour attributes and complexity, such as number of tours per day, as well as tour-related mode choices, to examine the relationship between land use patterns and activity-travel behavior. Instead of studying the tour, trip chaining, and mode choice behavior of workers in an aggregated manner, this study employs distinct disaggregated population groups. We adopt a daily time-use and activity pattern recognition framework to group workers into heterogeneous clusters. We then estimate the group utility functions for tour propensity and mode choices. Compared to earlier works on tour complexity and mode choice modeling, this study contributes by further exploring how, for example, a 9 to 5 worker behaves differently from a shorter-day worker, or from workers who typically avoid travel in the peak hours. It is hypothesized that there exist significant differences in tour frequency, trip chaining, and mode choices expressed in terms of parameter estimates and magnitude of covariates between worker groups with different daily activity patterns.

This study utilizes data drawn from the STAR household travel survey data, conducted in Halifax Regional Municipality. Hafezi (2017a, b) first identified five clusters of out-of-home workers, from their activity-travel patterns. The groups were identified a posteriori from similarities in the time-use and activity patterns of the individuals. The group labels reflect our interpretation of the most salient feature of the activity patterns exhibited within each group. They were labeled as extended workers, 8 to 4 workers, shorter work-day workers, 7 to 3 workers, and 9 to 5 workers. For these five groups, tour complexity, trip chaining, and tour mode choices are used as dependent variables, and their relationships to individual's socio-demographic characteristics, and neighborhood land use patterns, are examined. The results of this study are expected to be implemented within an activity-based travel demand model named Scheduler for Activities, Locations, and Travel (SALT). In the overall SALT modeling framework, initially, clusters with homogeneous time-use and activity pattern are identified. In the next step the activity type choices and activity sequences are predicted (Daisy 2018a). Next, we form the tours and stop frequency and assign tour mode choices as modeled in the current study. We analyze stop-frequency as a step to better understand the trip chaining behavior. Finally, temporal attributes of predicted activities in the agenda are inferred and 24-hour schedules of individuals are constructed.

7.2 Literature Review

Numerous studies have examined both travel distances and trip chaining along with their explanatory socio-demographic and urban-form determinants. Trip chaining was defined in the early work of McGuckin and Murakami (1999). More recently Primerano et al. (2008) defined trip chaining as the mixture of one or more intermediate activities with the main activity, where home is the start and end of trips. Several studies utilized trip chaining as a measure of tour complexity, using stop frequency within the tours (McGuckin and Murakami 1999; Frank et al. 2008; Liu 2012; Wang 2014). Individual characteristics and socio-demographic attributes, including gender, the age, and employment status have been identified as significant predictors of trip chaining (McGuckin and Murakami 1999; Liu 2012; Kitamura and Susilo 2005; Susilo and Kitamura 2008). Moreover, these studies also confirmed that women, the elderly, and workers are more likely to use trip chains compared to others. Other household characteristics, for instance, household income level and the presence of children, are positively related to trip chaining (Bhat et al. 1999; Krizek 2003; Kitamura and Susilo 2005; Susilo and Kitamura 2008; Acker and Witlox 2011; Liu 2012; Wang 2014). Conversely, household size and number of vehicles in the household are negatively related to trip chaining behavior (Kitamura and Susilo 2005; Susilo and Kitamura 2008; Bricka 2008; Acker and Witlox 2011). However, the association between population density and tour complexity is still not well-understood. Several studies suggest that higher densities are associated with trip chaining (Maat and Timmermans 2006; Bricka 2008). In contrast, other studies suggest that individuals living in low-density neighborhood are more likely to trip-chain tours due to their locational deficiencies (Kitamura and Susilo 2005; Noland and Thomas 2007).

Number of tours per day is found to be associated with employment-related variables, household structure variables, accessibility, location variables, and mobility-related variables (Bhat et al. 1999). Household socio-demographic characteristics such as number of adults, number of employees, number of vehicles, household income, home-to-work distance, work neighborhood accessibility, and work residential accessibility were also found to be associated with tour frequency (Krizek 2003). Yet, most of these studies examined tours for workers and non-workers under similar frameworks, suppressing the differences between these two groups of individuals in terms of daily-activity patterns and schedules.

The relationship between trip chaining and mode choice has also received considerable study. Some studies suggest that, in general, complexity of the tour as measured by stop frequency has positive association with the choices of auto and walk modes (in the latter if the trip is short and a sidewalk is available), but a negative association with the transit mode choice (Bhat et al. 1999; Yun et al. 2014). To our knowledge, an analysis of workers' tour complexity and mode choice under a cluster-based framework has not been undertaken. Hence, this study aims to deliver a comprehensive econometric investigation for worker clusters, to better understand their trip chaining, tour complexity, and mode choices for HRM commuters. The results of this study are expected to be incorporated into the Scheduler for Activity, Location, and Travel (SALT) model for Halifax.

7.3 Data

7.3.1 Data Source

Time-use and travel data employed in this study were obtained from the Halifax Space-Time Activity Research (STAR) project conducted between April 2007 and May 2008. STAR was the world′s first large survey to use global positioning system (GPS) technology to augment and verify household activity-travel diary data. Halifax is a medium-sized city (c.400,000 population) and the capital of the Canadian province of Nova Scotia. It has a growing economy with diversified population segments. Population growth per year averages 0.5%. The STAR project collected travel and household information of all members of randomly selected households in HRM, and the total sample consists of 1,971 households. In each sample household, one individual 15 years of age or older was selected as the primary respondent, and completed a two-day GPS-validated time-diary. The STAR survey responses were distributed equally through all days of the week and 11 months of the year (December was excluded owing to unusual activity patterns related to Christmas). We used only weekday diaries for cluster identification and empirical model estimation (Hafezi 2017a, b).

The time diary coding and questionnaire on household characteristics were based on the General Social Survey (GSS) Cycle 19 of Statistics Canada (2006). Primary respondents carried a GPS-device (Hewlett Packard iPAQ hw6955) for all out-of-home activity, programmed to collect GPS data every 2 s. The GPS data provided precise start and end times, and travel routes for all “stops” with more than 2 minutes stopping duration. These GPS data were used with CATI software in day-after interviews with respondents, to verify and enhance the time-diary data. The land use characteristics and built environment variables used in this study were obtained from work by Neatt et al. 2017.

7.3.2 Cluster Description

In our study we used distinct population groups with similar distributions of start time, activity type, and frequency. As activity generation modules have a direct effect on prediction accuracy, it is preferable that populations with similar characteristics are grouped into distinct homogeneous clusters. Therefore, initially, five clusters of out-of-home workers were identified from the STAR dataset based on their homogeneous time-use activity patterns. To identify the daily-activity pattern, a pattern recognition model was applied (Hafezi et al. 2017a, b). A subtractive clustering algorithm was utilized to initialize the total cluster number and cluster centroids. Identification of individuals with homogeneous activity patterns was accomplished using a fuzzy c-means clustering algorithm, and sets of representative activity patterns were identified using a multiple sequence alignment method. The group labels reflect our interpretation of the most salient feature of the activity patterns within each group. Advanced decision tree models were used to explore inter-dependencies in each identified cluster, and characterization of cluster memberships through their socio-demographic attributes was achieved by use of the CART algorithm. Table 7.1 shows the personal attributes and time allocation to activities for the five identified worker clusters. These clusters are labeled as follows: (1) extended day workers, (2) 8:00 am to 4:00 pm workers, (3) shorter work-day workers, (4) 7:00 am to 3:00 pm workers, and (5) 9:00 am to 5:00 pm workers.

Table 7.1 Cluster characteristics

Socio-Demographic variables		Sample mean (%)	Mean of cluster (%) #1	#2	#3	#4	#5
Gender	Female	0.53	0.53	0.44	0.52	0.47	0.53
Age	Young adults (ages 15-35 years)	0.10	0.12	0.10	0.10	0.05	0.09
	Middle-aged adults (ages 36-55 years)	0.49	0.67	0.66	0.71	0.72	0.70
	Older adults (aged older than 55 years)	0.41	0.20	0.24	0.19	0.23	0.22
Education Level	Diploma or university certificate	0.67	0.76	0.76	0.85	0.53	0.80
Occupation	Regular shift	0.53	0.73	0.93	0.87	0.93	0.89
	Irregular schedule	0.10	0.22	0.03	0.09	0.07	0.08
	Student	0.03	0.01	0.00	0.01	0.00	0.01
	Retired	0.23	0.02	0.02	0.01	0.00	0.00
	Work at home	0.15	0.23	0.13	0.30	0.06	0.26
Flexible schedule	Have no flexibility in a work schedule	0.50	0.55	0.54	0.44	0.63	0.40
Job number	Have more than one job	0.07	0.09	0.04	0.05	0.07	0.08
Income	Low-income (<= $ 40,000)	0.39	0.28	0.22	0.32	0.29	0.26
	Middle-income ($ 40,000 - $ 100,000)	0.53	0.60	0.68	0.55	0.64	0.59
	High-income (> $ 100,000)	0.09	0.12	0.10	0.13	0.08	0.15
Total cluster membership			137	401	171	229	348
Average Number of tours per person per day			1.56	1.67	1.87	1.41	1.53
Percentage in total (number of person-days)			4.93	14.43	6.16	8.24	12.53
Activity categories	**Descriptions**		**Share of daily activity engagement (%)**				
Home chores (H)	Working at home, eating/meal preparation, indoor or outdoor cleaning, interior or exterior home maintenance, child care, other in-home activities.		25.09	27.81	33.78	30.14	29.47
Home leisure (L)	Watching TV/listening to radio, reading books/newspapers, etc.		12.04	16.57	16.22	17.73	15.32
Night sleep (N)	Night sleep		62.88	55.62	50.00	52.13	55.21
Total in-home (%)			100.0	100.0	100.0	100.0	100.0
Workplace (W)	Work/job, all other activities at work, work related (conferences, meetings, etc.).		**89.26**	**82.61**	**79.53**	**89.31**	**86.56**
Shopping & services (P)	Shopping for goods and services, routine shopping.		1.61	3.16	5.56	2.41	2.98
School/college (S)	Class participation, all other activities at school.		0.00	0.14	0.40	0.29	0.15
Organizational/ hobbies (G)	Organizational, voluntary, religious activities. Hobbies are done mainly for pleasure, cards, board games, all other hobbies activities.		2.34	3.24	5.11	1.76	2.62
Entertainment (E)	Eat meal outside of home, all other entertainment activities.		4.67	6.92	6.48	3.73	4.30
Sports (T)	Walking, jogging, bicycling, all sports related activities.		2.12	3.93	2.93	2.50	3.40
Total out-of-home (%)			100.0	100.0	100.0	100.0	100.0

The details of each cluster are as follows:

Cluster#1 is the extended day workers group. The individuals belonging to this cluster typically participate in work activity for a longer duration, extending from 8:00 am to 8:00 pm. This cluster predominantly comprises middle-aged females aged between 36 and 55 years old (67.0%). Almost 76.0% of them are high-school graduates, and 73.0% are full-time workers. Individuals from this group are mostly middle income (60.0%), and the majority of the workers (55.0%) had no flexibility in their work schedule.

Cluster#2 is the 8:00 am to 4:00 pm worker cluster. This cluster mostly comprises middle-aged males with high-school graduation or better. More than 92% of the workers in this cluster work full-time, and their income level is middle-income. Workers of this cluster participate in discretionary activities typically in the evening.

Cluster#3 is the shorter work-day workers, who work less than 5 hours a day and who typically finish their work in the early afternoon before 2:00 pm. The majority of this cluster are middle-aged females between 36 to 55 years old (71.0%). Additionally, 85.0% of individuals in this group are high school graduates, and 56.0 % had some flexibility in their work schedule.

Cluster#4 consists mostly of 7:00 am to 3:00 pm workers. The majority of individuals from this group are middle-aged males between 36 to 55 years old, and 47.0% have middle-income. Nearly all individuals in this cluster are full-time workers (93.0%), and 63.0% of them had no flexibility in work schedule.

Cluster#5 mostly comprises individuals who work from 9:00 am to 5:00 pm. Unlike cluster 4, individuals from cluster 5 usually travel to and from work during the morning and evening peak hours. A large proportion of individuals from this cluster are middle-aged females between 36 to 55 years old (53.0%) with middle-income, and most are high school graduates. The majority of the workers (60.0%) indicate that they have some flexibility in their work schedule.

7.3.3 Tour Formation Behavior

Home-based tours are defined as home-to-home journeys comprising a sequence of out-of-home trip-stops (destinations) with no intermediate home stops (Shiftan 1998). The intermediate stops within a tour are measured to identify the trip chaining behavior. In this study, the cutoff point for number of stops in a tour was four trips, as very few workers undertake more than four trips in a tour. For each respondent, the tour frequency per day, stops per tour, and mode choices for each tour were identified.

After forming the tours, based on the purpose of the primary activity, the tour purpose for work and school tours was assigned based on the priority order by Stopher et al. (1996), which gives highest priority to work, followed by education. If an individual makes complex tours with more than one stop, then the second activity purpose is also identified. As all the clusters in this study are worker clusters, the highest participated activity is work activity. With the exception of cluster 3, more than 60% of total tours are simple or complex work tours. Other discretionary activity purposes for non-work tours were identified based on the longest activity duration. Simple tours were classified into six categories: home-work-home (HWH), home-school-home (HSH), home-shopping-home

(HPH), home-organizational/hobbies-home (HGH), home-entertainment-home (HEH), and home-sports-home (HTH). In this study, we simplified the complex tours with three or more stops to two stops with the highest duration. Complex tours (with two non-home stops) were categorized as shown in Table 7.2. Variables utilized in the empirical model estimation are presented in Table 7.3. We added HWH and HPH as explanatory variables with an assumption that if the number of single stop work (HWH) or shopping (HPH) tours increases then the total number of tours undertaken in a given day would increase.

The travel mode/modes for each tour were selected based on the longest in-mode travel time. The choice set generates ten choices, which are car drive, car passenger, transit, bike, walk, car drive & walk, car passenger & walk, transit & walk, bike & walk, and car & transit. The number of stops and average duration of time spent at each stop was calculated through data mining methods.

Table 7.2 Tour typology

Tour Type	Description
H-W-H	Home-Work-Home
H-P-H	Home-Shopping & Services-Home
H-S-H	Home-School/College-Home
H-G-H	Home-Organizational/Hobbies-Home
H-E-H	Home-Entertainment-Home
H-T-H	Home-Sports-Home
H-W-P-H	Home-Work & Shopping-Home
H-W-S-H	Home-Work & School-Home
H-W-G-H	Home-Work & Organizational-Home
H-W-E-H	Home-Work & Entertainment-Home
H-W-T-H	Home-Work & Sport-Home
H-S-G-H	Home-School & Organizational-Home
H-P-G-H	Home-Shopping & Organizational-Home
H-P-E-H	Home-Shopping & Entertainment-Home
H-P-T-H	Home-Shopping & Sports-Home
H-E-G-H	Home-Entertainment & Organizational-Home
H-S-E-H	Home-Entertainment & School-Home
H-T-G-H	Home-Sports & Entertainment-Home
H-T-E-H	Home-Sports & Organizational-Home

Table 7.3 Details of exploratory variables used in the empirical analysis

Abbreviation	Variable Description
MDDURA	Total Travel Time per tour
HPH	Home-Shopping-Home tour (1, if the tour is H-P-H tour, 0 otherwise)
HGH	Home-Organizational/hobbies-Home tour (1, if the tour is H-G-H tour, 0 otherwise)
HEH	Home-Entertainment-Home tour (1, if the tour is H-E-H tour, 0 otherwise)
HTH	Home-Sport-Home tour (1, if the tour is H-T-H tour, 0 otherwise)
HPGH	Home-Shopping-Organizational-Home tour (1, if the tour is H-P-G-H tour, 0 otherwise)
HPEH	Home-Shopping-Entertainment-Home tour (1, if the tour is H-P-E-H tour, 0 otherwise)
HEGH	Home-Entertainment-Organizational-Home tour (1, if the tour is H-E-G-H tour, 0 otherwise)
TOURN	Number of tours per day
STOPN	Number of stops in a tour
MALE	Male (1, if the gender of the individual is male, 0 otherwise)
MARRID	Married (1, if the individual is married, 0 otherwise)
SINGLE	Single (1, if the individual is single, 0 otherwise)
HHSIZE	Household Size
AGE	Age of the individual
DIPLOMA	Diploma or university certificate (1, if the individual has obtained grade 12 graduation level education, 0 otherwise)
HISCHL	High school graduate (1, if the individual has obtained high school degree, 0 otherwise)
DRIVRLIC	Driver License (1, if the individual has a valid driver license, 0 otherwise)
BUSPASS	Bus Pass (1, if the individual has a valid Bus pass, 0 otherwise)
CAR	Car Mode (1, if individual choose car mode for the tour, 0 otherwise)
WLK	Walk Mode (1, if individual choose walk mode for the tour, 0 otherwise)
LOWINCOM	Low income level (1, if the individual belongs to low income level, 0 otherwise)
HIGHINCOM	High Income (1, if the gender of the individual has high income level, 0 otherwise)
DWLOWN	Home Ownership (1, if the individual owns a home, 0 otherwise)
HV1	Number of car in the household
COMTYM	Mean commute time in minutes
URBCR	Urban Core (1, if the individual is living in urban core, 0 otherwise)
SUBRB	Suburban (1, if the individual is living in suburban area, 0 otherwise)
URBFR	Urban Fringe (Commuter shed) (1, if the individual is living in Urban Fringe area, 0 otherwise)
RURAL	Rural (1, if the individual is living in rural, 0 otherwise)
POPDEN	Population density of the home neighborhood
LNDUS	Land use mix in the home neighborhood
RESID	Residential area density in the home neighborhood
COMM	Commercial area density in the home neighborhood
SIDEWALK	Sidewalk density in the home neighborhood
INTERSC	Intersection density in the home neighborhood
RETAIL	Retail floor area ratio in the home neighborhood

7.3.4 Descriptive Statistics

From the descriptive statistics in Table 7.1, it is found that on average travelers make more than one home-based tour per day across all the clusters. The average tours per day is highest for cluster 3 (1.87 tours/day) and lowest for cluster 4 (1.41 tours/day). The tours are categorized in 19 categories. From Figure 7.1, across all the clusters, nearly 25% of all tours are simple home-work-home (HWH) tours. Among the simple tours, home-shopping-home (HPH) is the second highest in all clusters except cluster 1. Nearly 40% of tours made by clusters 1 and 2 are simple tours, whereas more than 50% of the tours made by clusters 3, 4, and 5 are simple tours. Other frequently undertaken tours are home-work-entertainment-home (HWEH) (more than 10%), home-work-shopping-home (HWPH) (more than 10%), and home-work-organization/hobbies-home (HWGH) (more than 9%).

From Figure 7.2(a), we can see that all individuals from all the clusters make at least one home-based tour per day. On weekdays, almost 40% of individuals from all the clusters except cluster 4 make a second tour. Individuals from cluster 3 make more tours per day compared to other clusters. Less than 15% of individuals from all the clusters except cluster 3 make a third tour in each weekday. From Figure 7.2(b), we can see that the most frequently used mode by far is car drive, followed by car drive and walk, and car passenger. Presumably, this is because individuals making complex tours find car and walk to be the most convenient mode of choice. Other frequently used modes/multi-modes are walk, car passenger & walk, and transit & walk. These six mode/multi-mode categories were used for further empirical modeling, while transit, bike, bike & walk, and car & transit were excluded from the choice set.

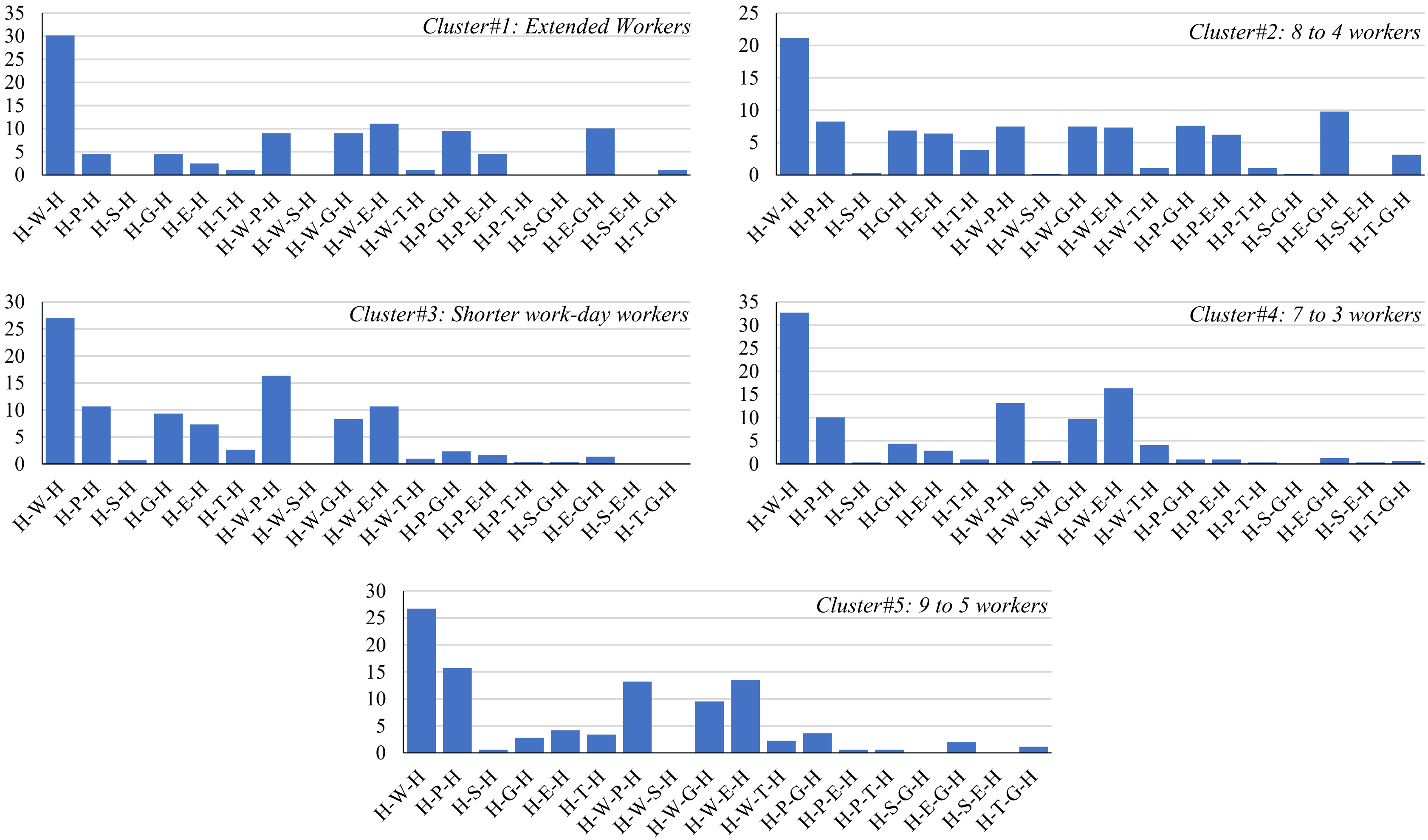

Figure 7.1 Cluster-wise share of different tour types of individuals

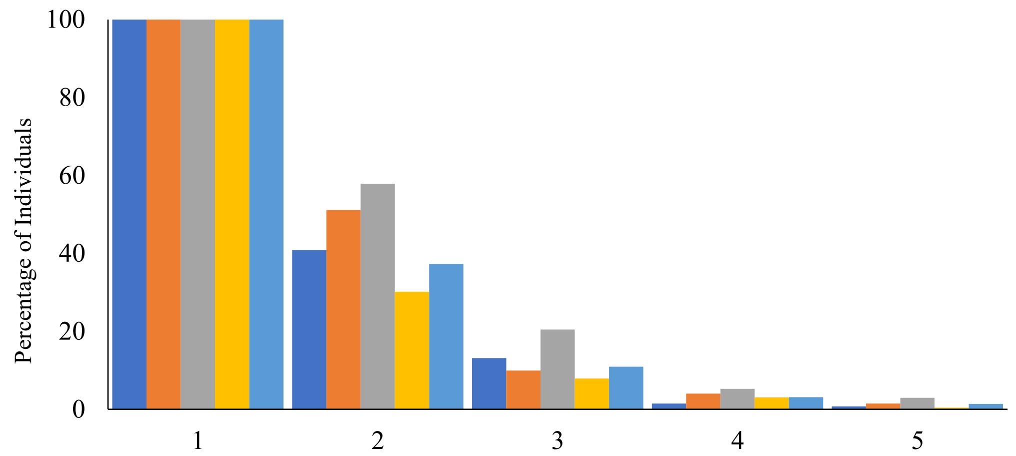

Figure 7.2 (a) Number of tours per day

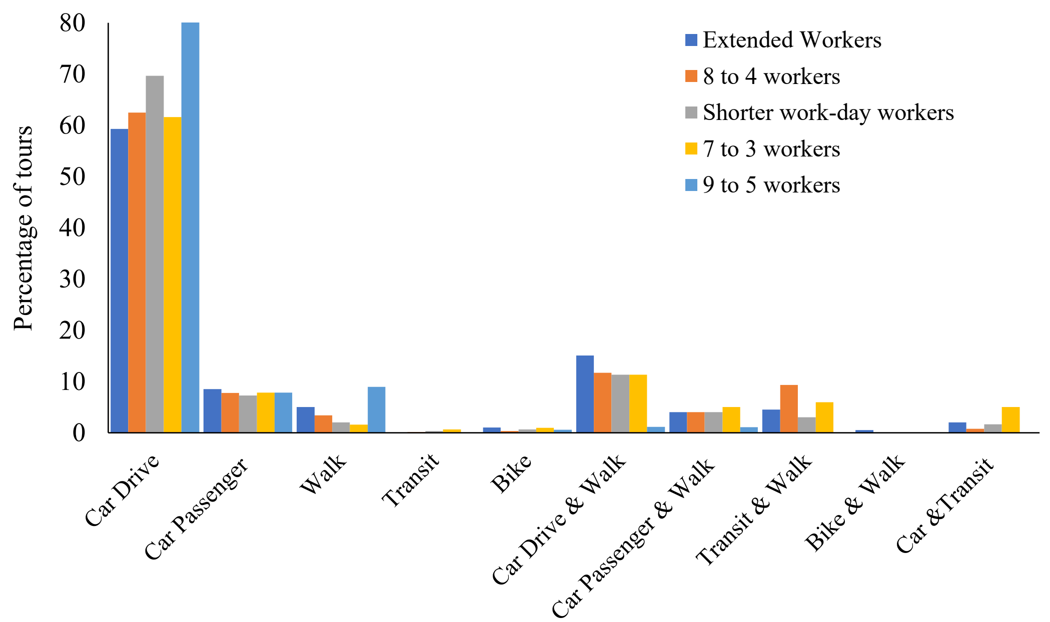

Figure 7.2 (b) Distribution of Modal Share

Figure 7.2 Cluster-wise distribution of modal share of tours and number of tours per day

Figure 7.3 depicts the modal shares for all tour categories and displays significant variation among all the clusters. However, the use of car drive is the dominant mode of transportation across all the clusters for all activity categories.

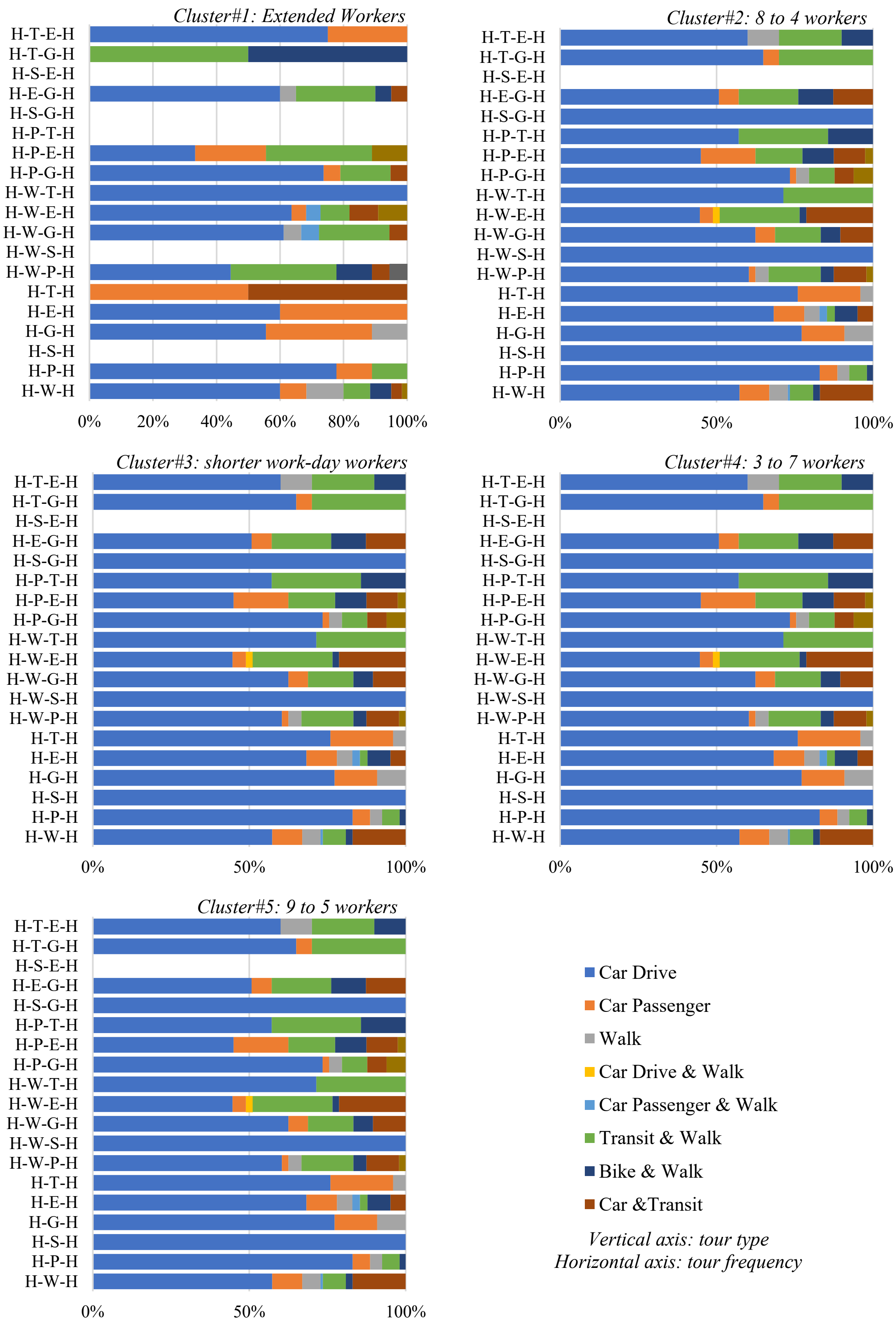

Figure 7.3 Cluster-wise distribution of modal share of tours by different tour types

7.4 Methods

7.4.1 Poisson Regression Model

The Poisson regression model has been utilized earlier for trip and tour study mostly in the aggregated manner. However, to the best of the authors' knowledge, no investigation has applied this model to modeling of tour frequency with distinct disaggregated population clusters. Each cluster contains populations with similar distributions of start time, activity type, and frequency. Therefore, it is assumed that tour frequency per day for clusters with homogeneous daily activity patterns is more likely to have a Poisson distribution. One of the advantages of using the Poisson regression model compared to other alternative econometric models is that tours are countable and number of tours may be treated as a non-negative integer. Thus, the number of tours per day was modeled by utilizing Poisson regression. To illustrate the modeling approach taken in this study, let $\mathrm{Y_r}$ be the frequency of home-based tours completed in a given day by an individual R. Then, a Poisson model can be written as:

$$\Pr(\mathrm{Y_r}) = \frac{e^{-\mu_r}\mu_r^{Y_r}}{Y_r}, r = 1,2,3 \ldots \ldots R \tag{1}$$

where, $Pr(Y_r)$ is the probability of Y_r tours performed by the R th individual, and μ_r is the expected value of Y_r, which can be represented as:

$$\mathrm{E(Y_r)} = \mu_r = \alpha^t X_r \tag{2}$$

Here, α represents a vector of regression parameters to be estimated, and X_r is a vector of variables describing individual's personal and household characteristics, land use attributes, and tour attributes. The partial effects in this non-linear regression are obtained from $\mu_r \alpha$ (Greene 2006).

7.4.2 Ordered Probit Model

For trip chaining, it is hypothesized that with the increase in trips per chain, the tour complexity increases. Therefore, an ordered probit model was developed and applied to the distinct population clusters, instead of other alternative count models such as negative binomial and Poisson regression. Our main contribution is in analyzing the workers with different time-use and daily activity pattern separately, and estimating different empirical models for each worker cluster. The number of stops per tour observed in this study is ordered. So, it is rational to assume that there exists some level of utility associated with trip chaining. As a result, an ordered probit modeling structure was applied to investigate the underlying factors that influence the trip chaining behavior of home-based tours. Let us assume, $\mathrm{q_r}$ is the latent and continuous measure of tour stop frequency for non-work tours for individual R. Assuming ε_r as an error term and α as a vector of parameters associated with explanatory variables $\mathrm{X_r}$, the ordered probit model can be written as follows:

$$\mathrm{Y_r^*} = \alpha \mathrm{X_r} + \varepsilon_\mathrm{r} \tag{3}$$

Where $\mathrm{X_r}$ is a vector of observed explanatory variables including personal characteristics, household socio-demographic characteristics, land use characteristics, and tour attributes. The ordered probit model estimates $\mathrm{k} - 1$ threshold values that horizontally divide the underlying continuous variable to forecast the observed count values, where k is the largest possible count value. Therefore, trip chain Y_r takes on values starting from 0 through N to generate an ordered segment of the latent trip chaining propensity into the observed stop frequencies. To convert Y_r^* to Y_r the cut points θ are introduced as written:

$$Y_r^* \begin{cases} 0 \text{ if } Y_r \leq \theta_0 \\ 1 \text{ if } \theta_0 \leq Y_r \leq \theta_1 \\ 2 \text{ if } \theta_1 \leq Y_r \leq \theta_2 \\ \vdots \\ k \text{ if } \theta_{k-1} \leq Y_r \leq \theta_k \end{cases} \tag{4}$$

Where θ_k is the cutoff points for the possible outcomes. The threshold frequency for this study ranges from 0 to 4. The estimation process of the ordered probit model is straightforward. It is assumed that θ and ε are independently distributed, the likelihood of the trip chaining can be attained as a mixture of a selected distribution at zero and assumed distribution of the response variable Y_r. Maximum likelihood is used to estimate the parameters of the model. This model is an extension of a probit model for a binary outcome where the projected probability of observing a particular ordinal outcome can be generated as follows:

$$\Pr_r(k) = \Pr(Y_r = k) = \Pr(\theta_k < Y_r^*) = 1 - \varphi(\theta_k - \alpha X_r) \tag{5}$$

Assuming an indicator variable ω that equals one if the traveler makes k stops in a tour, and 0 otherwise, the log likelihood can be written as:

$$\ln L = \sum_{i=1}^{k} \sum_{k=0}^{k} \omega \ln[\varphi(\theta_k - \alpha X_r) - \varphi(\theta_{k-1} - \alpha X_r)] \tag{6}$$

The log-likelihood was estimated by using the maximum likelihood method. Finally, Alkaline Information Criteria (AIC) is used to evaluate the goodness of fit of the estimated models.

7.4.3 Multinomial Logit Model (MNL)

We generated the mode choice set through machine learning and data mining techniques and selected the one or two modes with the highest in-mode time in each tour. For example, an individual may make his first tour in the day through auto drive only, and then a second tour using both auto drive and walk modes for the second tour. Empirical investigation using correlation among the modes of the generated choice set showed that a nested framework for modeling was not appropriate. Thus, we developed a random utility-based Multinomial Logit MNL model instead of a nested logit model. For choice models to avoid correlation between the error terms across the observations for the same individual, we defined the varying number of periods (PDS) column and utilized in model estimation as required in NLOGIT software. Six modes, including three multi-modes, were considered in the choice set: auto drive, auto passenger, walk, transit & walk, car drive & walk, and car passenger & walk. Transit and bike were excluded from the choice set as their proportions were very small. For cluster 7, the choice set includes five modes except transit & walk. The random utility theory of McFadden (1974) postulates that utilities can be expressed as the sum of measured attractiveness and a random term as follows:

$$U_{ir} = \beta_{ir} Z_{ir} + \epsilon_{ir} \tag{7}$$

Where, U_{ir} is the systematic utility, Z_{ir} is a vector of observed attributes of the alternative mode i and individual R and ϵ_{ir} is the random error. If Pr is the probability of an individual

R choosing mode i in our given choice set, then the MNL model can be written as follows (McFadden 1974):

$$\mathrm{Pr}_{\mathrm{ir}} = \frac{e^{U_{\mathrm{ir}}}}{\sum_{i=1}^{m} e^{U_{\mathrm{ir}}}} = \frac{e^{\beta_{\mathrm{ir}} Z_{\mathrm{ir}}}}{\sum_{i=1}^{m} e^{\beta_{\mathrm{ir}} Z_{\mathrm{ir}}}} \tag{8}$$

Where, $\beta_{ir}$is the corresponding parameter of vector Z_{ir}. The number of alternative choices is expressed as m in the above equation where $m = 6$ for this study.

If f_{ir} is the choice indicator (=1 if i is chosen by individual R and 0, otherwise) and Pr_{ir} is the probability that individual R chooses alternative i, then the log-likelihood function can be written as follows:

$$\mathrm{LL}(\beta) = \sum_i \sum_R d_{ir} \ln(\mathrm{Pr}_{ir}(\beta)) \tag{9}$$

Parameters, β_{ir} are calibrated by using the maximum-likelihood estimation. Parameter values were obtained by maximizing the likelihood function obtained by equating the first derivative of the likelihood function to zero.

7.5 Discussion of Results

7.5.1 Poisson Regression Models for Tour Frequency

Poisson regression model results of tour frequency for all five worker clusters are presented in Table 7.4. Among the tour attributes, trip chaining has significant negative association with tour frequency, as expected. This is because trip chaining allows individuals to combine more activities in a single tour, and reduces the need for additional tours. Our model also sheds light on the tour types associated with tour frequency. HWH and HPH simple tours are positively associated with tour frequency. This is because simple tours initiate the need for new tours. In contrast, complex tours are negatively associated with tour frequency. This is because complex tours combine or bundle more activities into each tour that an individual wishes to complete in the day, avoiding the need to return home and generate new tours.

Various personal characteristics, household characteristics, and land use characteristics are found to be significant in the model. Among the personal characteristics, gender has significant positive signs for cluster 1 and 5, which suggests that males generate more tours per day in comparison to female counterparts. Conversely, there is a negative coefficient for cluster 3. Age is found to be significant for cluster 1 and 5, with a positive effect. As expected, household size is statistically significant for cluster 1, 2, and 5, and has a positive sign for tour generation (Krizek 2003). Also, it is interesting to note that for clusters 2, 3, and 5 single workers are less likely to make more tours per day compared to married individuals. Among other household characteristics, the number of cars in the household has positive association with tour frequency for clusters 2 and 4, presumably because access to more cars offers greater opportunity and convenience to participate in more tours.

Some of the land use and built environment characteristics are also found to be significant in the models, and the results are consistent with the existing literature. The location of residence has significant influence on tour frequency. Individuals from urban core and suburban areas make more tours compared to those who live in urban fringe and rural areas, particularly for clusters 2, 3, and 4. HRM is a mid-sized city and individuals living in urban core and suburban areas have greater access to travel to activity destinations in 15-20 mins compared to those who live in rural and urban fringe areas. Nevertheless, land use mix has significant negative relationships with tour frequency for clusters 1, 2, and 3. Presumably, if the land use mix in the home neighborhood is higher, then it offers individuals greater accessibility to activity destinations, which will motivate them to trip chain rather than generating new tours (Chen and Akar 2017). Note, however, that land use mix has a positive association with tour frequency for cluster 5 (9 to 5 workers). Intersection density and retail density have both positive and negative impacts across the clusters. However, population density has a positive sign across all the clusters, and is significant for clusters 3 and 5. Presumably, areas of higher population density are more likely to have higher land use mix and greater accessibility to activity locations, which encourage individuals to trip chain rather than making new tours.

Table 7.4 Estimation results of Poisson regression models for tour frequency

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coef.	**Coef.**	**Coef.**	**Coef.**	**Coef.**
Constant	1.27	0.82*	1.04	0.73	-0.28
STOPN	-0.11*	-0.03**	-0.18*	-0.14*	-0.02
HWH	0.27*		0.62*	0.72*	
HPH		0.26*			0.27*
HWPH		-0.25*		-0.16*	-0.10**
HWGH		-0.21			-0.05
MDDURA		-0.04*			
MALE	0.10**		-0.08	0.01	0.05**
AGE	0.05**		0.02	-0.02	0.04**
SINGLE		-0.10**	-0.02*	0.08	-0.11*
HHSIZE	0.06**	0.02**	0.01	-0.01	0.08*
DRIVRLIC	-0.35		0.27	0.10	0.16
BUSPASS		-0.08			
CAR			0.12*	0.09**	-0.19
WLK	0.27**		0.16		0.19**
PAIDWRKR	-0.04		-0.12	-0.04	0.08
SLFEMP	0.26	-0.05			
DAYTYM		0.03			
HIGHINCOM	-0.05*	0.04	-0.07*	-0.04**	0.12
DWLOWN	-0.16		-0.16**	0.16**	-0.22
MR5YR	0.13		0.07	0.10	0.11**
LES1YR		-0.16			
HV1		0.04**		0.03*	
COMTYM	-0.09	-0.08	-0.01	0.02	0.02**
URBCR		0.02*	0.10**	0.20**	0.13
SUBRB	0.32**				0.06**
URBFR	0.35	-0.06**	-0.16*	-0.01*	
POPDEN			0.06**	0.05	0.06*
LNDUS	-0.22**	-0.02**	-0.05	-0.05	0.06**
INTERSC	0.05	-0.01	-0.08	-0.05	-0.05
SIDWLK	0.23	0.10**			-0.11**
RETAIL	0.19	-0.22	0.35	-0.12	-0.14
SERVC	-0.44	-0.17	0.10	-0.03	0.24**
Log likelihood (Constant only)	-258.66	-850.62	-420.03	-401.56	-495.88
Log likelihood (Full model)	-246.73	-811.00	-388.70	-373.58	-481.48
AIC	2.69	2.60	2.74	2.49	2.85

**Represents the significant parameters at 98% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

7.5.2 Ordered Probit Model for Trip Chaining

The trip chaining behavior of workers was studied by using ordered probit models. The parameter estimates of the models are presented in Table 7.5. Among the tour attributes, complex tours are positively associated with trip chaining, as expected. Among the complex tours, HWPH, HPGH, HWGH are more likely to have more trip segments than other tour types. Tour number per day is negatively associated with trip chaining, since trip chaining reduces the need for additional tours.

Personal attributes such as gender, age, and marital status were found to be significant in the final model. Male individuals of clusters 1 and 3 are negatively associated with trip chaining, whereas being male is positively associated with trip chaining for cluster 2 and 5. Presumably, men from cluster 2 and 5 make longer tours compared to their female counterparts, which is consistent with previous findings (Chen and Akar 2017). The existence of variation among the clusters indicates the latent heterogeneity among workers regarding their activity patterns and schedules, which would not be captured if they were modeled altogether. Age of the individual is found to be positively associated with trip chaining for cluster 1 (extended workers), but is not significant for other clusters. Household size is positively associated with trip chaining for workers of cluster 1, but negatively associated for cluster 3 (shorter work-day). Perhaps shorter work-day workers with family commitments have tight schedules and lack time for chaining.

Possession of a driver license is negatively associated with trip chaining, and significantly so for cluster 1. The ability to drive may provide greater convenience to conduct multiple tours instead of engaging in multiple trips in the same tour. For similar reasons, number of cars in the household has a negative association with trip chaining, significantly so for clusters 1 and 2. Mode of transportation for the tour (car or walk) is, however, significant only for cluster 2, which shows a positive association between trip chaining and the car mode. Among the work-related attributes, paid workers in cluster 1 (extended workers) have positive association with trip chaining, perhaps related to their time budget constraints. Consistently, individuals who work at night are more likely to trip chain compared to others, but significantly so only for cluster 1.

Residential location affects trip chaining, though significantly only for clusters 4 and 5. As expected, individuals living in the urban core and suburbs of Halifax are less likely to make complex tours, whereas those living in fringe areas are more likely to trip chain. Population density in the home neighborhood, residential density, and land use mix are positively associated with trip chaining, though seldom significantly. Tours undertaken from and to compact and densified residential areas tend to be short and complex due to the nearness of activity destinations, which is consistent with the findings of Kitamura and Susilo (2005). Similarly, intersection density has positive associations with trip chaining. Presumably, intersection density offers a better transportation network as well as travel options which may motivate people to make complex tours. Retail floor area ratio is found to be negatively associated with trip chaining, which is consistent with the findings of Chen and Akar (2017).

Table 7.5 Estimation results for ordered Probit model for trip chaining

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
	Coef.	**Coef.**	**Coef.**	**Coef.**	**Coef.**
Constant	-2.36	1.29	-0.09	-1.69	-3.12**
TOURN	-0.55*	-0.48*	-0.28*	-0.50*	-0.09**
HWPH	2.83*	2.40*	2.67*	2.19*	2.19*
HWGH	1.90*	2.35*	1.98*	2.30*	1.67*
HPGH	1.74*	1.97*	1.85*	1.30**	1.40*
MDDURA	0.01*	0.01*	0.01*	0.04*	0.01*
MALE	-0.01*	0.23	-0.37**	0.19	0.08**
AGE	0.04*	0.02		-0.01	0.01
SINGLE	1.78*	0.14	-0.03	0.27	-0.65**
HHSIZE	0.17**	-0.07	-0.10*	0.07	-0.10
COLLGLS			-0.39		
DRIVRLIC	-0.18**			-0.33	0.15
BUSPASS		-0.09			
CAR	0.42	0.29*	-0.15		0.04
WLK				-0.08	
PAIDWRKR	2.00*		-0.36	-0.20	-0.14
SLFEMP		-0.41			
DAYTYM		-0.30	0.02		
NYTTYM	3.88*				
HIGHINCOM	0.54	0.19			0.24
DWLOWN	0.26	-0.95**	-0.48		0.62
MR5YR	-1.37*			0.54	0.32
LES1YR		-1.14**	-0.99**		
HV1	-0.28**	-0.02*		-0.23	
COMTYM	0.01	-0.03**	0.01	-0.01	0.05
URBCR	-0.61		0.28		-1.14*
SUBRB	-0.12	-0.31	-0.16	-0.03	-0.96*
URBFR		0.09		1.14*	
POPDEN	0.26		0.08	0.12	0.07
LNDUS	0.01	0.07		0.20	0.06
RESID	0.05	0.03*	0.01	0.01	0.04*
INTERSC	0.01*	-0.03		0.35*	-0.03
SIDWLK		-0.20	-2.36	-0.46	
RETAIL	-0.13**		-0.90	-0.19*	0.66
SERVC	-1.79*	0.34	0.66		0.61
Threshold parameters index					
Mu(01)	0.07	0.01	0.10*	0.01	0.04
Mu(02)	0.27*	0.06*	0.33*	0.09*	0.11*
Mu(03)	0.36*	0.17*	0.43*	0.25*	0.44*
Mu(04)				0.27*	0.61*
Log likelihood (Constant only)	-195.35	-627.39	-285.04	-303.78	-400.99
Log likelihood (Full model)	-127.22	-439.89	-168.89	-216.44	-287.91
AIC	1.58	1.46	1.31	1.54	1.78

**Represents the significant parameters at 98% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

7.5.3 Tour Mode Choice Results of Multinomial Logit (MNL) Model

We conducted a robust descriptive analysis to create the choice set for empirical mode choice modeling. Initial analyses considered ten modes/multi-modes, which were car drive, car passenger, walk, transit, bike, car drive & walk, car passenger & walk, transit & walk, bike & walk, and car & transit. However, modal shares for transit, bike, bike & walk, and car & transit were less than 2%, and these were therefore excluded for the empirical modeling. The remaining six mode/multi-modes were modeled, utilizing a series of MNL models. In the case of worker cluster 4, the transit and walk mode was excluded due to its low mode-choice percentage.

Table 7.6, Table 7.7, Table 7.8, Table 7.9, Table 7.10 and Table 7.11 presents results from the mode choice models for all five clusters. Among the tour attributes, the sign of in-vehicle travel time is negative and significant as expected. The frequency of tours per day has positive association with the car drive, car passenger, transit & walk, car drive & walk, and walk mode choices. However, there exist both positive and negative associations across the clusters indicating heterogeneity among clusters. Interestingly, trip chaining (number of stops) has positive association with the car drive mode choice. This is consistent with earlier studies where it has been found that complex tours are more auto-oriented in comparison to simple tours (Strathman et al. 1994; Ye et al. 2007). On the other hand, trip chaining has negative association with the walk mode choice. This can be because walk mode is mostly limited to the comfortable walking distance and if the activity destinations are not nearby, then this mode is unlikely to be chosen. Type of tour as well as purpose of the tour are also found to be significant in the model for some clusters.

Several personal and household socio-demographic characteristics were found to be significant in the final models, though usually only for one or two of the clusters rather than for all of them. Male individuals are more likely to drive a car rather than to be a car passenger in a tour. Males are also more likely to choose walk or transit & walk mode for a home-based tour compared to their female counterparts. Age is significantly positive for the car-drive mode for cluster 4 (7 to 3 workers). It indicates that older aged workers are more likely to drive a car for their tour mode choice. Driving a car provides independence, convenience, and comfort for older aged travelers, and older workers are also more likely to afford this mode. However, age has both positive and negative association among clusters for car passenger, walk, transit & walk, and car drive & walk mode choices for cluster 4, indicating variation among clusters in terms of mode choice decisions. However, age does not have any significant effect on other clusters. Larger household size significantly increases the propensity of choosing car drive, car passenger, and car drive & walk mode for home-based tour for cluster 3 (shorter work-day). This suggests that co-commuting with other household members is particularly attractive for this group.

Individuals are more likely to drive to accomplish complex work-shopping and work-entertainment tours, probably because choosing the car mode increases the convenience of carrying shopping goods, and perhaps individuals travel longer for entertainment activities. Definitely, after working 4-8 hours, the car is more convenient to accomplish after-work entertainment and shopping activities. Similarly, the walk mode also shows a significant positive association with work-shopping tours for cluster 3 (8-to-4 group). On the other hand, transit & walk, and car drive & walk are found to have both positive and negative associations across the clusters for complex work-shopping and work-entertainment tours.

Among the other built environment and land use attributes, residential location is a significant factor for tour mode choices for some clusters. Individuals living in the urban core are less likely to choose car drive, car passenger, and drive & walk modes, whereas they are more likely to choose the walk mode. Clearly, individuals living in the core are within desirable walking distance to the Halifax downtown and a number of traditional shopping streets which offer a high density of activity locations. On the other hand, individuals living in suburban areas are somewhat more likely to choose the car drive or drive & walk modes for home-based tours, and less likely to choose the walk mode. Individuals in the suburbs need to travel longer distances for both work and non-work trips, and the walk option is often not viable.

Land use mix, population density, sidewalk density, retail density, service density, commercial density, and intersection density are found to be significant for tour mode choices, but typically only for one or two worker clusters. As expected, higher land use mix, intersection density, and sidewalk density have positive effects on walk mode choice and negative effects on the car drive mode choice. Greater land use mix offers a higher number of activity destinations in closer proximity, and greater sidewalk density offers perceived feeling of safety and security among travelers to choose the walk mode. Population density has negative coefficients with most mode choices other than transit and walk. Somewhat unexpectedly, both population and residential dwelling densities are negatively related to the walk choice for most worker clusters. Other land use variables show both positive and negative association across the clusters, indicating the need for a daily-activity pattern based clustering approach to model workers' activity-travel behavior.

Table 7.6 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Car drive

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
	Coef.	**Coef.**	**Coef.**	**Coef.**	**Coef.**
CONSTANT	-0.81	3.24*	0.49	-1.31	2.13
MIDDURA	-0.03*	-0.01	-0.02	-0.15	-0.01*
TOURN	0.64**	0.26	3.08	-0.17	
STOPN	0.80*	0.27*	0.30	0.02	0.09*
HWH	-2.75	-0.13	1.12		
HWEH			1.09**		
HWPH		0.90*			
MALE	-0.96	1.42*			0.40
AGE				0.11*	
HHSZE	0.08	-0.23	0.85*	-3.37	0.13
SINGLE		-0.26	-1.07		
HIGHINCOM		0.53**		-2.36**	
LOWINCOM	-1.13		-0.19		-1.44**
HV1		0.20	-0.43**		
DIPLOMA					-1.43
MARRID					1.54
PAIDWRKR			-1.07		
URBCR	-3.91	-0.55	-0.55	0.73	-1.30
SUBRB	2.03**		0.18*		
POPDEN			-0.21	-0.94	-0.55
LNDUS	-2.34**		-0.47	-1.71	0.24
RESID		-0.03		0.10	
COMM	0.05	-0.01	-0.01	0.01	-0.01
INTERSC	-0.28		0.12	1.30	0.52
RETAIL			-0.44*	5.20	7.37
SIDWLK	-1.37		0.14		-0.83
SERVC	2.72**	0.21		-1.23**	-2.68

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 7.7 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Car passenger

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coef.	**Coef.**	**Coef.**	**Coef.**	**Coef.**
CONSTANT	0.55	2.42	-1.02	1.22**	1.41
MIDDURA	-0.04	-0.07**	-0.05*	-0.14	-0.02**
TOURN	0.82**	0.10	2.63	1.07	
STOPN	-1.34*	-0.41*	0.23	-0.18	0.01
HWH	-4.27**	-0.34	0.96		
HWEH			1.44		
HWPH		-0.31			
MALE	-2.82**	-0.43			-1.46
AGE				-0.23**	
HHSZE	0.10	-0.22	1.09*	-3.82	-0.21
SINGLE		-0.25	-0.53		
HIGHINCOM		-0.08		-5.64*	
LOWINCOM	0.31		-0.53		-1.49
HV1		0.11	-0.90		
DIPLOMA					-1.31
MARRID					1.56
PAIDWRKR			-0.69		
URBCR	-0.37	0.04	-0.99	-3.91	-1.49**
SUBRB	-2.77		0.38		
POPDEN			-0.50**	-0.46	-0.09
RESID		-0.03		0.03	
COMM	-0.01	-0.10	-0.01	0.09	-0.07
LNDUS	2.03		-0.59	1.61	-0.75**
INTERSC	-0.41		0.21	-0.78	0.34
RETAIL			0.23	4.28*	-6.11
SIDWLK	-1.29		-0.04		-0.19
SERVC	4.19	0.01		-1.28	-2.88*

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 7.8 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Walk

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coef.	**Coef.**	**Coef.**	**Coef.**	**Coef.**
CONSTANT	12.06	3.50	-2.95	-2.87	-2.78
MIDDURA	-0.16	-0.01	-0.01	-0.10**	-0.03
TOURN	-1.35	0.18	3.16	-0.33	
STOPN	-0.13	-0.74*	-0.36	0.16	-2.57
HWH	-0.23	-0.08	2.89**		
HWEH			1.15		
HWPH		3.20*			
MALE	5.51**	1.78*			-0.09
AGE				-0.03	
HHSZE	-1.35	-0.35	1.05	-4.02	0.05
SINGLE		-1.15**	-0.43		
HIGHINCOM		-0.85**		-3.77	
LOWINCOM	0.44**		0.04		-2.21
HV1		0.23	-0.51		
DIPLOMA					-0.70
MARRID					0.49
PAIDWRKR			-2.08		
URBCR	0.46	0.98**	1.05	-0.15	0.50
SUBRB	-0.15		-0.12		
POPDEN			-0.11	-1.37	-0.77
LNDUS	3.99		0.91**	3.18	0.90**
INTERSC	1.50		0.13	1.14	-0.01
RETAIL			1.04**	5.97	5.26
SIDWLK	1.80**		4.68		0.47
SERVC	-1.65	2.43		-1.30	3.20**
RESID		-0.03*		0.03	
COMM	-7.53	-0.87	-0.01	0.06	-0.04

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 7.9 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Transit and walk

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Coef.	**Coef.**	**Coef.**	**Coef.**	**Coef.**
CONSTANT	-2.69	3.87**	-3.61		0.77
MIDDURA	-0.03	-0.01	-0.01		-0.02
TOURN	-0.34	-1.17	2.05**		
STOPN	-0.60**	-0.14	0.12		0.24**
HWH	-2.88	1.21**	2.78		
HWEH			-3.07*		
HWPH		1.03			
MALE	0.19	0.72			0.35
AGE	0.03	0.01	0.52		0.05
HHSZE		-0.04	-0.60		
SINGLE		0.75			
HIGHINCOM	0.38		-0.06		-0.43
LOWINCOM		-1.53*	0.18		
HV1					-1.17**
DIPLOMA					-0.97
MARRID			-1.61		
PAIDWRKR	-0.53	-0.32	-2.56		0.28
URBCR	-0.79		-2.25		
SUBRB			0.04		0.52
POPDEN	0.76		-0.19		0.04
LNDUS	1.17		0.49**		-0.12
RESID	0.06	0.05	-0.02		0.01**
COMM	-0.16	0.12**	-0.03		0.11*
INTERSC			0.75		3.96
RETAIL	-2.98**		-1.42		-0.51
SIDWLK	2.78	-0.05			-7.88
SERVC		-0.06			

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 7.10 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Car drive and walk

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
	Coef.	**Coef.**	**Coef.**	**Coef.**	**Coef.**
CONSTANT	-2.17	0.80	-2.85	-4.99	-2.37
MIDDURA	-0.04	-0.01**	-0.01**	-0.16	-0.01**
TOURN	1.62**	-0.30	2.83	1.43**	
STOPN	-0.16	0.04	0.76	0.25	0.53*
HWH	-1.66	0.38	2.99		
HWEH			1.29		
HWPH		0.41			
MALE	2.08	0.52			0.01
AGE				0.02	
HHSZE	-0.52	-0.32	0.96**	-4.42	0.26**
SINGLE		0.15	-2.22		
HIGHINCOM		0.95		-2.27	
LOWINCOM	-3.59*		-0.52		-1.50
HV1		0.27**	0.01		
DIPLOMA					-0.55
MARRID					0.44
PAIDWRKR			-1.30		
URBCR	-7.80*	-0.16	0.28	-5.94	0.51
SUBRB	-4.57**		0.28		
POPDEN			-1.05	-3.32**	-0.21
LNDUS	2.32		-0.90**	3.30	0.30
RESID		-0.74		0.02	
COMM	0.04	-0.09	-0.01	0.01	-0.03
INTERSC	-0.02		0.46**	1.92**	0.58
RETAIL			-0.32	2.13	1.07
SIDWLK	-2.91*		-0.40		-2.19*
SERVC	2.33	0.07		-0.15	-3.55

**Represents the significant parameters at 99% confidence level (P-value<0.01)*

***Represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 7.11 Tour mode choice parameter estimation results of Multinomial Logit (MNL) model: Fitness parameters

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
Log likelihood (Constant only)	-140.07	-769.23	-338.70	-458.86	-219.41
Log likelihood (Full model)	-236.86	-613.27	-277.54	-244.82	-292.82
AIC	2.24	2.17	2.51	2.93	2.05

7.6 Conclusions

This study contributes to the existing literature discussion on the connection between land use and travel behavior. Transportation planners and land use policy makers need to better understand how urban form and land use combine with socio-demographic characteristics to shape the travel behavior of individuals. Many empirical studies have focused on simple trips to study this association (Acker and Witlox 2011), but it is increasingly recognized that much travel takes place as trip chaining tours. Another under-recognized issue in the literature is that activity-based modeling of travel behavior is typically far too aggregated, and treats all out-of-home travelers as an undifferentiated group. Time-use studies show that daily activity schedules vary significantly among the total worker population, and workers contribute to both peak hour urban traffic as well as off-peak travel. Therefore, this study offers an approach to tour modeling for worker population clusters, including prediction of tour formation, tour type, tour frequency, trip chaining, and tour mode choice.

Five worker clusters were drawn from the Halifax Space-Time Activity (STAR) dataset using a daily activity pattern recognition method: Cluster 1: extended workers, Cluster 2: 8 to 4 workers, Cluster 3: shorter work-day workers, Cluster 4: 7 to 3 workers, Cluster 5: 9 to 5 workers. Only 25% of the home-based work tours made by workers are simple home-work-home tours, indicating that workers are more likely to make complex work tours. Across all the clusters there exists significant variation in terms of tour type, tour mode choice, and tour frequency. However, extended workers and 8 to 4 workers make nearly 40% of total tours as simple one stop tours, whereas other workers make 50% of total tours as simple one stop tours. It is evident that workers have a strong tendency to add non-work trips to the after-work commute, and particularly trips for shopping, organizational, and entertainment activities. Given that the socio-demographic characteristics and daily activity pattern of the clusters are explicitly identified, the effects of travel demand and congestion management policies can be magnified or muted for each cluster. Moreover, since the complexity of work commute chains varies significantly among population groups, the distributional effects of alternative traffic and congestion management policies may warrant closer analyses.

Tour frequency of the five worker clusters was modeled by the Poisson regression model. Consistent with earlier studies, socio-demographic variables and household characteristics were found to be significant in the model. For extended workers, with the increase in household size, the propensity to trip chain increases, which suggests that larger households might require increasingly complex work commutes that can resist peak spreading programs (e.g., flex-time and staggered work hours). Additionally, urban form attributes were also found significant in the final model. As expected, individuals from urban core and suburban areas make more tours per day compared to those from urban fringe and rural areas. It is also interesting to note that land use mix impacts negatively on tour frequency. This suggest that higher land use mix in the home neighborhood offers individuals greater accessibility to activity destinations, which motivate individuals to trip chain rather than generate new tours.

An ordered probit model was utilized to study the trip chaining behavior of the five worker groups. Household size is positively associated with trip chaining for extended workers and 7 to 3 workers, whereas for shorter work-day workers the association is positive. Possession of a driver license and number of cars in the household both have negative association with trip chaining, indicating that availability of the car mode motivates more simple tours. Among the complex tour types, HWPH, HPGH, HWGH were the most frequent sequences. These findings suggest that policies which seek to increase ride-sharing or vehicle occupancy may provide better congestion relief, through reduction of vehicle trips. In addition, promotion of higher occupancy levels would encourage rescheduling of non-work trips that are currently linked to the work commute.

Tour mode choices were estimated using MNL models. The model results indicate that individuals living in the urban core are less likely to choose the car mode and more likely to choose the walk mode. In addition, better street connectivity, higher retail density, greater land use mix, and higher sidewalk density all motivate individuals to choose the walk mode compared to others. This warrants policies to densify suburban neighborhoods with better street connectivity and higher land use mixes which can increase the usage of active transportation. Complex tours are mostly auto-oriented. To counter the attraction of the auto mode, policy makers should improve existing transit routing, timing, and reliability of service, but the more complex the tour, the less likely that these improvements would be effective. These findings have implications for land use planning policy, and support current policies promoting more efficient public transport, greater land use mix, and densification. Our findings also justify the current ongoing 'integrated mobility plan' for Halifax.

This study sheds light on the role of land use and urban form variables in shaping travel patterns and tour complexity for a mid-sized Canadian city, but there are some limitations to the current study. This study can be improved and extended by analyzing the household interaction, accompaniment pattern, and joint travel of worker clusters. Also, a comparative study between worker and non-worker clusters (Daisy 2018a) would be interesting. Furthermore, we aim to employ the STAR GPS data to accurately locate activity destination choices, and incorporate them in the proposed modeling framework. In this study, we analyzed tour mode choice per tour instead of using a whole-day consideration. For example, an individual may make an initial tour in the day through auto drive, and then a second tour using both auto drive and walk modes. An alternative approach for further work is to utilize the daily percentage frequency of each mode to represent individuals' mode choices in nested logit or cross-nested logit models. Another extension of this study would be to investigate model estimation comparison between time-use activity pattern based population groups and a latent class model.

The results presented in this study clearly reveal that analysis of activity patterns for worker clusters can be a useful component in modeling and predicting complex travel behavior. This work forms part of a series by the authors, and the modeling framework is expected to be incorporated into the Scheduler for Activity, Location and Travel (SALT) model for Halifax.

Chapter 8 Understanding and Modeling the Activity-Travel Behavior of University Commuters at a Large Canadian University[6]

8.1 Introduction

Enrolment in post-secondary education is increasing rapidly for various reasons, such as population growth, growing numbers of educational institutes, and growing desire for higher education. The increase in university populations leads to an increase in the overall number of university commuters. Consequently, university campus authorities need to estimate and manage high volumes of automobile traffic and higher demand for on-campus parking. In addition, empirical data on travel demand is required to improve university travel demand management strategies, in order to reduce on-campus traffic and establish sustainable transportation on-campus (Toor and Havlick 2004; Black et al. 1999; Balsas 2003). However, despite the substantial impact of university populations on regional travel demand models, only a few studies to date have been carried out to understand the activity-travel behavior of university populations (Axhausen and Garling 1992; Bowman and Ben-Akiva 2001; Bhat et al. 1894). Notably, understanding the activity-travel behavior of university populations is more vital for cities, such as Halifax City, where one or more large universities are major trip generators of travel demand.

Metropolitan planning organizations and Transportation engineers recommended that university populations should be considered as a sub-population with special travel behavioral characteristics in regional travel demand models. Recent travel demand models

[6] A version of this chapter has been published:
Daisy, N. S., Hafezi, M. H., L. Liu., and Millward, H. Understanding and modeling the activity-travel behavior of university commuters at a large Canadian university. Journal of Urban Planning and Development. Vol. 144(2), 2018., DOI:10.1061/(ASCE)UP.1943-5444.0000442.

are more focused in disaggregated modeling such as activity-based models. This study aims to model the activity-travel behavior of students, staff, and faculty at a large Canadian university using a disaggregated modeling approach.

Though the importance of studying the activity-travel behavior of university populations has been increasingly recognized in the past decade, mostly in American universities for finer regional travel demand models, there has not been any travel diary survey conducted for a university population in Canada. The current study is also unique in analyzing the results of the first university-based travel diary survey across Canada. According to Statistics Canada (2015), post-secondary enrolment increased 1.2% in the 2013/14 academic year and there are now more than two million students enrolled in post-secondary education, which is more than 5% of the total population of Canada. The student population in the province of Nova Scotia is higher than the national average, at 5.90%, which suggests the need for greater attention to understand their activity-travel behavior. Unfortunately, the latest General Social Survey of Canada (GSS) conducted in 2010 captured samples of only 538 students with student status and 293 students with full or part-time employment across Canada, and the numbers for Nova Scotia are only 26 and 16 respectively. These small samples are inadequate for many purposes, and may indicate under-representation of university students in the Regional Travel Surveys. This may result from using landline or mailed tools to reach survey respondents, since university students are more frequent users of mobile phone (Wang et al. 2012). Hence, the current study contributes in the current literature by providing a rich data set on activity-travel behavior of large Canadian university population Dalhousie University that can be used in regional travel demand models of Nova Scotia. Hence, the Environmentally Aware Travel

Diary Survey (EnACT) was conducted among Dalhousie University commuters in Spring 2016. Dalhousie University is the largest university in the Maritime Provinces of Canada and it is one of the main trip generators in the province of Nova Scotia.

To date, travel demands of university students have not been properly modeled in regional travel demand models in Nova Scotia. Most often, travel demand models consider trip rates with socio-economic information (e.g., household size, household income, car ownership, etc.) generated for the general population, which do not appropriately reflect the travel behavior of the university population (Wang et al. 2012). Broadly, university campuses comprise mixed land uses with vibrant built environment characteristics, and offer more opportunities to participate in different types of activities within accessible distance. These unique features distinguish university travel behavior from the general population, and require special consideration for travel demand management strategies. The data of this study were collected by the authors through a unique online-based travel diary survey among all four Dalhousie University campuses in Spring 2016. The insights from this study will inform travel demand management strategies, and trip generation modeling for university-oriented cities like Halifax, Nova Scotia, where university populations impact traffic volumes and travel patterns significantly. Results obtained from this study are expected to be incorporated within the activity-based travel demand model for Halifax Regional Municipality (HRM), Nova Scotia, which is currently under development.

8.2 Literature Review

University communities can be assumed to be an under-represented group in travel surveys even though they represent a significant portion of total population. Previous studies confirm under-representation of university communities or student population in national travel surveys (Wang et al. 2012; Volosin et al. 2014). Still there are few peer reviewed articles on university populations, and those discussing their inclusion in regional travel models are even fewer. There are, however, several studies that have been focused on university-related commuting mode choices (Balsas 2003; Rodriguez and Joo 2004; Akar et al. 2012). For instance, Rodriguez and Joo (2004) studied the relationship between mode choices and spatial characteristics among University of North Carolina-Chapel Hill commuters. Akar et al. (2012) predicted mode choices for home to campus trips by Ohio State University (OSU) populations based on a travel survey conducted in 2011. The results of the study revealed that it is more likely that students will use the biking and transit modes for home to campus trips compared to faculty members and staff. In another study, Kamruzzaman et al. (2011) utilized trip diary information to investigate out-of-home travel and activities by university students. Another study by Balsas (2003) focused on the development of on-campus sustainable transport systems through policy actions at eight American universities. Interestingly, this study examined policy actions aimed at non-motorized transport, and showed that such actions produced significant changes in modes of travel for the commute to school.

A school-day activity-travel diary survey was conducted at North Carolina State University to explore university travel behavior characteristics (Eom 2007). Results showed that undergraduate students and graduate students who are living on campus participate more in out-of-home activities than students who are living off-campus. Similarly, a one-week travel diary survey conducted at a Thailand university examined the activity-travel patterns of university students at a rural university (Limanond et al. 2011). Another one-week travel diary survey conducted at the University of Western Australia was utilized to study the commuting patterns of students and staff, and the study examined the on-campus zones of car driving, walking, and biking (Shannon 2006). Virginia Department of Transportation sponsored two sets of web-based surveys in 2009 and 2010 to study university students' travel behavior. In 2009, the travel diary survey was conducted for four universities (i.e. University of Virginia, Virginia Commonwealth University, Old Dominion University, and Virginia Polytechnic Institute and State University). This study identified differences between student populations and general populations by comparing the collected data with the National Household Travel Survey (NHTS). The results showed that university students are a relatively lower income segment of the population and they have atypical travel behaviors (Khattak 2011). The study found significant differences between urban residents and students, where students from urban campuses were found to participate in more out-of-home activities than the general population.

Despite the presence of the above few studies on university populations, there exists no activity-travel behavior study on a university population in the context of Canadian universities. Thus, this study aims to fill this gap in knowledge by exploring the activity-travel characteristics of population groups at Dalhousie University, the largest university in the Maritime Provinces of Canada. It is expected that results of the empirical models of this study can be utilized to quantify trip generation for university populations throughout Canada, and can be incorporated into the regional travel demand model for Nova Scotia.

8.3 Modeling Approach

A count-data modeling approach is appropriate for modeling the trip frequency since it shows the characteristics of a discrete distribution which is non-negative and has integer values (W. Greene, “Accounting for excess zeros and sample selection in Poisson and negative binomial regression models,” Working Paper EC-94-10, Stern School of Business, New York University, New York; Jang 2005). The Poisson model has been utilized in several studies for trip frequency modeling as trips occur randomly and independently over time (Bhat et al. 1998; Misra 1999; Ma and Goulias 1999; Wallace et al. 1999; Jang 2005; Habib and Daisy 2013). However, the Poisson model postulates the equi-dispersion theorem where the conditional variance of the dependent variable is equal to the conditional mean. Mullahy (1997) claimed that if there is over-dispersion in the data, then a negative binomial model should be utilized instead of the Poisson model. A number of studies have utilized Zero-Inflated Negative Binomial (ZINB) models for modeling trip frequencies when there is over-dispersion and excess zeros in the dataset (for instance, Jang 2005; Khattak et al. 2011; Xu et al. 2015). Thus, this study also employs a ZINB model for empirical analysis as there are excess zeros and over dispersion in the data. The model is utilized to report and analyze trip frequencies of automobile, Active Transportation (AT), and transit trips. In this study, Active Transportation (AT) includes both walk and bike trips.

For frequency analysis, because the dependent variables are non-negative and integers, alternative modeling approaches are the negative binomial model, zero-inflated model, and Poisson regression model. According to the Poisson regression model, the probability $P(y_k)$ of having y_k number of trips by a given mode for observation k can be written as:

$$P(y_k) = \frac{\exp(-\lambda_k)\lambda_k^{y_k}}{y_k!} \tag{1}$$

Where

λ_k is the Poisson parameter for observation k,

y_k is equal to the expected number of trips by the mode.

The mean and variance for the number of trips is assumed to be equal (i.e. $E[y_k] = Var[y_k]$) in the Poisson model. If the expected mean number of trips is not equal to the variance, then the data are assumed to be under-dispersed or over-dispersed. In such situations, parameter estimation with the Poisson regression model will be incorrect (Lambert, 1992; W. Greene, "Accounting for excess zeros and sample selection in Poisson and negative binomial regression models," Working Paper EC-94-10, Stern School of Business, New York University, New York). To address this over-dispersion or under-dispersion the Negative Binomial (NB) form can be utilized, since it adds a random error term in the parameter estimation. Hence, the form of λ_k calculation for NB can be written as follows:

$$\lambda_k = \exp(\beta X_k + \varepsilon_k) \tag{2}$$

However, in the case of trip making by a specific mode, there exist zero and non-zero values among the observations. This dual-state process requires dual-state count models, such as the Zero Inflated Negative Binomial model (ZINB) or the Zero-Inflated Poisson (ZIP), which explicitly separate the true zero-state process and count data process and allow explanatory variables to impact both occurrences. Thus, $1 - p_k$ is estimated as the probability that an individual actually makes zero automobile or AT or transit trips and follows an NB distribution. Given this,

$$y_k \begin{cases} = 0 \text{ with probability } p_k + (1 - p_k)[\frac{\theta}{\theta+\lambda_k}]^\theta \\ = K \text{ with probability } (1 - p_k)[\frac{\Gamma(\theta+K)u_k^\theta(1-u_k)^K}{\Gamma(\theta)K!}] \end{cases} \quad (3)$$

Where

K is the number of trips made by an individual k by automobile or AT or transit,

$\theta = 1/\alpha$ (with α being the dispersion parameter),

$u_k = \frac{\theta}{\theta+\lambda_k}$ with λ_k being the mean,

$\Gamma(.)$ is a gamma function.

In the NB model, the Poisson assumption requires the mean to be equal to the variance by letting $Var[y_k] = E[y_k]\{1 + \alpha E[y_k]\}$ that is relaxed by the parameter estimation of α. And for ZINB the dispersion $Var[y_k]$ follows the following equation:

$$Var[y_k] = U[y_k]\{1 + \frac{p_k}{1-p_k} U[y_k]\} \quad (4)$$

Thus for the ZINB model, $\frac{p_k}{1-p_k}$ can be interpreted as α. Regarding the zero-inflated model application, there arises the problem of distinguishing whether the NB or Poisson distribution is the source of over-dispersion. The probability density function (PDF) for the random variable y_k :

$$P(y_k) = (1 - p_k)\left[\frac{\Gamma(\theta+k)u_k^{\theta}(1-u_k)^K}{\Gamma(\theta)K!}\right] + Z_k p_k \tag{5}$$

Where

$$y_k \begin{cases} = 0 \text{ when } Z_k = 1 \\ \neq 0 \text{ when } Z_k = 0 \end{cases}$$

The application of the indicator variable Z_k is utilized for the maximization of the log-likelihood function. Where the log likelihood function can be written as follows:

$$\sum_k \log(p(y_k)) \tag{6}$$

However, for identifying better fitted distribution, Vuong (1989) proposed the Vuong Test which provides proper power using count-data (W. Greene, "Accounting for excess zeros and sample selection in Poisson and negative binomial regression models," Working Paper EC-94-10, Stern School of Business, New York University, New York). The Vuong statistic is calculated as follows:

$$G = \frac{\bar{q}\sqrt{R}}{S_q} \tag{7}$$

Where

$\bar{q}$ is the mean with $q = \log[\frac{f_1(.)}{f_2(.)}]$

$f_1(.)$ indicates the distribution function of the ZINB distribution,

$f_2(.)$ indicates the distribution function of the parent NB distribution,

S_q represents the standard deviation,

R represents the sample size.

The ZINB model better represents the data where the value of G is greater than 1.96 (the 95% confidence level in the t-test). On the other hand, a value of G less than -1.96 favors the parent NB. The next section continues with a description of the survey and descriptive statistics.

8.4 Travel Diary Survey

8.4.1 Dalhousie Environmentally Aware Travel Diary Survey (EnACT) Survey

The Transportation and Environmental Simulation Studies (TESS) group at Dalhousie University conducted an online web-based one-day travel diary survey, Environmentally Aware Travel Diary Survey (EnACT), in Spring 2016 across the entire population of Dalhousie University commuters. The survey covered all Dalhousie commuters, comprising undergraduate students, graduate students, faculty members, and staff from all four campuses. After designing the survey, a pilot study was conducted to understand the fill-out timing, understanding of questions, and user friendliness. Following receipt of feedback and comments the survey was modified. Through the co-operation of the university administration, a survey link was circulated to all students, faculty members, and staff. The EnACT survey included six sections: (I) household information, (II) individual information, (III) environmental attitudes and behavior, (IV) attitudes toward transportation, (V) use of information and communications technology (ICT), and (VI) 24-hour travel log.

The typology of the survey was consistent with the General Social Survey of Canada (GSS) (Statistics Canada 2011) and the Halifax Space Time Activity Research survey (STAR) (Millward and Spinney 2011), for comparison purposes, and so that findings and results from this survey can be utilized for disaggregated regional travel demand modeling. Following survey data collection, and after rigorous cleaning, error-checking and geo-coding, the survey yielded a sample of 346 fully completed 24-hour travel logs with demographic and socio-economic information for the city campuses. The sample demographic characteristics were compared with those of the total university population (using the Dalhousie Commuter Survey and the Dalhousie Analytics Data), to investigate the representativeness of the sample. It was found that the EnACT sample is roughly representative with respect to age, employment status, gender, travel mode and commute time in comparison to those of the Dalhousie University population. However, female individuals are slightly over-represented and higher age individuals are slightly under-represented. In the current study, we utilized the sample directly obtained from the survey without sample weighting or population synthesis. Future studies will include matching with the entire university population and expanding the sample size using population synthesis technique. A comprehensive explanatory analysis of all six sections of EnACT survey can be found in Liu et al. (2016).

8.4.2 Summary Statistics of Variables

Only the three Dalhousie University campuses located in Halifax Regional Municipality (HRM) were considered for this study, since one campus (Truro) has different campus and locational characteristics. The sample consists of 37.28% undergraduate students, 36.42% graduate students, 6.94% faculty members and 19.36 % staff members. Summary statistics of the response variables are shown in Table 8.1. It is encouraging that the mean number of AT trips per day for the sample is more than 1 trip per person, whereas the automobile mode has a mean of less than 1 trip. The transit mode, perhaps surprisingly, is used much less than either AT or automobile. The high standard deviations for all three modes, along with zero trips per day as the minimum, explains the over- or under-dispersion of the data. The mean number of automobiles per household is 1.02 and the mean household size is 2.88.

Table 8.1 Summary statistics of variables used in the empirical model

Variable	Description	Mean	Stan. Dev.	Min	Max
	Dependent Variable				
At Trips	Number AT trips per day	1.2	1.49	0	7
Automobile Trips	Number automobile trips per day	0.87	1.28	0	6
Transit Trips	Number transit trips per day	0.56	0.94	0	5
	Independent Variable				
AGE	Age of the respondent	30.96	12.08	18	72
FEMALE	Gender of respondent (dummy, 1 if the respondent is a female, 0 otherwise)	62.42%			
LESS15K	Annual personal income (dummy, 1 if the income is less than $15,000 per year, 0 otherwise)	38.44%			
BT50T75K	Annual personal income (dummy, 1 if the income is between $50,000 to $75,000 per year, 0 otherwise)	9.54%			
Q6HHSIZE	Household size	2.88	1.42	1	7
STAFF	Staff (dummy, 1 if the respondent belongs to staff segment of the population, 0 otherwise)	19.36%			
FACULTY	Faculty (dummy, 1 if the respondent belongs to Faculty segment of the population, 0 otherwise)	6.94%			
GRAD	Graduate (dummy, 1 if the respondent belongs to Graduate segment of the population, 0 otherwise)	36.42%			
UNGRAD	Undergraduate student (dummy, 1 if the respondent belongs to undergraduate segment of the population, 0 otherwise)	37.28%			
HIGHFLEX	Highly flexible (dummy, 1 if the school/work schedule of the respondent is flexible, 0 otherwise)	21.68%			
ABV10YR	Above 10 years of living (dummy, 1 if the respondent is living in the current house for more than 10 years, 0 otherwise)	16.47%			
Q9HHCAR	Number of automobiles at household	1.02	0.95	0.00	6.00
DRVLNS	Driving license (dummy, 1 if the respondent has a driver license, 0 otherwise)	88.73%			
HOMOWN	Home ownership (dummy, 1 if the respondent owns a house, 0 otherwise)	33.24%			
DISTHW	Total home to campus distance	7.06	14.22	0.18	114.74

8.5 Characteristics of Daily Travel Behavior

This study aimed to investigate similarities and dissimilarities of activity-travel characteristics between different segments of the university sample. To that end, Table 8.2 shows the travel mode shares of undergraduate students, graduate students, faculty members, and staff by trip purpose. The highest mode share percentage for the undergraduate and graduate student segments is for the walk mode, whereas for faculty and staff segments the most used mode is auto drive. For work and school related trips, the most used modes for graduate students are walk, automobile (drive), and transit, whereas for undergraduate students they are walk and auto drive. For faculty members, the most used modes for work and school are automobile (drive) and walk. In contrast, staff members choose automobile (drive), automobile (passenger), and walk modes more than other modes.

For shopping activities, the most used mode is automobile driver for graduate students, undergraduate students, and staff, whereas faculty members preferred the walk mode for shopping. This may be explained by the neighborhood built environments and the presence/absence of shopping opportunities close to home. However, each sample segment makes significant numbers of walk and transit trips, which is consistent with findings from other university based studies (Miller 2012). It is interesting to note that, for entertainment related activities and sports and hobbies related activities, all segments prefer automobile (drive). Presumably, this occurs due to convenience and comfort, and the preference to travel in company with friends or team mates.

Table 8.2 Mode share (%) by trip purpose for various university market segments

		Automobile (driver)	Automobile (passenger)	Walk	Transit (Bus/Ferry)	Bicycle	Total
Graduate Students	Household related works	0.0	25.0	25.0	25.0	25.0	100.0
	Paid work	26.3	0.0	38.6	26.3	8.8	100.0
	Entertainment related activities	45.0	5.0	20.0	15.0	15.0	100.0
	Organizational, voluntary and religious activity	0.0	25.0	25.0	12.5	37.5	100.0
	Personal care related activities	14.7	0.0	47.1	17.6	20.6	100.0
	School & education related activities	17.2	4.1	39.1	32.0	7.7	100.0
	Shopping activities	44.2	4.7	30.2	18.6	2.3	100.0
	Care giving activities	100.0	0.0	0.0	0.0	0.0	100.0
	Sports and hobbies	35.7	0.0	21.4	28.6	14.3	100.0
	All activities	23.7	3.7	36.3	26.3	10.0	100.0
Undergraduate Students	Household related works	50.0	0.0	33.3	16.7	0.0	100.0
	Paid work	26.8	0.0	39.3	23.2	10.7	100.0
	Entertainment related activities	41.7	0.0	33.3	25.0	0.0	100.0
	Organizational, voluntary and religious activity	0.0	0.0	85.7	14.3	0.0	100.0
	Personal care related activities	18.2	9.1	68.2	0.0	4.5	100.0
	School & education related activities	24.3	4.1	49.1	18.9	3.6	100.0
	Shopping activities	35.3	20.6	35.3	8.8	0.0	100.0
	Sports and hobbies	42.9	28.6	28.6	0.0	0.0	100.0
	All activities	27.1	5.5	46.2	17.2	4.0	100.0
Faculty	Paid work	35.3	3.9	29.4	15.7	15.7	100.0
	Entertainment related activities	100.0	0.0	0.0	0.0	0.0	100.0
	Organizational, voluntary and religious activity	50.0	0.0	50.0	0.0	0.0	100.0
	Personal care related activities	66.7	0.0	33.3	0.0	0.0	100.0
	School & education related activities	100.0	0.0	0.0	0.0	0.0	100.0
	Shopping activities	0.0	0.0	100.0	0.0	0.0	100.0
	Sports and hobbies	36.4	0.0	18.2	27.3	18.2	100.0
	All activities	40.8	2.6	28.9	14.5	13.2	100.0
Staff	Household related works	33.3	33.3	33.3	0.0	0.0	100.0
	Paid work	30.8	5.6	30.1	24.5	9.1	100.0
	Entertainment related activities	14.3	42.9	14.3	0.0	28.6	100.0
	Media & Communication	0.0	0.0	100.0	0.0	0.0	100.0
	Organizational, voluntary and religious activity	71.4	0.0	0.0	0.0	28.6	100.0
	Personal care related activities	85.7	0.0	14.3	0.0	0.0	100.0
	School & education related activities	0.0	50.0	50.0	0.0	0.0	100.0
	Shopping activities	62.5	0.0	18.8	6.3	12.5	100.0
	Care giving activities	100.0	0.0	0.0	0.0	0.0	100.0
	Sports and hobbies	46.7	0.0	33.3	13.3	6.7	100.0
	All activities	37.4	6.8	27.7	18.4	9.7	100.0

Table 8.3 presents the mean trip travel times, trip distances, and trip rates by trip purpose for each university sample segment. Among all segments of the Dalhousie community, the mean travel time for a trip is approximately 24 minutes, which is consistent with other studies (Volosin et al. 2014). However, there is variation among different segments in terms of mean trip durations: the faculty mean trip travel time is slightly shorter than others, whereas the staff mean trip travel time is highest among all segments. In contrast, the mean trip travel time for paid work is longer for both faculty members and staff compared to student segments. Overall, staff segment trip duration is higher for most of the activities compared to other sample segments. Furthermore, trip travel time for graduate students is longer than for undergraduate students.

Trip rates are estimated by dividing the total number of trips by the number of respondents in respective segments, which implies that zero-trip makers are being considered as well. From the mean total number of trips per day, it is found that undergraduate students make the lowest number of trips compared to other population segments. In contrast, faculty members have the highest mean trip rate, at 2.92 trips. For student population segments, the trip rates for school and education related purposes is more than one, whereas for faculty and staff segments, the paid work trip rates are more than one.

For trip distance, the overall mean trip distance is highest for faculty members (7.51 km), whereas graduate students had the shortest mean travel distance (3.84 km). On mean, faculty members and staff travel 13.81 km and 16.55 km respectively for paid work, whereas undergraduate students and graduate students travel only 4.32 km and 7.81 km respectively for work/school trips. Clearly, students live nearer to the campus compared to faculty members and staff. Even though faculty members travel longer distances and have higher trip rates their travel time is lower compared to other groups. That suggests that faculty members use faster modes of travel to reach the activity locations.

Table 8.3 Mean trip travel times, distances, and trip rates for various activity types

		Undergraduate Students	Graduate Students	Faculty	Staff
Trip Rate per Day	Household related works	0.05	0.03	0	0.04
	Paid work	0.44	0.47	1.96	2.01
	Entertainment related activities	0.19	0.17	0.08	0.1
	Media & Communication	0	0	0	0.01
	Organizational, voluntary and religious activity	0.06	0.07	0.08	0.1
	Personal care related activities	0.17	0.28	0.23	0.1
	School & education related activities	1.34	1.4	0.08	0.06
	Shopping activities	0.27	0.36	0.08	0.23
	Care giving activities	0	0.01	0	0.04
	Sports and hobbies	0.06	0.12	0.42	0.21
	All Trips	2.58	2.89	2.92	2.9
Trips duration (mins/day)	Household related works	30.83	17.5	0	36.67
	Paid work	21.96	19.47	25.49	30.77
	Entertainment related activities	23.75	26.5	27.5	45
	Media & Communication	0	0	0	30
	Organizational, voluntary and religious activity	37.14	18.13	15	22.86
	Personal care related activities	20.45	25.59	22.5	24.29
	School & education related activities	20.83	23.55	20	28.75
	Shopping activities	23.24	24.3	10	21.25
	Care giving activities	0	10	0	18.33
	Sports and hobbies	27.14	32.14	26.36	24
	All Trips	25.67	21.91	20.98	28.19
Trip distance (km)	Household related works	4.08	2.17	0	7.67
	Paid work	5.4	4.2	13.81	16.55
	Entertainment related activities	10.4	4.96	8.52	6.27
	Media & Communication	0	0	0	2.4
	Organizational, voluntary and religious activity	2.37	1.86	11.57	2.08
	Personal care related activities	2.83	2.55	7.03	6.44
	School & education related activities	4.32	7.81	4.4	4.58
	Shopping activities	4.47	4.37	4.75	3.21
	Care giving activities	0	1.36	0	2.03
	Sports and hobbies	5.68	4.72	2.49	6.19
	All trips	4.94	3.84	7.51	5.74

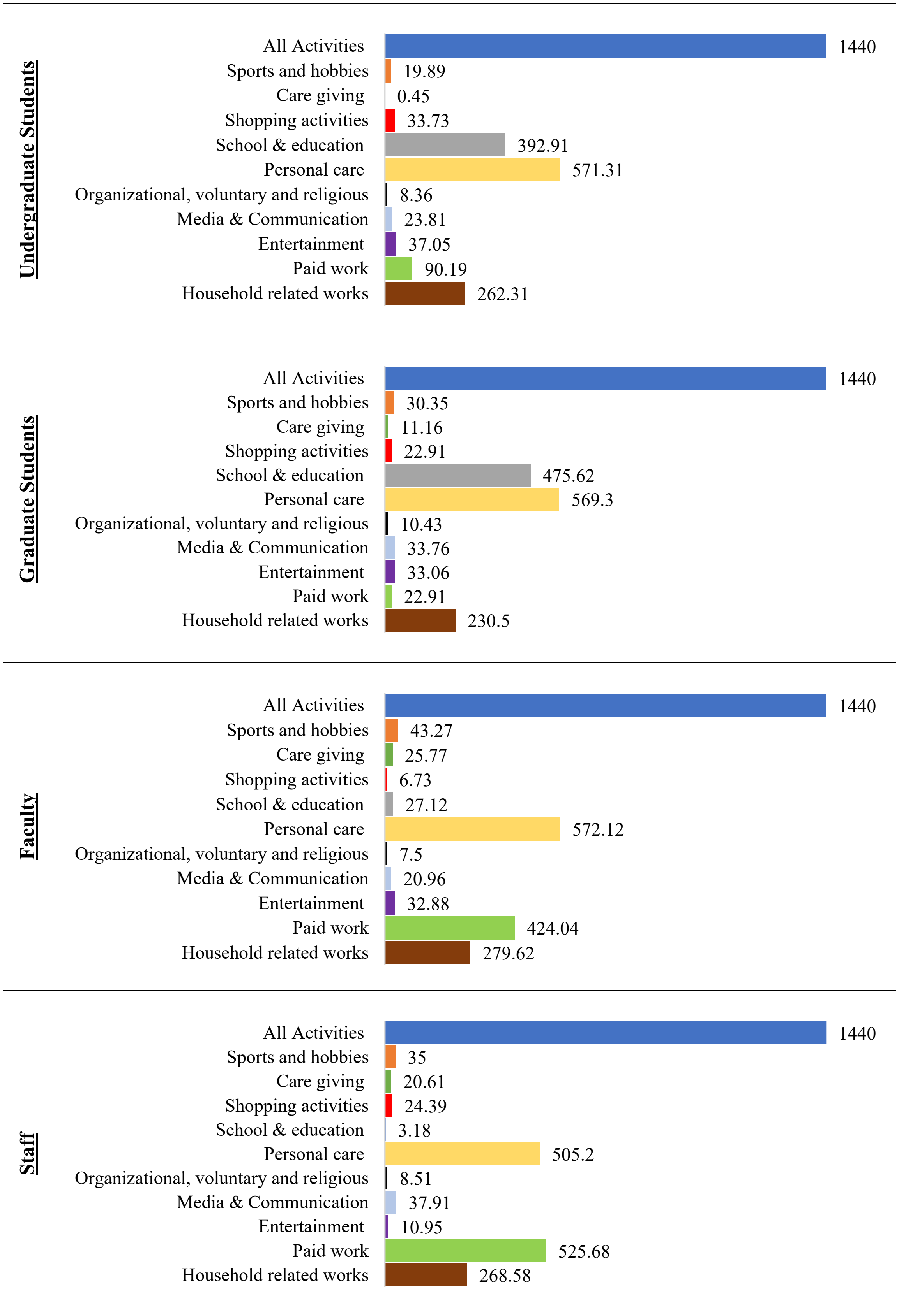

Figure 8.1 Activity participation segment (minutes)

Figure 8.1 presents the activity participation behavior of Dalhousie University community sample segments. Across all the segments the highest activity duration per day is allocated for personal care related activities. Much of this is accounted for by sleep, meals, and grooming. As expected, the second longest activity duration for undergraduate students and graduate students is school and education related activities, whereas paid work is the second longest activity duration for faculty members and staff. Compared to all the segments, staff spend the shortest time on entertainment related activities, whereas undergraduate students spend the shortest time on sports and hobbies related activities.

8.6 Model Results

Parameter estimation results of the maximum likelihood estimation of Poisson, NB and ZINB models described in the modeling approach section of this study are presented in Table 8.4, Table 8.5 and Table 8.6. For all the modes, including daily automobile trips, AT trips, and transit trips, the ZINB specification best describes the underlying two-state process. The Poisson and NB models provide very similar coefficients for both daily AT and automobile trip models. The over-dispersion parameter α in the NB model is greater than 1.0 for all the models, indicating that there is significant over-dispersion in the data. The ZINB estimates for daily auto trips, AT trips, and transit trips shown in Table 8.4, Table 8.5 and Table 8.6 have Vuong statistics much greater than 1.96; the values of 3.56, 5.11 and 3.1 for automobile, AT, and transit respectively indicate that there is a higher probability that a two-state process is present. Thus, ZINB is the best estimator of the zero-trip and non-zero trip states, with plausible signs.

From the ZINB model for automobile trips presented in Table 8.4, it is found that the respondent's gender is a significant factor. The sign for female gender is negative, which shows that male respondents are more likely to employ automobile trips than female students. If the respondent's duration of living in current residence is less than one year, then it is more likely that they would make fewer automobile trips compared to those who are staying longer. As expected, more automobiles in the respondent's household is related to a higher probability that the respondent will make auto trips. Similarly, having a driver license has a positive coefficient, showing that the respondent is more likely to make auto trips per day. It is also found that household size has a positive sign, indicating that respondents in larger households will make more automobile trips compared to those from smaller households. Similarly, individuals with own home are likely to undertake more automobile trips, in comparison to others.

Table 8.4 Parameter estimation results of trip frequency models for Dalhousie university population for automobile trips

Variable	**Poisson**		**NB**		**ZINB**	
	Coefficient	***t*-stat**	**Coefficient**	***t*-stat**	**Coefficient**	***t*-stat**
Constant	-1.1	-2.3	0.67	1.98	1.21	**2.41**
AT	-1.00	-4.65	-1.00	-6.07	-1.00	**-7.34**
BUS	-1.36	-3.08	-1.36	-3.68	-1.36	**-4.87**
FEMALE	-0.22	-1.36	-0.21	-1.18	-0.33	**-2.15**
LES1YR	-0.30	-2.59	-0.30	-1.37	-0.30	**-2.24**
ABV10YR	0.42	1.91	0.42	1.38	0.42	1.21
Q9HHCAR	0.32	1.03	0.25	1.31	0.41	**2.01**
DRVLNS	0.46	1.85	0.46	2.51	-0.46	**2.42**
STAFF	-0.42	-1.19	-0.42	-1.72	-0.42	-1.48
UNGRAD	0.08	1.67	0.08	1.31	0.08	1.26
HIGHFLEX	-0.17	-1.32	-0.20	-1.1	-0.27	-1.58
Q6HHSIZE	0.06	1.49	0.06	1.26	0.06	**2.18**
HOMOWN	0.13	1.39	0.13	1.15	0.13	**2.09**
Dispersion Parameter						
Alpha			0.2	2.52	0.51	**3.26**
Tau					-1.11	**-3.31**
Log likelihood (Null)		-505.17		-505.17		-505.17
Log likelihood (Full)		-471.11		-411.23		-401.55
Vuong Statistic						3.56

[*]Bold in Table 4 represents the significant parameters at 99% confidence level (P-value<0.01)

From the parameter estimates for AT trips per day reported in Table 8.5, it is more likely that younger aged commuters will undertake more AT trips per day, compared to older ones. In contrast, gender is not a significant variable in the model. Among other socio-demographic characteristics, it is evident that if the annual income is lower than $15,000 then the probability of making AT trips increases. This would apply primarily to students. Also, housing tenure of less than one year has a positive coefficient value, indicating the transient student groups who change housing with each new academic year. Among the four university segments, staff members are less likely to employ AT trips compared to others. Another interesting finding is that if the daily school/work schedule is flexible then it can be expected that the individual will make more AT trips. The coefficient values for faculty members and staff have negative signs, whereas for graduate students the sign is positive. These variables have been retained in the final model, with the assumption that if the sample size were larger these variables would have been significant.

Table 8.5 Parameter estimation results of trip frequency models for Dalhousie university population for AT trips

Variable	Poisson		NB		ZINB	
	Coefficient	*t*-stat	Coefficient	*t*-stat	Coefficient	*t*-stat
Constant	0.30	1.89	0.31	1.54	0.67	**2.11**
BUS	-0.76	-2.15	-0.76	-2.81	-0.76	**-4.15**
CAR	-0.86	-2.27	-0.86	-3.70	-0.86	**-4.72**
FEMALE	-0.54	-1.35	-0.01	-1.43	-0.01	-1.37
AGE	-0.21	-0.55	-0.11	-0.16	-0.21	**-2.14**
LESS15K	0.20	1.09	0.43	1.43	0.55	**2.53**
BT50T75K	-0.23	1.50	0.23	1.58	0.23	1.50
LES1YR	0.46	2.29	0.46	1.16	0.46	**2.06**
STAFF	-0.43	-1.56	-0.43	-1.18	-0.43	**-1.84**
FACULTY	-0.10	-1.70	-0.10	-1.79	-0.10	-1.62
HIGHFLEX	0.14	-1.29	-0.21	-1.11	-0.32	**-2.09**
Dispersion parameter						
Alpha			1.21	1.86	0.63	**2.51**
Tau					-0.31	-1.68
Log likelihood (Null)		-587.12		-587.12		-587.12
Log likelihood (Full)		-556.71		-534.01		-498.34
Vuong Statistic						5.11

**Bold in Table 5 represents the significant parameters at 95% confidence level (P-value<0.05)*

Table 8.6 also shows the ZINB parameter estimates for transit trip frequencies. The ZINB model for transit trips shows that the number of daily automobile and AT trips have significant negative impact on daily transit trip frequencies. Among the personal characteristics variables, age of the individual has a negative sign, which indicates that individuals with younger ages will undertake more transit trips, compared to older ones. The coefficient for annual income less than $15,000 is positive, but negative for annual income between $15,000 to $25,000. This shows that individuals with less than $15,000 will undertake more transit trips per day than those with an annual income of more than $15,000. It is also found that individuals with no flexibility in work will undertake more transit trips than others. In addition, distance between home and work has a negative coefficient value, showing that individuals living far from the campus employ fewer transit trips than those who live nearer.

Table 8.6 Parameter estimation results of trip frequency models for Dalhousie university population for AT trips

Variable	Poisson		NB		ZINB	
	Coefficient	***t*-stat**	**Coefficient**	***t*-stat**	**Coefficient**	***t*-stat**
Constant	0.04	1.11	0.10	1.51	0.43	**1.85**
AT	-0.96	-4.80	-0.96	-4.07	-0.96	**-5.13**
CAR	-1.44	-3.16	-1.44	-3.71	-1.44	**-4.14**
FEMALE	-0.15	-2.50	0.00	-1.39	0.00	-1.08
AGEGRP	-0.27	-1.41	-0.23	-1.29	-0.27	**-2.13**
HOMOWN	-0.23	-1.59	-0.23	-1.60	-0.23	-1.76
LESS15K	0.25	1.85	0.20	1.90	0.31	**2.35**
BT15T25K	-0.10	-1.35	-0.12	-1.18	-0.14	**-2.11**
UNGRAD	-0.20	-1.95	-0.20	-1.57	-0.20	-1.38
NOFLEX	0.16	1.24	0.16	1.13	0.21	**2.12**
DISTHW	0.20	1.58	-0.13	-1.69	-0.17	**-1.86**
Dispersion parameter						
Alpha			0.43	1.37	0.71	**2.30**
Tau					-0.21	**-1.98**
Log likelihood (Null)		-452.27		-452.27		-452.27
Log likelihood (Full)		-411.31		-393.41		-376.33
Vuong Statistic						3.1

**Bold in Table 6 represents the significant parameters at 95% confidence level (P-value<0.05)*

8.7 Conclusions

This study explored the trip characteristics, activity characteristics, and travel behavior of students and workers at the largest university in the Maritime Provinces of Canada. The target population comprised undergraduate students, graduate students, faculty members, and staff at Dalhousie University, which is a significant generator of travel demand in Halifax Regional Municipality (HRM). This is the first study of activity-travel behavior for a large Canadian university. The findings of this study provide insights which can be employed in TDM strategies at other large urban campuses, and empirical data that can be utilized to represent university populations in regional travel demand models.

The study presented a detailed tabulation of travel mode share, trip characteristics, trip duration, trip length, and activity duration by purpose. Staff members reside farther away from campus and they travel a longer distance for home to work commuting, which is consistent with previous findings (Volosin 2014). In contrast, students reside nearer to the campus and their mean trip length from home to school is significantly shorter than those of faculty members and staff. Although faculty members, on mean, travel a longer distance they use faster modes of travel, and therefore experience shorter trip durations per trip. In general, the most used mode for undergraduate and graduate students is walking, whereas for faculty and staff segments, the most used mode is automobile (drive). It is interesting to note that, for entertainment related activities and sports and hobbies, all four sample segments choose automobile (drive). Another interesting finding is that the mean number of AT trips per day for the overall sample is more than 1.0 trip, whereas automobile has less than 1.0 trip. These are encouraging findings for promoting active transportation for commuting trips in the case of university population. Discouragingly, however, transit usage was lower than for automobile usage, even for the student groups.

Statistical analysis showed the existence of over-dispersion and excess zeros among the dependent variables, thus suggesting the use of a zero-inflated negative binomial structure for modeling activity-travel demands. The ZINB model for AT trips per day suggests that number of AT trips per day is positively associated with older age commuters, annual income below $15,000, housing tenure less than one year, and highly flexible school/work schedules. In contrast, the ZINB model for automobile trips suggests that frequency of automobile trips per day is negatively associated with female respondents, housing tenure less than 1 year, and larger household size, and positively associated with household auto ownership and having a driver license. The ZINB model suggests that number of daily transit trips is positively associated with annual income less than $15,000 and with no flexibility in work schedule, whereas number of daily automobile trips, number of daily AT trips, age of the individual, annual income between $15,000 to $25,000, and home to work/school distance are negatively associated with the daily transit trips.

In summary, data collected through the EnACT survey provide a rich data set on a large Canadian university population, which can be considered as one of the major trip generators that affect regional traffic. The empirical model presented in this study can improve disaggregated regional travel demand models with more accuracy and better precision. This study analyzes and compares the activity-travel demand for all of the four university population segments, which have received very limited attention in the literature. The analysis shows that there exists significant heterogeneity among different market segments of the university population, which needs more attention from transportation planning. Further studies will investigate the association between built environment, land-use, and mode choice across the different university segments using econometric modeling and ArcGIS 10.5.1. The findings on university population segments can assist the development and implementation of more practical and strategic planning solutions to promote a walking- and bicycle-friendly environment near and on campus, and enhance management of on-campus travel demand. The results of this study are expected to be incorporated within the activity-based travel demand model, Scheduler for Activities, Locations and Travel (SALT), for Halifax Regional Municipality (HRM), Nova Scotia, which is currently under development.

Chapter 9 A Pseudo Panel Investigation of Out-of-Home Discretionary Activity Participation[7]

9.1 Introduction

Activity based theories assume that a household or individual's activity-travel dynamics and spatial patterns change over the duration of life cycle stages. These stages are triggered over time due to socio-demographic changes and economic factors. Interestingly, short-term changes in household or individual travel dynamics contribute to the long-term behavioral variability (Clarke 1982). As a result, the overall change in travel behavior, at a point in time, depends on long-term individual experiences and responses. Moreover, the inter-temporal variation in the form of multi-day and multi-period surveys can provide information on the change in people's travel behavior over time (Goodwin 1997). The collection and analysis of multi-period surveys have been introduced into the transport literature to link the longitudinal changes of a traveler's behavior. Undoubtedly, panel surveys are the best approach to study inter-temporal (longitudinal) changes, since they gather information at the individual level over time. Presently, however, limited use has been made of panel surveys in transportation research. The panel surveys employed to date are for relatively short periods of time, and cover only certain demographics. This can be attributed to their high costs and sample attrition problem (Tsai et al. 2014). Repeated cross-sectional data is a viable alternative, which provides data for longer periods of time with more detailed information at the individual level. In comparison to genuine panel

[7] A version of this chapter has been published:
Daisy, N. S. et al. A pseudo panel investigation of out-of-home discretionary activity participation. Peer reviewed proceedings of the 94th Annual Meeting of Transportation Research Board (TRB), Washington, D.C., USA., 2015.

data, repeated cross sections draw different individuals within each wave. Thus, it is not possible to explore the travel behavior changes of individuals over time. However, an alternative solution is to conduct a pseudo panel approach by utilizing the repeated cross-sectional surveys. The pseudo panel approach, introduced by Deaton (1985), offers the advantage of using repeated cross section data for panel data analysis by constructing cohorts and dynamic variables.

The application of the pseudo panel data model is not new to the field of transportation research. There is a significant number of repeated cross section data sets available in the literature. The pseudo panel data approach is mostly applied for forecasting car ownership and car travel demand (Dargay 2002; Huang 2007; Weis and Axhausen 2009). Recently this method has been applied in forecasting public transport demand at a cohort level (Tsai et al. 2014). The application of the pseudo panel is still limited in other areas of transportation research due to the associated technical difficulties in applied research. This study extends the pseudo panel method to participation in out-of-home discretionary activity for Nova Scotia, Canada, by utilizing the repeated cross section data of the General Social Survey (GSS) conducted by Statistics Canada from 1992 to 2010. This study employs cohort data construction which is defined in terms of the gender and birth year of the respondent. The frequency of participation in out-of-home discretionary activities (i.e. shopping, grocery, social, recreational, entertainment, and organizational) is constructed from data describing the participation of an individual among these sub categories of activities.

The level of traffic congestion during off-peak periods and on weekends requires transportation professionals and policy makers to rethink the generation and participation decision of activities for forecasting travel demands (Bernard et al. 2011). Peak hour congestion can be assumed to be a result of work and school trips. However, off-peak hour congestion is significantly affected by discretionary activities. A longitudinal study on out-of-home discretionary activity participation is useful for making long-term planning and policy decisions to address the off-peak hour travel demand of individuals.

The remainder of the study begins with a literature review of pseudo panel data modeling in transportation research. This section is followed by the description of the construction of the pseudo panel data. The methodology of the dynamic pseudo panel data model is then described, and the descriptive statistics of the cohorts are presented. Then the empirical results of a random coefficient model are presented and discussed. Finally, the study concludes by outlining future research possibilities.

9.2 Literature Review

Predicting the future travel demand as well as an individual's response to a policy change is governed by an individual's travel behavior. Previously, researchers focused on daily travel through static models and cross-sectional data. Static models or cross-sectional data cannot accommodate the dynamic effects and changes (Dargay 2002) of an individual. However, dynamic changes are important elements of travel behavior (Goodwin 1987). Additionally, changes such as age, family structure, and life cycle changes (e.g. single to married, birth of a child) trigger the creation of new travel behavior patterns as a result of new demand for activities. Thus, a dynamic model that recognizes time as a dimension and travel choices as dependent on changing preferences, is also important to study travel behavior patterns. However, dynamic models require longitudinal data, most commonly collected in other fields (e.g. health and employment). Travel research is mostly limited to the use of cross sectional rather than longitudinal data, but the concept of multi-day multi-period surveys is not new. Worrall (Garrison and Worrall 1966; Worrall 1967) addressed the value and importance of longitudinal data collection. Following Worrall's early advocacy, longitudinal data was not employed for another two decades, when it was implemented in Dutch panel surveys (Van Wissen and Meurs 1989) and in the Puget Sound transportation panel survey (Murakami and Watterson 1990). The data was in these surveys was collected using a multi-day, multi-period survey concept. Currently a number of multi-day and multi-period surveys exist in the literature, including for the following locations: Uppsala, Sweden, where travel information was collected for a 35-day period (Hanson and Huff 1988); Reading, England, where activity information was collected for one week (Pas 1986); and in the Mobidrive survey in Germany, where travel diaries were kept for several weeks (Axhausen et al. 2001). The long duration of data collecting, high costs, and loss of respondents limit the ability of high frequency panel studies. Thus, pseudo panel data can be utilized as an alternative to explore the behavioral change of individuals over time, when panel data are not available.

Panel data and pseudo panel data are constructed by pooling repeated comparable cross-section data that has been collected over time. Both data types require individual's responses to similar questions that have been posed in a similar manner, in order to maintain comparability over time. As a result, maintaining comparability over time is difficult in constructing true panel data for many reasons. These reasons include failure to answer the same questions over several years, and abandoning the survey due to death, migration, or non-participation (Russel and Fraas 2005). For the above-mentioned reasons, pseudo panel data can be a suitable alternative. Pseudo panel data, introduced by Deaton (1985), is more commonly applied in economic behavior research within the last decade. The application of pseudo panel data in transportation research has been limited to forecasting car ownership and car travel demand (Tsai et al. 2014; Dargay 2002; Huang 2007; Weis and Axhausen 2009). Pseudo panel data is created by repeating cross-sectional data through grouping individuals or households into cohorts based on time invariant variables. These variables include birth year, gender, region, and household location (i.e. the characteristics of data points that are stable over time). Additionally, time invariant variables or characteristics should be available for all the individuals or households in each cross-sectional data period. Each individual or cohort is considered as the individual panel unit (Verbeek and Nijman 1992). For every cohort, the mean value of each variable for each time period is calculated to represent the observations. Thus, aggregations of similar cohorts across time are produced. This allows the grouping of cohorts of individuals born in the same birth period to be done. Such grouping can be empirically estimated as if the individuals were genuine panel data, which has been demonstrated by recent studies (Dargay and Vythoulkas 1999; Gardes et al. 2005; Dargay 2007; Warunsiri and Mcnown 2010; Bernard et al. 2011).

Pseudo panel data creates cohorts that are considered as unique panel units, and offers less measurement error at the individual level. The main aim of cohort formation is to reduce the intra cohort variation, to increase the inter cohort variations, and to trace the cohorts over time (Verbeek and Nijman 1992). As a result, the creation and size of the cohorts is also important. Studies show that the formation of cohorts is conducted based on time invariant individual or household characteristics (birth year, gender, race, residential location, and house size). Birth year is associated with travel behavior, so it is the most commonly used time invariant variable. Other time invariant variables pertinent to travel studies include gender and residential location (Weis and Axhausen 2009), household location and house size (Bernard et al. 2011), and household distance to the central business district (Tsai et al. 2014). Repeated cross-section surveys are not the primary concern for longitudinal studies. Therefore, cohort formation based on time-invariant characteristics needs to be completed so that the cohort size is large enough to provide the representativeness as well as number of cohorts. Generally speaking, pseudo panel data formation is a trade-off between number of cohorts and size of the cohorts. A small number of observations within the cohort increases the measurement error of the population (Deaton 1985; Verbeek and Nijman 1992), whereas a higher number of cohorts are required to study the inter-cohort variation (Verbeek and Nijman 1992). Typically, the pseudo panel has a relatively small number of observations.

In transportation research, due to the unavailability of genuine panel surveys, researchers have usually resorted to repeated cross-sectional data for travel behavior analysis. Geographic modeling by Madre (1990) is an early example of repeated cross section data usage. Dargay and Vythoulkas (1999) utilized pseudo panel data for the first study of car ownership and its determinants for short and long run elasticity. More recently, Dargay (2002) studied the difference in car ownership between rural and urban areas. Simultaneously, Huang (2007) studied household car ownership by using the linear dynamic econometric model, while Weis and Axhausen (2009) studied induced travel by utilizing a structural equation model (SEM).

However, travel behaviors are derived from demand which is generated from an individual`s decision to partake in a particular activity, in a given day and area. The allocation of time to various activities (i.e., mandatory, maintenance, discretionary) determines the number of trips as well as travel mode choice. It is interesting to note that in the past two decades, travel time expenditures and trip making have both increased, mostly due to the increase in discretionary activity-travel engagement (Toole-Holt et al. 2005). Participation and time allocation of discretionary activities have been studied rigorously by Kitamura (1984). Kraan (1996) investigated the total weekly time allocation to in-home time, out-of-home time, and travel for discretionary activities. Bhat (1998) examined activity participation and time allocation to in-home and out-of-home discretionary activities. In another study, Bhat and Misra (1999) modeled the time allocation in in-home and out-of-home discretionary activities for weekdays and weekends. Meloni (2004) examined allocation of time to discretionary in-home and out-of-home activities and trips. Temporal changes in activity choice and duration for maintenance and discretionary activities was investigated by Kundori et al. (2010). Additionally, Chen and Mokhtarian (2006) studied the trade-off in time allocation between maintenance activities/travel and discretionary activities/travel. Bhat (1998) studied modeling in out-of-home recreation, social, and non-maintenance shopping activities. Finally, Meloni et al. (2007) investigated activity time allocation in discretionary activities.

Given the increased participation in discretionary activities over time and its contribution to trip generation, researchers have studied activity participation and activity time allocation in discrete-continuous based utility models, such as a discrete continuous model, nested logit model, nested tobit model, and random utility based microeconomic model. Out-of-home discretionary activities can be planned within a short time and are flexible in time and location. Thus, it is very important to investigate participation in out-of-home discretionary activities over time. This allows the effects of socio-demographic and travel behavior related variables to be incorporated in the trip generation stage of discretionary related travel demand forecasting. Long run variation of participation in discretionary activities has not previously been studied; therefore, this study aims to investigate this variation through the application of a pseudo panel data approach. Participation in discretionary activities over time in Nova Scotia, Canada, are studied by utilizing the GSS episode and household data to construct pseudo-panel data.

9.3 Methods

9.3.1 Model Specification

Discretionary activities are those based on an individual's choice to participate, and which are flexible regarding time and location. As a result, the number and duration of discretionary activities is typically smaller than those for mandatory activities, within a total sample surveyed in a cross sectional survey. Though panel data for multi-day or multi-period are not available for the province of Nova Scotia, the repeated cross sectional data are produced on a regular basis. The simple static genuine panel data model can be written as follows:

$$Y_{nt} = \alpha_0 + \alpha_1 X_{nt} + u_{nt},\ u_{nt} = \beta_n + e_{nt} \tag{1}$$

Where n is the panel units (e.g. individuals, households, country), t is the time period, u_{nt} is the composite error term that includes the fixed individual effects β_{nt} and error term e_{nt}. In repeated cross sections, n is not necessarily the same from one period to another. As a result, writing $n(t)$ would be more appropriate, but to simplify the notation, the index has been kept as n.

As with true panel data, a set of t independent cross sections represented by equation (1) is pooled in pseudo panel data. Unlike true panel data, in pseudo panel data, n is most likely to be a different set of individuals sampled in each cross sectional survey. When individuals are aggregated into cohorts and mean values or proportions for each cohort have been calculated, the above equation (1) can be expanded to a pseudo panel data model as follows:

$$\overline{Y_{ct}} = \alpha_0 + \alpha_1 \overline{X_{ct}} + \overline{u_{ct}}, \qquad \overline{u_{ct}} = \overline{\beta_c} + \overline{e_{ct}} \tag{2}$$

In the above equation, the subscript c instead of i is used to denote the analyst-created cohorts in the pseudo panel data set. $\overline{Y_{ct}}$ and $\overline{X}_{ct}$ in equation (2) denote the mean value for all individuals classified into cohort c at time period t. The unobserved group effect $\overline{\beta_c}$ is time varying because the cohorts of the same group consist of different individuals over time. Regarding the error terms of two different periods, it is assumed that the error terms of different cohorts are not correlated. In this study, the assumption is made that ($\overline{\beta_c} = \overline{\beta_{ct}}$

) for every t and the fixed effect $\overline{\beta_c}$ is treated like a fixed individual effect (β_c), resulting in the basic pseudo panel equation (2).

9.3.2 Estimation Method

The availability of variances and covariance obtained from the cohort's sample means are the focus of much of Deaton's work on pseudo panel data (Deaton, 1985). If the cohort size increases the measurement errors can be considered as zero (Baltagi 1995). The model in equation (2) assumes that the $\overline{\beta_c}$ error term is identically and independently distributed (IID) with a mean of zero and it also represents possible bias from unobserved and fixed cohort heterogeneity. This assumption is not needed for fixed effects model.

The detailed formulation and estimation of the random coefficient model depends on the specific assumptions about the parameter variation. Now subtracting (2) from (1), it must be equally true that:

$$(Y_{nt} - \overline{Y_{ct}}) = (X_{nt} - \overline{X}_{ct})\alpha_1 + (u_{nt} - \overline{u}_{ct}) \tag{3}$$

The above three equations provide the basis of estimating α_1 and is a weighted average of the estimates produced by the between and within estimators. Particularly, random-effects estimator turns out to be equivalent to the estimation of:

$$(Y_{nt} - \theta\overline{Y_{ct}}) = (1-\theta)\alpha + (X_{nt} - \theta\overline{X}_{ct})\alpha_1 + (u_{nt} - \theta\overline{u}_{ct}) \tag{4}$$

Where θ is a function of σ_{β}^{2} and σ_{e}^{2}. The random coefficient model assumes no correlation between β_c and $\overline{X}_{ct}$. The between estimator is less efficient compared to the random coefficient model because the between estimator discards information in the data in favor of sample means.

In this study, it is assumed that $\overline{Y_{ct}}$ depends on some other exogenous variables and the equation will be:

$$\overline{Y_{ct}} = \gamma_0 + \gamma_1 \overline{Z_c} + \gamma_2 \overline{W_{ct}} + \overline{v_{ct}}, \qquad \overline{v_{ct}} = \overline{\beta_c} + \overline{e_{ct}} \tag{5}$$

Where $\overline{W}_{ct}$ is a vector of time variant variables, and $\overline{Z}_c$ is the vector of time invariant variables that are cohort fixed effects. Additionally, γ_1 and γ_2 are vectors of parameters and $\overline{v_{ct}}$ is an error term that is assumed to be independent and identically distributed.

This simulated log-likelihood function is maximized to obtain parameter estimates. Finally, the goodness-of-fit of the estimated models is evaluated in terms of Rho-square, which is calculated by subtracting the ratio of log-likelihood of the full model and the null model (constant only model) from one.

9.4 Data Used for Empirical Application

9.4.1 Data Source and Sample Size

The metadata of activity episodes of 24 hours, as well as the respondent's household and socio-demographic characteristics, are obtained from the General Social Surveys (GSS) of

1986, 1992, 1998, 2005, and 2010. From the family files of each survey period's GSS data, the socio-demographic variables of the respondents of Nova Scotia (NS) have been extracted. From these data the total respondents for each surveyed year were obtained for NS.

9.4.2 Reclassification of Activities

In the GSS activity episode dataset, there are 188 sub-categories of activities under 10 major categories. First, based on the place of the activity, all the activities were classified into two categories: in home and out-of-home activities. All activities were then classified into three broad categories (i.e., mandatory, maintenance, and discretionary) based on the fixity and flexibility of location, duration, frequency, and purpose. All out-of-home discretionary activities were classified into five broad categories (i.e., recreational, voluntary and social, shopping, entertainment, and media). Performing specific steps associated with each category, the frequency datasets for the out-of-home discretionary activity were obtained. Since not all the respondents of the surveyed year participated in out-of-home discretionary activities for the given day, a total of 645, 672, 1105 and 962 participants were generated for 1992, 1998, 2005 and 2010 respectively.

9.4.3 Cohort Construction

GSS time use data is available from 1986 to 2010 constituting five repeated cross-sectional data surveys. Data for 2015 are also now available, but were published too late for inclusion in this study. This study considered four repeated cross-sectional surveys conducted in 1992, 1998, 2005 and 2010. This was a result of the significant variation of the age cohort classification in 1986 in comparison to later years, which violated a prerequisite of the panel or pseudo panel data models. Repeated cross-sectional data of these surveyed years was utilized to create pseudo panel data following the method introduced by Deaton (1985). This pseudo panel then classified individuals into analyst-defined cohorts based on time-invariant criteria such as birth year. Stable cohorts were defined by gender and year of birth of the respondent. The gender characteristics consisted of a male cohort and a female cohort. The generation characteristics consisted of fourteen cohorts, where each cohort represented a five-year span. First and fourteenth generation cohorts contained individuals born between 1906 to 1910, and 1991 to 1995, respectively. Note that all survey respondents were aged 15 or over at the time of survey, so that the 14th cohort was only present in 2010.

The gender and generation cohort definitions describe 28 potential (2*14=28) cohorts (Table 9.1). Repeated over the four survey years, there are a total of 112 cells of cohort mean data. Since the survey interval is more than five years between 1992 and 1998, and between 1998 to 2005, this study assumes that individuals in the 1992 and 1998 samples fall in the same age cohort or birth cohort as if they had been surveyed in 1995 and 2000 respectively. Another issue of data availability results from age being obtained as a categorical variable rather than a continuous one. Although categorical classification reduces the chance of missing data information for age of the respondent, it increased the complexity in cohort formation for pseudo panel data. This study considers that respondents would fall in the same age category as if they had been surveyed in 1995 and 2000 instead of 1992 and 1998 respectively. Therefore, cohorts aged 80 years and older for each cross-sectional survey were discarded for the next cross sectional survey. The secular movement of younger cohorts into the dataset and older cohorts out of the data set generated a total of 260 cohorts for this study.

Table 9.1 Constructing pseudo panel by respondent`s date of birth (mean age and sample size for all cohorts)

Birth Year	1995		2000		2005		2010	
	Mean Age	Sample Size	Mean Age	Sample Size	Mean Age	Sample Size	Mean Age	Sample Size
1991-1995							17	36
1986-1990					17	79	22	44
1981-1985			17	42	22	59	27	45
1976-1980	17	55	22	53	27	67	32	62
1971-1975	22	56	27	53	32	78	37	72
1966-1970	27	79	32	64	37	91	42	85
1961-1965	32	82	37	70	42	114	47	79
1956-1960	37	78	42	39	47	96	52	107
1951-1955	42	53	47	64	52	109	57	111
1946-1950	47	45	52	56	57	99	62	101
1941-1945	52	29	57	55	62	70	67	92
1936-1940	57	36	62	30	67	57	72	40
1931-1935	62	33	67	39	72	32	77	44
1926-1930	67	26	72	40	77	45	82	44
1921-1925	72	28	77	45	82	109		
1916-1920	77	22	82	22				
1911-1915	82	23						
Total		**645**		**672**		**1105**		**962**

9.4.4 Variable Information

Three types of variable can be used to represent the various characteristics of the pseudo panel data cohorts. Based on how individual information has been collected in the surveys and its relationship with cohorts, any given characteristics can be represented by continuous variables, one or more dummy variables, or one or more proportional variables. Gender and generation characteristics were used as dummy variables. The generation characteristics specified whether an individual was or was not a member of a given generation. Generation characteristics consisted of fourteen categories or levels: (a) born 1906-1910, (b) born 1911-1915, (c) born 1916-1920, (d) born 1921-1925, (e) born 1926-

1930, (f) born 1931-1935, (g) born 1936-1940, (h) born 1941-1945, (i) born 1946-1950, (j) born 1951-1955, (k) born 1956-1960, (l) born 1961-1965, (m) born 1966-1970, (n) born 1971-1975, (o) born 1976-1980, (p) born 1981-1985, (q) born 1986-1990, and (r) born 1991-1995. While the information related to age would have allowed calculation of a continuous mean age for each cohort, this study used dummy variables defined by cohort birth years, as the repeated cross-sectional surveys provided age as a categorical variable. These fourteen dummy variables, with names corresponding to the cohort labels, were constructed to represent the generation characteristics.

Some individual characteristics vary from person to person over the cohort, such as educational qualification, marital status, or home ownership. It is important to note that some individuals possess these characteristics and others do not. Thus, cohorts will contain both the people with and without a specific characteristic. For such variables, a value equal to the proportion of individuals in the cohort with the characteristics is reported as the cell value for the whole cohort. These are termed proportional variables in the literature (Russel and Fraas 2005). For this study, the pseudo panel data contained a few characteristics which required the formation of one or more proportional variables, such as educational status, marital status, home ownership, employment status, and income. Mean household size for each cohort was calculated and became a continuous variable in the pseudo panel dataset.

The indicators for travel behavior and time budget were also calculated for the pseudo panel data sets. The mean of all the travel behavior variables was calculated and the derived independent variables of travel behavior are continuous. Variables calculated from activity episodes are: total duration of discretionary activities per day, total duration of in-home activities per day, total duration of out-of-home activities per day, total duration of paid work per day, total daily trip duration, number of out-of-home activity episodes per day, and number of trips per day. The independent variable, number of separate types of discretionary out-of-home activity, has been calculated for each cohort as well.

9.4.5 Descriptive Statistics

Table 9.2 provides the basic descriptive statistics of the pseudo panel: socioeconomic and geographical characteristics of synthetic individuals of the constructed cohorts. Among the individual characteristics, the age of the respondent was reported as categories. Recall that these data were used in the pseudo panel data set as dummy variables and a total of 14 dummy variables were created for the cohorts. So, the group of individuals who were in the mean age 17 cohort in 1992 were in mean age 22, mean age 27 and mean age 32 in 1998, 2005 and 2010 respectively. Gender of the respondent is a constant cohort formation variable over the years so the dummy of gender is made for the pseudo panel data set. Other characteristics, such as marital status, educational status, university completion or incompletion, and secondary completion or incompletion, are reported as the average percentage of the respondents in a cohort. For example, marital status single in Table 9.2 shows that on an average within the cohorts 29.49% of respondents are single and in some cohorts, all the respondents are single (since maximum is 100% for some cohorts). The average percentages of university and secondary school completed among cohorts are 20.13% and 39.77%, respectively. In some cohorts, there are no respondents who have attended university or have an education level higher than high school graduation. However, in some cohorts a maximum of 50% of the respondents have university degree or have attended university.

Table 9.2 Summary statistics of explanatory variables used in pseudo panel data model

Variable	Number of cohorts	Mean/proportion	Standard Deviation	Minimum	Maximum
Respondents Socio-demographic Characteristics					
Gender of the respondent (female=1)	260	50%			
Mean age 17	260	3.08%			
Mean age 22	260	5.38%			
Mean age 27	260	6.92%			
Mean age 32	260	7.69%			
Mean age 37	260	8.46%			
Mean age 42	260	8.46%			
Mean age 47	260	8.46%			
Mean age 52	260	7.69%			
Mean age 57	260	7.69%			
Mean age 62	260	7.69%			
Mean age 67	260	7.69%			
Mean age 72	260	7.69%			
Mean age 77	260	7.69%			
Mean age 82	260	5.38%			
Full time employment	260	0.270269	0.251552	0	0.92
Part time employment	260	0.354654	0.297265	0	1
Secondary complete	260	0.397692	0.162505	0.03	0.88
University complete	260	0.201269	0.097646	0	0.5
Marital status, Single	260	0.294923	0.318376	0	1
Personal income of the respondents between \$40,000 to \$49,000	260	0.050923	0.060179	0	0.21
Household Characteristics					
Home Ownership	260	0.729154	0.158265	0.32	1
Household Size of the respondent (1, if the HH size is 2 or more than 2, 0 otherwise)	260	81.54%			
Travel Behavior Characteristics					
Number of Discretionary out-home activities per day	260	2.0000	0.696889	1	4
Total duration of-discretionary activities	260	186.5228	43.43766	90	335
Number out-of-home activities per day	260	4.502115	1.163049	2.12	9
Number of trips per day	260	4.512115	0.854361	2.14	6.76
Total duration of traveling per day	260	85.68138	22.05137	38.57	160.3
Total duration of work per day	260	434.2165	150.5328	0	912
Total duration of in-home activities	260	1024.304	132.3523	775.7	1333.

For male and female cohorts, the variation over the time is presented in Figure 9.1 and Figure 9.2. However, on an average the respondents from each cohort are participating in at least 2 discretionary activities. Among the travel behavior characteristics, the average number of all out-of-home discretionary activity participation per day for all the cohorts is 2.09. On the other hand, the average duration of the discretionary activities among the respondents of the cohorts is 186.52 minutes. The standard deviation is high (43.44 minutes), implying that respondents from some cohorts are spending higher amounts of time compared to those of other cohorts. Among other travel behavior characteristics, the respondents of the cohorts have an average of 4.5 out-of-home activity episodes per day whereas the average number of trips among cohorts is 4.5 as well. However, on an average, people spend 85.68 minutes per day on travelling and spend on an average 1024 minutes at home. The average duration of work is 434 minutes, but some cohorts have mean time spent as zero (0). This is because some cohorts (those with mean age 17 years, 77 years, and 82 years) are not in the workforce, as is evident from the full time and part time employment percentages as well.

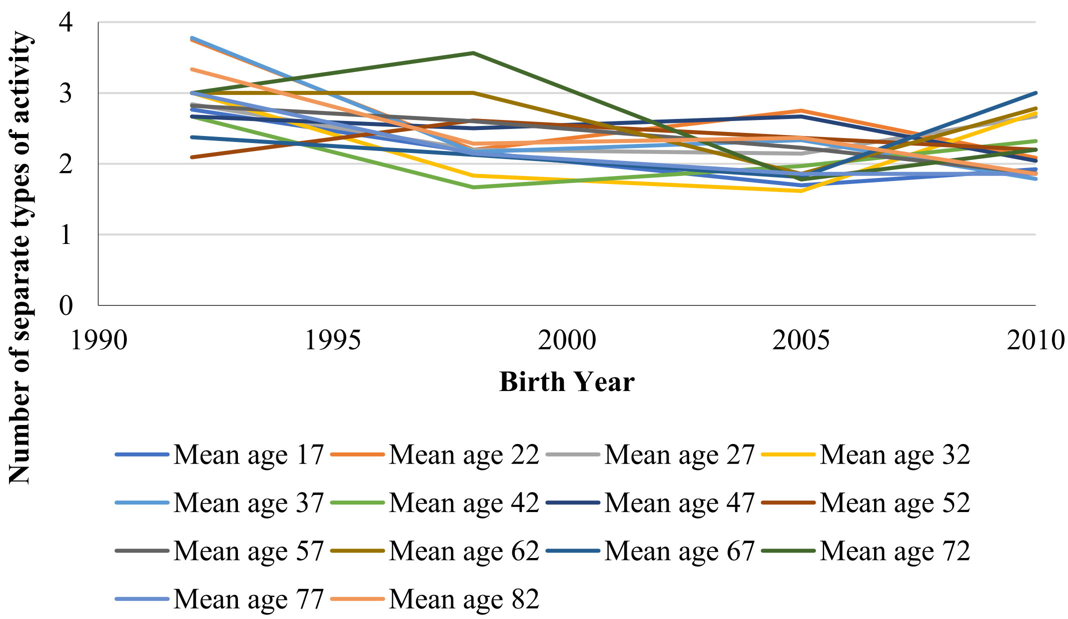

Figure 9.1 Average out of home discretionary activity participation of cohorts (male)

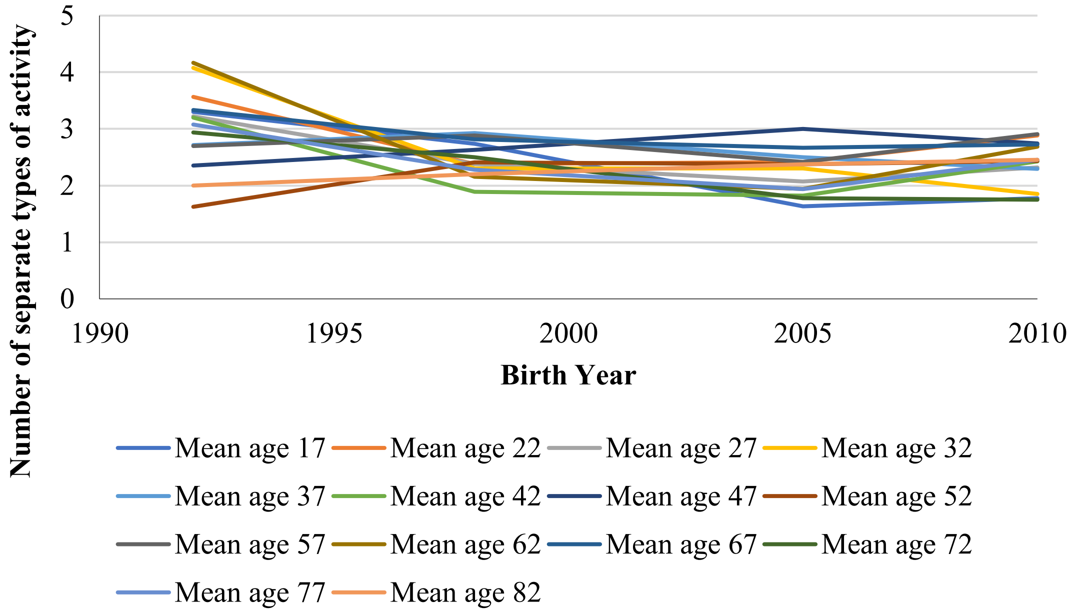

Figure 9.2 Average out of home discretionary activity participation of cohorts (female)

9.5 Discussion of the Results

Table 9.3 shows the parameter estimates of all the variables retained in the random coefficient effects model. The majority of the variables exhibit statistical significance within the 95% confidence interval (t-statistic greater than 1.96). For some of the retained variables, the t-statistic is less than the threshold value; however, they have been retained in the final model specification since intuitively they can be posited to be have behavioral effects, with an assumption that if a larger data set were available, these parameters might show statistical significance.

9.5.1 Random Coefficient Model

The time-invariant variable of birth year is expected to affect the initial status of individuals and their travel behavior over the time. Among the demographic characteristics, gender and age are the most likely determinants of changes in discretionary activity participation. The positive coefficient value for a respondent's gender (female =1) implies that female respondents are participating in more discretionary activities than their male counterparts over the years. However, this coefficient value is small and not significant in the model. Consequently, there is no significant difference between male and female cohorts over the years in terms of out-of-home discretionary activity participation.

Table 9.3 Parameter estimation results of random coefficient model for the pseudo panel data

Variables	**Coeff.**	***t*-stat.**	**P-value**
Constant	0.9118	1.33	0.183
Respondents Characteristics			
Gender of the respondent (1 if the cohort is female, 0 otherwise)	-0.0744	-0.82	0.415
Mean Age 22	-0.0028	-0.04	0.97
Mean age 27	0.2616	1.72	0.086
Mean Age 32	0.0191	0.26	0.793
Mean age 37	0.0222	0.26	0.792
Mean age 42	0.0458	0.49	0.625
Mean age 47	0.0798	0.78	0.434
Mean age 52	0.1254	1.12	0.264
Mean age 57	0.1461	1.22	0.222
Mean age 62	0.1725	1.36	0.172
Mean age 67	0.1839	1.38	0.169
Mean age 72	0.2207	1.58	0.114
Mean age 77	0.2433	1.67	0.096
Marital status, Single	-0.4939	-3.12	0.002
Full time employment	-0.3988	-2.05	0.041
Part time employment	0.0635	0.45	0.652
Superior incomplete or complete	0.0612	0.14	0.885
Household Characteristics			
Personal income of the respondents between $40,000 to $49,000	0.2711	0.52	0.605
Home Ownership	-0.4694	-1.37	0.169
Household Size of the respondent (Dummy variable =1, if the cohort is female, otherwise 0)	0.0878	0.65	0.513
Travel Behavior Characteristics			
Total duration of-discretionary activities	0.0035	3.7	0
Total duration of traveling per day	-0.0075	-3.29	0.001
Total duration of work per day	-0.0008	-2.24	0.025
Total duration of in-home activities	-0.0006	-1.35	0.179
Number out-of-home activities per day	0.2241	5.3200	0.0000
Number of trips per day	0.3344	4.76	0
LL likelihood (Constant only model)	-91.015798		
LL likelihood (Full model)	-12.172692		
Rho Square	0.7842177		

To understand how the participation rate in out-of-home discretionary activities changes with age, the coefficients of adjacent age cohorts can be compared. For example, the respondents within the youngest age cohort (with a mean age of 22) have a negative coefficient value whereas respondents within the second youngest age cohort (with a mean age of 27) have a positive coefficient value. This shows that an individual entering into a higher age cohort is likely to participate in more out-of-home discretionary activities. The model results show that participation in out-of-home discretionary activities increases consistently with age.

To control for the changes in a cohort's proportion of married individuals, the random coefficient model included single and married variables that measured the proportion of a cohort that was single and married respectively. The married variable was not significant but the single variable was highly significant with a negative coefficient value. Although the personal characteristic of educational qualification (which represents the proportion of university completion or attendance) was found to be insignificant at the 0.05 significance level, the coefficient is positive; this implies that participation in out-of-home discretionary activities will be higher if a respondent is more educated. Assuming that education and income are correlated, higher levels of education would lead to opportunities for a greater variety of out-of-home discretionary activities. As expected, the model results show that higher incomes are also associated with higher levels of participation in out-of-home discretionary activities.

Individuals with part time work are likely to participate in more out-of-home discretionary activities, whereas individuals with full time work are likely to participate in fewer out-of-home discretionary activities (with significance at the 0.05 level). The model results show that household ownership is a statistically significant determinant of participation in out-of-home discretionary activities, with household owners being likely to participate in fewer out-of-home discretionary activities. This result could be due to a number of factors including greater time requirements for household maintenance. An individual living in a household with more than two people is likely to participate in more out-of-home discretionary activities than individuals living in smaller households. Among the travel behavior characteristics, model results show that the duration of in-home activities and the duration of work activities reduces the level of discretionary activity participation. Further, the duration of travel variable has a negative coefficient value, which implies a lower participation in out-of-home discretionary activities as trip times increase. These results are logical since increasing the time allocated to one activity reduces the time available to participate in other activities. The model results also show that if the number of participation in out-of-home discretionary activities increases, subsequently, the total duration of discretionary activities will increase. An individual that takes more trips is likely to participate in more out-of-home discretionary activities. This positive correlation is logical given that trips are required to participate in out-of-home activities.

9.6 Conclusion

Despite the existence of substantial research on participation and time allocation in out-of-home discretionary activities, the longitudinal variation in out-of-home discretionary activities is still under-researched. This study identifies the determinants affecting the discretionary activity participation over the years 1992 to 2010. The estimation of the random coefficient model of the pseudo panel data suggests that gender and age characteristics of an individual have insignificant long run effects over the lifespan on out-of-home discretionary activity participation. Other personal characteristics such as educational qualification, marital status, and personal income, however, have significant effects on out-of-home discretionary activity participations. The duration of out-of-home activities, in-home activities, and work duration have significant negative effects on participation in discretionary activities. Among the travel behavior characteristics, higher trips increase the level of participation in out-of-home discretionary whereas increase in duration of traveling decreases the participation rate.

This study has some limitations, such as small sample size and use of only one methodological approach. It would be interesting if the models could be tested for a larger dataset. Moreover, this study does not address the latent heterogeneity issues. Future models should examine alternative methods, including a latent class model. Nevertheless, the study contributes in several ways. Particularly, it explores the influence of individual and socio-demographic characteristics, household characteristics, and travel behavior characteristics over time. The behavioral insights obtained from this study could be useful for investigating longitudinal variations in behavior, to better forecast future travel demands for NS, Canada.

Chapter 10 Population Synthesis based Pseudo Panel Modeling of Out-of-Home Discretionary Activity Duration[8]

10.1 Introduction

This study presents a novel approach to pseudo panel based out-of-home discretionary activities modeling. Although the need for longitudinal studies in transportation has been recognized for some time, panel data has not been available for longer periods due to high survey costs and sample attrition problems (Tsai et al. 2014). Nevertheless, it is vital to investigate longitudinal changes in activity/travel behavior through panel investigation, particularly on a multi-year scale, to better understand changing behavioral dynamics. Repeated cross-sectional surveys for activity/travel behavior are available in many regions, which could be a viable option that provides data for a longer time horizon. However, repeated cross-sectional surveys draw different individuals in each wave, making comparison through time problematical. Pseudo panel techniques offer the advantage of utilizing the repeated cross-sectional information for panel analysis by constructing comparable cohorts of individuals in each wave (Deaton 1985). Pseudo panel activity data could essentially circumvent the unavailability of multi-year panel data. Therefore, this study develops a pseudo panel based method by utilizing the repeated cross sectional Canadian time-use surveys. The method is employed to study time allocation to the four-major out-of-home discretionary activities: social, recreational, shopping, and entertainment activities.

---

[8] A version of this chapter has been published:
Daisy, N. S. et al. Population synthesis based pseudo panel modeling of out-of-home discretionary activity duration. Peer reviewed proceedings of the 95th Annual Meeting of Transportation Research Board (TRB), Washington, D.C., USA., 2016.

There is a growing interest in employing pseudo panel investigation in travel behavior research (Yang and Timmermans 2012) to develop dynamic activity based travel models. Most panel-based activity analysis is however restricted to short-term (e.g. weekly panel) studies. It is evident from literature review that multi-year pseudo-panel based activity-travel studies are limited (Daisy et al. 2015). One of the major difficulties in applying the pseudo panel approach in activity based analysis is the small sample size in most multi-year repeated cross sectional surveys. Since a pseudo panel generates panel units by consistent aggregation of cross sectional data into time invariant criteria based groups, it requires a larger dataset for forming a sufficient number of cohorts for meaningful empirical analysis. Though sample weights and regression models are sometimes utilized to expand the sample population, it often becomes computationally intensive (Chung and Goulias 1995; Gelman and Carlin 2002). To avoid associated complexities, recent studies are experimenting with population synthesis approaches for sample expansion (Auld et al. 2009; Chung and Goulias 1997). Thus, this study used a population synthesis technique to expand the sample into 5% of the total population for each survey year considered in this study. The population synthesis technique is utilized in both the individual and household level marginal tables that are able to simultaneously synthesize the sample data into the target population. In this study, the synthetic algorithm is implemented by utilizing individual level control tables and seed data. Since time-invariant variables are required for generating cohort averages, time-invariant individual level variables such as birth year, gender, and birth place of the respondent have been used for creating the synthetic population. This expanded synthetic population is then converted into cohorts for duration modeling.

There have been numerous studies conducted regarding participation in and time allocation to discretionary activities (Kitamura 1984; Kraan 1996; Bhat and Misra 1999). However, there is a growing need to study longitudinal changes for time allocation to different types of out-of-home discretionary activities on a multi-year basis, since understanding long-term changes in activity patterns is becoming vital for long range planning. Given the importance of studying time allocation for discretionary activities over time and the lack of genuine time use panel data for longer periods, cohorts from pseudo panel data were utilized to estimate accelerated hazard-based duration models for four major types of out-of-home discretionary activities (i.e. shopping, social, entertainment, and recreational). Since the pseudo panel data in this study resembles repeated choices of duration, thus, the study takes a latent class modelling approach to account for unobserved heterogeneity due to panel effects.

10.2 Literature Review

In the past two decades, discretionary activities have gained paramount importance for their relevance to the quality of life and wellbeing of individuals (Toole-Holt 2005). Activity participation and time allocation for discretionary activities have been studied rigorously in the time use and travel behavior literature (Kitamura 1984). For instance, Kraan (1996) investigated the total weekly time allocation to in-home, out-of-home, and travel for discretionary activities. Bhat (1998) examined activity participation and time allocation to in-home and out-of-home discretionary activities. In another study, Bhat and Misra (1999) modelled the time allocation for in-home and out-of-home discretionary activities for weekdays and weekends. Meloni et al. (2004) examined the allocation of time to discretionary in-home and out-of-home activities and trips. Modeling activity choice and duration with history dependency for maintenance and discretionary activities were investigated by Kundori et al. (2010). Additionally, Chen and Mokhtarian (2006) studied the trade-off in time allocation between maintenance activities/travel and discretionary activities/travel. Recently, Meloni et al. (2007) investigated activity time allocation in discretionary activities.

However, study of longitudinal changes in activity patterns is important for improving our understanding of long-term trends in activity participation and time allocation. Though panel surveys are available, they are limited to short time scales, for example, multi-weeks or 1-2 years (Van-Wissen and Meurs 1989; Murakami and Watterson 1990; Hanson and Huff 1988; Pas 1986; Axhausen et al. 2001). For instance, a multi-day multi-period survey was administered in Dutch panel surveys (Van-Wissen and Meurs 1989) and in the Puget Sound transportation panel survey (Murakami and Watterson 1990). Currently a number of multi-day and multi-week surveys and modelling research exist in the literature, including: Uppsala, Sweden, where travel information was collected over a 35-day period (Hanson and Huff 1988); Reading, England, where activity information was collected for one week (Pas 1986); and the Mobidrive survey in Germany, where travel diaries were reported for several weeks (Axhausen et al. 2001). However, panel data for a longer time horizon is almost absent. This absence of genuine panel data limits the study of differences among age groups or so-called “generations” in terms of time allocation to discretionary activities. Thus, an investigation using a pseudo panel approach is critical to explore time allocation behavior of individuals.

Pseudo panel data is constructed by pooling repeated comparable cross-section data that has been collected over time. The pseudo panel method, introduced by Deaton (1985), is more commonly applied in economic behavior research. Madre (1990) is an early example of the application of repeated cross sections for the pseudo panel approach in travel analysis. Dargay and Vythoulkas (1999) utilized pseudo panel techniques for modeling car ownership and its determinants for short and long run elasticity. More recent studies include the difference in car ownership between rural and urban areas (Dargay, 2002), household car ownership (Huang, 2007), and induced travel analysis (Weis and Axhausen, 2009) etc.

Pseudo panel analysis considers the cohort as the individual panel unit (Verbeek and Nijman, 1992). For every cohort, the mean value of each variable for each period is calculated to represent the panel observations. Thus, aggregations of individuals having similar personal characteristics in cohorts across time are produced. Such grouping can be empirically estimated as if the individuals were genuine panel data, which has been demonstrated by recent studies (Dargay and Vythoulkas 1999; Dargay 2007). The main aim of cohort formation is to reduce the intra-cohort variation, to increase the inter-cohort variations, and to trace the cohorts over time (Verbeek and Nijman 1992). Time invariant characteristics are utilized to form the cohorts. Since the birth year and gender are associated with travel behavior, these are the most commonly used time invariant variables. Other variables considered in several studies include residential location (Weis and Axhausen 2009), household location and house size (Bernard et al. 2011), and household distance to the central business district (Tsai et al. 2014).

Repeated cross-section surveys are not primarily designed for longitudinal studies. Therefore, cohort formation based on time-invariant criteria needs to be carefully designed to increase the number of cohorts as well as cohort representativeness. Essentially, pseudo panel data formation is a trade-off between the number of cohorts and size of the cohorts. A small number of observations within the cohort increases the measurement error of the population (Deaton 1985), whereas a higher number of cohorts is required to study the inter-cohort variation (Verbeek and Nijman 1992). Since the pseudo panel approach generates a smaller number of observation units, a further econometric application becomes challenging when the sample size in repeated cross sections is small, and generally speaking this is the case. For example, the highest sample size for repeated cross sections of the General Social Survey of Canada (GSS) survey for Nova Scotia from 1986 to 2010 was only 1105.

A small sample size generates fewer panel cohorts. There are, however, a number of methods to expand the sample, such as sample weights, expansion weights, etc. This study utilized a time invariant multi-criterion based synthetic population generation method to expand the samples of repeated cross sectional Canadian time use surveys. The goal of population synthesis is to create or expand disaggregate sample data through an iterative process to the target 5% of Nova Scotia population obtained from the Canadian census. There are several studies that have examined population synthesis methods such as Iterative Proportional Fitting (IPF) (Bowman and Ben-Akiva 2001), Combinatorial Optimization (CO) (Voas and Williamson 2010), Iterative Proportional Updating (IPU) (Ye et al. 2009), and Fitness Based Synthesis (FBS) (Ma 2011). Major differences in population synthesizer methods are the direct effect on precision of the synthesized population, and the ability to simultaneously control both the individual and household characteristics. This study uses a population synthesis technique that utilizes a fitness value for replicating the target population (Ma 2011). This study extends the application of the population synthesis technique for sample expansion to multiple years using time-invariant information.

As indicated earlier, numerous studies analyze the duration of discretionary activities (Kitamura 1984; Kraan 1996; Meloni et al. 2004; Chen and Mokhtarian 2006; Meloni et al. 2007). The majority of the studies focus on the investigation of activities for a typical day. A few studies have used short term panel data (Guo and Bhat 2007), but studies on longer time horizons are not available. Moreover, in most cases panel effects were ignored (Yee and Niemeier 2000; Berg et al. 2012). In order to capture panel effects due to repeated choices, this study utilizes a Latent Class Accelerated Hazard (LCAH) model for duration analysis. The method used in this study allows incorporation of latent heterogeneity within the modeling framework, as evident in Lee and Timmermans (2007) and Berg et al. (2012). This study extends the method for pseudo panel investigation.

10.3 Methods

10.3.1 Population Synthesis Based Pseudo Panel Construction

10.3.1.1 Population Synthesis Technique

For the pseudo panel construction, a population synthesis based method is utilized to generate the synthetic sample population for an 18 year time horizon. Two groups of the dataset are used to synthesize population through a population synthesis technique in this study: seed data and control tables. The seed data are extracted from 1992, 1998, 2005 and 2010 General Social Survey (GSS) public use micro datasets, and control tables are obtained from the Canadian Census tabulations. For sample expansion, three time-invariant individual variables of birth year, gender, and education level are employed in the population synthesizer (cross tabulation of birth year against gender and educational level against gender). The population synthesis begins by calculating a fitness value, followed by selecting the appropriate individual from the seed data, who is then added into the synthesized list (Ma 2011). Details of the improved algorithm and computational framework can be found in Hafezi et al. (2018). If a is the selected individual, b is the iteration number, u is the index representing both count and control tables, U is the total number of both count and control tables, and s is the index representing the various cells in the count table, the equation for fitness used to generate synthetic population is as follows:

$$F_U^{ab} = n_u * [(T_{us}^{b-1})^2 - (T_{us}^{b-1} - JK_{us}^{a})^2] \tag{1}$$

$$T_{us}^{b-1} = M_{us} - PM_{us}^{b-1} \tag{2}$$

$$g^u = find(F_u^{ab} > 0) \tag{3}$$

$$lm_b^{us} = find(g_u^b >= 0) \tag{4}$$

$$km_b = intersect(lm_b^{1:us}) \tag{5}$$

Where:

F_U^{ab} is the fitness value for control table U.

M_{us} represents the amount of cell s in control table u.

PM_{us}^{b-1} represents the value of cell s in the count table u.

T_{us}^{b-1} is the difference value between control and count tables for cell s in control table u.

JK_{us}^{a} is the contribution of the a^{th} individual in the seed data to the s^{th} cell in control table u.

g^{u} is the selected individual type according to the fitness value

lm_{b}^{us} is the selected individual for the cell s in the count table u in the iteration b

km_{b} is a set of selected individuals for adding into the count tables and synthesized population list.

F_{U}^{ab} is the fitness of a specific individual from the seed population to be selected for inclusion into the synthetic population. g^{u} is referred to find all seed individuals where the criteria is fulfilled (positive F_{U}^{ab}) for each specific control table. In the next step, alternative individuals for adding into the count tables and synthesized population list are found by intersecting between the selected list of individuals among all control tables. The fitness calculation and an individual's selection are repeated until there are no positive fitness values left in the procedure.

10.3.2 Duration Modeling

The hazard-based duration model considers the conditional probability of a time duration ending at time t, density $f(t)$ and survival function $S(t) = \Pr ob[T \geq t]$, where the duration continues until time t. For survival time, the accelerated hazard model takes a log-linear model, where the error variable has the density function $f(t)$. Hence, a parametric hazard model for discretionary activity duration can be denoted as:

$$Z = \varphi \overline{Y_{ct}} + p\varepsilon \tag{6}$$

Where $Z = \log(t)$

t, is a random variable representing the continuous activity duration

φ, is an unknown parameter vector of the covariates vector $\overline{Y_{ct}}$

ε, is an error variable with density $f(t)$

Based on the distribution function of the error term, different types of accelerated hazard models can be considered, such as Weibull, log-normal, log-logistic, or exponential. Recent works recommend the Weibull distribution as a major descriptor of travel behavior analysis, particularly for duration of activities (Hensher and Mannering 1994; Bhat 1996). Thus, this study assumes a Weibull distribution for the baseline hazard distribution. The hazard function of the Weibull distribution can be written as follows:

$$\lambda(t) = \lambda p(\lambda t)^{p-1} = p\lambda^{p} \times t^{p-1} \tag{7}$$

Where, λ is the scale parameter and p is the shape parameter. This hazard function is monotonically decreasing for p <1, monotonically increasing if p >1, and reduces to a constant if p =1. For the Weibull distribution, the survival function will be:

$$S(t) = \exp\left(-(\lambda t)^{p}\right) \tag{8}$$

Density function of the Weibull distribution:

$$f(t) = \lambda p(\lambda t)^{P-1} \exp\left[-(\lambda t)^{P}\right] \tag{9}$$

To incorporate heterogeneity, a LCAH model is formulated. Let's assume, φ_k is the unknown parameter vector for each class k and ε_k has a Weibull distribution with scale parameter p . Thus, for a cohort of individuals belonging to latent class k, the LCAH model for the logarithm of social/recreational/shopping/entertainment activity with duration Z given covariate vector Y is:

$$Z = \varphi_k \overline{Y_{ct}} + p_k \varepsilon_k \tag{10}$$

If the number of observations is N and for R random draws, the simulated log-likelihood function would be:

$$LogL = \sum_{c=1}^{N} \frac{1}{R} \sum_{r=1}^{R} \left\{ \log f\left[Z_c, \lambda_c\left(\varepsilon_{cr}\right), p\right] + \log S[Z_c, \lambda_c, (\varepsilon_{cr}), p] \right\} \tag{11}$$

Where, the unconditional log-likelihood is maximized by integrating ε_k out of the conditional log likelihood (Greene 2007). Finally, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values are calculated to examine the model degree of fit, due to the hierarchical structure of the latent class models.

10.4 Construction of the Pseudo Panel Dataset

10.4.1 Data Source and Sample Size

The micro-data of activity episodes for 24 hours, as well as the respondent's household and socio-demographic characteristics, were obtained from the GSS. The expanded population amounts for 1992, 1998, 2005 and 2010 are 46000, 46600, 46900 and 47105 respectively. The error-percentage method was used to validate the synthesized population for each specific year. For each of the control tables (birth year, gender, and education level), error percentage was calculated by comparison of the target and synthesized population. Table 10.1 presents the error percentages of the synthesized population. As highlighted in Table 10.1, the population synthesis technique synthesized population at the individual level with high precision.

Table 10.1 Error percentages of synthesized population

Explanatory Variables	Error Percentages			
	1992	1998	2005	2010
Birth Year	1.65	0.47	0.60	0.44
Gender	2.22	1.74	1.86	1.13
Education level	1.53	0.42	0.28	0.15

10.4.2 Reclassification of Activities

In the GSS activity episode dataset, there are 188 sub-categories of activities, under 10 major categories. First, based on the place of the activity, all the activities were classified into two categories: in home and out-of-home activities. Next, all out-of-home activities were classified into three broad categories (e.g. mandatory, maintenance, and discretionary) based on the fixity and flexibility of location, duration, frequency, and purpose (Bhat and Koppelman 1993; Srinivasan 2004). Lastly, all out-of-home discretionary activities were reclassified into four broad categories, namely shopping, recreational, social, and entertainment.

10.4.3 Cohort Construction

Pseudo panel data is created from repeated cross sections by classifying individuals into cohorts based on three time invariant criteria, which in this study are gender, birth year, and educational level. Gender characteristics consist of a male cohort and a female cohort. The birth year characteristics consist of fourteen cohorts, where each cohort represents a five-year span. First and fourteenth birth year cohorts contained individuals born between 1906 to 1910, and 1991 to 1995, respectively. The education level characteristics consist of five educational levels of the respondents. Hence, gender, education level, and birth year characteristics describe 140 potential (2×5×14=140) cohorts. Since the survey wasn't normally distributed, however, fewer than 140 cohorts were expected and retained.

10.4.4 Variable Generation

In the pseudo panel data, three types of variables can be used to represent the various characteristics of the cohorts. Gender, education level, and birth year characteristics were used as dummy variables. The birth year range is specified by whether an individual was or was not a member of the cohort for a given period of time. Birth year characteristics consisted of a total fourteen birth years for each cross-section: (a) born 1906-1910, (b) born 1911-1915, (c) born 1916-1920, (d) born 1921-1925, (e) born 1926-1930, (f) born 1931-1935, (g) born 1936-1940, (h) born 1941-1945, (i) born 1946-1950, (j) born 1951-1955, (k) born 1956-1960, (l) born 1961-1965, (m) born 1966-1970, (n) born 1971-1975, (o) born 1976-1980, (p) born 1981-1985, (q) born 1986-1990, and (r) born 1991-1995.

While the information related to age would have allowed the calculation of a continuous mean age for each cohort, this study used dummy variables defined by cohort birth years, since the repeated cross-sectional surveys collected age as a categorical variable. These fourteen dummy variables, with names corresponding to the cohort labels, were constructed to represent the age attributes of the cohort.

Some individual characteristics can be unique for each individual, which cannot be considered as uniform over the cohort, such as educational qualification, marital status, or home ownership. For such variables, a value equal to the proportion of individuals in the cohort with the characteristics is reported as the cell value for the whole cohort, which is termed as proportional variables in literature (Dargay and Vythoulkas 1999). In this study, the pseudo panel data contained a few characteristics that required the formation of one or more proportional variables such as educational status (primary and high school), marital status, home ownership, employment status, income (less than 10k, between 10k to 20k, between 20k to 30k, and, between 30k to 40k), etc. Mean household size for each cohort was calculated and transformed into a continuous variable in the pseudo panel dataset.

The indicators for travel behavior and time budget are also calculated for pseudo panel data sets. Variables calculated from activity episodes are: total duration of in-home activities per day (DuraInHom), total duration of out-of-home activities per day (DuraOutHm), total duration of shopping activities per day (DuraShop), total duration of paid work per day (DuraWork), total duration of entertainment activities per day (DuraEnter), total duration of recreational activities per day (DuraRec), total duration of sedentary activities per day (DuraSeden), total duration of social activities per day (DuraSocial), total duration of travelling per day (TTDura), total commute time per day

(TTWork), and total number of trips per day (TTrips) etc. The dependent variables, duration of each of the four discretionary out-of-home activities per day, was calculated for each cohort as well.

10.4.5 Descriptive Statistics

Activity attributes, socio-demographic and personal characteristics were selected as the exogenous (independent) variables for the model. Summary statistics of the continuous and proportional variables upon which the final empirical studies were carried out are listed in Table 10.2. The average duration of shopping, entertainment, recreation, and social activities per day was 28.50 minutes, 95.44 minutes, 66.16 minutes, and 60.00 minutes respectively for 1992. The average durations for all the four years are listed in Table 10.2. Marital status, monthly personal income, employment status, home ownership status, household size, and living arrangement were found to be consistent over the years, and therefore utilized for pseudo panel construction. For example, average household sizes for 1992, 1998, 2005 and 2010 were 2.54, 2.32, 2.23 and 2.36 respectively.

In addition, activity durations, number of trips per day, total duration spent on travelling, etc. were used in the final model. Note that, for each category of discretionary activity the number of cohorts was different due to the different activity participation rates and patterns observed in the multi-year samples.

Table 10.2 Summary statistics of explanatory variables used in duration models

Variable	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
	1992		1998		2005		2010	
No. of cohorts	106		115		116		114	
Dependent Variables (minutes)								
DuraShop	28.50	46.68	30.44	46.31	24.10	31.68	38.58	59.03
DuraEnter	95.44	119.85	112.5	73.43	125.94	86.74	114.76	87.52
DuraRec	66.16	65.55	100.05	100.19	119.23	96.02	85.81	84.30
DuraSocial	60	71.34	79.17	54.93	30.12	48.13	72.99	60.30
Explanatory Variables (%)								
MAge27	9.43	-	7.83	-	6.03	-	5.26	-
MAge32	9.43	-	7.83	-	6.90	-	7.89	-
MAge37	8.49	-	8.70	-	7.76	-	7.02	-
MAge42	3.77	-	7.83	-	7.76	-	7.02	-
MAge47	6.60	-	9.57	-	8.62	-	9.65	-
MAge52	7.55	-	6.09	-	7.76	-	7.02	-
MAge57	7.55	-	6.96	-	9.48	-	8.77	-
MAge62	5.66	-	8.70	-	6.90	-	7.89	-
MAge67	6.60	-	6.09	-	6.03	-	7.89	-
MAge72	7.55	-	7.83	-	6.03	-	6.14	-
MAge77	6.60	-	4.35	-	6.03	-	7.02	-
MAge82	5.66	-	5.22	-	6.90	-	7.02	-
Male	45.28	-	47.83	-	47.41	-	45.61	-
EduPrim	17.92	-	18.26	-	22.41	-	21.93	-
EduCollg	21.70	-	17.39	-	20.69	-	21.05	-
PSingle	0.26	0.39	0.07	0.21	0.10	0.21	0.11	0.25
PMarried	0.52	0.42	0.26	0.37	0.24	0.32	0.29	0.36
INCM3	0.27	0.36	0.19	0.30	0.10	0.20	0.11	0.22
INCM5	0.12	0.29	0.07	0.19	0.07	0.18	0.08	0.17
INCM6	0.04	0.15	0.07	0.20	0.07	0.18	0.06	0.14
INCM7	0.04	0.16	0.02	0.11	0.07	0.20	0.16	0.29
PPartTym	0.05	0.15	0.05	0.19	0.05	0.16	0.03	0.14
PFullTym	0.38	0.41	0.05	0.16	0.08	0.20	0.07	0.18
PTenant	0.27	0.35	0.70	0.22	0.73	0.16	0.79	0.11
PHomOwn	0.73	0.35	0.30	0.35	0.27	0.34	0.21	0.29
HH Size	2.54	1.20	2.32	1.01	2.23	0.88	2.36	1.00
LAAlone	0.26	0.38	0.23	0.33	0.19	0.27	0.23	0.35
LASpouse	0.25	0.36	0.22	0.33	0.22	0.30	0.27	0.34
DuraInHom	1025.85	294.57	1020.16	271.85	1009.71	300.06	1085.77	209.22
DuraOutHm	370.15	297.39	383.13	244.62	360.01	278.76	293.29	192.99
DuraWork	114.75	148.43	107.28	172.47	110.04	165.14	109.54	160.57
DuraSeden	84.841	101.37	94.48	150.05	112.51	135.96	86.34	131.77
TTWork	8.71	11.73	9.98	18.50	10.15	17.57	9.37	14.54
TTDura	55.64	65.74	67.89	54.55	69.69	57.01	75.33	51.20
TTrips	3.82	2.47	4.37	3.35	3.76	2.53	3.73	2.08

10.5 Discussion of the Duration Model Results

10.5.1 Major Findings

Parameter estimates of four duration models for shopping, entertainment, recreation, and social activities are presented in Table 10.3, Table 10.4, Table 10.5 and Table 10.6. The models were evaluated in terms of AIC and BIC. The AIC and BIC values of the full models are considerably lower than the null model. The model results suggest that baseline hazard is monotonically increasing for all types of discretionary activities. The effects of the explanatory variables are briefly discussed below.

10.5.1.1 Personal Characteristics

As shown in Table 10.3, among the personal characteristics, age, gender, and education level affect the time allocation to different types of discretionary activities significantly. The age of the individuals is found as a significant factor for time allocation to all the four types of discretionary activities; shopping, entertainment, recreational and social. It is found that the higher the age, the higher the time commitment for shopping activities, recreational activities, and social activities. There is a trend of increasing coefficient values with the increase in the age of the cohorts for these three activities. But age has a negative effect on the duration of entertainment activities over the years. As expected, negative coefficient values increase with the increase in age for entertainment activities. Comparatively, male respondents are found to spend more time in recreational activities but less time in shopping activities than their counterparts. Individuals with up to a higher tertiary are less likely to spend time on recreational activities. On the other hand, individuals who obtained a high school degree, allocate higher time for entertainment activities.

10.5.1.2 Socio-Demographic Characteristics

As shown in Table 10.4, with regard to socio-demographic characteristics the following are found as significant factors: marital status, monthly personal income, full time or part time worker status, household size, home ownership, and living arrangement. The proportion of single and married individuals in the cohorts have a significant positive effect on shopping activity duration. On the other hand, married individuals of the pseudo panel spend less time in recreational activities. Personal monthly income has a significant effect on shopping activities, recreational activities, and social activities. As expected, people with lower monthly income spend less time in shopping activity but they spend more time on social activity. Middle category monthly income has a negative impact on recreational and social activity duration, whereas low to medium categories spend less time in both recreational and social activities. Part time workers spend less time in shopping activities, whereas full time workers spend less time in recreational activities. Similarly, a larger household size is associated with shorter entertainment activity duration but longer social activity duration. An individual with tenant home ownership status will spend less time on entertainment activities but have higher time commitments for recreational activity. On the other hand, individuals with home ownership have a longer shopping activity duration but spend less time on social activities. Furthermore, living with spouse is associated with longer shopping, recreational, entertainment, and social activity durations.

10.5.1.3 Activity Attributes

Among the activity attributes, duration of in-home activities and work significantly affect the time allocation to all types of discretionary activities. As shown in Table 10.5, it is evident from the model results that there are trade-offs in terms of time budget for daily activities. Among other activity attributes, duration of all out-of-home activities is found to negatively affect the duration of social activities. On the other hand, time allocation on shopping activities negatively affects the duration of entertainment and recreation activities. Greater time allocation to entertainment activities decreases the probability of an individual spending longer on shopping and social activities. Conversely, with increased duration of recreational activities, shopping and entertainment activities will decrease. It is interesting to note that time spent on sedentary activities decreases the duration of recreational activities. Time spent on social activities decreases the duration of entertainment activities. Among the travel attributes, higher commute times increase the probability of lower shopping activity duration. As expected, longer travelling time per day decreases the activity duration for shopping activities, recreational activities, and social activities for the higher membership classes. On the other hand, the number of trips per day negatively affects the duration for entertainment activities, recreational activities, and social activities.

10.5.2 Latent Panel Effects

Since this study used longer time periods and different generations, heterogeneity can inherently exist. The model results confirm that heterogeneity exists among panel cohorts for each type of activity. Result shows variations of effects for age-cohorts as well as the explanatory variables across two Latent Classes (LCs). Two latent classes can be interpreted as two different groups of individuals, with different values, needs, constraints, and capabilities, and hence different time use behavior. For example, male respondents from LC1 spend a longer time in entertainment activities but male respondents from LC2 spend a shorter duration. Individuals from LC1 with a larger household size spend a longer time in social activities, whereas a shorter time is observed for LC2. Similarly, with increase in travelling duration per day, individuals from LC1 spend a longer time in recreational activities, whereas individuals from LC2 spend less time. Class specific effects for each activity types are briefly discussed below:

10.5.2.1 Shopping Activity Duration Model

In the case of the shopping activity model, the majority of the members (56.3%) belong to LC1, and 43.7% of the cohorts fall in LC2. Individuals from LC1 continue spending a longer time in shopping activities even after 65 years old, though the coefficient value for single individuals is higher. Similarly, individuals from LC1 with both low and high monthly personal income and having home ownership spend a longer duration in shopping activities. However, if the total duration of travelling per day increases or if the individuals have a part time job, then the duration of shopping activities is less in LC1. In contrast, individuals from LC2 with both single and married marital status spend less time in shopping activity after 65 years old. As expected, individuals from LC2 with low or higher monthly personal income spend a shorter duration in shopping. In the case of LC2, individuals spend a longer duration if they are a part time worker and if their total duration of travelling is higher.

Table 10.3 LCAH models results for personal characteristics

Variable	Shopping				Entertainment				Recreation				Social			
	LC 1		LC 2		LC 1		LC 2		LC 1		LC 2		LC 1		LC 2	
	β	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***
Constant	7.516	0.000	5.926	0.030	7.014	0.000	7.362	0.000	5.963	0.000	6.153	0.000	14.052	0.000	5.233	0.000
MAge27	0.090	0.385	-0.087	0.561	0.521	0.171	-0.789	0.000	0.066	0.674	0.007	0.929	1.666	0.000	-1.218	0.000
MAge32					0.346	0.116	-0.470	0.000					0.568	0.017	0.187	0.387
MAge37	0.191	0.060	0.125	0.441	-0.301	0.883	-0.368	0.000	-0.335	0.040	0.235	0.007	0.972	0.000	0.155	0.478
MAge42	0.233	0.020	0.029	0.871	0.350	0.014	-0.273	0.013	-0.118	0.394	0.345	0.001	1.044	0.000	0.000	0.999
MAge47	0.345	0.001	0.157	0.368	-0.478	0.949	0.879	0.127	0.125	0.512	0.312	0.104	0.990	0.000	-0.092	0.702
MAge52	0.339	0.001	0.184	0.273	0.086	0.943	-0.637	0.000	-0.338	0.013	0.458	0.000	1.168	0.000	-0.475	0.071
MAge57	0.270	0.023	0.219	0.190	0.198	0.366	-0.486	0.017	-0.030	0.831	0.447	0.000	1.314	0.000	-0.561	0.071
MAge62	0.330	0.007	0.085	0.633	0.434	0.699	-0.965	0.000	-0.127	0.325	0.347	0.000	1.139	0.000	0.136	0.596
MAge67	0.330	0.007	-0.026	0.893	-0.215	0.644	-1.065	0.000	-0.076	0.582	0.312	0.001	1.346	0.000	-0.162	0.858
MAge72	0.308	0.011	-0.196	0.371	-0.062	0.000	-0.708	0.000					1.294	0.000	-0.289	0.530
MAge77	0.274	0.029	-0.012	0.973	0.569	0.812	-1.144	0.000	-0.487	0.005	0.146	0.159	1.353	0.000	-0.262	0.493
MAge82	0.241	0.060	-0.176	0.611	0.405	0.000	-1.143	0.000					1.374	0.000	-0.892	0.005
Male	-0.337	0.000	-0.443	0.000	0.052	0.483	-0.036	0.486	0.089	0.268	0.204	0.007	0.123	0.281	-0.030	0.912

Table 10.4 LCAH models results for socio-demographic characteristics

Variable	Shopping				Entertainment				Recreation				Social			
	LC 1		LC 2		LC 1		LC 2		LC 1		LC 2		LC 1		LC 2	
	β	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***
EduPrim									0.261	0.004	-0.266	0.016				
EduCollg	0.050	0.375	0.045	0.654	0.104	0.289	0.024	0.065					-0.104	0.223	0.034	0.793
PSingle	0.618	0.000	-0.739	0.049	0.083	0.177	0.009	0.123	-0.103	0.236	0.048	0.646	0.163	0.208	-0.227	0.254
PMarried	0.378	0.006	-0.535	0.172					0.226	0.160	-0.474	0.000				
INCM3	0.434	0.000	-0.748	0.024	0.148	0.135	0.017	0.342	0.140	0.256	0.177	0.037				
INCM5	0.340	0.002	-0.324	0.046					0.366	0.027	-1.123	0.000	0.770	0.000	-0.617	0.007
INCM6					1.031	0.000	0.009	0.750	0.652	0.000	-0.699	0.000	-0.307	0.071	-0.046	0.873
INCM7	0.215	0.157	-0.183	0.597												
PPartTym	-0.141	0.236	0.744	0.000												
PFullTym									-0.281	0.033	0.214	0.039				
PTenant					-0.457	0.002	-0.142	0.026	-0.034	0.751	0.499	0.000				
PHomOwn	0.351	0.001	-0.516	0.002									-0.271	0.069	-0.094	0.585
HH Size					-0.286	0.000	-0.183	0.000	0.021	0.632	0.021	0.498	0.126	0.004	-0.046	0.524
LAAlone	0.461	0.000	0.562	0.009					-0.488	0.008	0.128	0.170				
LASpouse	0.406	0.000	0.209	0.250	0.489	0.000	0.214	0.000	-0.613	0.000	0.291	0.003	0.376	0.002	-0.310	0.372

Table 10.5 LCAH models results for activity and trip attributes

Variable	Shopping				Entertainment				Recreation				Social			
	LC 1		LC 2		LC 1		LC 2		LC 1		LC 2		LC 1		LC 2	
	β	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***
DuraInHom	-0.003	0.001	-0.001	0.664	-0.001	0.000	-0.001	0.000	-0.002	0.061	-0.001	0.000	-0.008	0.000	0.003	0.609
DuraOutHm	-0.001	0.422	-0.002	0.301									-0.005	0.000	0.001	0.356
DuraShop					-0.003	0.590	-0.001	0.001	-0.003	0.013	-0.002	0.000				
DuraWork	-0.002	0.000	-0.001	0.003	-0.001	0.000	-0.006	0.000	-0.002	0.000	-0.001	0.000	-0.002	0.000	-0.001	0.017
DuraEnter	-0.001	0.000	0.004	0.021									-0.004	0.000	-0.001	0.216
DuraRec	-0.005	0.000	0.001	0.841	0.004	0.000	-0.007	0.000								
DuraSeden									-0.003	0.365	-0.002	0.000				
DuraSocial					0.005	0.000	-0.001	0.000								
TTWork	-0.008	0.000	-0.005	0.032												
TTDura	-0.002	0.018	0.005	0.083					0.003	0.003	-0.001	0.010	-0.007	0.000	-0.002	0.204
TTrips					0.001	0.327	-0.001	0.000	-0.093	0.001	-0.082	0.000	-0.012	0.000	-0.011	0.719

Table 10.6 LCAH models configuration

Variable	Shopping				Entertainment				Recreation				Social			
	LC 1		LC 2		LC 1		LC 2		LC 1		LC 2		LC 1		LC 2	
	β	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***	**β**	***p-value***
$1/P$	0.402	0.000	0.519	0.000	0.405	0.000	0.462	0.000	0.596	0.000	0.309	0.000	0.533	0.000	0.487	0.000
Class membership	**0.563**	0.000	0.437	0.000	0.481	0.000	**0.519**	0.000	0.493	0.000	**0.507**	0.000	**0.616**	0.000	0.384	0.000
LL Function	-940.39				-928.78				-833.03				-784.91			
LL (b=0)	-1289.08				-1069.12				-1116.85				-947.59			
Parameters	63				59				59				57			
AIC	2.088				1.918				1.853				2.365			
BIC	2.407				2.201				2.151				2.731			

10.5.2.2 Entertainment Activity Duration Model

The class membership for LC1 and LC2 for entertainment activity duration is 48.1% and 51.9% respectively. Up to age 45 years, the coefficient values are positive in LC1. From 50 years to 65 years old and 75 years to 82 years old, individuals from LC1 spend a longer time in entertainment activities. However, male respondents spend a longer time for entertainment activities in LC1. Individuals from LC1 spend a longer duration in entertainment activities if they spend a longer duration in recreational and social activities. If the number of trips per day increases, individuals from LC1 have higher time allocation to entertainment activities. In the case of LC2, the coefficient values are negative up to 45 years. From 50 years to 65 years old and 75 years to 82 years old, individuals spend less time in LC2. Male respondents from LC2 spend a shorter duration. Time spent on recreational activities and social activities is found to negatively affect the entertainment activity duration in LC2. Similarly, if the number of trips per day increases, individuals from LC2 have less time for entertainment activities.

10.5.2.3 Recreational Activity Duration Model

In the case of recreational activities, the class membership shows the existence of two almost equal groups in the sample population: LC1 (49.3%) and LC2 (50.7%). In the case of LC1, the duration of out-of-home recreational activity decreases with an increase in age. Individuals with primary education only spend more time in recreational activity, whereas full time workers spend a shorter duration in LC1. Married individuals and individuals with higher personal monthly income spend a longer duration in recreational activity in LC1. Similarly, individuals from LC1 spend less time in recreational activities with both alone and spouse living arrangements. In contrast, duration of out-of-home recreational activity increases with increase in age in LC2. Individuals from LC2 with primary education only spend less time in recreational activities. However, individuals spend a longer time in recreational activity if they are full time workers in LC2. Individuals spend a shorter duration with an increase in personal monthly income or if the individuals are married in LC2. In the case of LC2, individuals spend a longer duration with both living arrangements, alone and with spouse.

10.5.2.4 Social Activity Duration Model

With regard to social activity duration, the class memberships for LC1 and LC2 are 61.6% and 38.4% respectively. Individuals from LC1 spend a longer duration in social activities with an increase in age and also if they are single. Both low and higher personal monthly income is associated with a longer duration in social activities in LC1. However, if the individuals obtained high school degree, then the duration of social activity is less for LC1. Individuals from LC1 spend more time in social activities if the household size increases and also if they live with a spouse. In contrast, individuals from LC2 spend less time in social activities with an increase in age and if they are single. Individuals having low or higher monthly personal income spend a shorter duration in social activities for LC2. But if the individuals obtained high school degree, then it is more likely that they will spend a longer duration in social activities. Individuals from LC2 spend less time in social activities if the household size increases and also if they live with a spouse. Compared to the earlier work of the authors (Daisy et al. 2015), the model fits increased substantially due to the application of population synthesis for sample expansion.

10.6 Conclusion

This study contributes significantly to travel behavior analysis research by introducing an innovative population synthesis approach for pseudo-panel based duration modeling of out-of-home discretionary activities over a long time period. The population synthesizer expanded the sample of each year to 5% of the population, yielding a larger sample size. Consequently, a higher number of cohorts could be generated to construct the pseudo panel datasets for empirical analysis. Model results demonstrate the effects of key socio-economic characteristics and activity attributes for individuals' activity patterns spanning an 18 year period. It was found that up to age 45 years, the time allocations to shopping, recreation, and social activities increase, whereas the duration of entertainment activities decreases. In contrast, the middle age group (50 years to 65 years) allocated more time for shopping, recreation, and social activities but less time for entertainment activities, compared to other age groups. Both single and married individuals spend a longer duration in shopping activity. Personal monthly income has a significant influence on all types of out-of-home discretionary activities. It is interesting to note that time spent on social activities decreases as entertainment activity increases. Also, time spent on sedentary activities decreases as duration of recreational activities increases. Longer travel time to work is negatively associated with shopping duration. Most importantly, a higher number of trips per day is negatively associated with all three of entertainment, recreational, and social activities.

The models' results confirm the existence of a substantial amount of latent heterogeneity across the sample population. To summarize, the main contributions of this study are two-fold: a) development of a population synthesis based method of pseudo panel modeling of activity duration, and b) examination of intrinsic heterogeneity in pseudo panel data using the Latent Class Accelerated Hazard (LCAH) procedure to capture panel effects. This study demonstrates that the pseudo panel methodology can be utilized for exploring long-term longitudinal dynamics of activity behavior when a panel travel survey is absent.

Chapter 11 Conclusion

11.1 Summary

Over the past few decades, activity-based modeling approaches have received more attention relative to traditional travel demand models. The intent of this paradigm shift is largely to capture the more realistic travel behavior of individuals. Specifically, why, where, and how people make those travel decisions is determined by a set of available resources (such as car ownership) and constraints (such as coordinating with family members). Where people travel and undertake their activities is also related to the land use pattern and urban form of a city, as well as the location of the home, work, shopping, and recreational opportunities. Activity-travel behavior is difficult to estimate, yet important to understand, because, in combination, individuals' travel decisions affect the performance of a transportation system. Even though numerous econometric models have been developed for travel behavior analysis, compared to other alternative approaches, there is still uncertainty related to model prediction, behavioral changes, and model input. This study aimed to develop a new modeling framework to better understand and model activity-travel behavior so that the relative uncertainty associated with mobility decisions of travelers can be captured. To this end, this dissertation presented the Scheduler for Activities, Locations, and Travel (SALT) travel demand model, including a disaggregated advanced econometric modeling framework to micro-simulate behavioral mechanisms and mobility decisions of population groups in the region. The proposed modeling framework can be used for exploring the short-term and longitudinal activity-travel behavior decisions to undertake an activity in-home or out-of-home. Also, the factors that contribute to activity-travel decisions are modeled through the development of econometric models and synthetic pseudo-panel formation.

The SALT model adopts behaviorally realistic daily-activity pattern-based population segmentation techniques in its core modeling system. The first phase of this study developed an activity generation model that can estimate the activity type choices for use in agenda formation along with correlation matrix estimation. The second phase of the study entailed the development of trip chaining, tour frequency, and multimodal tour mode choice models for distinct worker and non-worker clusters within the SALT model. The next phase of the study contributed by modeling the activity-travel behavior of university population segments as a special trip generator in the regional travel demand models. The final phase of the study offered an innovative synthetic pseudo-panel methodology to study the longitudinal travel behavior with its application to activity participation and activity duration.

11.2 Conclusions of Research Findings

The findings of this research offer deeper insights for modeling the activity-travel behavior mechanism and mobility decisions of travelers to advance travel demand management. Conclusions drawn from the results of this research are outlined in the following.

11.2.1 Activity Participation

This phase aimed to model activity type choice of worker and non-worker clusters through a Cluster-based Multivariate Probit Model (C-MVP). The C-MVP models included five worker clusters, which were labelled as extended day workers, 8 am to 4 pm workers, shorter work-day workers, 7 am to 3 pm workers, and 9 am to 5 pm workers, and the non-student and non-worker clusters, which were labelled as non-worker with mid-day activities, non-worker with evening activities, non-worker with stay-at-home pattern, non-worker with morning shopping activity, and non-worker with afternoon shopping activity. For the model estimation, it was assumed that a non-zero correlation exists between the types of activities in which individuals participate. It was hypothesized that the choices of activity types in a day are not mutually exclusive, and therefore the error terms of the activity type choices may be mutually inclusive and correlated. Activity participation in one activity is associated with both the previous and succeeding activity. This interdependency between activities can be captured with a correlation matrix in the C-MVP model. In addition, the trip chaining behavior of worker and non-worker clusters was explored using transition matrices. Model results revealed that using the MVP model with GHK estimator, instead of a randomized sample, resulted in a superior convergence of the model. Another finding was that using the homogeneous daily-activity pattern as an input for C-MVP modeling can increase the accuracy of model predictions in terms of the utility function generation. This suggested that an explicit unifying framework with homogeneous daily-activity time-use patterns could be a possible pathway to model activity-travel behavior of travelers, instead of latent segmentation assumptions. The empirical model results indicate that each cluster has a separate utility function for each activity type. In addition to the socio-demographic and household attributes, it has been found that built-environment and land use characteristics have a significant influence on the decision-making process of participating in a particular activity. Furthermore, the C-MVP model, compared to alternative econometric models such as multinomial logit and multinomial probit, can estimate the separate utility function for individuals and simultaneously capture the correlation between activities in the traveler agenda. Additionally, transition matrices were estimated to understand the correlation among in-home and out-of-home activities.

11.2.2 Tour Complexity and Mode Choice

Given the predicted set of activities in the traveler agenda, this phase aimed to model trip chaining, tour frequency, and multimodal tour mode choice of distinct population groups through advanced econometric models within the SALT model. In total, 19 categories of tour types were defined to form the simple and complex tours in the travelers' activity agenda. Furthermore, 10 mode choices, including unimodal and multimodal choice sets, were derived from the SR survey for modeling tour mode choice. A series of Poisson regression models, Ordered Probit models, and Multinomial Logit (MNL) models were developed to better understand the tour complexity and mode choice behavior of population clusters. These model results provided deeper insights regarding the complex relationship among trip chaining, tour frequency, and multimodal tour mode choice phenomena and socio-demographics, residential locations, and built environment characteristics. The Poisson regression model has been utilized earlier for trip and tour study mostly in the aggregated manner; however, to the best of the author's knowledge, no investigation has explored the application of this model for modeling tour frequency with distinct disaggregated population clusters. One of the advantages of using the Poisson regression model compared to other alternative econometric models is that tours are countable and it treats the number of tours as a non-negative integer. For trip chaining, it is hypothesized that with the increase in more trips per chain, the tour complexity increases. Therefore, an ordered probit model was developed instead of other alternative count models such as negative binomial and Poisson regression. Furthermore, to model multimodal choices, the MNL model was developed instead of the nested logit model. Compared to earlier works on tour complexity and mode choice modeling, this study contributed by further exploring how, for example, an extended-day worker behaves differently than a shorter-day worker or workers who typically try to avoid travel in the off-peak hours. Most importantly, results showed the utility function and sign of the covariates in each model to signify the presence of heterogeneity between different population groups, thus confirming the utility of clustering homogeneous daily activity time-use patterns to increase the accuracy of the model predictions.

11.2.3 Mode Specific Trip Frequency for University Population

This stage aimed to model the activity-travel behavior of students, staff, and faculty at a large Canadian university using a two-state zero-inflated negative binomial model. To date, the activity-travel behavior in the context of Canadian universities has not yet been explored. The university population can be considered as a sub-population with special travel behavioral characteristics in the regional travel demand models. However, in each cycle of the General Social Survey (GSS), the university populations, especially students, are under-represented. Therefore, this research contributed by designing and collecting an inventory of activity diaries of the university population, along with the activity-travel behavior analysis and empirical model building. Model results suggested that the mean trip length of home to school trip is significantly shorter for students in comparison to the faculty and staff. It also found that auto-drive dominates over other modes for entertainment activities, sports, and hobby-related activities. These results provide deeper insights for campus-based sustainable transportation system provision and policy implementation. The empirical models built in this study can improve the travel demand prediction for university commuters in the regional travel demand model with better accuracy and precision.

11.2.4 Longitudinal Synthetic Pseudo-Panel Framework

Longitudinal panel travel diary information is mostly unavailable due to sample attrition problems and the costs. Therefore, this stage of study aimed to develop an innovative synthetic pseudo panel framework to better understand and analyze the longitudinal pattern of activity-travel behavior using public data available from the Canadian General Social Survey (GSS). Before the cohort formation, the samples from each GSS cycle were expanded to 5% of the base year population by utilizing a population synthesis technique. The time-invariant characteristics of birth year, gender, and educational level of individuals were utilized to form the cohorts. Synthetic pseudo panels were then formed and an advanced latent-class accelerated hazard-based duration model with two latent classes was utilized to incorporate the heterogeneity among population groups over the years. Pseudo panels have been utilized earlier for other transport related research areas such as transit ridership study or emissions, but to the best of the author's knowledge, no investigation has explored the application of this technique to longitudinal modeling of activity participation and activity duration. The proposed modeling framework can be applied for longitudinal activity-travel behavior analysis, where genuine panel data are absent. For instance, comparison among activity-travel patterns of population groups over longer time horizons using the synthetic pseudo panel within the SALT modeling framework can be done.

11.3 Model Implementation

In this study, numerous advanced micro-behavioral models were developed within the Scheduler for Activities, Locations, and Travel (SALT) travel demand model system. Each of the models developed can be implemented independently and can be utilized to simulate the activity-travel decisions of individuals representative of specific population cohorts. The SALT system can be implemented at a fine disaggregated geographic scale if the required activity-travel diary data are available. Furthermore, the developed models can be applied to predict the traveler decisions for transit route choices, mode choices, time of the day choices, etc. The activity generation module can be straightforwardly applied to predict the daily trip generation of different demographic cohorts. The accurate prediction of both tour mode choice and trip-chaining have important applications in policy making and decisions, particularly related to the modal shift from auto to transit or active transportation. This module can also be interlinked with further modules within the SALT system to estimate the emission factors and energy consumption of travelers for a 24-hour period.

Overall, the SALT system is a robust and comprehensive activity-pattern based model that can be utilized for further advanced travel demand management at the regional municipal, or district level, with further required estimations and validations. It offers a unifying framework for modeling the travel demands of homogeneous population cohorts derived from the activity-travel decision-making process. The process centers on behavioral relationships between independent socio-demographic and built-form variables and activity participation, trip chaining, tour frequency, and tour mode choice. Micro-behavioral models developed in this study can be applied to other areas such as residential location and vehicle ownership modeling. The SALT system has the potential to make important contributions to active travel research, policy, planning, and practice.

11.4 Recommendations for Future Work

This dissertation and the models developed within it are not without their limitations. Therefore, the following recommendations are offered for the continuation of the work presented herein. The models developed in this study are mostly derived from one or two-day travel diaries. As traveler decision making is a complex phenomenon, the activity diary of one full week would be ideal to better model and predict the short-term decisions of individuals. The pseudo panel study can be improved by adding the locational data of individuals, which are related to long-term choices such as work location and home location, and analyzing how their characteristics influence the activity participation or time allocation to discretionary activities over the longer period. Even though this study considered in-home activities, certain activities formerly completed out-of-home may increasingly be completed in the home, such as online shopping, and these can be explicitly incorporated into the activity generation module for further improvement of the research.

In addition to future work, there are several prospects to extend the current work. For instance, extending the activity type choice models to activity time allocation would improve the scheduling module in the activity-based travel demand modeling system. Moreover, initially, travel diary data can be expanded through the population synthesis technique, before modeling activity generation and tour formation. Another extension of this work could be to incorporate the e-shopping, e-banking etc. time uses in activity participation models or to add weather components (visibility, daylight etc.) and the availability of autonomous or self-driving vehicles in the mode choice models. This study can also be extended by incorporating the activity location choices for different activity purposes. Furthermore, household interactions can be considered in tour level and trip level modeling. Models developed in this study can be connected to route choice algorithms to determine the actual travel times and costs for home-based tours. Another way would be data fusion between STAR and EnACT to better understand lifestyle choices, which might influence activity generation, trip chaining, and mode choices. This thesis fills the gap in travel behavior knowledge by offering a unifying disaggregated modeling framework to model short-term and long-term activity and travel decisions of representative cohorts. To this end, this thesis adds both substantively and methodologically to the literature.

References

Aas, D. Designs for large scale, time use studies of the 24-hour day. Z. Staikov, (ed.), It's about time: International Research Group on time budgets and social activities (pp. 17-53). Sofia: Institute of Sociology at the Bulgarian Academy of Sciences, Bulgarian Sociological Association, 1982.

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

Akar, G., Chen, N., and Gordon, S. I. Influence of neighborhood types on trip distances: Spatial error models for central Ohio. International Journal of Sustainable Transportation, No. 10, 2016, pp. 284-293.

Akar, G., Flynn, C., and Namgung, M. Travel choices and links to transportation demand management. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2319, 2012, pp. 77-85.

Allahviranloo, M., and Recker, W. A multivariate Probit model of activity participation behavior. Proceedings of the 93rd Annual Meeting of Transportation Research Board (TRB), Washington, D.C., USA., 2014.

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

Arentze, T. A., and Timmermans, H. J. P. A need-based model of multi-day, multi-person activity generation. Transportation Research Part B: Methodological, Vol. 43(2), 2009, pp. 251-265.

Arentze, T. A., and Timmermans, H. J. P. Albatross version 2: A learning-based transportation oriented simulation system. Eindhoven, The Netherlands: European Institute of Retailing and Services Studies, 2005.

Arentze, T. A., and Timmermans, H. J. P. ALBATROSS: A learning based transportation oriented simulation system. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1706, 2000, pp. 136-144.

Arentze, T. A., and Timmermans, H. J. P. ALBATROSS: A learning-based transportation-oriented simulation system. ElRASS, Eindhoven University of Technology, The Netherlands, 2000.

Arentze, T., Hofman, F., Joh, C., and Timmermans, H. The development of ALBATROSS: Some key issues. Brilon, W., Huber, F., Schreckenberg, M., Wallentowitz, H. (Eds.), Traffic and Mobility. Berlin: Springer, 1999, pp. 57-72.

Auld, J., and Mohammadain, A. Activity planning processes in the agent-based dynamic activity planning and travel scheduling (ADAPTS) model. Transportation Research Part A: Policy and Practice, Vol. 46(8), 2012, pp. 1386-1403.

Auld, J., and Mohammadian, A. Framework for the development of the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. Transportation Letters, Vol. 1 (3), 2009, pp.245-255.

Auld, J., and Mohammadian, A., and Wies, K. Population synthesis with sub-region-level control variable aggregation. ASCE Journal of Transportation Engineering, Vol. 135, No. 9, 2009, pp.632-639.

Axhausen, K. and Garling, T. Activity-based approaches to travel analysis: Conceptual frameworks, models, and research problems. Transport Reviews, Vol. 12, No. 4, 1992, pp. 323-341.

Axhausen, K. W., A. Zimmermann, S. Schonfelder, G. Rindsfuser, and T. Haupt. Observing the rhythms of daily life: A six-week travel diary. Transportation, Vol. 29, No. 2, 2001, pp.95-124.

Azari, K. A., Arintono, S., Hamid, H., and Davoodi, S.R. Evaluation of demand for different trip purposes under various congestion-pricing scenarios. Journal of Transport Geography, Vol. 29, 2013, pp. 43-51.

Bagley, M. N., and Mokhtarian, P. L. The impact of residential neighborhood type on travel behavior: a structural equation modeling approach. Annual Regional Science, Vol. 36, 2002, pp.279-297.

Balsas, C. J. L. Sustainable transportation planning on college campuses. Transport Policy, Vol. 10, No. 1, 2003, pp.35-49.

Baltagi, H. B. Econometric analysis of panel data, New York: John Wiley, 1995.

Becker, G. S. A Theory of the allocation of time. Economic Journal, Vol. 75, 1965, pp.493 517.

Becker, G. S. The economic approach to human behavior. The University of Chicago Press, Chicago, IL, 1976.

Ben-Akiva M. E. and Lerman, S. R. Discrete choice analysis: Theory and application to travel demand. MIT Press, Cambridge, MA, 1985.

Berg, P. V. D., Arentze, T., and Timmermans, H. A latent class accelerated hazard model of social activity duration. Transportation Research Part A: Policy and Practice, Vol. 46, 2012, pp.12-21.

Bernard, L. T., Bolduc, D., and Yameogo, N. D. A Pseudo-panel data model of household electricity demand. Resource and Energy Economics, Vol. 33, No. 1, 2011, pp.315-325.

Bhat, C. R. A hazard-based duration model of shopping activity with nonparametric baseline specification and nonparametric control for unobserved heterogeneity. Transportation Research Part B: Methodological, Vol. 30, No. 3, 1996, pp.189-207.

Bhat, C. R. A model of post home-arrival activity participation behavior. Transportation Research, Vol. 32, No. 6, 1998, pp.387-400.

Bhat, C. R. A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. Transportation Part B, Vol. 39, 2005, pp. 679-707.

Bhat, C. R. and Koppelman, F. S. A retrospective and prospective survey of time-use research. Transportation, Vol. 26(2), 1999, pp.119-139.

Bhat, C. R. and Koppelman, F. S. A conceptual framework for individual activity program generation. Transportation Research Part A: Policy and Practice, Vol. 27, No. 6, 1993, pp.443-446.

Bhat, C. R. and Koppelman, F. S. Activity-based modeling of travel demand. R.W. Hall (ed.) The Handbook of Transportation Science, Kluwer Academic Publishers, Norwell, Massachusetts, 1999, pp. 35-61.

Bhat, C. R. and Misra, R. Comprehensive activity travel pattern modeling system for non-workers with empirical focus on organization of activity episodes. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1777, 2001, pp. 16-24.

Bhat, C. R. and Misra, R. Nonworker activity-travel patterns: Organization of activities. Presented in the 79th Annual Meeting of the Transportation Research Board (TRB), Washington, D.C., USA., 2000.

Bhat, C. R., and Misra, R. Discretionary activity time allocation of individuals between in-home and out-of-home and between weekdays and weekends. Transportation, Vol. 26, No. 2, 1999, pp. 193-209.

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

Bhat, C. R., and Srinivasan, S. A multidimensional mixed ordered-response model for analyzing weekend activity participation. Transportation Research Part B: Methodological, Vol. 39(3), 2005, pp. 255-278.

Bhat, C. R., Carini, J. P. and Misra, R. On modeling the generation and organization of household activity stops. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1676, 1999, pp. 153-574.

Bhat, C. R., Guo, J. Y., Srinivasan, S., and Sivakumar, A. A comprehensive econometric micro-simulator for daily activity-travel patterns (CEMDAP). Transportation Research Record: Journal of the Transportation Research Board, Vol. 1894, 2004, pp. 57-66.

Bhat, C. R., Srinivasan, S., and Axhausen, K.W. An analysis of multiple interepisode durations using a unifying multivariate hazard model. Transportation Research Part B: Methodological, Vol. 39(9), 2005, pp.797-823.

Black, J., Mason, C., and Stanley, K. Travel demand management: Policy context and an application by the university of New South Wales (UNSW) as a large trip generator. Transport Engineering in Australia, Vol. 5, No. 2, 1999, pp.1-11.

Bowman, J. L. The day activity schedule approach to travel demand analysis. PhD thesis, Massachusetts Institute of Technology, 1998.

Bowman, J. L., and Ben-Akiva, M. Activity-based disaggregate travel demand model system with activity schedules. Transportation Research Part A: Policy and Practice, Vol. 35, No. 1, 2001, pp.1-28.

Bowman, J. L., and Bradley, M. A. Activity-based travel forecasting model for SACOG: Technical Memos Numbers 1-11, 2005-2006. URL http://jbowman.net.

Bradley, M. A., Portland Metro, Bowman, J. L., and Cambridge Systematics. A system of activity-based models for Portland, Oregon. USDOT report number DOT-T-99-02, produced for the Travel Model Improvement Program of the USDOT and EPA, Washington, D.C, 1998.

Bradley, M., Bowman, J. L., and Griesenbeck, B. Activity-based model for a medium-sized city: Sacramento. Traffic Engineering and Control, Vol. 50(2), 2009, pp. 73-79.

Bradley, M., Bowmanm, J. L., and Griesenbeck, B. Development and application of the SACSIM activity-based model system. Presented in the 11th World Conference on Transport Research, Berkeley, USA, 2007.

Bradley, M., Outwinter, M., Jonnalagadda, N., and Ruiter, E. Estimation of an activity based micro simulation model for San Francisco. Proceedings of the 80th Annual Meeting of the Transportation Research Board (TRB), Washington, D.C., USA., 2001.

Bricka, S. Trip chaining: Linking the influences and implications. PhD thesis, University of Texas, 2008, Austin, TX, USA.

Castiglione, J., Bradley, M., and Gliebe, J. Activity-based travel demand models: A primer. TRB's second Strategic Highway Research Program (SHRP 2) Report S2-C46-RR-1, 2015.

Castro, M., Eluru, N., Bhat, C. R., and Pendyala, R. M. Joint model of participation in non-work activities and time-of-day choice set formation for workers. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2254, 2011, pp. 140-150.

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

Chapin, F. S. Jr. Human activity patterns in the city: Things people do in time and space. John Wiley and Sons, London, 1974.

Chen, C., and Mokhtarian, P. L. Tradeoffs between time allocations to maintenance activities/travel and discretionary activities/travel. Transportation, Vol. 33, 2006, pp.223-240.

Chen, Y. J., and Akar, G. Using trip chaining and joint travel as mediating variables to explore the relationships among travel behavior, socio-demographics, and urban form. The Journal of Transport and Land Use, Vol. 10(1), 2017, pp. 573-588.

Chu, Y. L. Modeling workers' daily non-work activity participation and duration. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1926, 2005, pp. 10-18.

Chu, Y. Modeling workers' daily out-of-home maintenance activity participation and duration. Proceedings of the 96th Annual Meeting of the Transportation Research Board (TRB), Washington, D.C., USA., 2017.

Chung, J., and Goulias, K. G. Sample selection bias with multiple selection rules: An Application with residential relocation, attrition, and activity participation in the Puget sound transportation panel. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1493, 1995, pp.128-135.

Chung, J., and Goulias, K. G. Travel demand forecasting using microsimulation: Initial results from a case study in Pennsylvania. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1607, 1997, pp.24-30.

Clarke, M. I. Activity modeling: A research tool or a practical planning technique? Behavioral Research for Transport Policy, pp. 3-15, 1986, VNU Science Press, Utrecht, The Netherlands.

Clarke, M., Dix, M., and Goodwin, P. Some issues of dynamics in forecasting travel behavior-A discussion paper. Transportation, Vol. 11(2), 1982, pp.153-172.

Crane, R. The influence of urban form on travel: An interpretive review. Journal of Planning Literature, Vol. 15, 2000, pp.3-23.

Cullen, I. and Godson, V. Urban networks: The structure of activity patterns. Progress in Planning, Vol. 4, 1975, pp.1-96.

Cullen, I. and Phelps, E. Diary techniques and the problems of urban life. Final Report submitted to the Social Science Research Council, Grant HR 2336, London, 1975.

Daisy, N. S. et al. A pseudo panel investigation of out-of-home discretionary activity participation. Presented in 94th Annual Meeting of Transportation Research Board, Washington D.C., 2015.

Daisy, N. S. et al. Population synthesis based pseudo panel modeling of out-of-home discretionary activity duration. Peer reviewed proceedings of the 95th Annual Meeting of Transportation Research Board (TRB), Washington, D.C., USA., 2016.

Daisy, N. S. Modeling activity-travel behavior for activity-based travel demand modeling. Doctoral Research in Transport Modeling. Presented at the 97th Annual Meeting of Transportation Research Board (TRB), Washington, D.C., USA., 2018a.

Daisy, N. S., Hafezi, M. H., Liu, L. and Millward, H. Housing location and commuting mode choices of university students and employees: An application of bivariate Probit models. Peer reviewed ASCE proceedings of the International Conference on Transportation and Development (ICTD 2018). Pittsburgh, Pennsylvania, USA., 2018b.

Daisy, N. S., Hafezi, M. H., Liu, L., and Millward, H. Understanding and modeling the activity-travel behavior of university commuters at a large Canadian university. Journal of Urban Planning and Development. 144(2)., 2018c.

Daisy, N. S., Liu, L. and Millward, H. Analyzing tours: Application of a traveler grouping based cluster analysis. Presented at the 53rd Canadian Transportation Research Forum (CTRF). Winnipeg, Canada., 2017a.

Daisy, N. S., Liu, L., and Millward, H. Optimizing daily travel sequences and time-use patterns of individuals. Presented at the 53rd Canadian Transportation Research Forum (CTRF). Winnipeg, Canada., 2017b.

Daisy, N. S., Millward, H., Hafezi, M. H., and Liu., L. Trip-chaining and tour complexity: contrasts between worker and non-worker groups in Halifax, Nova Scotia. Presented at the 28th Atlantic Division of the Canadian Association of Geographers (ACAG/ACGA) Conference: Saint Mary's University, Halifax, Nova Scotia., 2017c.

Daisy, N. S., Millward, H. and Liu, L. Analyzing time windows and time allocation to in-home and out-of-home activities in workers' activity patterns. Proceedings of the 53rd Canadian Transportation Research Forum (CTRF). 2018a, Ottawa, Canada.

Daisy, N. S., Millward, H., and L. Liu. Individuals' activity-travel behavior in travel demand models: A review of recent progress. Peer reviewed ASCE proceedings of the 18th COTA Conference International Conference of Transportation Professionals (CICTP). 2018b, Shanghai, China.

Daisy, N. S., Millward, H. and Liu, L. Out-of-home activity choices and activity transitions for non-worker population groups. Proceedings of the 15th International Conference on Travel Behavior Research, 2018c, Santa Barbara, California USA.

Damm, D. Interdependencies in activity behavior. Transportation Research Record: Journal of the Transportation Research Board, Vol. 750, 1980. pp. 33-40.

Dargay, J. M. Determinants of car ownership in rural and urban areas: A pseudo-panel analysis. Transportation Research E, Vol. 38, 2002, pp.351-366.

Dargay, J. M. The Effect of prices and income on car travel in the UK. Transportation Research A, Vol. 41, No. 10, 2007, pp.949-960.

Dargay, J. The effect of income on car ownership: evidence of asymmetry. Transportation Research Part A: Policy and Practice, Vol. 35, 2001, pp.807-821

Dargay, J., and Vythoulkas, P. Estimation of a dynamic car ownership model: A pseudo-panel approach. Journal of Transport Economics and Policy, Vol. 33, No. 3, 1999, pp.287-302

Davidson, W., Donnelly, R., Vovsha, P., Freedman, J., Ruegg, S., Hicks, J., Castiglione, J., and Picado, R. Synthesis of first practices and operational research approaches in activity-based travel demand modeling. Transportation Research Part A: Policy and Practice, Vol. 41(5), 2007, pp.464-488.

Deaton, A. Panel data from time series of cross-sections. Journal of Econometrics, Vol. 30, 1985, pp.109-126.

Dharmowijoyo, D. B., Susilo, Y. O., and Karlstrom, A. The day-to-day variability in travelers' activity-travel patterns in the Jakarta Metropolitan Area (JMA). Presented in the 94th Annual Meeting of Transportation Research Board (TRB), Washington, D.C., USA., 2014.

Duncan, M. How much can trip chaining reduce VMT? A simplified method. Transportation, No. 43, 2015, pp. 1-17.

Ellegard, K., and Svedin, U. Torsten Hagerstrand's time-geography as the cradle of the activity approach in transport geography. Journal of Transport Geography, Vol. 23, 2012, pp. 17-25.

Eom, J. Incorporating activity-based special generator data into a conventional planning model. PhD thesis, North Carolina, United States: North Carolina State University, 2007.

ESRI 2017. ArcGIS desktop: Release 10.5.1. Redlands, CA: Environmental Systems Research Institute. 2017.

Ettema, D. F., Borgers, A. W. J., and Timmermans, H. J. P. A simulation model of activity scheduling heuristics (SMASH). Transportation Research Record: Journal of the Transportation Research Board, Vol. 1413, 1993, pp.1-11.

Ettema, D., Bastin, F., Polak, J., and Ashiru, O. Modeling the joint choice of activity timing and duration. Transportation Research A, Vol. 41(9), 2007, pp. 827-841.

Ettema, D., Schwanen, T., and Timmermans, H. Task Allocation patterns: An assessment of household-level strategies. Presented in the EIRASS conference on Progress in Activity-Based Analysis, Maastricht, The Netherlands, 2004.

Evans, A. On the theory of the valuation and allocation of time. Scottish Journal of Political Economy, Vol. 19, 1972, pp.1-17.

Ewing, R., and Cervero, R. Travel and the built environment. Journal of American Planning Association, Vol. 76 (3), 2010, pp.265-294.

Frank, L. D., Bradley, M., Kavage, S., Chapman, J., and Lawton, T. Urban form, travel time, and cost relationships with tour complexity and mode choice. Transportation, Vol. 35 (1), 2008, pp.37-54.

Frank, L. D., Saelens, B. E., Powell, K. E., and Chapman, J. E. Stepping towards causation: do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Social Science and Medicine journal, Vol. 65, 2007, pp.1898-1914.

Gardes, F., Duncan, G., Gaubert, P., Gurgand, M., and Starzec, C. Panel and pseudo-panel estimation of cross-sectional and time series elasticities of food consumption: the case of U.S. and Polish data. Journal of Business and Economic Statistics, Vol. 23, No. 2, 2005, pp. 242-253.

Garling, T., Brannas, K., Garvill, J., Golledge, R. G., Gopal, S., Holm, E., and Lindberg, E. Household activity scheduling. Transport Policy, Management and Technology Towards 2001. Vol. 4. Ventura, CA: Western Periodicals, 1989, pp. 235-248.

Garling, T., Kwan, M. P., and Golledge, R. G. Computational-process modeling of household travel activity scheduling. Transportation Research Part B: Methodological, Vol. 25, 1994, pp.355-364.

Garrison, W. L. and Worrall, R. D. Monitoring urban travel. Report submitted to NCHRP, September 1996.

Geerken, M. and Gove, W. R. At home and at work: The family's allocation of labor. Sage Publications, Beverly Hills, 1983.

Gelman, A., and Carlin, J. B. Post-stratification and weighting adjustments in survey nonresponse. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little (eds.), pp.289-302, Wiley, New York, 2002.

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

Gliebe, J. P., and Koppelman, F. S. Modeling Household activity-travel interactions as parallel constrained choices. Transportation, Vol. 32(5), 2005, pp.449-471.

Goodwin, P. B. Family changes and public transport use 1984-1987: A dynamic analysis using panel data. Transportation, Vol. 16, 1987, pp.121-154.

Goodwin, P. Have panel surveys told Us anything new? Golob, T. F., Kitamura, R. and Long, L. (eds.), Panels for Transportation Planning, Boston, 1997.

Goulias, K. G, Chen, Y., Bhat, C. R., and Eluru, N. Activity-based microsimulation model system in Southern California: Design, implementation, preliminary findings, and future plans. Presented in the 3rd Conference on Innovations in Travel Modeling, TRB, 2010, Tempe, AZ.

Goulias, K. G. and Kim, T. An analysis of activity type classification and issues related to the with whom and for whom questions of an activity diary. Chapter 14 in Progress in Activity Based Analysis. Harry Timmermans (ed.), Elsevier, 2005, pp. 309-334.

Goulias, K. G., and Kitamura, R. A Dynamic model system for regional travel demand forecasting. Panels for Transportation Planning: Methods and Applications. Golob, T., R. Kitamura, and L. Long (eds.), Kluwer Academic Publishers, Boston, Chapter 13, 1996, pp. 321-348.

Greene, W. Econometric analysis, 6th Edition, 2008, Englewood Cliffs, Prentice Hall.

Greene, W. H. Econometric analysis. 5th edition. Upper Saddle River, NJ: Prentice- Hall, 2003.

Greene, W. H. Econometric Analysis. Pearson, Noida, 2006.

Greene, W. LIMDEP version 9.0: Reference Guide, 2007, Plainview, New York, Econometric Software, Inc.

Guo, J. Y., and Bhat, C. R. Population synthesis for micro-simulating travel behavior. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2014, 2007, pp.92-101.

Habib, K. M. N., El-Assi, W., Hasnine, S., and Lamers, J. Activity-travel behavior of non-workers in the national capital region of Canada: application of a comprehensive utility maximizing system of travel option modeling, Proceedings of the 95th Annual Meeting Transportation Research Board (TRB), Washington, D.C., USA., 2016.

Habib, M. A., and Daisy, N. S. Examining frequency and duration of out-of-home physical activity participation of school-going children. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2357, 2013, pp. 116-124.

Hafezi, M. H., Daisy, N. S., Millward, H., and Liu, L. Commuting to campus: Findings from the Dalhousie EnACT travel survey. Department of Civil and Resource Engineering, Dalhousie University, Halifax, Canada., 2018a.

Hafezi, M. H., Liu, L., and Millward, H. Modeling activity scheduling behavior of travelers for activity-based travel demand model. Presented at the 97th Annual Meeting of the Transportation Research Board (TRB), Washington, D.C., USA., 2018b.

Hafezi, M. H., Liu, L., and Millward, H. A time-use activity-pattern recognition model for activity-based travel demand modeling. Transportation, 2017a, pp.1-26.

Hafezi, M. H., Liu, L., and Millward, H. Identification of representative patterns of time use activity through fuzzy c-means clustering. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2668, 2017b, pp. 38-50.

Hafezi, M. H., Daisy, N. S., Liu, L., and Millward, H. Daily time-use activity patterns at a large Canadian university. Presented at the 96th Annual Meeting of the Transportation Research Board (TRB), Washington, D.C., USA., 2017c.

Hagerstrand, T. Reflections on "What about people in regional science?" Papers in Regional Science, Vol. 66(1), 1989, pp. 1-6.

Hagerstrand, T. What about people in regional science? Papers in Regional Science, Vol. 24, 1970, pp. 7-24.

Hajivassiliou, V. and Ruud, P. Classical estimation methods for LDV models using simulation. Handbook of Econometrics. R. Engle and D. McFadden (eds.), No. 5, 1994, Amsterdam: North-Holland.

Hamed, M. M., and F. L. Mannering. modeling travelers' post-work activity involvement: Toward a new methodology. Transportation Science, Vol. 27(4), 1993, pp. 381-394.

Handy, S. Smart growth and the transportation-land use connection: What does the research tell us? International journal of Regional Science Review, Vol. 28, 2005, pp. 146-167.

Hanson, S., and Huff, J. O. Systematic variability in repetitious travel. Transportation, Vol. 15, 1988, pp.111-135.

Harding C., Miller, E. J., Patterson, Z., and Axhausen, K.W. Multiple purpose tours and efficient trip chaining: An analysis of the effects of land use and transit on travel behavior in Switzerland. Proceedings of the 94th Annual Meeting of the Transportation Research Board, Washington D.C., USA., 2015.

Harvey, A. S., and Taylor, M. E. Activity settings and travel behavior: A social contact perspective. Transportation, Vol. 27(1), 2000, pp. 53-73.

Heideman, C. Spatial behavior studies: Concepts and contexts. P.R. Stopher, A.H. Meyberg and W. Brog (eds.), New Horizons in Travel Behavior Research, 1981, pp. 289-315, Lexington Books, Lexington, Massachusetts.

Hensher, D. A., and Mannering, F. Hazard-based duration models and their application to transport analysis. Transport Reviews, Vol. 14, 1994, pp.63-82.

Hensher, D. A., and Reyes, A. J. Trip chaining as a barrier to the propensity to use public transport. Transportation, No. 27, 2000, pp. 341-361.

Ho, C. Q., and Mulley, C. Multiple purposes at a single destination: A key to a better understanding of the relationship between tour complexity and mode choice. Transportation Research Part A: Policy and Practice, No. 49, 2013, pp. 206-219.

Huang, B. The use of pseudo panel data for forecasting car ownership. PhD thesis, London, University of London, 2007.

Jang, T. Y. Count data models for trip generation. Journal of Transportation Engineering, Vol. 131, No. 6, 2005, pp.444-50.

Joh, C. H., Arentze, T. A. and Timmermans, H. J. P. Measuring and predicting adaptation behavior in multi-dimensional activity-travel patterns. Transportmetrica, Vol. 2, 2006, pp. 153-173.

Jones, P. M. New approaches to understanding travel behavior: The human activity approach. Behavioral travel modeling. Hensher DA, Stopher PR (ed.) Redwood Burn Ltd., London, 1979, pp. 55-80.

Jones, P. M., Dix, M. C., Clarke, M. I., and Heggie, I. G. Understanding travel behavior. Aldershot: Gower, 1983.

Jones, P. M., Koppelman, F. S., and Orfeuil, J. P. Activity analysis: State of the art and future directions. In Developments in Dynamic and Activity-Based Approaches to Travel Analysis, pp. 34-55, Gower, Aldershot, England, 1990.

Jovicic, G. Activity based travel demand modeling-A literature study. Danmarks Transport Forskning, 2001.

Kamruzzaman, M., Hine, J., Gunay, B., and Blair, N. Using GIS to Visualize and evaluate student travel behavior. Journal of Transport Geography, Vo. 19, No. 1, 2011, pp.13-32.

Khattak, A., Wang, X., Vandecar-Burdin, T., and Wilson-John, W. Old Dominion University student travel survey, Final Report: Virginia Department of Transportation. Transportation and Mobility Planning Division, 2011.

Kitamura, R. A model of daily time allocation to discretionary out of- home activities and trips. Transportation Research Part B: Methodological, Vol. 18, No. 3, 1984, pp. 255-266.

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

Kitamura, R., and Susilo, Y. O. Is travel demand insatiable? A study of changes in structural relationships underlying travel. Transportmetrica, No. 1, 2005, pp. 23-45.

Kitamura, R., Lula, C., and Pass, E. I. AMOS: An activity-based flexible and behavioral tool for evaluation of TDM measures. Presented in the 21st PTRC Summer Annual Meeting, University of Manchester, United Kingdom, 1993.

Kitamura, R., Pas, E. I., Lula, V., Lawton, K., and Benson, E. The sequenced activity mobility simulator (SAMS): An integrated approach to modeling transportation, land use and air quality. Transportation, Vol. 23, 1996, pp. 267-291.

Kitamura, R., Yamamoto, T., Fujii, S., and Sampath, S. A discrete-continuous analysis of time allocation to two types of discretionary activities which accounts for unobserved heterogeneity. Lesort JB (ed.) Transportation and Traffic Theory, 1996, pp 431-453. Oxford: Elsevier.

Konduri, K.C., Ye, X., and Pendyala, R.M. Probit-based discrete-continuous model of activity choice and duration with history dependency. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2156, 2010, pp.17-27.

Koohsari, M. J., Sugiyama, T., Sahlqvist, S., Mavoa, S., Hadgraft, N., and Owen, N. Neighborhood environmental attributes and adults' sedentary behaviors: review and research agenda. Preventive Medicine, Vol. 77, 2015, pp. 141-149.

Koohsari, M., Owen, N., Cole, R., Mavoa, S., Oka, K., Hanibuchi, T., and Sugiyama, T. Built environmental factors and adults' travel behaviors: role of street layout and local destinations. Preventive Medicine, Vol. 19, 2017, pp. 124-128.

Kraan M. Time to travel? A model for the allocation of time and money. PhD thesis, Department of Civil Engineering, The Netherlands: University of Twente, 1996.

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

Kuppam, A. R., and Pendyala, R. M. A structural equations analysis of commuter activity and travel patterns. Transportation, Vol. 28(1), 2001, pp. 33-54.

Kurani, K. S. and Lee-Gosselin, M. E. H. Synthesis of past activity analysis applications. Presented in the Travel Model Improvement Program (TMIP) Conference on activity-based travel forecasting, New Orleans, June 2-5, 1996.

Lambert, D. Zero-inflated Poisson regression with an application to defects in manufacturing. Technometrics, Vol. 34, 1992, pp.1-14.

LaMondia, J. J., Blackmar, C. E., and Bhat, C. R. Comparing transit accessibility measures: a case study of access to healthcare facilities. Proceedings of the 89rd Annual Meeting of Transportation Research Board, Washington, DC, USA., 2010.

Lawe, S. DaySim activity-based model implementation for Jacksonville, FL. Presented to the Panel on Activity-based Models, Florida Department of Transportation, 2010.

Lawson, C. T. Household travel/activity decisions. PhD thesis at Portland State University, USA, 1998.

Lee, B., and Timmermans, H. A latent class accelerated hazard model of activity episode durations. Transportation Research Part B: Methodological, Vol. 41, 2007, pp.426-447.

Lee, M., and McNally, M. G. An empirical investigation on the dynamic processes of activity scheduling and trip chaining. Transportation, Vol. 33(6), 2006, pp. 553-565.

Limanond, T., Butsingkorn, T., and Chermkhunthod, C. Travel behavior of university students who live on campus: A case study of a rural university in Asia. Transport Policy, Vol. 18, No. 1, 2011, pp.163-171.

Liu C. Exploring the influence of urban form on travel and energy consumption, using structural equation modeling. PhD thesis, College Park, Maryland: University of Maryland, 2012.

Liu, C., Susilo, Y. O., and Karlstrom, A. Investigating the impacts of weather variability on individual's daily activity-travel patterns: a comparison between commuters and non-commuters in Sweden. Proceedings of the 93rd Annual Meeting of Transportation Research Board, Washington, D. C., USA., 2014.

Liu, L., Hafezi, M. H., and Daisy, N. S. Results of the 2016 Dalhousie environmentally aware travel diaries survey. Dalhousie Transportation and Environmental Simulation Studies (TESS) group, Department of Civil and Resource Engineering, Dalhousie University, Halifax, Canada.

Lu, X., and Pas, E. I. Socio-demographics, activity participation and travel behavior. Transportation Research Part A: Policy and Practice, Vol. 33(1), 1999, pp. 1-18.

Ma, J., and Goulias, K. G. Application of Poisson regression models to activity frequency analysis and prediction. Transportation Research Record: Journal of Transportation Research Board, Vol. 1676, 1999, pp. 86-94.

Ma, Lu. Generating disaggregate population characteristics for input to travel-demand models. PhD thesis. University of Florida, 2011.

Maat, K., and Timmermans, H. Influence of land use on tour complexity: A Dutch case. Transportation Research Record: Journal of Transportation Research Board, Vol. 1977, 2006, 234-241.

Maat, K., Van Wee, B., and Stead, D. Land use and travel behavior: expected effects from the perspective of utility theory and activity-based theories. Environment and Planning B, Vol. 32, 2005, pp. 33-46.

Madre, J. L. Long term forecasting of car ownership and use. Jones, P. (ed.) Developments in Dynamic and Activity Based Approaches to Travel Analysis, Aldershot: Gower Publishing, Oxford Studies in Transport, 1990.

Manheim M. L. Fundamentals of transportation systems analysis, Vol. 1: Basic Concepts. MIT Press. Cambridge, MA, 1979.

Manoj, M., and Verma, A. Activity-travel behavior of non-workers from Bangalore city in India. Transportation Research Part A: Policy and Practice, No. 78, 2015, pp. 400-424.

McFadden, D. Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics, Academic Press, New York, 1974.

McGuckin, N., and Murakami, E. Examining trip-chaining behavior: Comparison of travel by men and women. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1693, 1999, pp.79-85.

McNally, M. G. An Activity-based microsimulation model for travel demand forecasting. D. Ettema and H. Timmermans (eds.), Activity-based Approaches to Transportation Modeling, Elsevier Science, pp.37-54, 1995.

Meloni, I., Guala, L., and Loddo, A. Time allocation to discretionary in home, out of home activities and to trips. Transportation, Vol. 31, No. 1, 2004, pp. 69-96.

Meloni, I., Spissu, E., and Bez, M. A model of the dynamic process of time allocation to discretionary activities. Transportation Science, Vol. 41, 2007, pp.15-28.

Miller, E. J., and Roorda, M. J. A prototype model of 24-hour household activity scheduling for the Toronto area. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1831 (1), 2003, pp.114-121.

Miller, J. Results of the 2011-12 campus travel survey. University of California, Davis Institute of Transportation Studies, June 2012, Research Report - UCD-ITS-RR-12-08, URL http://www.its.ucdavis.edu/?page_id=10063&pub_id=1644.

Millward, H., and Spinney, J. Time use, travel behavior, and the rural-urban continuum: results from the Halifax STAR project. Journal of Transport Geography, Vol. 19(1), 2011, pp. 51-58.

Misra, R. Toward a comprehensive representation and analysis framework for nonworker activity-travel pattern modeling. PhD thesis, Department of Civil Engineering, University of Texas, 1999, Austin, TX.

Misra, R., and Bhat, C. R. Activity-travel patterns of non-workers in the San Francisco Bay area: exploratory analysis. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1718, 2000, pp. 43-51.

Mullahy, J. Heterogeneity, excess zeros, and the structure of count data models. Applied Economics, Vol. 12, 1997, pp.337-350.

Murakami, E. and Watterson, W. T. Developing a household travel panel survey for the Puget Sound region. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1285, 1990, pp. 40-46.

National Household Survey. NHS in Brief: Commuting to work (2011) URL http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-012-x/99-012-x2011003_1-eng.pdf .

Neatt, K., Millward, H., and Spinney, J. E. L. Neighborhood walking densities: A multivariate analysis in Halifax, Canada. Journal of Transport Geography, Vol. 61, 2017, pp.9-16.

Nichols, B., Childress, S., and Coe, S. Sound casting at PSRC: Activity-based model development with EMME. Presented in the INRO Conference, Seattle, USA, 2014.

Noland, R. B. and Thomas, J. V. Multivariate analysis of trip chaining behavior. Environment and Planning B, Vol. 34, 2007, pp. 953-970.

Ortuzar, J. D. and Willumsen, L. G. Modeling transport. John Wiley and Sons: Chichester, UK, 2011.

Ortuzar, J. D. and Willumsen, L. G. Modelling transport, 1994, Wiley, Chichester.

Paleti, R., Pendyala, R. M., Bhat, C. R. and Konduri, K. C. A joint tour-based model of tour complexity, passenger accompaniment, vehicle type choice, and tour length. Presented in the 91st Annual Meeting of Transportation Research Board, Washington D.C., USA., 2011.

Pas, E. I. Multi-day samples, parameter estimation precision, and data collection costs for least squares regression trip-generation models. Environment and Planning A, Vol. 18, 1986, pp. 73-87.

Pas, E. I. Recent advances in activity-based travel demand modeling. Proc., Activity-Based Travel Forecasting Conference. USDOT, New Orleans, Louisiana, 1996, pp. 79-102.

Pas, E. I. State of the art and research opportunities in travel demand: Another perspective. Transportation Research Part A: Policy and Practice, Vol. 19, 1985, pp. 460-64

PB Consult. The MORPC travel demand model validation and final report. Prepared for the Mid-Ohio Region Planning Commission, 2005.

PB, J. Bowman, and Bradley, M.A. Regional Transportation plan major update project for the Atlanta regional commission, General Modeling Task 13 (Activity/Tour-Based Models). Progress Report for the Year 2005, 2005.

Pendyala, R. M., and Bhat, C. R. An exploration of the relationship between timing and duration of maintenance activities. Transportation, Vol. 31(4), 2004, pp. 429-456.

Pendyala, R. M., and Goulias, K. G. Time use and activity perspectives in travel behavior research. Transportation, Vol. 29(1), 2002, pp.1-4.

Pendyala, R. M., and Ye, X. Contributions to understanding joint relationships among activity and travel variables. Progress in Activity-based Analysis. The Netherlands: Elsevier, 2005, pp. 1-24.

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

Pendyala, R., Kitamura, R., Kikuchi, A., Yamamoto, T., and Fujii, S. Florida activity mobility simulator: Overview and preliminary validation results. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1921, 2005, pp. 123-130.

Pendyala, R.M., Yamamoto, T., and Kitamura, R. On the formulation of time-space prisms to model constraints on personal activity-travel engagement. Transportation, Vol. 29(1), 2002, pp. 73-94.

Pinjari, A. R., Bhat, C. R., and Hensher, D. A. Residential self-selection effects in an activity time-use behavior model. Transportation Research B, No. 43, 2009, pp. 729-748.

Pinjari, A. R., Eluru, N., Copperman, R., Sener, I. N., Guo, J. Y., Srinivasan, S., and Bhat, C. R. Activity-based travel-demand analysis for metropolitan areas in Texas: CEMDAP models, framework, software architecture and application results. Research Report 4080-8, Center for Transportation Research, The University of Texas at Austin, 2006.

Pratt, J. Concavity of the log likelihood. Journal of the American Statistical Association, Vol. 76, 1981, pp. 103-116.

Primerano, F., Taylor, M. A., Pitaksringkarn, L., and Tisato, P. Defining and understanding trip chaining behavior. Transportation, No. 35, 2008, pp. 55-72.

Rajagopalan, B. S., Pinjari, A. R., and Bhat, C. R. Comprehensive model of worker non-work-activity time use and timing behavior. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2134, 2009, pp. 51-62.

Rasouli, S., and Timmermans, H. Activity-based models of travel demand: promises, progress and prospects. International journal of urban science, Vol. 18(1), 2013, pp. 31-60.

Recker, W. W. A bridge between travel demand modeling and activity-based travel analysis. Transportation Research Part B: Methodological, Vol. 35(5), 2001, pp. 481-506.

Recker, W. W. The household activity pattern problem: general formulation and solution, Transportation Research Part B: Methodological, Vol. 29, 1995, pp.61-77.

Recker, W. W., McNally, M. G., and Root, G. S. A model of complex travel behavior: Part I. Theoretical development. Transportation Research Part A: General, Vol. 20, 1986a, pp. 307 318.

Recker, W. W., McNally, M. G., and Root, G. S. A model of complex travel behavior: Part II. An operational model. Transportation Research Part A: General, Vol. 20, 1986b, pp. 319 330.

Reichman, S. Travel adjustments and life styles: A behavioral approach. P. Stopher and A. Meyburg (eds.), Behavioral Travel Demand Models, Lexington Books, Lexington, Massachusetts, 1976.

Renalds, A., Smith, T., and Hale, P. A systematic review of built environment and health. Family and Community Health, Vol. 33, 2010, pp. 68-78.

Rodriguez, D. A., and Joo, J. The relationship between non-motorized mode choice and the local physical environment. Transportation Research Part D: Transport and Environment, Vol. 9, No. 2, 2004, pp. 151-173.

Rossi, T., Lemp, J., Komaduri, A., and Ehrlich, J. Comparison of activity-based model parameters between two cities. Proceedings of the 14th Transportation Research Board National Transportation Planning Applications Conference, Columbus, USA., 2013.

Ruiz, T., and Roorda, M. J. Analysis of planning decisions during the activity-scheduling process. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2054, 2008, pp. 46-55.

Russel, J. E., and Fraas, J. W. An application of panel regression to pseudo panel data. Multiple Linear Regression Viewpoints, Vol. 31, No. 1, 2005, pp.1-15.

Sabina, E. E., and Rossi, T. Using activity-based models for policy decision making. Presented in the Innovations in Travel Demand Modeling Conference, Transportation Research Board, Austin, USA. pp. 177-180, 2006.

Saelens, B. E., and Handy, S. L. Built environment correlates of walking: A review. Medicine and Science in Sports and Exercise, Vol. 40, 2008, pp. S550-S566.

Sallis, J. F., Saelens, B. E., Frank, L. D., Conway, T. L., Slymen, D. J., Cain, K. L., Chapman, J. E., and Kerr, J. Neighborhood built environment and income: examining multiple health outcomes. Social Science and Medicine, Vol. 68, 2009, pp. 1285-1293.

Scott, D. M., and Kanaroglou, P. S. An activity-episode generation model that captures interactions between household heads: Development and empirical analysis. Transportation Research Part B: Methodological, Vol. 36, 2002, pp. 875-896.

Shannon, T., Giles-Corti, B., Pikora, T., Bulsara, M., Shilton, T., and Bull, F. Active commuting in a university setting: Assessing commuting habits and potential for modal change. Transport Policy, Vol. 13, No. 3, 2006, pp. 240-253.

Shiftan, Y. Practical approach to model trip chaining. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1645, 1998, pp. 17-23.

Shiftan, Y., and Golani, A. Effect of auto restraint on travel behavior. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1932, 2005, pp. 156-163.

Spissu, E., Pinjari, A. R., Bhat, C. R., Pendyala, R. M., and Axhausen, K. W. An analysis of weekly out-of-home discretionary activity participation and time-use behavior. Transportation, Vol. 36 (2), 2009, pp. 483-510.

Srinivasan, S. Modeling household interactions in daily activity generation. PhD thesis, Faculty of the Graduate School, University of Texas at Austin, 2004.

Srinivasan, S., and Bhat, C. R. Modeling household interactions in daily in-home and out-of-home maintenance activity participation. Transportation, Vol. 32(5), 2005, pp. 523-544.

Statistics Canada. General social survey 2011-cycle 24: Time-stress and well-being, 2011.

Statistics Canada. Table number 477-0019-Postsecondary enrolments, by registration status, Pan-Canadian Standard Classification of Education (PCSCE), Classification of Instructional Programs, Primary Grouping (CIP_PG), gender and student status, 2015, CANSIM (database). Accessed on June 01, 2016.

Steed, J. L., and C. R. Bhat. On modeling the departure time choice for home-based social-recreational and shopping trips. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1706, 2000, pp. 152-159.

Stopher, P. R., D. T. Hartgen, and Y. Li. Smart: Simulation model for activities, resources and travel. Transportation, No. 23, 1996, pp. 293-312.

Strathman, J. G., and Dueker, K. J. Understanding trip chaining. Chapter 1, No. 3, Special reports on trip and vehicle attributes-IN: 1990 NPTS special reports.

Strathman, J. G., Dueker, K. J., and Davis, J. S. Effects of household structure and selected travel characteristics on trip chaining. Transportation, Vol. 21, 1994, pp. 23-45.

Susilo, Y. O., and Kitamura, R. Analysis of day-to-day variability in an individual's action space: exploration of 6-week Mobidrive travel diary data. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1902, 2005, pp.124-133.

Susilo, Y. O., and Kitamura, R. Structural changes in commuters' daily travel: The case of auto and transit commuters in the Osaka metropolitan area of Japan, 1980-2000. Transportation Research Part A: Policy and Practice, No. 42, 2008, pp. 95-115.

Szalai, A. (ed.). The use of time. Mouton, The Hague, 1972.

Timmermans, H. J. P., Arentze, T., and Joh, C. H. Analyzing space-time behavior: New approaches to old problems. Progress in Human Geography, Vol. 26(2), 2002, 175-190.

Timmermans, H., Middelkoop, M. V., and Borgers, A. Microsimulation system for predicting leisure activity-travel patterns. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1894, 2004, pp. 20-27.

Toole-Holt, L., Polzin, S. E., and Pendyala, R. M. Two minutes per person per day each year: Exploration of growth in travel time expenditures. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1917, 2005, pp.45-53.

Toor, W., and S. Havlick. Transportation and sustainable campus communities: Issues, examples, solutions. Washington DC: Island Press; 2004.

Tsai, C. P., C. Mulley, and G. Clifton. A review of pseudo panel data approach in estimating short-run and long-run public transport demand elasticities. Transport Reviews, Vol. 34, No. 1, 2014, pp. 102-121.

TURP. TURP (Time Use Research Program). Retrieved from Halifax regional space time activity research (STAR) survey: A GPS-assisted household time-use survey, survey methods. Halifax: Saint Mary's University, 2008.

USDOT. Activity-based Travel forecasting conference proceedings. US Department of Transportation, 1997, Washington, DC, Report DOT-T-97-17.

Van Acker, V., and Witlox, F. Commuting trips within tours: How is commuting related to land use? Transportation, Vol. 38, 2011, pp. 465-486.

Van-Wissen, L. J. G., and Meurs, H. The Dutch mobility panel: Experiences and evaluations. Transportation, Vol. 16, 1989, pp.99-119.

Verbeek, M. Pseudo Panel Data. Matyas, L. and Sevestre, P. (eds.) The econometrics of panel data: Fundamentals and recent developments in theory and practice. Kluwer Academic Publishers, the Netherlands, 1992.

Verbeek, M., and Nijman, T. Can cohort data be treated as genuine panel data? Empirical Economics, Vol. 17, 1992, pp.9-23.

Voas, D., and P. Williamson. Evaluating goodness-of-fit measures for synthetic microdata. geographical and environmental modeling, Vol. 5, No. 2, 2010, pp. 177-200.

Volosin, S. E. A Study of university student travel behavior. PhD thesis. Arizona State University. 2014.

Volosin, S. E., Paul, S., Pendyala, R. M., Livshits, V., and Maneva, P. Activity-travel characteristics of a large university population. Presented in 93rd annual meeting of the Transportation Research Board (TRB), Washington D.C., USA., 2014.

Vovsha P., Petersen, E. and Donnelly, R. A model for allocation of maintenance activities to the household members. Proceedings of the 83rd Annual Meeting of the Transportation Research Board (TRB), Washington, D.C., USA., 2004.

Vovsha P., Petersen, E. and Donnelly, R. Donnelly. Micro-simulation in travel demand modeling: Lessons learned from the New York best practice model. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1805, 2002, pp. 68-77.

Vuong, Q. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica, Vol. 57, 1989, pp. 307-334.

Wallace, B., Mannering, F., and Rutherford, G. Evaluating effects of transportation demand management strategies on trip generation by using Poisson and negative binomial regression. Transportation Research Record: Journal of the Transportation Research Board. Vol. 1682, 1999, pp.0-7.

Wang, R. The stops made by commuters: Evidence from the 2009 US National Household Travel Survey. Journal of Transport Geography, No. 47, 2014, pp. 109-118.

Wang, X., Khattak, A. J. and Son, S. What can be learned from analyzing university student travel demand? Transportation Research Record: Journal of the Transportation Research Board, Vol. 2322, 2012, pp. 129-137.

Warunsiri, S. and Mcnown, R. The returns to education in Thailand: A pseudo-panel approach. World Development, Vol. 38, No. 11, 2010, pp. 1616-1625.

Weis, C. and Axhausen, K. W. Induced travel demand: Evidence from a pseudo panel data based structural equations model. Research in Transportation Economics, Vol. 25, 2009, pp. 8-18.

Wen, C. Development of stop generation and tour formation models for the analysis of travel/activity behavior. PhD thesis, Department of Civil Engineering, Northwestern University, Evanston, Illinois, 1998.

Worrall, R. D. The urban panel as a longitudinal data source. Highway Research Record, Vol. 194, 1967, pp. 62-77.

Xu, C., Wang, W., Jiang, X., Li, X., and Xiang, Y. Analyzing travel time and frequency of e-bike trips by men and women using hazard-based duration and zero-inflated negative binomial models. Proceedings of the 94th Annual Meeting of Transportation Research Board (TRB), Washington, D.C., USA., 2015.

Yamamoto, T., Fujii, S., Kitamura, R., and Yoshida, H. Analysis of time allocation, departure time, and route choice behavior under congestion pricing. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1725, 2000, pp. 95-101.

Yang, D., and Timmermans, H. Estimation of influence of fuel price on individual dynamic travel behavior: Evidence from a pseudo panel data. Travel Behavior Research: Current Foundations, Future Prospects, Chapter 19, 2012, pp. 433-448.

Ye, X., Konduri, K., Pendyala, R. M., Sana, B., and Waddell, P. A methodology to match distributions of both household and person attributes in the generation of synthetic populations. Presented in 88th Annual Meeting of the Transportation Research Board (TRB), Washington, D.C., USA., 2009.

Ye, X., Pendyala, R. M., and Gottardi, G. An exploration of the relationship between mode choice and complexity of trip chaining patterns. Transportation Research Part B: Methodological, Vol. 41(1), 2007, pp. 96-113.

Ye, X., Pendyala, R. M., and Yang, X. Exploring activity-travel patterns in Xiamen, China. Presented in the 1st International Conference on Transportation Infrastructure. International Society for Maintenance and Rehabilitation of Transport Infrastructure, 2008, Beijing, China.

Yee, J. L., and Niemeier, D. A. Analysis of activity duration using the Puget sound transportation panel. Transportation Research Part A: Policy and Practice, Vol. 34, No. 8, 2000, pp. 607-624.

Yun, M. P., Chen, Z. H., and Liu, J. Y. Comparison of mode choice behavior for work tours and non-work tours considering trip chain complexity. Proceedings of the 93rd Annual Meeting of the Transportation Research Board, Washington D.C., USA., 2014.

Zhang, J., Timmermans, H., and Borgers, A. Model structure kernel for household task allocation incorporating household interaction and inter-activity dependency. Proceedings of the 83rd Annual Meeting of the Transportation Research Board, Washington D.C., USA., 2004.

Appendix A Copyright Permission

March 15, 2018

Journal of Urban Planning and Development

I am preparing my Ph.D. thesis for submission to the Faculty of Graduate Studies at Dalhousie University, Halifax, Nova Scotia, Canada. I am seeking your permission to include a manuscript version of the following paper(s) as a chapter in the thesis:

[Naznin Sultana Daisy, Mohammad Hesam Hafezi, Lei Liu and Hugh Millward, "Understanding and Modeling the Activity-Travel Behavior of University Commuters at a Large Canadian University", Journal of Urban Planning and Development, Vol. 144, Issue No. 02, 2018, DOI:10.1061/(ASCE)UP.1943-5444.0000442.]

Canadian graduate theses are reproduced by the Library and Archives of Canada (formerly National Library of Canada) through a non-exclusive, world-wide license to reproduce, loan, distribute, or sell theses. I am also seeking your permission for the material described above to be reproduced and distributed by the LAC(NLC). Further details about the LAC(NLC) thesis program are available on the LAC(NLC) website (www.nlc-bnc.ca).

Full publication details and a copy of this permission letter will be included in the thesis.

Yours sincerely,

Naznin Sultana Daisy

From: PERMISSIONS
Sent: March 16, 2018 11:19 AM
To: Naznin Sultana Daisy
Subject: RE: Request for Copyright Permission

Dear Naznin Sultana Daisy,
Thank you for your inquiry. As an original author of an ASCE journal article or proceedings paper, you are permitted to reuse your own content (including figures and tables) for another ASCE or non-ASCE publication, provided it does not account for more than 25% of the new work. This email serves as permission to reuse your work, Understanding and Modeling the Activity-Travel Behavior of University Commuters at a Large Canadian University, *Journal of Urban Planning and Development*, 144, Issue No. 02, 2018.

A full credit line must be added to the material being reprinted. For reuse in non-ASCE publications, add the words "With permission from ASCE" to your source citation. For Intranet posting, add the following additional notice: "This material may be downloaded for personal use only. Any other use requires prior permission of the American Society of Civil Engineers. This material may be found at [URL/link of abstract in the ASCE Library or Civil Engineering Database]."

Each license is unique, covering only the terms and conditions specified in it. Even if you have obtained a license for certain ASCE copyrighted content, you will need to obtain another license if you plan to reuse that content outside the terms of the existing license. For example: If you already have a license to reuse a figure in a journal, you still need a new license to use the same figure in a magazine. You need separate license for each edition.

For more information on how an author may reuse their own material, please view: http://ascelibrary.org/page/informationforasceauthorsreusingyourownmaterial

Sincerely,

Leslie Connelly
Senior Marketing Coordinator
American Society of Civil Engineers
1801 Alexander Bell Drive
Reston, VA 20191

PERMISSIONS@asce.org

703-295-6169
Internet: www.asce.org/pubs | www.ascelibrary.org |
http://ascelibrary.org/page/rightsrequests