A COMPREHENSIVE STUDY ON HOUSEHOLD TRANSPORTATION EXPENDITURE, VEHICLE OWNERSHIP, AND USAGE PATTERNS: TALE OF CANADIAN CITIES

by

Sabreena ANOWAR

Department of Civil Engineering and Applied Mechanics
MCGILL UNIVERSITY
Montreal, Canada

August 2015

A Dissertation
submitted to McGill University
in partial fulfillment of the requirements of the degree of
DOCTOR OF PHILOSOPHY

ABSTRACT

A hierarchy of decision exists across varying temporal scale in the context of household automobile fleet ownership and usage decisions. In the long term, the hierarchy includes the allocation of monetary resources to various consumer expenditure categories and long term vehicle fleet choices (number and type) while in the short-term, conditional on the availability of the fleet, decisions regarding vehicle usage (type and mileage) and travel destinations are considered. Thus, a combined exploration of these aspects would provide a more complete and clear understanding of the factors associated with these decision processes. Given these considerations, the objective of this dissertation is to contribute to the growing body of travel behavior literature by focusing on transport expenditure, vehicle fleet choice, and usage decisions of Canadian households. Specifically, the current dissertation aims to bridge the gaps in the standard literature along five directions: (1) transport expenditure, (2) population heterogeneity, (3) appropriate econometric models - ordered vs unordered logit models for modeling vehicle ownership while accounting for population heterogeneity, (4) multiple time point data pooling (pseudo panel analysis), and (5) short term vehicle usage. As part of the contributions in the dissertation, several econometric models are formulated, estimated, and validated to address the aforementioned issues through five different studies.

Chapter 2 presents a comprehensive literature review on vehicle ownership modeling (methodology, empirical findings and estimation methods) depending on different representation of the ownership decision process (number, type, usage, and duration of holding). In Chapter 3, to shed light on the important factors affecting expenditure of households and its evolution, variants of multiple discrete continuous extreme value (MDCEV) models are formulated. The results indicate that socio-economic and demographic attributes along with residential location characteristics were found to impact household expenditure significantly.

In addressing population heterogeneity issue in the context of vehicle ownership, latent segmentation based ordered and multinomial logit models are formulated, estimated and validated in Chapter 4. The estimation results indicate that there is population heterogeneity in vehicle fleet size decision of households, and latent class models are an elegant method of capturing it. A comparison between alternate discrete outcome frameworks for modeling vehicle ownership while accounting for population heterogeneity is carried out in Chapter 5. Specifically, the empirical

comparison is undertaken by estimating several ordered and unordered models and the performance of the alternative frameworks is examined in the context of both model estimation and validation (at the aggregate and disaggregate level) by using a host of comparison metrics. The results from the exercise point towards the superiority of the generalized ordered framework in comparison with its unordered counterpart in modeling vehicle ownership. In Chapter 6, a simple yet efficient cross-sectional data usage technique is proposed for tackling the longitudinal data unavailability issue. The study demonstrates that formation of pseudo-panel by pooling multiple year cross-sectional datasets allows us to identify how the impact of exogenous variables has altered with time. Chapter 7 of the dissertation contributes to the growing literature on short term vehicle usage decisions by examining four activity travel choice processes: spatial flexibility of the activity, temporal flexibility of the activity, activity vehicle type, and primary driver (for auto users). The analysis results revealed that several individual and household socio-demographic characteristics, residential location, and activity attributes as well as contextual variables influence the packaged choice. Moreover, error correlation between spatially planned and temporally planned alternatives are also identified from the analysis.

Although the domain of travel behavior literature is continuously evolving and being enriched, there is a paucity of literature in the context of Canadian urban regions. The current dissertation aims to bridge the gap by presenting studies using Canadian household data. The econometric models developed in the current dissertation are estimated using Survey of Household Spending (SHS) of Canada, Origin-Destination (O-D) surveys of both Greater Montreal Area (GMA) and Quebec City, and Quebec City Travel and Activity Panel Survey (QCTAPS) of Quebec City. In addition to model formulation, a variety of policy exercises are conducted and presented to illustrate the applications of the models developed.

Several policy recommendations can be made based on the results obtained from the empirical analyses of the dissertation. First and foremost, we need a combination of short and long range policies to tackle the problem of over dependence on automobiles. Providing efficient, reliable, and convenient transit and car sharing alternatives, particularly for work trips, should be given higher priority. It can be coupled with advertisement campaigns symbolizing transit as a "green alternative" to encourage usage. For long term, in addition to the densification and diversification of land use, internalization of the societal costs of private modes of transportation and urban sprawl to influence household choices need to be considered as well.

RÉSUMÉ

La composition de la flotte automobile des ménages (en termes du budget alloué pour l'achat de véhicules, du nombre et du type de véhicules) consiste de décisions à long et à court termes. À long terme, la hiérarchie de décision inclut l'allocation des ressources financières aux différentes catégories de dépenses du consommateur et aux décisions à long terme des choix de flottes de véhicules tandis qu'à court terme, conditionnelle à la disponibilité de la flotte, les décisions concernant l'utilisation du véhicule (type et kilométrage) et la destination des déplacements sont considérées. Ainsi, une exploration combinée de ces aspects permettrait une compréhension plus complète et plus claire des facteurs associés à ces processus de décision. Ceci dit, l'objectif de cette thèse est de contribuer à la masse croissante de la littérature de comportement des déplacements en mettant l'accent sur les dépenses de transport, le choix de la flotte de véhicules et les décisions d'utilisation des ménages canadiens. Plus précisément, cette thèse vise à combler les lacunes dans la littérature dans cinq catégories: (1) les dépenses en transport (2) l'hétérogénéité de la population (3) les questions économétriques telles que l'utilisation de modèles logit ordonnés et non-ordonnées pour la modélisation de la possession de véhicules en tenant compte de l'hétérogénéité de la population (4) mutualisation des points de temps multiples (pseudo-analyse de panel) et (5) l'utilisation des véhicules à court terme. Dans le cadre de la contribution de la thèse, plusieurs modèles économétriques sont formulés, estimés et validés pour adresser les problèmes mentionnés ci-dessus à travers cinq études différentes.

Le chapitre 2 présente une revue exhaustive de la littérature sur la modélisation de la possession de véhicules (méthodologie, faits saillants et méthodes d'estimation) en fonction des représentations différente du processus de décision de la possession (le nombre, le type, l'utilisation, et la durée de possession). Au chapitre 3, pour faire la lumière sur les facteurs importants affectant les dépenses des ménages et l'évolution de ces dépenses, des variantes des modèles de valeurs discrètes continues multiples sont formulées. Les résultats indiquent que les attributs socio-économiques et démographiques ainsi que les caractéristiques de l'emplacement résidentiels impactent les dépenses des ménages de manière significative.

En abordant la problématique de l'hétérogénéité de la population dans le contexte de la possession de véhicules, les modèles logit à base segmentée de type ordonné et multinomial sont formulés, estimés et validés dans le chapitre 4. Les résultats de indiquent qu'il existe une

hétérogénéité de la population en ce qui concerne la décision de la taille de la flotte de véhicules des ménages, et les modèles de classes latentes sont une méthode élégante de le capturer. Une comparaison entre les cadres alternatifs de résultats discrets pour la modélisation de la possession de véhicules tout en tenant compte de l'hétérogénéité de la population est effectuée au chapitre 5. Plus précisément, la comparaison empirique est effectuée en estimant plusieurs modèles à résultats ordonnés et non ordonnée basée sur la segmentation latente et la performance des cadres alternatifs sont examinés dans le contexte de l'estimation et la validation des modèles (aux niveaux agrégés et désagrégés) en utilisant une série de mesures de comparaisons. Les résultats de l'exercice indiquent une supériorité du cadre ordonné généralisé en comparaison avec son homologue non ordonnée dans la modélisation de la possession de véhicules. Au chapitre 6, une technique d'utilisation transversale simple, mais efficace des données est proposée dans l'une des études de la dissertation pour adresser l'indisponibilité longitudinale de données. L'étude démontre que la formation de pseudo-panel en mettant en commun des sources de données transversales d'années multiples nous permet de déterminer comment l'impact des variables exogènes a changé avec le temps. Chapitre 7 de la thèse contribue à la littérature croissante sur les décisions d'utilisation des véhicules à court terme en examinant quatre processus de choix d'activité de déplacement : la flexibilité spatiale de l'activité, la flexibilité temporelle de l'activité, l'activité type de véhicule, et le conducteur primaire (pour les utilisateurs d'automobiles). Les résultats d'analyse ont révélé que plusieurs des caractéristiques individuelles et des ménages sociodémographiques, l'emplacement et les attributs de l'activité ainsi que les variables contextuelles influencent le choix agencé. En outre, l'erreur de corrélation entre alternatives prévues spatialement et temporellement a également été identifiée à partir de l'analyse.

Bien que le domaine de la littérature sur le comportement des déplacements soit en constante évolution et enrichissement, il y a une pénurie de la littérature dans le contexte des régions urbaines du Canada. La présente thèse vise à combler cette lacune en présentant des études utilisant des données sur les ménages canadiens. Les modèles économétriques développés dans le mémoire en cours sont estimés à partir de l'Enquête sur les dépenses des ménages (SHS) du Canada, les enquêtes origine-destination (OD) des régions de Montréal et de Québec, et le Québec City Voyage et Panel Survey Activité (QCTAPS) de la ville de Québec. En plus de la formulation du modèle, une variété d'applications de politiques sont menées et présentées pour illustrer l'utilisation des modèles développés.

Plusieurs recommandations de politiques peuvent être faites sur la base des résultats obtenus à partir des analyses empiriques de la thèse. Tout d'abord, nous avons besoin d'une combinaison intelligente de politiques de courtes et de longues portées pour lutter contre le problème de la dépendance à l'automobile. Une priorité plus élevée doit être attribuée aux systèmes de transport en commun et d'auto partage, en les rendant plus efficaces, fiables et pratiques, en particulier pour les déplacements de travail. Il peut être couplé à des campagnes publicitaires symbolisant le transport en commun en tant que «symbole environnemental" pour encourager son utilisation. À long terme, en plus de la densification et la diversification de l'utilisation du sol, on a besoin de considérer l'internalisation des coûts sociaux des modes de transport privés et de l'étalement urbain dans le but d'influencer les choix des ménages.

ACKNOWLEDGEMENTS

According to the celebrated web comic Piled Higher and Deeper (PhD), I entered the path of doctoral education being a “foolish young” damsel and the journey was destined to end in “bitter remorse”. On the contrary, I am writing this acknowledgement, meaning I have survived grad school and soon the journey is hopefully going to “culminate in a ceremony” where I am going to “walk down the aisle dressed in a gown”. The journey, one of enlightenment and self-discovery, like any other, had its ups and downs, got tumultuous at times, but was rewarding and memorable, nonetheless. The integral part of the journey were the people surrounding me, some I already knew and some I met during my stay in Montreal, who contributed directly or indirectly, willingly or unknowingly in every step of the way and therefore, the successful completion of it is as much their credit as it is mine. So, I would like to take this opportunity to humbly acknowledge all those people who have supported, inspired, motivated, helped, and most importantly endured me throughout the journey of my doctoral education.

First, my deepest appreciation goes to the Faculty of Engineering of McGill University for conferring upon me the McGill Engineering Doctoral Award (MEDA) which provided me with the needed financial support all throughout my graduate studies at McGill. It was a huge honor as well to receive the MEDA award in the form of two prestigious fellowships – Max Stern Recruitment Fellowship along with Graduate Excellence Fellowship (GEF). I give my sincere thanks to all the generous donors.

This dissertation would not have come to fruition without both of my supervisor’s timely guidance and support. So, I would like to express my sincere gratitude to my supervisor, Dr Luis F Miranda-Moreno for his supervision and financial support. My dream of becoming a McGillian wouldn’t be possible had he not accepted me as one of his supervisees. Additional gratitude is also in order for the help with travel and conference expenses. I owe my deepest and heartfelt gratitude to my co-supervisor, Dr. Naveen Eluru, for his invaluable words of advice, support and encouragement, both professionally and personally. His passion for research is contagious, dedication to his students is unparalleled, and above all, his compassion for others is pervasive. Be it divine intervention or stroke of sheer good luck, I got the opportunity to know him in person and work under his supervision which I would forever cherish. I am also thankful to Dr. Saeed Mirza, Prof. Luc E. Chouinard, Dr. Jinxia Liu, Dr. Kevin Manaugh, and Dr. Zachary Patterson for serving on my dissertation committee and for their valuable suggestions which helped better shape the dissertation proposal.

I would like to convey my very special thanks to Dr Richard Tay and Bonu (Shamsunnahar Yasmin) – two immensely important persons in my life, who had been, still are, and will always be my friend, philosopher, and guide all throughout my journey of life. Thank you Richard for introducing me to the amazing world of discrete choice modeling, for the reassuring words which kept me afloat in times of crisis, and the kind advice which guided me in times of indecisiveness. I feel truly blessed to have you as my mentor. Bonu, since being introduced to you on the BEVCO bus, we have come a long way. Thank you for making my entire graduate life (both in Calgary and Montreal) extraordinarily memorable and fun-filled. I would reminisce each and every moment of it – the lame jokes and giggles, the heated arguments, the late night cooking bonanza, the long hours of card games, and of course, the crazy shopping sprees, both on-line and in-store.

My acknowledgement would be incomplete without mentioning my friends with whom I share great memories. Thank you all, from the core of my heart. Sabrina, for our long walks and talks which I dearly miss – I look forward to taking that walk to the top of the Mount Royal with you someday; Adham, for being a wonderful company always – be it playing games or dining out; Frederic, for being my Ping-Pong buddy (Go TeamTranspo) and for all the exciting board game fights; Sandra, for being the nicest neighbor in the office – I really enjoyed all our chitchats; Miguel, for the handsomest killer smile. Golnaz, Ahmad, Will, Tim, Joshua, Ting, Junshi, and Jill for being my friend. Thank you Annie and Snigdha for helping me overcome my GIS fears.

I am forever indebted to my family for their endless love and care, relentless support, and continuous encouragement in all of my life pursuits. Ammu, without her sacrifice and unconditional love, I would not be where I am today. It is from her I have learnt not to give up and fight till the end, against all odds and that too with a broad smile. Abbu, thank you for being the exemplar of integrity, honesty, patience and indomitable perseverance and instilling in me the importance of these principles in life. Thank you Putu for being the greatest baby brother in the entire universe. My deepest gratitude to my beloved country, Bangladesh, for giving me the strong educational foundation to face the world. Thanks to Bangladesh Cricket Team (Go Tigers) for giving me numerous occasions to celebrate! Finally, thanks to the Almighty for blessing me with such wonderful and supportive family and friends, for providing me with the strength to endure through trying times, and for enabling me to complete this dissertation work successfully.

DEDICATION

To my mother *Selina Anowar,* my father *Anowar Hossain,* and my brother *Wahid*

TABLE OF CONTENTS

LIST OF FIGURES xvi

LIST OF TABLES xvii

AUTHOR CONTRIBUTIONS xix

CHAPTER 1 INTRODUCTION 1

1.1 Auto-centrism 1

1.1.1 Auto-centrism in Canada 2

1.2 The Impacts of Auto-Centrism 2

1.3 The Importance of Analyzing Automobile Ownership, Usage, and Expenditure 4

1.4 Scope of the Dissertation 4

1.5 Gaps in the Literature 6

1.5.1 Transport Expenditure 7

1.5.2 Population Heterogeneity 7

1.5.3 Ordered or Unordered? The Eternal Debate 9

1.5.4 Multiple Time Point Data Pooling 10

1.5.5 Short Term Vehicle Usage 12

1.6 Objectives of the Dissertation 13

1.7 Outline and Contributions of the Dissertation 14

CHAPTER 2 ALTERNATIVE MODELING APPROACHES USED FOR EXAMINING AUTOMOBILE OWNERSHIP: A COMPREHENSIVE REVIEW 17

2.1 Introduction 17

2.1.1 Vehicle Ownership Representation 17

2.1.2 Contribution and Organization of the Chapter 19

2.2 Methods 20

2.2.1 Exogenous Static Models 21

2.2.2 Endogenous Static Models 26

2.2.3 Exogenous Dynamic Models 32

2.2.4 Endogenous Dynamic Models 35

2.3 Parameter Estimation Methods 37

2.4 Empirical Findings 38

2.4.1 Individual and Household Demographics 39

2.4.2 Residential Location and Built Environment Attributes 40

2.4.3 Others...41

2.5 Issues and Challenges in Vehicle Ownership Modeling 43

2.5.1 Data Issues 43

2.5.2 Spatial Correlation 44

2.6 Summary and Conclusions 44

CHAPTER 3 ANALYSIS OF THE EVOLUTION OF TRANSPORTATION EXPENDITURE: A LONG TERM PERSPECTIVE 54

3.1 Background 54

3.1.1 Modelling Budgetary Allocations...54

3.1.2 Contribution and Organization of the Chapter 55

3.2 Literature Review 56

3.3 Econometric Framework 58

3.3.1 Basic Structure and Traditional MDCEV Model 58

3.3.2 Scaled MDCEV Model...60

3.3.3 Mixed MDCEV Model 60

3.4 Empirical Data 61

3.4.1 Data Source...61

3.4.2 Dependent Variable Compilation 61

3.4.3 Descriptive Analysis...63

3.5 Empirical Analysis **64**

3.5.1 Variables Specification 64

3.5.2 Model Specification 65

3.5.3 Model Fit Measures 65

3.5.4 Empirical Results 66

3.5.5 Policy Simulation Results 71

3.6 Summary and Conclusions **72**

3.7 Link between Chapter 3 and Chapter 4 .... **73**

CHAPTER 4 ADDRESSING POPULATION HETEROGENEITY IN DISCRETE CHOICE MODELS **80**

4.1 Introduction **80**

4.1.1 Contribution and Organization of the Chapter 80

4.2 Econometric Framework **81**

4.3 Data **83**

4.4 Empirical Analysis **84**

4.4.1 Variables Considered 84

4.4.2 Model Specification and Performance Evaluation 85

4.4.3 Behavioral Interpretation 86

4.4.4 Estimation Results 86

4.4.5 Validation Results 89

4.4.6 Elasticity Effects 90

4.5 Summary and Conclusions **91**

4.6 Link between Chapter 4 and Chapter 5 .... **92**

CHAPTER 5 APPROPRIATE FRAMEWORK OF ANALYSIS WHILE ADDRESSING POPULATION HETEROGENEITY IN DISCRETE CHOICE MODELS **99**

5.1 Introduction **99**

5.1.1 Contribution and Organization of the Chapter 99

5.2 Econometric Framework **100**
5.2.1 Post Estimation Qualitative Analysis 102
5.3 Empirical Context **103**
5.3.1 Study Area 103
5.3.2 Earlier Work 103
5.4 Data **104**
5.4.1 Data Source 104
5.4.2 Descriptive Statistics 105
5.5 Empirical Analysis **105**
5.5.1 Variables Considered 105
5.5.2 Model Specification and Performance Evaluation 107
5.5.3 Behavioural Interpretation 109
5.5.4 Estimation Results 110
5.5.5 Validation Results 115
5.6 Summary and Conclusions **116**
5.7 Link between Chapter 5 and Chapter 6 .... **117**
CHAPTER 6 ANALYSIS OF VEHICLE OWNERSHIP EVOLUTION USING PSEUDO PANEL ANALYSIS **128**
6.1 Introduction **128**
6.1.1 Contribution and Organization of the Chapter 128
6.2 Econometric Framework **130**
6.3 Data **131**
6.4 Empirical Analysis **132**
6.4.1 Variables Considered 132
6.4.2 Estimation Results 133
6.4.3 Elasticity Effects and Policy Analysis 138
6.5 Summary and Conclusions **140**

6.6 Link between Chapter 6 and Chapter 7 .... **141**

CHAPTER 7 A JOINT ECONOMETRIC ANALYSIS OF ACTIVITY FLEXIBILITY, VEHICLE TYPE CHOICE AND PRIMARY DRIVER SELECTION **148**

7.1 Introduction **148**

7.1.1 Earlier Work 148

7.1.2 Methods 149

7.1.3 Contribution and Organization of the Chapter 150

7.2 Data **151**

7.2.1 Survey Instrument 151

7.2.2 Choice Set Formation and Descriptive Statistics 152

7.3 Econometric Framework **154**

7.4 Empirical Analysis **155**

7.4.1 Variable Specification 155

7.4.2 Estimation Results 155

7.5 Summary and Conclusions **161**

CHAPTER 8 CONCLUSIONS AND RECOMMENDATIONS **167**

8.1 Introduction **167**

8.2 Analysis of the Evolution of Transportation Expenditure: A Long Term Perspective **168**

8.2.1 Implications and Suggested Policy Measures 168

8.3 Addressing Population Heterogeneity in Discrete Choice Models **169**

8.3.1 Implications and Suggested Policy Measures 170

8.4 Appropriate Framework of Analysis while Addressing Population Heterogeneity in Discrete Choice Models **171**

8.4.1 Implications 172

8.5 Analysis of Vehicle Ownership Evolution Using Pseudo Panel Analysis **172**

8.5.1 Implications and Suggested Policy Measures 173

8.6 A Joint Econometric Analysis of Activity Flexibility, Vehicle Type Choice and Primary Driver Selection **174**

8.6.1 Implications and Suggested Policy Measures 175

8.7 Directions for Future Research **175**

8.7.1 Limitations of the Current Research 176

8.7.2 Research Extensions 176

APPENDIX A **180**

REFERENCES **187**

LIST OF FIGURES

Figure 5.1 Proportion of Households (HH) in Different Vehicle Ownership Categories in TAZs 118

Figure 6.1 Evolution of Car Ownership Levels across Years for Artificial Households in Ville-Marie Borough 142

Figure 6.2 Evolution of Car Ownership Levels across Years for Artificial Households in Plateau-Mont-Royal Borough 143

LIST OF TABLES

Table 2.1 Summary of Previous Studies on Automobile Ownership 46

Table 2.2 Decision Matrix for Vehicle Ownership Model Selection 53

Table 3.1 Sample Characteristics 74

Table 3.2 Summary Profile of Expenditure Categories and Savings 75

Table 3.3 Estimation Results (N=13000) 76

Table 3.4 Policy Simulation Results 79

Table 4.1 Summary Statistics of Variables 93

Table 4.2 Segment Characteristics and Mean Values of Segmentation Variables (N = 5218) 94

Table 4.3 Parameter Estimates (N=5218) 95

Table 4.4 Measures of Fit in the Validation Sample (N = 1326) 97

Table 4.5 Elasticity Effects of Important Variables 98

Table 5.1 Summary Statistics of Variables* 119

Table 5.2 Segment Characteristics and Mean Values of Segmentation Variables 120

Table 5.3 Importance Rating of Variables in the Segment Specific Choice Model 121

Table 5.4 LSGOLII Estimation Results (N=4913) 122

Table 5.5 Measures of Fit in Validation Sample (N=1972) 125

Table 6.1 Summary Statistics of Socio-demographic and Land Use Variables 144

Table 6.2 MGOL Estimation Results (N=12000) 145

Table 6.3 Elasticity Effects 147

Table 7.1 Distribution of Perceived Temporal and Spatial Flexibility by Vehicle Type 162

Table 7.2 Distribution of Individual Characteristics across Temporal Flexibility, Spatial Flexibility and Vehicle Type 163

Table 7.3 Estimation Results (N = 378 individuals and 8098 activity records) 164

AUTHOR CONTRIBUTIONS

The dissertation contains six full length journal manuscripts. Among these six manuscripts, four have already been accepted and/or published in several distinguished journals and the other two manuscripts are under review for publication consideration. I, Sabreena Anowar, am the first author of all of the manuscripts documented in the dissertation. My contribution to the articles include data cleaning and preparation, model estimations, and writing. The co-authors contributed to the manuscripts by sharing their data, providing valuable insights and comments, as well as editing the manuscripts. The publication details of the manuscripts are presented below.

Chapter 2 is based on the article: **S. Anowar,** N. Eluru and L. F. Miranda-Moreno, "Alternative Modeling Approaches Used for Examining Automobile Ownership: A Comprehensive Review". *Transport Reviews,* Vol. 34 (4), 2014, pp. 441-473.

Chapter 3 is based on the article: **S. Anowar,** N. Eluru and L. F. Miranda-Moreno, "How Household Transportation Expenditures Have Evolved in Canada: A Long Term Perspective". The paper is under review in *Transportation*.

Chapter 4 is based on the article: **S. Anowar,** S. Yasmin, N. Eluru and L. F. Miranda-Moreno, "Analyzing Car Ownership in Quebec City: A Comparison of Traditional and Latent Class Ordered and Unordered Models". *Transportation,* Vol. 41 (5), 2014, pp. 1013-1039.

Chapter 5 is based on the article: **S. Anowar,** S. Yasmin, N. Eluru and L. F. Miranda-Moreno, "Incorporating Population Heterogeneity in Vehicle Ownership: A Montreal Case Study". The paper is under review in 14th International Conference on Travel Behavior Research (IATBR).

Chapter 6 is based on the article: Chapter 6 is based on the article: **S. Anowar,** N. Eluru and L. F. Miranda-Moreno, "Analyzing Vehicle Ownership Evolution in Montreal, Canada Using Pseudo-Panel Analysis". The paper is accepted for publication in *Transportation.*

Chapter 7 is based on the article: **S. Anowar,** N. Eluru, L. F. Miranda-Moreno and M. Lee-Gosselin, “A Joint Econometric Analysis of Temporal and Spatial Flexibility of Activities, Vehicle Type Choice and Primary Driver Selection”. The paper is accepted for publication in *Transportation Research Record (TRR).*

CHAPTER 1 INTRODUCTION

1.1 Auto-centrism

In today's fast-paced world, mobility is the key to success and private automobiles are the key mobility tool for individuals and households. To many, owning a private vehicle is not only a utilitarian necessity but also a symbol of "power, status, control and freedom" (Yamamoto, 2009). The combination of the "symbolic perceived utility" (self-satisfaction, increased social esteem or higher status symbol) along with the tangible utility (increased mobility, greater comfort, higher access to opportunities) has resulted in increased auto-dependency both in the occidental (Caulfield, 2012) and the oriental worlds (Wu et al., 1999; Li et al., 2010)[1].

The term auto-centrism or "automobile dependency" is formally defined as "high levels of per capita automobile travel, automobile-oriented land-use patterns, along with reduced and unattractive transport alternatives" (Litman, 2002; Newman and Kenworthy, 1999, pp.1). In the early 20th Century, suburbs developed and expanded along streetcar and/or tramway routes and remained relatively compact, where active transportation kept its importance and was actively used. However, the post-World War II era marks the emergence of the auto-centric culture when North America underwent a massive urbanization phase of decentralization of the urban cores into suburban sprawls that eventually led to the popularization of private cars as travel mode. The magic-vicious-circle of automobile oriented culture was thus born and since then, it has not stopped turning; on the contrary, it has escalated to such an extent that it is now viewed with more concern than ever (Dupuy, 1999).

The cycle begins with disperse land use developments and creation of single use zones with generous parking supply. These land use planning choices along with decreasing fuel costs and heavily subsidized highway constructions by governments ensure that private vehicles become not just practical, but an essential part of most people's everyday transportation needs (Komanoff and Pucher, 2003). As a result, people make more trips, travel more miles, and purchase and own more vehicles. The observed increase in automobile traffic results in chronic road congestion which the governments try to counter by building more roads. New road development enables

[1] In recent years, a reversal in vehicle ownership levels in developed countries is being reported; highlighting a possible "peak" in ownership levels (Kuhnimhof et al., 2013; Millard-Ball and Schipper, 2011).

urban areas to expand even further (far beyond the reach of non-motorized mode of transport) and facilitates growth of new residential and/or commercial areas. Eventually, newer and longer trips are generated, once again followed by the growth of the road network, and so on and on (Handy, 1993; Weinberger, 2007).

1.1.1 Auto-centrism in Canada

Auto-centric culture is most prevalent in the cities in United States followed by Australian and Canadian cities. Majority of the households in the North American and Australian communities own automobiles and rely on them for trips (Sen, 2006; Soltani, 2005; Litman, 2002; Kenworthy and Laube, 1999). According to recent statistics, in Canada, 84.4 percent of households owned or leased at least one vehicle in 2007 (Canada NR, 2009). Nationally, between 2000 and 2009, average number of vehicles per household increased from 1.43 to 1.47, which represents only a 3 percent increase. However, during this time, the number of Sport Utility Vehicles (SUV) almost doubled, increasing markedly from 6.9 percent to 12.8 percent (Canada NR, 2009). Alarmingly, in addition to ownership, the use of personal vehicles by Canadians for daily trips is also increasing. A recent study showed that 82 percent of Canadian commuters currently drive to work, compared to only 12 percent who take public transit and 6 percent who walk or bike (Turcotte, 2011). In fact, the proportion of individuals using the auto mode for travel increased from 68 percent in 1992 to 74 percent in 2005 as observed from the time-use data from the Canadian General Social Survey (CGSS) (Turcotte, 2008). It appears that personal vehicles are an essential household commodity in Canada.

The remainder of the chapter is organized as follows. Section 1.2 provides an overview of the impacts of auto-centrism while Section 1.3 presents the importance of analyzing different dimensions of household automobile choice decisions. In Section 1.4 we discuss the scope of the dissertation. The gaps in the existing literature are discussed in detail in Section 1.5. In light of the discussions in Section 1.5, the objectives of the dissertation are presented in Section 1.6 and finally, Section 1.6 outlines the rest of the dissertation.

1.2 The Impacts of Auto-Centrism

Admittedly, private automobiles have given individuals unprecedented mobility and have catered the accessibility needs of individuals and households for performing their activities. In many cases, automobiles are "more reliable, comfortable, and convenient" way of travel. Nevertheless, the extent of the concomitant negative consequences of auto-centrism outweigh its seemingly positive benefits. In fact, the negative impacts (environmental, social and economic) are manifold and are of significance at the household, community, regional, and global level (Sen, 2006). For instance, households incur large amount of transportation costs in the form of fuel expenditure, vehicular acquisition, maintenance expenses, parking fees, and roadway taxes (Bhat et al., 2009). These costs heavily influence the overall monetary budget allocation process of households and consequently, households might end up allocating less resources towards savings in order to accommodate excessive transportation costs. From personal health and well-being perspective, sedentary lifestyles of individuals belonging to auto-dependent households result in rising health risks (Handy et al., 2005). At the community level, building of motorways and expressways divides the landscape into separate zones and as a result, the neighboring communities suffer from the "social exclusion" phenomenon along with transport poverty, and noise pollution issues (Litman, 2003; Mohan, 2002). In addition, per capita traffic related crashes increase with high levels of per capita automobile travel. Deaths and injuries resulting from road crashes are acknowledged to be a serious global health concern in the standard literature. Moreover, increased vehicular traffic on roadways causes acute traffic congestion leading to travel time delays and financial losses (excess fuel usage and lost time work). For example, in Canada, the total annual cost of congestion (in 2002 dollars) ranged from $2.3 billion to $3.7 billion for the major urban areas and more than 90 percent of this cost represented the value of the time lost to auto travelers (drivers and their passengers) in congestion (Canada T, 2006).

Automobile dependency is viewed primarily as a major detriment towards environmental sustainability due to the consumption of non-renewable resources and production of greenhouse gases (GHG) responsible for air quality deterioration and global warming (Chapman, 2007). It has been reported that globally transport sector accounts for about 13 percent of overall GHG emissions (IPCC, 2007) while road transportation alone accounts for about 74 percent of the total transportation carbon dioxide (CO_2) emissions (Rodrigue et al., 2006). The largest sources of transportation-related GHG emissions include passenger cars and light-duty trucks, including Sport Utility Vehicles (SUVs), pickup trucks, and minivans. These emissions not only degrade the environment, but affect various aspects of human health adversely (Selander et al., 2009) and increased dependency on private automobiles for daily travel is exacerbating the situation.

1.3 The Importance of Analyzing Automobile Ownership, Usage, and Expenditure

It is evident from the discussions in the previous sections that increased automobile ownership and usage are the two telltale signs of auto-centrism and the impacts of such over reliance on automobiles are far-reaching and pervade most aspects of everyday life with serious policy implications. Private car ownership plays a vital role in the daily travel decisions of individuals and households influencing a range of long-term and short-term decisions. In the long-term, vehicle ownership decisions are strongly tied with residential location and residential tenure (Eluru et al., 2010b). In terms of short-term decisions, the level of car ownership influences the various aspects of activity travel patterns including activity frequency, activity duration, activity location, and travel mode choice for out-of-home work and non-work pursuits (Bhat and Lockwood, 2004; Pucher and Renne, 2003; Bhat and Castelar, 2002). In addition, vehicle fleet size and usage are an integral part of the conventional four step travel demand forecasting process, influencing three of the four steps-trip generation, trip distribution, and mode choice. Thus, understanding the decision process along with its associated factors has become all the more important.

Closely intertwined with the actual decision of vehicle fleet ownership (type and number) and usage is the decision of allocation of resources in the transportation expenditure category. In fact, household budgetary allocation in general and transport budgetary allocation in particular affect a whole range of travel behavior choice processes. Specifically, in the long run, along with number of vehicles, residential location and housing inventory are reliant on household budgetary decisions while, in the short term, daily vehicle type choice (from current household fleet) and usage decisions, activity participation, and location decisions are affected by household expenditure allocation decisions. Clearly, these long term and short term decisions are likely to impact activity travel patterns significantly. Hence, it is beneficial to identify the determinants of household budgetary allocations to understand how households respond to varying situations due to policy measures, environmental concerns, fuel price fluctuations, and economic challenges. On that note, given the strong influence on travel patterns, it would be useful to consider monetary allocation decisions as a precursor to modeling travel demand processes.

1.4 Scope of the Dissertation

Montreal is one of the largest cities in North America and the second largest Census Metropolitan Area (CMA) in Canada. The city is characterized by a diverse and relatively dense urban form (compared to most US cities of similar size) built up over many phases and a unique heterogeneous multimodal transportation system comprising of metro, commuter train, and an extensive bus service, and boasts to be one of the two least vehicle dependent regions in Canada. Encouragingly, the city has a relatively high share of transit ridership (for a North American city) (Eluru et al., 2012a). Moreover, with 500 km of recreational and on-street paths, Montreal is also one of North America's bicycle friendly cities . In addition, the city also offers car sharing (such as Communauto) which is deemed as an ideal complement to the assortment of alternative transportation options already available as mentioned before (de Lorimier and El-Geneidy, 2013). However, recent years have seen a rapid extension of the suburban sprawl further into the surrounding regions. In fact, in the Greater Montreal Area (GMA), the average household vehicle ownership has increased from 1.06 in 1987 to 1.25 in 2008 (Roorda et al., 2008).

On the other hand, Quebec City, the provincial capital of Quebec does not have a lively and economically active urban core as that of Montreal. In fact, unlike Montreal, the city is characterized by a more sprawled urban form, lower population densities, and land use mix. According to Statistics Canada (2010), the population of the Quebec City census metropolitan area (CMA) increased by 62% in the period between 1971 and 2006. At the same time, the built area increased by 261% which is four times higher than the population increase. Moreover, reduced transit options and other transportation alternatives coupled with well-developed highway system have resulted in high rate of motorization, particularly among the inhabitants of the suburbs.

Canada, as a country has diverse socio-demographic composition and economic traits. From the discussion above, it is readily understandable that its urban regions offer reasonable contrast in terms of size and land use mix as well as public transportation system and culture, thus making them ideal subjects to study vehicle ownership and usage decision of households. However, vast majority of the extant literature is US centric, concerning US urban areas and unfortunately, there is a paucity of literature in the context of Canadian urban regions (Potoglou and Kanaroglou, 2008a; Roorda et al., 2000; Potoglou, 2008). The concerns are twofold. First, transferability of US evidence to the Canada/Quebec context may not be adequate given socio-cultural, vehicle fleet, urban form, and mobility pattern differences. To be sure, the land use and transportation patterns of the Canadian urban regions were found to fall in the middle spectrum of the overly auto reliant culture of the US cities and heavy transit orientation of the European cities (Gilbert, 2003; Kenworthy, 1991). Second, most of the studies are restricted to a particular urban area. This makes it difficult to assess how overall built environment and regional characteristics can affect the vehicle ownership, type, and usage decisions. For example, the impact of land use mix may differ across cities of different sizes or that have different transportation systems and/or spatial configurations.

Within the Canadian context, the current dissertation is focused on the overall structure of vehicle ownership with emphasis on household budget allocation, vehicle holdings, and usage process. The vehicle decision process considered consists of long term and short term process. In the long term, the decision hierarchy includes the allocation of monetary resources to various consumer expenditure categories and long term vehicle fleet choices. Transportation expenditure represents a large portion of the household budget, being second only to housing expenses incurred by households (Haas et al., 2009; Thakuriah and Liao, 2005). Needless to say, the amount of resource that households allocate to transportation has a direct impact on their vehicle fleet size and usage decisions. In fact, in developed countries like Canada, private vehicle costs (costs associated with purchase/lease, maintenance and operation of vehicles) dominate the household transportation expenses. In conjunction with transportation expenditure, vehicle fleet size decisions are considered.

In the short-term, conditional on the availability of the fleet, decisions regarding vehicle usage (type and mileage) and travel destinations are considered. Thus, a combined exploration of these aspects would provide a more complete and clear understanding of the factors associated with these decision processes. Such understanding is needed to develop policies for creating choice environments conducive towards sustainable transportation practices. However, it is not possible to cover every single aspect associated with these decision processes. Hence, the scope of the dissertation is limited to the examination of transport expenditure of households across Canada along with the vehicle ownership and usage analysis for two unique Canadian cities of Quebec province, Montreal and Quebec City.

1.5 Gaps in the Literature

In the past few decades, a considerable number of studies have examined travel choice decisions and the associated travel costs of households (see de Jong et al., 2004; Bunch, 2000). As a result of this continued research effort, travel behavior literature is replete with studies on various forms of transportation expenditure and auto-ownership modeling. Despite the extensiveness of literature, several methodological and empirical gaps still exist. The next objective(s) of the dissertation is to bridge these gaps in the standard literature. In the following subsections, the gaps are discussed in detail along five specific directions: (1) transport expenditure, (2) population heterogeneity, (3) econometric techniques (the debate on ordered vs. unordered regression models), (4) multiple time point data pooling, and (5) short term vehicle usage.

1.5.1 Transport Expenditure

Earlier transport research has predominately focused on household travel time budgets while transportation expenditure has received much less attention (see for example, Golob, 1990; Moriarty, 2002). Extant studies suffer from one or more of the following shortcomings as well. First, the focus of traditional travel behavior literature is on transportation related expenditure alone. There has been little research to analyze transport expenditure in conjunction with the array of commodities, goods, and services that households incur expenses on, thus limiting the ability to investigate the potential substitution or complementarity amongst the different expenditure categories (except Ferdous et al., 2010). Second, earlier studies have developed quantitative models almost exclusively with single year cross-sectional databases (except Thakuriah and Mallon-Keita, 2014). As a result, they are able to provide only a snapshot of the transportation expenditure pattern and not able to capture patterns that evolve with time due to technological advances or temporal factors. While such analysis is very useful, there is no consideration of household evolution and global socio-economic factors in the decision process. For example, how households respond to various temporal shocks – such as recession or a sudden spike in gas prices cannot be accommodated within the budgetary process unless household budgetary allocation decision framework is developed for a longer duration. Moreover, all of the earlier studies ignore the multiple discreteness of the expenditure categories (except Ferdous et al., 2010), hence, they do not represent behavior appropriately.

1.5.2 Population Heterogeneity

In the travel behavior literature, the widely applied discrete choice models in the earlier studies assume that the influence of exogenous variables remains the same for the entire population of households. It is very restrictive assumption. To illustrate the importance of varying impact of exogenous variables, let us consider the car ownership decision outcomes of two households (H1 and H2) with the same attributes except for transit accessibility variable; H1 has low accessibility and H2 has high accessibility. Now let us consider the influence of "number of employed adults" variable in these households. H1, with low transit accessibility, is inclined to have higher vehicle ownership with increased number of employed adults. On the other hand, for H2, the household with high transit accessibility, the increasing number of employed adults might not increase vehicle ownership (at least not at the same magnitude as for H1). This is an example of how transit accessibility moderates the influence of "number of employed adults" in determining vehicle ownership. If population homogeneity is imposed on the "number of employed adults" variable, the resulting coefficient would be incorrect. The illustration provided is a case of one variable (transit accessibility) moderating the influence of another variable (number of employed adults). However, in the context of vehicle ownership, it is possible that multiple variables might serve as a moderating influence on a reasonably large set of exogenous variables. The phenomenon is termed as "systematic heterogeneity" or heterogeneity corresponding to observed components. Surprisingly, the role of systematic heterogeneity in the vehicle ownership context has not been investigated in the vast existing literature.

Several econometric approaches can be adopted to relax the assumption to allow for population heterogeneity. The simplest method employed for addressing it is to introduce interaction effects of various exogenous variables (see Tang and Mokhtarian, 2009). While this will definitely improve the model, it might not always be adequate to capture the variability in the data. A second approach is to employ mixed or random coefficient versions of the ordered and unordered models (Eluru and Bhat, 2007; Bhat, 1998; Nobile et al., 1997; Nolan, 2010). In this approach, although the mean of the random coefficients can be allowed to vary across households based on observed variables, the variance and the distributional form of a random coefficient is restricted to be the same across all households. Besides, an *a priori* distributional form has to be imposed on the random coefficients, and the normal distribution assumption is usually chosen even though there is no reason why other distribution forms may not be more appropriate. Additionally, these approaches are more focused on the error component of the model and require extensive simulations for model parameter estimation. Employing simulation based approaches for model estimation has significant implications for statistical inferences because of loss of accuracy in the estimation of the variance-covariance matrix (Bhat, 2011). The advances in simulation have resulted in the widespread use of these approaches. However, prior to enhancing our understanding of the unobserved component, it is necessary to focus our attention on the systematic component (observed variables).

Another approach to allow for heterogeneity effects is to consider segmenting the population based on exogenous variables and estimate separate models for each segment. However, because there may be many variables to consider for the segmentation scheme, the number of segments to be considered can increase rapidly. Moreover, some segments might have very small sample size and this can cause problems in estimation (see Bhat, 1997). To overcome these issues, clustering techniques that allow us to segment the population based on a multivariate set of factors have also been used by the researchers (Depaire et al., 2008). However, the approach still requires allocating data records exclusively to a particular segment leading to efficiency loss in estimation. Further, the possible effects of unobserved factors that may moderate the impact of observed exogenous variables are not considered.

An elegant alternative approach to accommodate heterogeneity is to undertake an endogenous (or sometimes also referred to as a latent or finite mixture modeling) segmentation approach (see Bhat, 1997). Since the segments are unobserved to the analyst, they are termed as latent or endogenous. Recent research in various transportation fields has seen a revival of interest in the latent class models (Walker and Li, 2007; Eluru et al. 2012a; Yasmin et al., 2014; Sobhani et al., 2013; Greene and Hensher, 2003; Xie et al., 2012; Wafa et al., 2015). This approach recognizes that decision makers can be probabilistically assigned to different behaviourally similar segments as a function of observed attributes (Bhat, 1997; Srinivasan et al., 2009). Within each segment, separate discrete choice models predict decision maker behavior. This approach may be viewed as a combination of the above approaches. The approach considers a multivariate set of exogenous variables in the segmentation. Moreover, there is no need to specify a distributional assumption for the coefficients (Greene and Hensher, 2003). Therefore, vehicle ownership analysis could benefit from the application of the latent class versions of the ordered and unordered models.

1.5.3 Ordered or Unordered? The Eternal Debate

Household vehicle ownership variable is often compiled in travel surveys as an ordinal discrete variable. However, ordered and unordered, both model frameworks are equally applicable to analyze the decision process. Many approaches exploit the inherent ordering of the discrete variable by employing ordered response (OR) models. But the traditional OR formulations such as ordered logit (OL) or ordered probit (OP) impose a restrictive and monotonic impact of the exogenous variables on the vehicle fleet size alternatives. To overcome this issue, researchers have resorted to the unordered response (UR) models such as multinomial logit (MNL) that allow the impact of exogenous variables to vary across observed vehicle ownership levels (Potoglou and Kanaroglou, 2008a; Potoglou and Susilo, 2008). As a result, these models offer greater explanatory power because of the additional exogenous effects that can be explored. However, the increased flexibility from the unordered models is obtained at the cost of neglecting the inherent ordering of the vehicle ownership levels. The applicability of the two different frameworks has evoked considerable debate on what is the appropriate framework of analysis.

Recently, generalized ordered logit (GOL) model (proposed by Terza, 1985) has emerged as a mathematical equivalent of the MNL model. This model relaxes the monotonic effect of exogenous variables of the traditional ordered models and allows the analyst to estimate the same number of parameters as the MNL for an ordinal discrete variable while still recognizing the inherent ordered nature of the variable (Eluru et al., 2008). Importantly, recent evidence comparing the performance of the GOL model with its unordered counterparts has established that the GOL model is indeed an appropriate framework to study ordered variables (Eluru, 2013; Yasmin and Eluru, 2013). Bhat and Pulugurta (1998) demonstrated how an unordered model offered superior data fit by undertaking a comparison exercise using multiple datasets. However, their comparison exercise was limited to traditional OL and MNL models only. It can be argued that an exercise comparing alternative frameworks is incomplete without considering the generalized ordered logit (GOL) model. Unfortunately, such comparison exercise, particularly in examining how the population heterogeneity capture varies across ordered and unordered models in the context of vehicle ownership has not been undertaken so far.

1.5.4 Multiple Time Point Data Pooling

Earlier vehicle ownership studies rely mainly on cross-sectional data collected for a single point in time. As a result, these research efforts are only able to provide a snapshot of vehicle fleet size decision process and offer useful insights on the role of exogenous variables on vehicle ownership/type decision processes. However, they are unable to capture the evolution of vehicle ownership/type over the years. In order to study the evolution, that is affected by life cycle changes (such as the birth or moving out of a child) and/or land use and urban infrastructure and perception (such as introduction of improved transit facilities), longitudinal databases that track vehicle ownership decisions of the same households across multiple years are likely to be more informative (Woldeamanuel et al., 2009). Unfortunately, compiling such detailed data is prohibitively expensive and provides many challenges associated with respondent fatigue and retention (Hanly and Dargay, 2000; Yee and Niemeier, 1996).

One innovative approach to overcome the data availability challenge is to compile several cross sectional data sources over multiple time points. The availability of multiple cross sectional datasets for different years provides a useful compromise between a single year cross sectional dataset and a truly longitudinal dataset compiled across multiple years. Though the multiple waves are not compiled based on the same set of households, they still provide a welcomed opportunity to examine the impact of technology, altering perceptions of road and transit infrastructure, changing social and cultural trends on vehicle ownership (Dargay, 2002; Dargay and Vythoulkas, 1999). Further, pooled datasets allow identifying how the impact of exogenous variables has altered with time. For example, with improved perception of public transit, impact of a metro stop near the household might affect vehicle ownership reduction more in 2010 compared to its corresponding impact in 2000. Examination of vehicle ownership using pooled cross-sectional data (thus forming a pseudo-panel) is an overlooked arena of research although the technique has been successfully used for examining other travel behavior dimensions (see for example Sanko, 2013).

Data pooling of different respondents across multiple waves offers unique methodological challenges. The methodology should recognize the differences across multiple time points adequately. Specifically, the choice process for the respondents in a particular year might be influenced by various observed and unobserved attributes (Train, 2009; pp. 40-42). For example, if there is a significant spike in households with multiple employed individuals (from say 1995 to 2005) the vehicle ownership pattern might alter substantially across these two databases. This is an instance of how observed attributes affect vehicle ownership decision process. The outcome based models can accommodate such transitions reasonably through appropriate model specification ("number of workers in a household" variable). However, say we are interested in measuring the impact of growing environmental consciousness between 2000 and 2010 on vehicle ownership or in measuring the impact of psychological stress due to uncertainty in the job sector between 2000 and 2010 on household monetary expenditures. This is the case of unobserved variables (as it will be very hard to define exogenous variable of this type) specific to the study time period on the decision process. The accommodation of such unobserved effects becomes crucial in the analysis process. Hence, modeling approaches which are capable of simultaneously accommodating the influence of observed and unobserved attributes needs to be formulated and implemented.

1.5.5 Short Term Vehicle Usage

In addition to vehicle fleet size, the importance of the vehicle type choice decision is well recognized in the travel behavior literature. Traditionally, vehicle fleet decisions are examined as a long term choice with annual usage metrics in the previous research studies. These studies mostly aimed at examining the mix of vehicle holdings, and overall vehicle use, at the household level (Bhat and Sen, 2006; Bhat et al., 2009) and explored the association of individual and household demographics, residential and employment location attributes, and urban land-use and neighborhood characteristics with the decision process. Only recently, travel behavior models have started examining vehicle usage decisions (type and mileage) as a short-term decision in the context of activity travel analysis (see Faghih-Imani et al., 2014).

The emphasis of the literature on short term vehicle usage is on exploring the interaction of activity participation behavior with the vehicle type chosen on a per activity basis. The long-term vehicle usage observed (as studied in literature) is an aggregation of the household's yearly vehicle type and usage behavior. Thus by examining short term vehicle usage we explore, at a disaggregate level, the interaction of activity behavior and vehicle type choice. For activity based models, the long-term models will serve as control totals for vehicular usage while the short-term models will allow for enhanced prediction of daily vehicle type choice and usage. With the growing emphasis on emission modeling based on daily travel patterns it is important to accurately predict vehicle type choice at an activity level. Moreover, with the push toward integrated modeling approaches, there is a growing demand in the travel behavior literature on accommodating the possible interdependency across the choice dimensions in the modeling framework. Incorporation of interdependencies would provide a better understanding of the short term vehicle usage phenomena.

1.6 Objectives of the Dissertation

The dissertation aims to explore a wide spectrum of vehicle ownership related decision processes of Canadian households ranging from transportation expenditure (long term) to vehicle fleet size (long term) and usage (short term) while using comprehensive set of exogenous variables with particular focus on land use and urban form characteristics. Specifically, it is primarily motivated from the need to address the methodological and empirical issues as discussed in the previous section, thereby, contribute to the existing body of transportation and travel behavior literature. The following are the specific research problems that the research project intends to investigate elaborately.

The *first objective* is to examine household budgetary allocation and its evolution with a particular focus on transportation budget. In doing so, a multiple discrete continuous framework which recognizes that households choose to allocate budgets to multiple alternatives simultaneously will be employed.

The *second objective* of this research is to formulate, estimate and validate econometric models (ordered and unordered) accounting for systematic heterogeneity in the context of vehicle ownership. The performance of the formulated models with its traditional counterparts will also be evaluated to demonstrate the advantages of accommodating the effect of both observed and unobserved heterogeneity in examining vehicle ownership.

The *third objective* is to further enhance the vehicle ownership models and evaluate the performance of alternate (ordered and unordered) outcome frameworks for modeling vehicle ownership while accommodating for potential population heterogeneity.

The *fourth objective* of this dissertation project is to examine the evolution of vehicle ownership decisions using a pseudo-panel approach, thereby demonstrate a potential workaround for model development in the absence of detailed longitudinal data. Modeling frameworks capable of simultaneously accommodating the influence of observed and unobserved attributes on the fleet size decision of households across multiple time points will be implemented.

The *fifth objective* is to develop a joint econometric framework to investigate vehicle type choice at the individual activity level. In the unified framework, along with vehicle type categories, a parsimonious yet useful surrogate categorization of activity purpose and time of day will also be considered. Such a joint model can be used to analyze the interconnections among multitude of short term travel choices.

1.7 Outline and Contributions of the Dissertation

The remainder of the dissertation is structured in seven additional chapters in the following order.

A comprehensive review of the existing vehicle ownership literature is presented in *Chapter 2*. Different forms of vehicle ownership representation as well as the advantages and limitations of the methodological alternatives applied in the existing literature for examining vehicle ownership in the past two and a half decades (since 1990) are discussed in detail. In addition to the econometric modeling frameworks, parameter estimation methods, empirical findings, and existing data issues and challenges are also systematically documented. Important guidelines regarding methodology and contributory factors are gleaned from this review.

Chapter 3 investigates the evolution of transportation expenditure in relation to other household expenditure categories in Canada using a multiple discrete continuous framework. More specifically, the chapter proposes a methodology based on the scaled multiple discrete continuous extreme value (SMDCEV) model to simultaneously accommodate for the influence of observed and unobserved attributes on the budget allocation decisions of households across multiple time points. Public-use micro-data extracted from the Survey of Household Spending (SHS) for the years 1997 – 2009 is used for the model development purpose. Further, the SHS data is augmented with several annual economic indicators such as inflation rate, unemployment rate, gross domestic product (GDP), and wage rate. The applicability of the proposed model is demonstrated using a policy simulation exercise. This chapter contributes towards the first objective of this dissertation.

Chapter 4 proposes the use of latent class versions of traditional multinomial (LSMNL) and ordered logit (LSOL) models for examining vehicle fleet size decision of households. Specifically, the proposed models probabilistically allocate households into different segments based on land use and socio-demographic characteristics to recognize that the impacts of exogenous variables on vehicle ownership levels can vary across households based on both observed and unobserved factors. Moreover, the results of the comparison exercise (based on both aggregate and disaggregate measures of fit) of the developed latent class models with their traditional counterparts in the choice context examined is also presented. The proposed models are estimated using data derived from the Origin-Destination (O-D) surveys of Quebec City for the year 2001. This chapter contributes towards the second objective of this dissertation.

Chapter 5 addresses the question raised in the third objective by empirically comparing ordered and unordered models. In doing so, it uses as well as extends the modeling frameworks developed in *Chapter 4.* Latent class ordered logit (LSOL) model is extended to formulate latent class generalized ordered logit (LSGOL) model in order to create an equal footing for the comparison exercise. Several data fit comparison metrics (aggregate and disaggregate) are employed to evaluate the performance of the models and determine the appropriate model structure. The results from the exercise demonstrate the superiority of the generalized ordered framework in comparison with its unordered counterpart in modeling vehicle ownership decisions of households. In addition, the findings lend credence to the hypotheses that there is preference heterogeneity and that the heterogeneity can be explained in part by the observable land use attributes – thereby implicitly capturing the residential self-selection bias in the vehicle fleet size decisions. Another novel element of the research present this chapter is the identification procedure of the important variables affecting the class specific choice models. For the model development purpose, data extracted from the O-D surveys of Greater Montreal Area (GMA) for the year 2008 is used.

Chapter 6 contributes to the fourth objective of the dissertation by addressing the longitudinal data unavailability issue by stitching together multiple year cross-sectional datasets of vehicle ownership. Two variants of the generalized ordered logit (GOL) model – scaled GOL (SGOL) and mixed GOL (MGOL) models are applied to appropriately capture the impact of observed and unobserved attributes on vehicle ownership levels across the years. The applicability of the developed models is illustrated by computing elasticity effects and disaggregate level probability profiles. The proposed models are estimated using data derived from O-D surveys of Greater Montreal Area (GMA) for the years 1998, 2003, and 2008.

Chapter 7 contributes to the fifth objective by investigating the short term vehicle usage decisions. Specifically, a methodology for jointly examining the activity travel choice processes is proposed. Four decisions are considered: spatial flexibility of the activity, temporal flexibility of the activity, activity vehicle choice (characterized as vehicle type for auto users and other for non-auto users), and primary driver (for auto users). In terms of the econometric approach, a panel mixed multinomial logit (MMNL) model is applied to examine the simultaneity of these choice dimensions. The model specification accounts for the possible presence of common unobserved attributes among the joint choice alternatives. The empirical study is based on the longitudinal panel survey data of households in Quebec City.

Finally, *Chapter 8* concludes the dissertation by summarizing the findings from, and contributions of, this research along with a discussion on the potential implications of the important findings and suggestions for future avenues of research in this area.

CHAPTER 2 ALTERNATIVE MODELING APPROACHES USED FOR EXAMINING AUTOMOBILE OWNERSHIP: A COMPREHENSIVE REVIEW

2.1 Introduction

Given the wide ranging implications of over-reliance on private automobiles, household vehicle ownership and the associated dimensions including fleet size, vehicle type, and usage has been a topic of great interest to policy makers. Historically, models to investigate car ownership and usage have been under development since the 1930's (Whelan, 2007). The earlier literature has been focused on examining car ownership at an aggregate level (Holtzclaw et al., 2002; Clark, 2007). Since these studies analyze the ownership decision process at the national, regional or zonal level, they are effective in capturing the overall impact of urban form on the level of private automobiles in a city or region. Aggregate analysis is also considered cost efficient because of reduced data collection requirements (Potoglou and Kanaroglou, 2008a). Despite these advantages, the approach fails to capture the underlying behavioral mechanisms that actually guide the household decision process. Thus, their accuracy and policy sensitivity in practical applications is very limited (Kitamura and Bunch, 1990). On the other hand, disaggregate models, in which the "unit of observation" are individual households, alleviate many of these difficulties and can lead to more precise, detailed and policy relevant model findings (Eluru and Bhat, 2007). Therefore, more recent studies have focused on examining the car ownership decision at a disaggregate level (household level). We will focus on such household-level studies.

The methodological approaches applied to model car ownership range from simple linear regression to complex econometric formulations taking into account a rich set of covariates (Brownstone and Golob, 2009). The choice of model structure and functional form are typically driven by the objectives and context of the study. It is in this context that we undertake our review to examine the various methodological approaches employed in vehicle ownership modeling depending on the vehicle ownership representation.

2.1.1 <u>Vehicle Ownership Representation</u>

The dimension of crucial interest in vehicle ownership analysis is how to represent the ownership in the decision process. The methodological framework and policy analysis components are heavily reliant on the characterization of this decision process. In the extant transport and travel behavior literature, several representations of the automobile demand of households have been employed. In fact, the vehicle ownership representation provides us a clear framework for classifying the various research efforts examining vehicle ownership decision processes as highlighted in the subsequent discussion.

The simplest of the vehicle representation decision processes is the decision of how many vehicles to own or “auto ownership level” at a particular point of time (for example, see Manski and Sherman, 1980; Mannering and Winston, 1985; Golob, 1990; Hensher, 1992; Bhat and Pulugurta, 1998; Potoglou and Susilo, 2008). With the growing emphasis on vehicular emission modeling, there has been considerable work on modeling household fleet composition in terms of the mix of vehicle types (such as sedan, van, pick-up truck, Sport Utility Vehicle (SUV)) owned by a household (for example, see Mohammadian and Miller, 2003a; Choo and Mokhtarian, 2004). This group of studies are referred to as exogenous static models in our review i.e. studies that treat vehicle ownership as independent of other decisions.

Another line of inquiry is focused on examining the influence of one component of vehicle ownership on another component of vehicle ownership. For instance, it is plausible that individuals that have unobserved inclination for purchasing a pick-up truck are likely to have a positively influencing unobserved component for accumulating mileage with it. In fact, there is growing evidence to indicate that unobserved factors that influence household’s vehicle type purchasing decisions also impact the usage decisions for that vehicle. The examination of vehicle ownership models also reveals significant influence of land use and urban form on the vehicle fleet decision process (Schimek, 1996; Yamamoto, 2009; Li et al., 2010; Zegras, 2010). However, recent studies have demonstrated that incorporating land use and built environment as mere exogenous variables is not accurate as households have inherent preferences for residential location decisions thus leading to self-selection (Pinjari et al., 2008; Pinjari et al., 2011). There have been research efforts that attempt to capture the influence of other decision processes on vehicle ownership decisions. The process of accommodating for influence of additional dimensions is along the same lines of accounting for influence of unobserved components in the joint modeling of various components of vehicle ownership. In our review, these set of studies are together referred to as the endogenous static models.

The vehicle ownership representations discussed above are based on a snapshot of the vehicle ownership profiles. However, behaviorally households pass through a vehicle fleet decision process over time that includes vehicle purchase and vehicle disposal/sale. The changes to household vehicle fleet might be triggered by many events such as the birth of a child, changes to marital status affecting the vehicular requirements of the household. To elaborate, consider two households – first household with two older adults (male and female aged between 50 and 60) and second household with two younger adults (male and female aged between 25 and 30) with exactly identical vehicle fleet – a sedan and a coupe. In the static representation of vehicle ownership, these two households have the same dependent variable i.e. 2 cars (or if you consider vehicle type – 1 sedan and 1 coupe). However, the evolution process at play in the two household's would have been extremely different. For example, in the first household the couple might have moved to the current vehicle fleet as their kids have moved out and since they no longer required the SUV, they replaced it with a coupe. On the other hand, the second household might have just purchased their second vehicle as both members are now employed. The subtle differences in the evolution process have important implications for how the vehicle fleet might be altered in the future.

Naturally, research efforts have examined these decisions through a whole suite of models – vehicle holding duration, acquisition, disposal, and replacement models (Gilbert, 1992; Yamamoto et al., 1999). These studies consider the evolution of vehicle fleet i.e. they are not focused on the snapshot, but examine each vehicle fleet change decisions. This analysis allows analysts to see how life cycle changes in a household and existing fleet influence vehicle ownership decisions. These studies could examine vehicle ownership as a number or the more refined vehicle type characterization. The reader would recognize that all the vehicle ownership representations that consider vehicle ownership as a snapshot can be re-analyzed within this evolution framework giving rise to exogenous dynamic models and endogenous dynamic models. Of course, it is evident that the dynamic models require more detailed information on vehicle ownership decision processes compared to the static studies.

2.1.2 Contribution and Organization of the Chapter

The primary objective of this chapter is to provide a systematic overview and assessment of the methodological alternatives in the context of various potential representations of the vehicle ownership decision process as discussed in Section 2.1.1. To be sure, there have been earlier efforts to review the progress in modelling the vehicle ownership decision process (see de Jong et al., 2004; Potoglou and Kanaroglou, 2008b; Bunch, 2000). The last two studies focus on a small sample of methodological frameworks in their review. However, de Jong et al. (2004) provides a very comprehensive review of vehicle ownership models developed for the public sector. The study discusses both aggregate and disaggregate models developed prior to 2002. In recent years, owing to advances in computing, many advanced frameworks are being applied to model vehicle ownership. We review these recently developed modeling approaches and document their application in the context of the vehicle ownership representation discussed above. To summarize, the models found in the existing literature are classified as: exogenous static, exogenous dynamic, endogenous static, and endogenous dynamic. In this chapter, we present earlier research on vehicle ownership in the four categories identified.

The remainder of the chapter is structured in the following order. Section 2.2 contains elaborate review of the methodological approaches employed to model household vehicle ownership in the past two decades (beginning from the 90's). In Section 2.3, different parameter estimation techniques are discussed. Section 2.4 discusses the empirical findings of the reviewed studies providing an in-depth understanding of the broad range of factors that either increase or decrease the household demand for purchasing automobiles as well their type choice and subsequent usage. In light of the review, several prevailing data and methodological issues are discussed in Section 2.5. Finally, the chapter wraps up by summarizing the review in Section 2.6.

2.2 Methods

A summary of earlier studies (since 1990) classified based on the four vehicle ownership representations identified above is provided in Table 2.1. The table provides information on the study, data source, modeling methodology, vehicle demand form, what variables are considered including household demographics, individual, employment and life cycle attributes, built environment characteristics, transit attributes, policy scenarios, and unobserved effects. Several observations could be made from the table. First, most vehicle ownership studies are from North America (48 out of the 83 studies are from US and Canada). One quarter of the studies (22) is based on European data and a small number of studies are in the Asian (10), Australian (2) and South American (1) contexts. Second, for model estimation, the majority of studies (62 out of 83) rely on cross-sectional travel behavior surveys. Third, vehicle ownership decision has been mostly investigated as static exogenous choice using unordered choice mechanism with the most prevalent model structure being the multinomial logit (MNL). Fourth, household demographics and built environment characteristics (land use, urban form, and street network attributes) are the two most commonly examined exogenous variable groups. In recent years, the impact of transit attributes on the ownership decision process has also been investigated (30 out of 83)[2].

2.2.1 <u>Exogenous Static Models</u>

Within this group of models, the vehicle ownership decision process is considered in isolation of other choices. Based on the modeling approach employed, we have further sub-categorized the exogenous static models into standard discrete choice models, count models, advance discrete choice models, and other approaches.

2.2.1.1 <u>Standard Discrete Choice Models</u>

Researchers have most commonly applied binary logit or probit regression to represent binary car ownership levels of households, such as, owning a car vs. not owning a car. More specifically, these models capture the household's trade-off between the benefits (safety, privacy) of owning a private vehicle and disadvantages (higher travel time) of not owning it (Karlaftis and Golias, 2002; Li et al., 2010; Prillwitz et al., 2006). However, they do not distinguish the number of vehicles owned by households.

The issue of captive or loyal decision-making units (individual households) is another important aspect of car ownership modeling. In many instances, households, for one reason or another (financial constraints or environmental consciousness), will never own a car. If this captivity or loyalty to a particular choice alternative is not taken into account during model calibration, it can lead to biased estimation of coefficients (Swait and Ben-Akiva, 1986). To handle this problem, Gaudry and Dagenias (1979) proposed the dogit model, which considers choice set composition rather than considering a universal choice set. Specifically, it allows for two choice sets – (1) choice set with just the chosen alternative and (2) choice set involving all alternatives.

[2] For the literature review, we primarily focused on travel behavior literature while augmenting with research from marketing literature. The review process involved a two pronged approach. First, we employed the standard research databases for literature search on vehicle ownership. Second, a comprehensive cascading search of research based on the references in highly cited research articles on vehicle ownership was conducted. The two approaches ensured we covered the broad spectrum of literature on vehicle ownership.

Of course, the model forms a special case of full latent choice set consideration approach (Basar and Bhat, 2004). Whelan (2007) applied hierarchical binary dogit model by introducing a market saturation term for each level of household car ownership which would account for the range of reasons why some households are unable to acquire a vehicle or add to their existing stock. However, it should be noted that the dogit model structure, does not necessarily conform to the conventional utility maximizing theory (Whelan, 2007).

Household vehicle ownership variable is often compiled in travel surveys as an ordinal discrete variable. Naturally, many approaches exploit the inherent ordering of the discrete variable by employing ordered response models (ORMs). The most commonly used ORMs in the modeling of auto ownership are the traditional ordered logit (OL) (see, Bhat and Pulugurta, 1998; Hess and Ong, 2002; Kim and Kim, 2004; Potoglou and Susilo, 2008; Potoglou and Kanaroglou, 2008a) and ordered probit (OP) (see, Pendyala et al., 1995; Chu, 2002; Matas and Raymond, 2008; Potoglou and Susilo, 2008; Ma and Srinivasan, 2010) models. These models are motivated and derived in a latent variable framework. The reader would note that both OL and OP would produce similar results (Greene, 2003), therefore, either one of them could be chosen for the intended analysis.

The unordered multinomial discrete outcome models do not explicitly take into account the ordinal nature of the observed levels of car ownership. Rather, the mechanism is based on the random utility maximization (RUM) principle. According to this principle, the decision makers (households) associate a certain level of utility (or profit) with each car ownership level/type and choose the level/type that yields the maximum utility (see, Bhat and Pulugurta, 1998; Wu et al., 1999; Ryan and Han, 1999; Choo and Mokhtarian, 2004; Bento et al., 2005; Soltani, 2005; Potoglou and Kanaroglou, 2008a; Potoglou and Susilo, 2008; Potoglou, 2008; Zegras, 2010; Caulfield, 2012; Wong, 2013). Besides its closed form solution and computational simplicity, the standard multinomial logit model (MNL) also has the advantage of increased flexibility in model specification. That is, unlike OL or OP models, the MNL model does not place any restrictions on the effect of household characteristics across car ownership levels (Kim et al., 2007; Savolainen and Mannering, 2007). The additional flexibility, however, results in the estimation of substantially more parameters (Bhat and Pulugurta, 1998; Washington et al., 2011). Moreover, the traditional MNL model is also susceptible to the violation of independence of irrelevant alternatives (IIA) property.

In case the IIA property is not likely to be valid, the nested logit (NL) model structure has been suggested as an appropriate generalization of the MNL model. This model allows for correlation between the utilities of alternatives in common nests (Koppelman and Sethi, 2008). In order to estimate the model, car ownership levels or vehicle types that are similar to each other (due to unobserved preferences) are grouped into nests (see, McCarthy and Tay, 1998; Kermanshah and Ghazi, 2001; Mohammadian and Miller, 2002; Mohammadian and Miller, 2003a; Cao et al., 2006; Guo, 2013).). For instance, vehicle fleet decision can be partitioned into two levels, with vehicle availability (owning zero car vs owning car) being the first level while owning one car and owning two or more cars forming the second level and a two level NL model can be estimated (Kermanshah and Ghazi, 2001). In the context of vehicle type choice, McCarthy and Tay (1998) argued that vehicle makes/models can be nested according to their fuel efficiency, i.e. make/models in each fuel efficiency nest have similar unobserved characteristics and, accordingly, are likely to be correlated. Hence, they estimated a two level NL model for new vehicle purchase choices, where the first level contained three branches (low, medium and high fuel efficiency), and the second level contained all make-model combinations in the respective fuel efficiency category. Again, another possible correlation across alternatives is the correlation with adjacent alternatives – i.e. owning 2 cars is closely related to owning 1 car and 3 cars; an ordered generalized extreme value (OGEV) model (Small, 1994) can accommodate such structures. The assignment of alternatives to positions in the nesting structure and the number of nesting levels is the prerogative of the analyst. However, the NL model retains the restrictions that alternatives in a common nest have equal cross-elasticities and alternatives not in a common nest have cross-elasticities as for the MNL (Koppelman and Sethi, 2008).

2.2.1.2 <u>*Count Models*</u>

The observed automobile ownership levels of household are non-negative integers. Recognizing this property, several researchers have applied count data regression models to investigate the data. However, the application of count data regression models for modeling car ownership is not quite common.

The standard Poisson regression model assumes that the number of automobiles owned by household is independently Poisson distributed (see, Shay and Khattak, 2011). The Poisson model is based on the equal-dispersion assumption that the mean is equal to the variance. The assumption, however, is very restrictive because it does not hold in many cases, particularly when there is over or under-dispersion in the data. For instance, assuming a Poisson distribution for auto ownership data with problems of over-dispersion would result in underestimation of the standard error of the regression coefficients, which can lead to a biased selection of covariates. Moreover, the efficiency of the estimated parameters is also lost (Karlaftis and Golias, 2002).

The most common and extensively used approach to address the problem of inequality of mean and variance of the process is the negative binomial (Poisson-gamma model) model. Unlike Poisson model, the mean car ownership level is assumed to be random and follow a gamma probability distribution in this model (see, Shay and Khattak, 2005; Shay and Khattak, 2007). When the over-dispersion parameter is equal to zero, the negative binomial model reduces to Poisson regression model. The model has a closed-form solution; however, it is criticized by researchers for its incapability of handling under-dispersed data (Lord and Mannering, 2010).

In the car ownership literature, researchers have also used another modified version of the Poisson model termed as the Poisson-lognormal model (see, Karlaftis and Golias, 2002). In this model, the error term is assumed to be log-normal-rather than gamma-distributed. The model can account for unobserved heterogeneity and is more flexible than the negative binomial model. However, one important limitation of the model is that the marginal distribution of the model does not have a close form as the Poisson-gamma model (Winkelman, 2008).

The application of count models for household car ownership is quite restrictive because the household ownership variable rarely has values higher than 3 – thus allocating non-zero probability for a huge number of alternatives that are unlikely to be feasible for a large proportion of the population. Ideally, ordered response models are better suited to modeling vehicle ownership compared to the count models. In fact, in a recent paper (Castro et al., 2013) the authors' show that count models can be appropriately recast as ordered response models, providing further evidence that ordered models are more appropriate when the universal choice set is comprised of small number of categories.

2.2.1.3 <u>Advance Discrete Choice Models</u>

The traditional discrete choice models impose the restriction that the model parameters are same for the entire population – population homogeneity assumption. However, it is possible that the exogenous variable effects might vary across the population. Endogenous segmentation is an elegant approach for accommodating such systematic heterogeneity. The modeling technique has several appealing advantages. First, each segment is allowed to be identified with a multivariate set of exogenous variables, while also limiting the total number of segments to a number that is much lower than what would be implied by a full combinatorial scheme of the multivariate set of exogenous variables. Second, the probabilistic assignment of households to segments explicitly acknowledges the role played by unobserved factors in moderating the impact of observed exogenous variables. Third, within each segment, separate vehicle ownership representations can be estimated (unordered/ordered) to examine household choice behavior (see Beck et al., 2013). Finally, it circumvents the need to specify a distributional assumption for the coefficients (Greene and Hensher, 2003). It is important to note that latent class models are prone to stability issues in the estimation process. Such issues can be overcome by coding the log-likelihood function and its corresponding gradient function.

2.2.1.4 Other Approaches

In recent years, machine-learning techniques such as neural network or genetic algorithm (GA) are being applied to traffic and transportation problems. Mohammadian and Miller (2002) applied multilayer perceptron artificial neural network (ANN) for predicting household auto choices and also compared the results with the outcomes of traditional discrete choice method – the NL model. Typically, a neural network structure consists of a series of nodes. These are: input nodes for receiving the input signals, output nodes for giving the output signals, and hidden or intermediate nodes. Also, there are weight factors that link the various nodes together in hierarchical manner and these are assumed to be fixed in ANN (Lord and Mannering, 2010). This technique is capable of identifying associations among different variables in the database in a much quicker time than the traditional discrete choice models. However, integration of the ANN models in complex integrated modeling framework is difficult. Moreover, their application for policy and sensitivity is also very limited due to lack of explicit sensitivity measures (Mohammadian and Miller, 2002).

2.2.1.5 Synopsis

It was evident from the review that standard discrete choice models are by far the most commonly employed modeling approach in the exogenous static category and majority of the studies either applied the ordered or the unordered response mechanism. However, two of these studies attempted to compare the performance of the ordered and unordered response structures (Bhat and Pulugurta, 1998; Potoglou and Susilo, 2008). Based on several measures of data fit, these studies concluded that unordered response mechanisms such as MNL are more appropriate for auto ownership modeling. Further, recently, advanced models such as the latent segmentation models are found to outperform their traditional counterparts. They are also theoretically superior because they can accommodate systematic heterogeneity and thus allow for enhanced policy analysis.

2.2.2 Endogenous Static Models

In this section, we consider approaches that allow for modeling vehicle ownership in conjunction with other household choice outcomes. The joint modeling of multiple choices presents various methodological challenges. Broadly, two methods are employed to undertake such analysis. In the first approach, standard discrete choice methods described earlier are employed to analyze joint choices by defining choice alternatives as combination of various choices (such as residential location and vehicle ownership levels). The second approach, considers methods that incorporate unobserved correlations/dependencies across choice processes. The actual form of the model developed is based on the mechanism employed to accommodate these correlations. Based on these two approaches, the range of models applied in the context of vehicle ownership include: standard discrete choice models, mixed multidimensional choice modeling techniques, discrete continuous models, copula based models, Bayesian models, simultaneous equation models, and structural equation models (SEM).

2.2.2.1 Standard Discrete Choice Models

Standard discrete choice econometric frameworks (discussed in detail in the exogenous static section of the chapter) are also used to simultaneously model auto ownership choice with other decision processes of households, such as, mode choice, trip chaining or residential location. More specifically, in this type of modeling, all choice dimensions are considered as endogenous and are modelled as single joint choice. For instance, Dissanayake and Morikawa (2002) developed a two level NL model to analyze vehicle ownership, mode choice, and trip chaining behaviors of households in Bangkok metropolitan region, Thailand. Recently, Salon (2009) applied the traditional MNL model for investigating the choices of car ownership and commute mode along with the choice of residential location of households in New York City. However, it has to be recognized that combining choice alternatives of multiple choice dimensions into one compound choice bundle can lead to a dramatic increase in the number of choices to be modelled. Moreover, neither of the approaches can be used when the travel attribute is continuous (Pinjari et al., 2011).

Weinberger and Goetzke (2010) applied multinomial probit (MNP) model to jointly analyze the automobile ownership/residential location while capturing the effect of person's previous observations and experiences on the decision process. MNP model can also be derived following the random utility theory with the disturbance term assumed to be multivariate normally distributed. It allows for the relaxation of the IIA assumption, thus ensuring unbiased coefficient estimates despite possible correlation among different car ownership levels (Weinberger and Goetzke, 2010). However, the outcome probabilities are not closed form and hence, the estimation of the likelihood function requires numerical integration of multi-dimensional integrals making the model computationally difficult and time consuming (Washington et al., 2011)[3].

2.2.2.2 Mixed Multidimensional Choice Modeling

In the unified mixed multidimensional choice modeling approach, various decision processes are jointly modelled. More specifically, a series of sub-models are formulated for different choice dimensions. For example, Bhat and Guo (2007) developed an MNL model of residential location and an OL model of vehicle ownership. In another study, Pinjari et al. (2011) extended this approach and consequently developed an MNL model of residential location, OL models of vehicle ownership and bicycle ownership, and an MNL model of commute mode choice. Very recently, Paleti et al. (2013c) used the MNP model in order to jointly model residential location choice and vehicle ownership choice process while controlling for the immigration status of residents. Within the choice continuum, the sub-model components are econometrically joined together by using common stochastic terms (or random coefficients, or error components) and the parameters for each choice dimension are estimated simultaneously. The modeling framework is capable of incorporating a multitude of interdependencies among the choice dimensions of interest, such as: self-selection and endogeneity effects, correlation of error structures, and also unobserved heterogeneity (see Bhat and Guo, 2007; Pinjari et al., 2011 for more details). These types of models

[3] To be sure, recent techniques proposed and implemented by Bhat and his colleagues (Paleti *et al.*, 2013a; Paleti *et al.*, 2013b; Paleti *et al.*, 2013c; Bhat 2011), circumvent the need to employ simulation for the computation of MNP models. However, there are still challenges associated with the deployment of these techniques for traditional transportation models.

are well suited for modeling cross-sectional data sources and they also overcome the limitations of the standard MNL and NL approaches (as discussed before) for modeling multi-dimensional choice processes. Similarly, Yamamoto (2009) developed trivariate binary probit model of simultaneous ownership of car, motorcycle, and bicycle and Anastasopoulos et al. (2012) analyzed household automobile and motorcycle ownership with random parameters bivariate ordered probit model. Along similar lines, Konduri et al. (2011) proposed a probit-based discrete continuous model specification for jointly modeling vehicle type choice and tour length.

2.2.2.3 Discrete Continuous Models

In several situations, vehicle ownership decision of households may be associated with the choice of multiple alternatives simultaneously (number and types of vehicles), along with a continuous component (e.g. vehicle use/mileage) of choice for the chosen alternatives (Pinjari, 2011).To account for such multiple discrete-continuous choice situations, a parsimonious econometric framework termed as the multiple discrete continuous extreme value model (MDCEV) was proposed by Bhat (2005) and extended in Bhat (2008). Since its inception, several researchers have applied the model and its variants for investigating the household vehicle holdings and use by vehicle type (see, Bhat and Sen, 2006; Ahn et al., 2008; Bhat et al., 2009; Vyas et al., 2012).

Bhat and Sen (2006) applied the mixed version of the MDCEV model that can accommodate unobserved heteroscedasticity as well as error correlations across the vehicle type utility functions. However, it does not have a closed-form probability expression, hence, requires the application of computationally intensive simulation-based estimation methods. Recently, Ahn et al. (2008) employed conjoint analysis and employed the MDCEV framework to understand consumer preferences for alternative fuel vehicles. In another study, Bhat et al. (2009) extended the MDCEV formulation to joint nested MDCEV-MNL model structure that includes a MDCEV component to analyze the choice of vehicle type/vintage and usage in the upper level and an MNL component to analyze the choice of vehicle make/model in the lower nest. Vyas et al. (2012) also used the same model formulation to jointly estimate the household vehicle fleet characteristics and identify the primary driver for each of the vehicles.

The model has several attractive features in comparison with the conventional single discrete or discrete-continuous models. For instance, it is derived from the basic random utility theory with closed-form probability expressions and is practical even for situations with a large number of discrete consumption alternatives (Bhat and Sen, 2006; Bhat et al., 2009). In fact, the MDCEV model simplifies to linear-in-parameters MNL model if each of the household chooses only one alternative (Bhat et al., 2009; Pinjari, 2011). Nevertheless, the model is not without limitations. When applied to vehicle fleet composition analysis, the MDCEV model structure assumes that the process of acquiring vehicles is instantaneous, i.e. households choose to purchase the number of vehicles they want to own as well as the vehicle type and use decisions at a given instant. In fact, in reality, the existing household fleet ownership evolves over time with choices made in the past influencing choices in the future. Hence, it is fundamentally at odds with the more realistic process of household vehicle ownership and fails to capture the dynamics associated with vehicle transactions (Eluru et al., 2010a). Further, MDCEV assumes that the total utilization of vehicles (or continuous mileage component) is exogenous to the model. Similar to the MNL model, the MDCEV model also can be enhanced through nested and generalized extreme value variants to accommodate for common unobserved correlations across alternatives.

2.2.2.4 Copula based Joint Multinomial Discrete-Continuous Model

In recent years, the copula approach has been employed by several researchers for modeling joint distributions, such as, vehicle ownership/type and usage. One important advantage of the approach is that the resulting model has a closed-form probability expression allowing for maximum likelihood based estimation (Bhat and Eluru, 2009). The copula approach to discrete-continuous models is based on the concept of multivariate dependency form for the joint distribution of random variables, in which the dependency is independent of the pre-specified parametric marginal distributions for each random variable (Spissu et al., 2009; Bhat and Eluru, 2009).

Spissu et al., (2009) employed this approach to jointly analyze the type choice and utilization of the most recently purchased vehicle. The vehicle type choice component takes the familiar random utility formulation. Eventually, the vehicle type choice probability expressions would correspond to the MNL model probabilities. In the modeling framework, the vehicle mileage model component would take the form of the classic log-linear regression. The dependency between the underlying propensity of vehicle type choice and vehicle usage decisions for the household is determined by the type and the extent of the dependency between the stochastic terms of the individual components. In this model, the copulas are used to describe the joint distribution of the error terms. The authors applied different copula functions to test the presence of different forms of dependency and found that the Frank copula model yielded the best fit.

2.2.2.5 *Bayesian Multivariate Ordered Probit and Tobit Model (BMOPT)*

Fang (2008) developed a BMOPT model comprised of a multivariate ordered probit model with correlated covariance matrix for vehicle type choice and a multivariate Tobit model (Amemiya, 1984) for vehicle usage using data augmentation and Markov Chain Monte Carlo algorithms. The model is easy to implement and provides a simpler and more flexible framework for handling multiple-vehicle households. However, the model becomes computationally intensive with increasing vehicle categories. In another study, Brownstone and Fang (2009) extended the BMOPT model developed in Fang (2008) to treat local residential density as endogenous.

2.2.2.6 *Simultaneous Equation System*

The model system comprises of mutually dependent discrete choice models. For instance, Chen et al. (2008) proposed a two-equation simultaneous equation system comprising of two endogenous variables: car ownership and the propensity to use cars. In their specification, car use for commute trips was observed but the underlying propensity to use the car was unobserved. The authors assumed that the latent propensity includes the unobserved traits/attitudes towards car use. In another study, Schimek (1996) employed this modeling technique to explore individuals' residential choices and travel decisions, with auto ownership being an intermediating variable. Bhat and Koppelman (1993) developed an endogenous switching simultaneous equation model including husband's income, wife's income, wife's employment choice, and household car ownership as endogenous variables. More specifically, car ownership choice of household was modeled as a two equation switching ordered probit model system and the wife's employment was used as the endogenous switch. The model captures the unobserved behavioral factors influencing wife's employment choice and the resulting car ownership decisions. Additionally, the model can be extended to incorporate other long term household decisions such as residential location improving the travel demand forecasting capability (Bhat and Koppelman, 1993).

2.2.2.7 Structural Equation Model (SEM)

The structural equation modeling approach (SEM) has been in use in different research fields since the 70's because of its capability to handle hypotheses from a number of viewpoints (Senbil et al., 2009; van Acker and Witlox, 2010). In the car ownership context, these models are applied to untangle the role of car ownership in mediating (car ownership can be the outcome variable in one set of relationships and at the same time, it can be a predictor of other travel behaviors) the complex relationship between the built range environment and travel behavior (see, Golob et al., 1996; Golob et al., 1997; Giuliano and Dargay, 2006; Cao et al., 2007a; Gao et al,. 2008; Senbil et al., 2009; van Acker and Witlox, 2010; Aditjandra et al., 2012; de Abreu e Silva et al., 2012). Since, car ownership acts as an intermediate link between location decisions and travel behavior, including it in a single equation model will result in biased results (de Abreu e Silva et al., 2012).

Theoretically, SEM has two components, factor analysis/measurement model and structural equation/model (van Acker and Witlox, 2010; Washington et al., 2011; Aditjandra et al., 2012). The measurement models identify latent constructs underlying a group of manifest variables (or indicators) while the structural equations describe the directional relationship among latent and observed variables. SEM system enables us to separate out three types of effects. These are: total, direct and indirect effects of the explanatory variables. The direct effect can be interpreted as the response of the "effect" variable to the change in a "cause" variable while the indirect effect is the effect that a variable exerts on another variable through one or more endogenous variables (Gao et al., 2008). The total effect is the sum of the direct and the indirect effects of a variable. For example, in the model developed by Giuliano and Dargay (2006), it is possible to measure both the direct effect of income on travel decisions and also the indirect effect, through income's effect on car ownership, via the effect of car ownership on travel decisions.

2.2.2.8 Synopsis

There is a large body of literature on joint modeling in the vehicle ownership context. These models explore the joint nature of the relationship between vehicle ownership and other decision processes (such as residential location or level of vehicle usage), thus accommodating potential endogeneity issues. The models are typically estimated using traditional cross-sectional travel survey data. To summarize, the SEM system appears to be the most popular of the joint models discussed in this section. However, the modeling method cannot adequately handle multinomial choice variables. Thus, in recent years, multidimensional choice modeling technique is gaining prominence. We found that the number of choice dimensions considered varies from 2-6 in the studies reviewed.

2.2.3 Exogenous Dynamic Models

In this section we discuss the models that capture the dynamic nature of the automobile ownership decision. These models are estimated using panel data sets that possess both cross-sectional and time-series dimensions (Woldeamanuel et al., 2009). Panel or longitudinal data sets are formed when sample of households are observed at multiple points in time and the observations are separated by a certain interval of time (usually one year) (Gilbert, 1992). These datasets provide analysts with multiple records for each household allowing richer model specifications incorporating intra-household and inter-household correlations. It is important to note here that due to the lack of availability of panel data several researchers have considered the use of pseudo-panel datasets – a dataset formed by stitching together multiple cross-sectional datasets is referred to as pseudo-panel data. The models discussed in this section include: standard discrete models, duration models, and random effects models.

2.2.3.1 Standard Discrete Choice Models

Pendyala et al. (1995) investigated the changes in the relationship between household income and vehicle ownership using longitudinal data from the Dutch National Mobility Panel Survey. They developed OP models for six time points to monitor the evolution of income elasticities of car ownership over time. Their analysis results indicated that elasticity of car ownership changes over time. More recently, Matas and Raymond (2008) also developed OP model using a pseudo-panel dataset.

As discussed in the exogenous static section, one important limitation of the traditional ordered model (OL or OP) is that it constrains the impact of the exogenous variables to be monotonic for all alternatives. The recently proposed generalized ordered logit (GOL) model relaxes the monotonic effect of exogenous variables of the traditional ordered models while still recognizing the inherent ordered nature of the variable (Eluru et al., 2008). In other words, the GOL model is a flexible form of the traditional OL model that relaxes the restriction of constant threshold across population (Srinivasan, 2002; Eluru et al., 2008; Eluru, 2013). The scaled GOL model is a variant of GOL that accommodates the impact of unobserved time points in the modeling approach. Specifically, a scale parameter is introduced in the system that scales the coefficients to reflect the changes in variance of the unobserved portion of the utility for each time point.

2.2.3.2 Duration Models

Duration models provide an elegant alternative for handling the unique features (e.g. right-censoring, left-censoring) associated with vehicle holding duration data (Washington et al., 2011; Bhat and Pinjari, 2008). These models have clear methodological and conceptual advantages over the more traditional regression models when it comes to incorporating time-varying covariates. At the same time, they are sufficiently flexible to incorporate unobserved heterogeneity. Despite the advantages, the application of duration models in the transportation field is fairly uncommon. In the extant car ownership literature, the most common duration-model approach applied by the researchers is the hazard-based model. The model is used to investigate the automobile ownership duration as well as vehicle transaction behavior as a function of characteristics of the car, the household, and the economy (see, de Jong, 1996; Yamamoto and Kitamura, 2000). The hazard function gives the probability that the ownership spell will end immediately after time *t*, provided that it did not end before *t*. The shape of the hazard function which has important implication for duration dynamics can be chosen to be parametric, semi-parametric or non-parametric. Some examples of fully parametric functional form of hazard functions are: exponential, Weibull, log-logistic, Gompertz, log-normal, gamma, generalized gamma and generalized F. Yamamoto et al. (1997) found that Weibull distribution provides better likelihood estimates for vehicle holding duration compared to negative exponential, Weibull, gamma, log-logistic and log-normal distributions.

According to the traditional duration analysis, the automobile ownership spell would end as a result of a single event. However, in reality, several types of events may result in the termination of the car ownership spell (e.g. acquire a new or used vehicle, replace with a new or used vehicle, dispose of without replacement). In such cases competing risk duration model may be estimated by defining separate hazards for each particular exit state (see, Gilbert, 1992; Yamamoto et al., 1999; Mohammadian and Rashidi, 2007; Yamamoto, 2008). Then the overall hazard would be the sum of all the event-specific hazards since the risks are associated with mutually exclusive events. However, Hensher and Mannering (1994) argued that such assumption of independence among risks may not be appropriate.

2.2.3.3 Random Effects Models

Researchers have argued that unobserved temporal correlation might exist among the observations of panel car ownership data. These correlations are of two types: inter- and/or intra-temporal. For example, unobserved household specific preferences (e.g. acquired taste for a certain lifestyle) might result in persistence in car holding decisions of households that are invariant over time which is labeled as "spurious state dependence". On the other hand, if persistence is caused due to unobserved but time varying transaction cost (e.g. resistance to change in ownership levels due to search and information cost), it is termed as "true state dependence". Both these types of state dependence have different policy implications and failure to account for these might result in biased model results (Kitamura and Bunch, 1990; Nolan, 2010). To account for these unobserved factors, researchers have applied random effects models which are extensions of the traditional logit and probit models.

Nobile et al. (1997) proposed a random effects MNP model of household car ownership level. In their paper, the correlation is accounted for by using a general form for the error term covariance matrix. According to the authors, most of the variability in the observed choices could be attributed to between-household differences rather than within-household random disturbances. Unlike random effects MNP model, random effects MNL model is not restricted to normal distributions and the simulation of its choice probabilities is computationally easier (Train, 2009). Moreover, with panel data, the lagged dependent variable can be added without altering the probability expression or estimation procedure. Hence, random effects logit model is considered to be more convenient than its probit counterpart for representing state dependence (see, Mohammadian and Miller, 2003b; Bjorner and Leth-Petersen, 2007). The mixed logit model can be employed in two mathematically equivalent forms as random coefficients or error components (Train, 2009; Bhat et al., 2008). Both these specifications are formally equivalent, but differ in interpretation (Train, 2009). Mixed generalized ordered logit (MGOL) model allows the impact of observed attributes to vary across the population (in addition to accommodating impact of unobserved time points). This approach is analogous to splitting the error term into multiple error components.

2.2.3.4 Synopsis

In terms of the model structure, researchers mostly used hazard based duration models (single and/or competing) to analyze vehicle ownership duration or vehicle transaction decision while random or mixed models were mostly employed to analyze vehicle ownership over different time periods. In our review, we found two different types of dynamic model applications: purely dynamic and pseudo-dynamic. Unfortunately, literature in the domain of dynamic analysis of vehicle ownership decision is limited, presumably due to rigorous and expensive data collection requirements. Pooling of multi-year cross-sectional data might be a potential approach for overcoming the problems associated with unavailability of panel data.

2.2.4 Endogenous Dynamic Models

In this section, we focus on methods that bridge the advanced modeling techniques from endogenous static models with either panel or pseudo-panel data. In our extensive review, we found four types of endogenous dynamic modeling systems that endogenously analyzed the vehicle ownership decision. These are: copula based joint GEV based logit regression model, multinomial probit model, structural equation system and simultaneous equation system.

2.2.4.1 Copula based Model

Eluru et al. (2010a) and Paleti et al. (2011) proposed a joint discrete-continuous copula based framework to investigate the simultaneity of residential location choice, vehicle count and type choice, and vehicle usage decision characteristics of households. In this framework, the decision of residential choice, and choice of no vehicle purchase or one of several vehicle types, is captured using a GEV-based logit model, while vehicle utilization (as measured by annual vehicle miles of travel or VMT) of the chosen vehicle type is modeled using a classic log-linear regression model. Moreover, the number of vehicles owned is endogenously determined as the sum of the choice occasions when the household selects a certain vehicle type. In this particular case, the number of choice occasions is linked to the number of adults in households linked with the information on the vehicle purchase sequence. The model framework has several attractive and advantageous attributes associated with it. It can accommodate the many dimensions characterizing joint residential choice and vehicle fleet composition/usage decision system. The model also has a closed-form expression for most of the copulas available in the literature, thus for parameter estimation, computationally intensive simulation-based procedures are not required. The model is also capable of capturing the impacts of the types of vehicles already owned on the type of vehicle that might be purchased in a subsequent purchase decision and thereby, the dynamics of vehicle fleet ownership decisions.

2.2.4.2 *Multinomial Probit Model (MNP)*

Paleti et al. (2013b) investigated the spatial dependence effects in the fleet composition decision of households by using a MNP model. Similar to Eluru et al. (2010a) and Paleti et al. (2011), this model is capable of endogenously estimating the number of vehicles of each type that a household acquires by using a synthetic choice occasion approach where households are assumed to purchase vehicles over a series of choice occasions.

2.2.4.3 *Structural Equation Model (SEM)*

Golob (1990) developed a dynamic SEM linking four dependent travel behavioral variables: car ownership, travel time per week by car, travel time by public transit, and travel time by non-motorized modes. The model was developed to capture the dynamics of travel time expenditure while accounting for panel conditioning and period effects. More specifically, the model treats vehicle ownership as ordered-response probit variables and all travel times as censored (tobit) continuous variables.

2.2.4.4 *Simultaneous Equation System*

It is very likely that the previous choice or experience of owning a car may lead to a decision to acquire or dispose of a car, thereby influencing current or later levels and types of car ownership (Hanly and Dargay, 2000; de Jong and Kitamura, 2009). To test this hypothesis, Kitamura (2009) developed a dynamic simultaneous equation system of trip generation and modal split between private car and public transit in which household car ownership level was an endogenous variable. In the model, each of the three elements is assumed to be dependent upon its own value at the preceding time point and this dependence is introduced by incorporating lagged dependent variables. In the equation system, the car ownership model is formulated using the ordered-response probit model, linear regression is applied to model trip generation, while the logistic response curve (estimated using a weighted least-squares procedure) is used to represent modal split. Recently, Rashidi and Mohammadian (2011) proposed a dynamic hazard based system of equations for vehicle ownership, residential mobility, and employment relocation timing decisions. In their study, work location and residential relocation are included as endogenous variables.

2.2.4.5 Synopsis

As is evident from above, endogenous dynamic models are still a rarity in household ownership literature. These models endeavor to capture the evolution of household's preferences over time in their vehicle purchase and/or retention decisions while considering the impact of life cycle changes and/or existing vehicle fleet information. Among the different modeling types, the joint discrete-continuous copula based framework is attractive since it can simultaneously investigate vehicle count and type choice, and vehicle usage decision characteristics of households over time.

2.3 Parameter Estimation Methods

Maximum likelihood estimation (MLE) technique is the most common method used for developing car ownership models. This method begins with the assumption that the sample is the most likely sample to have been drawn from the population (Salon, 2009). It is generally the "most efficient estimation procedure" (Wooldridge, 2002). In addition, closed-form functions exist for the most common distributional assumptions (Lord and Mannering, 2010). However, estimation of flexible model structures, such as mixed MNL models using MLE technique requires the evaluation of multi-dimensional integrals that might prove to be computationally time consuming. In addition to MLE, limited information maximum likelihood (LIML) method has also been used by researchers. In this procedure, each equation is estimated individually after appropriately accounting for the limited dependent nature of the endogenous variable.

In recent years, Bayesian estimating methods have gained popularity in other fields of transport research because of its capability to handle complex models. However, the application of this technique in the area of car ownership modeling has been quite limited so far. In the Bayesian models, a prior distribution for the model parameters is specified first. Afterwards, the likelihood functions are used to update the prior distributions and obtain posterior distributions for the model parameters (Savolainen et al., 2011). Unlike classical methods, these models do not require a large sample to insure the adequacy of asymptotic approximations and the parameter estimates are reasonable as well. Nobile et al. (1997) used a hybrid Markov Chain Monte Carlo (MCMC) method which combines the Gibbs sampler and the Metropolis algorithm for model estimation. However, the run time required for simulation can be substantially high owing to a high number of draws to be made per record and this limits the application of the estimation technique.

Very recently, Bhat (2011) proposed another novel estimation technique labeled as the Maximum Approximated Composite Marginal Likelihood (MACML). One major difference with the conventional maximum likelihood estimation (MLE) approach is that the likelihood function only requires the computation of univariate and bivariate cumulative distribution function (Paleti et al., 2013a; Paleti et al., 2013b; Paleti et al., 2013c). Among other estimation methods, asymptotically distribution-free weighted least squares (ADF-WLS) estimating method has also been used by researchers in recent times (Senbil et al., 2009; Wu et al., 2009) to accommodate for the existence of the non-normal independent variables. The method is similar to simple least squares regression, and the estimated parameters are consistent and asymptotically normally distributed. Similar to MLE, this procedure is also scale invariant; however, it has the disadvantage of requiring a larger sample size.

2.4 Empirical Findings

In terms of explanatory variables, the variable groups considered in earlier research include household socio-demographic characteristics (such as income, number of children, workers, non-workers, adults, retirees, commuters and, licensed drivers, household size, household head characteristics, family type), residential location attributes (such as urban/rural location, distance to the central business district (CBD), and population centrality-a measure which plots cumulative percent of population living at various distances from the CBD against distance), and built environment variables (such as dwelling type, residential density, population density, employment density, land use mix, transit accessibility, and urban design), vehicle attributes (such as purchase, maintenance and parking costs, cargo space and existing vehicle holdings and their use). In addition to the observed variables, variability in vehicle holdings due to unobserved effects has been explored in several of the studies as well. A summary of the variables considered in the studies are presented in Table 1. The most significant empirical findings from these studies for the different variable groups are briefly summarized here.

2.4.1 Individual and Household Demographics

It is well established that among the different socio-demographic characteristics, household income dominates the choice process of household auto ownership. High income households, irrespective of country and region, always have a stronger preference to own higher number of private cars compared to middle-and low-income families (Karlaftis and Golias, 2002; Soltani, 2005; Cao et al., 2007b; Li et al., 2010). With respect to vehicle type choice, affluent households preferred new SUVs and were less likely to own and use pickup trucks and vans (Choo and Mokhtarian, 2004; Bhat and Sen, 2006; Bhat et al., 2009; Spissu et al., 2009). These households also tended to have shorter vehicle holding durations, presumably because they can afford to alter their vehicle holdings (Yamamoto and Kitamura, 2000). Spissu et al. (2009) reported that middle-income households accumulate more mileage on cars, while higher income households accumulate more miles on coupes and sedans.

In terms of other household socio-demographics, presence (and number) of employed adults, and license holders increased the probability of owning multiple cars (Bhat and Pulugurta, 1998; Ryan and Han, 1999; Chu, 2002; Karlaftis and Golias, 2002; Potoglou and Kanaroglou, 2008a; Bhat and Guo, 2007). Both attributes represent the household's daily commitments and increased mobility needs. A number of studies have also found that households tend to purchase vans/minivans as their household size increases (Bhat and Sen, 2006; Potoglou, 2008). In terms of household composition, single persons and lone parents were most unlikely to own multiple cars while couples with or without children were the most likely to own higher number of vehicles (Caulfield, 2012). In addition, families residing in their own homes are likely to own multiple private vehicles compared to families that rented or lived in apartments (Li et al., 2010; Zegras, 2010). As a result of the parking space flexibility, home owners had higher likelihood of purchasing pick-up trucks (Potoglou, 2008) and also had lower probabilities of disposing of their vehicles than do households who resided in rented homes (Yamamoto and Kitamura, 2000).

Researchers have also investigated the impact of several household head characteristics such as, age, marital status, level of education, job type and position, transit pass possession and use on private car ownership of households. For instance, Kermanshah and Ghazi (2001) reported that households with the head aged between 41-65 years, tended to own two or more cars. In the same study, they observed that the tendency to own second family car was higher when the heads of the households were either employers or retailers. Households with older household heads were generally more likely to own vehicles of an older vintage compared to younger households (Bhat et al., 2009). The probability of owning at least one car increased when the head of the household was a male (Matas and Raymond, 2008). In terms of vehicle type choice, male drivers were found to prefer bigger cars (Mohammadian and Miller, 2003a; Spissu et al., 2009) while females are found to be less inclined to drive pick-ups (Choo and Mokhtarian, 2004).

Apparently, when there are more children in the household or when the child is older, the mobility requirements of the households increase. This encourages households to acquire private vehicles (Yamamoto et al., 1999; Ryan and Han, 1999; Kermanshah and Ghazi, 2001; Soltani, 2005). In some studies, number of children was associated with reduced likelihood of owning cars. The result might seem counterintuitive at first glance. However, the negative effect of increased number of children on car ownership could be explained by the increased living expenses (food, clothing, and housing) that might curtail the amount of financial resources available for expenditures on acquiring and maintaining cars (Bhat and Koppelman, 1993). Presence of children in the household not only impacts vehicle fleet size, but also its composition. According to Bhat and Sen (2006), households with infants and young adults had a stronger preference for spacious vehicles such as the minivans. Few researchers investigated the impact of household race on vehicle ownership. Among them, Schimek (1996) and Hess and Ong (2002) found that white households had larger chances of owning multiple vehicles compared to non-white households. These households also had shorter expected ownership lengths (Gilbert, 1992).

2.4.2 Residential Location and Built Environment Attributes

In this group of exogenous variables, population and residential density are the two most researched urban form/built environment characteristics in the context of vehicle ownership of households. Increased population and residential density always has a negative impact on car ownership levels of households, presumably because these areas are relatively better served by alternative modes of transport (Schimek, 1996; Ryan and Han, 1999; Hess and Ong, 2002; Li et al., 2010; Zegras, 2010). Population density significantly impacts vehicle type choice of households as well. For instance, households residing in dense areas showed a stronger aversion towards pickup trucks and SUVs. In addition, several studies observed that car ownership decreased when the land-use mix increased (Hess and Ong, 2002; Chu, 2002; Soltani, 2005; Potoglou and Kanaroglou, 2008a; Yamamoto, 2009; van Acker and Witlox, 2010).

In addition to density, residential location variables also influenced car ownership and vehicle type choice decisions significantly. For instance, suburban and rural households were more likely to own pickup trucks compared to urban households (Cao et al., 2006; Bhat et al., 2009). Schimek (1996) and Bento et al. (2005) demonstrated that households had fewer cars when their locations was close to the center of the city. Interestingly, in Li et al. (2010), the distance from CBD was found to have a negative impact on car ownership levels of households of two Chinese cities. The result, though seemingly counterintuitive, might be pointing toward a divergence in attitudes of households towards auto ownership in developing and developed countries. Some studies also found association between employment density, job accessibility, and household auto ownership. Generally, households located in areas with higher employment density or job accessibility have lower level of vehicle ownership (Chu, 2002; Soltani, 2005; Gao et al., 2008) and lower preference for pick-up trucks (Spissu et al., 2009). Presumably, these households are less dependent on personal vehicles and more inclined to avail alternative modes of transport.

Another important determinant of car ownership is the transit accessibility measure usually captured as the proximity to transit stations (bus/rail), and transit supply. Increased transit access and high quality of transit service has a significant negative effect on the number of automobiles owned (Prevedouros and Schofer, 1992; Ryan and Han, 1999; Cullinane, 2002; Kim and Kim, 2004; Bento et al., 2005; Matas and Raymond, 2008; Potoglou and Kanaroglou, 2008a; Zegras, 2010) as well as purchase of larger vehicles (Spissu et al., 2009). Schimek (1996) and Hess and Ong (2002) illustrated that traditional neighborhoods with friendly walking and biking environments tend to reduce car ownership. In addition, Zegras (2010) demonstrated that grid street layout has a negative impact on vehicle ownership.

2.4.3 Others

In addition to the traditional socio-demographic and land use variables, some researchers have also explored the association of different life course events and attitudinal factors with the vehicle fleet size and type decisions of households. For example, change in residential location or moving increased the probability of fleet size alteration (Yamamoto, 2008). In another study, Weinberger and Goetzke (2010) found that former residence location of individuals has a significant influence on their choice of current automobile ownership level. People from non-metropolitan areas and other non-transit oriented cities were more likely to own higher number of cars, while people who moved from metropolitan areas with good transit system had a higher probability of owning fewer cars. Interestingly, individuals who are averse to travelling were more likely to drive luxury cars (Choo and Mokhtarian, 2004).

According to the findings of Karlaftis and Golias (2002), individuals who were more sensitive toward the value of their travel time were more inclined to own multiple cars and less likely to belong to an autoless household. Households, on average, preferred to purchase vehicle makes and models that are less expensive to purchase and operate (McCarthy and Tay, 1998). Additionally, the decision about keeping or disposing a car was mainly driven by income level of households and also by fuel expenditures (de Lapparent and Cernicchiaro, 2012). However, Wu et al. (1999) concluded that these costs had greater influences on mini-cars than on standard cars and that the preference toward standard cars was independent of their prices. Moreover, households with high incomes were found to be less sensitive to cost variables than were households with low incomes (Bhat et al., 2009). Dargay (2002) demonstrated that urban car owners were more sensitive to changes in motoring costs compared to their rural counterparts. This result suggests that car ownership in rural areas is a greater necessity. Mohammadian and Rashidi (2007) found that higher parking costs increased the willingness of a household to dispose of an extra vehicle or possibly trade it for a smaller vehicle.

Several researchers attempted to capture the common unobserved factors influencing car ownership and other decision processes as well. For instance, Pinjari et al. (2011) found a significant magnitude of common unobserved factors affecting auto ownership and auto mode choice as well as auto ownership and bicycle ownership propensity of households. In another study, unobserved lifestyle preferences of households were found to be associated with low density residential location and increased auto ownership levels of household (Paleti et al., 2013c). Interestingly, Bhat and Guo (2007) did not find any residential self-selection effect in car ownership propensity based on unobserved household specific factors. Moreover, similar lifestyle preferences might also deter individuals from choosing bicycle and walk as their commute mode (Paleti et al., 2013a). With respect to vehicle type and usage, Spissu et al. (2009) reported that unobserved factors that increased/decreased the likelihood of a household acquiring a specific type of vehicle also increased/decreased the vehicle mileage of that vehicle type. Among the different vehicle types considered in their study, the magnitudes of correlation were found to be higher for pick-up trucks and coupes.

2.5 Issues and Challenges in Vehicle Ownership Modeling

In spite of the advances described, there are issues that pose a formidable challenge to model vehicle ownership. In this section we highlight two main emerging issues that researchers need to consider in modeling vehicle ownership: (1) data and (2) spatial correlation.

2.5.1 Data Issues

The data used to model car ownership are limited due to the amount of information available from traditional household travel surveys or other data sources. Often times, not all variables affecting the decision process is collected; either because of survey length restrictions or because "measurement" is not possible at all, resulting in omitted variable problem. For example, let us consider an omitted variable (e.g. environmental consciousness) which is correlated with a measured variable (e.g. education) and the measured variable is found to be statistically significant in the car ownership model. The observed parameter might be spurious and the factor that is actually affecting the ownership decision might be the omitted variable. Omission of such relevant variables may lead to biased and inconsistent estimates of parameters and erroneous inferences (Kitamura, 2000; Lord and Mannering, 2010). Improved methods to consider such omitted variables in a revealed preference datasets need to be developed (see an example of such methods in the context of stated preference data for vehicle ownership in Daziano and Bolduc, 2013).

Another challenge with the data is the failure to recognize that travel behavior and urban form are evolving in continuous time. Rather than studying vehicle ownership as a snapshot using cross-sectional data, it is useful to consider the changes happening across time. Unfortunately, collection of panel data is prohibitively expensive, time consuming and has very low response retention rates. As an alternative, in recent times, a pseudo-panel approach that stitches together a series of cross-sectional datasets is used by the researchers to estimate dynamic car ownership models. These studies employ exogenous variable cohort averages in the analysis (Dargay and Vythoulkas, 1999; Dargay, 2002), thus resulting in a loss of data resolution. A more recent research effort that considers exogenous variables at a disaggregate level while explicitly accounting for unobserved correlation across each cross-section offers promise.

2.5.2 <u>Spatial Correlation</u>

The decision of vehicle fleet size and type might be heavily influenced by the choices made by neighboring households (Adjemian et al., 2010). If the neighbors own and drive hybrid electric vehicles, that household might become more environmentally conscious and purchase a hybrid electric vehicle (Chan et al., 2011; Paleti et al., 2013b). Spatial interdependence might also arise from unobserved attitudinal preferences such as peer pressure from social networks (Axsen and Kurani, 2012). That is, households who have a proclivity towards similar lifestyles might "cluster together" in neighborhoods that support their lifestyle preferences (Eluru et al., 2010a). Failure to account for such potential interdependence might result in biased parameter estimates. However, estimation of the discrete choice models accommodating spatial dependence effect requires evaluation of multidimensional integrals making the process intractable. A more recent effort proposed by Paleti et al., (2013b) is tractable and avoids simulation offering promise to incorporate spatial correlation in vehicle ownership studies.

2.6 Summary and Conclusions

This chapter reviews the disaggregate models examining household vehicle ownership that are developed over the last two decades (since 1990) using a four-way classification of the modeling frameworks. Specifically, the four model types discussed in detail are: exogenous static, endogenous static, exogenous dynamic, and endogenous dynamic. In each category, we begin by discussing the rudimentary models and then proceed on explaining the more complex models. Included in the discussion are the mathematical concepts behind the model development as well as the underlying behavioral reasoning, in the vehicle ownership context.

The research efforts using standard models in the exogenous static group offer useful insights on the role of exogenous variables (e.g. household socio-demographics, land use and urban form attributes, transit and infrastructure characteristics) on vehicle ownership decision processes. On the other hand, the endogenous models are motivated from the need to accurately analyze the interdependencies between different influential elements associated with vehicle ownership. Two major modeling streams can be found in the literature in this regard: joint discrete choice models involving nominal and/or ordinal endogenous variables, and structural equation models (SEM) involving continuous endogenous variables. The joint discrete choice models do not allow direct causality between their endogenous variables. Contrastingly, SEMs assumes direct mutual causality among endogenous variables. Simultaneous equation systems conceptually blend both these approaches, jointly modeling discrete endogenous variables as mutually dependent. More recent research on vehicle ownership has adopted dynamic models (exogenous and endogenous) that analyze vehicle ownership as a behavioral process that evolves over time. The common techniques employed in this domain include hazard based duration models, mixed effects model, and structural equation models.

In summary, the choice of model/s is guided by the objectives to be accomplished or issues to be addressed, data availability and most importantly, the nature of the dependent variable/s. In an attempt to aid researchers and practitioners, based on our extensive review and judgment, we provide a useful decision matrix table (see Table 2) for determining the appropriate model for various vehicle ownership contexts. We close with a cautionary advice that it is important to recognize that advanced models are not a substitute for accommodating observed heterogeneity in traditional models.

Table 2.1 Summary of Previous Studies on Automobile Ownership

Studies	Data Source & Type	Modeling Approach	Vehicle Demand Form	Variables Considered							Unobserved Effects
				Household Demographic	Individual Attributes	Employment Attributes	Life Cycle Attributes	Built Environment	Transit Attributes	Policy	
Exogenous Static (31)											
Karlaftis and Golias (2002)	Greece *Roadside Interviews*	Binary logit	VO	√	-	-	√	√	√	√	-
Li *et al.* (2010)	China *Household Survey*	Binary logit	VO	√	√	√	√	√	√	-	-
Ma and Srinivasan (2010)	USA *Census micro-data*	Binary probit	VO	√	√	-	√	√	-	-	-
Whelan (2007)	Great Britain *Travel Survey*	Binary dogit	VO	√	-	-	-	√	-	√	-
Bhat and Pulugurta (1998)	USA *Activity Survey* Netherlands *Travel Survey*	Ordered logit	VO	√	-	-	-	√	-	-	-
Hess and Ong (2002)	USA *Activity and Travel Survey*	Ordered logit	VO	√	√	-	-	√	√	-	-
Kim and Kim (2004)	USA *Travel Survey*	Ordered logit	VO	√	-	-	√	√	√	-	-
Potoglou and Susilo (2008)	USA Netherlands Japan *Travel Survey*	Ordered logit	VO	√	√	-	√	√	-	-	-
Potoglou and Kanaroglou (2008a)	Canada *Internet Survey*	Ordered logit	VO	√	-	√	√	√	√	-	-
Chu (2002)	USA *Travel Survey*	Ordered probit	VO	√	-	√	√	√	-	-	-
Cao *et al.* (2007b)	USA *Attitudinal Survey*	Ordered probit	VO	√	√	-	-	√	-	-	√
Potoglou and Susilo (2008)	USA Netherlands Japan *Travel Survey*	Ordered probit	VO	√	√	-	√	√	-	-	-
Ma and Srinivasan (2010)	USA	Ordered probit	VO	√	√	-	√	√	-	-	-

	Census micro-data										
Bhat and Pulugurta (1998)	USA *Activity Survey* Netherlands *Travel Survey*	Multinomial logit	VO	√	-	-	-	√	-	-	-
Wu *et al.* (1999)	China *Stated Preference Survey*	Multinomial logit	VT	√	√	√	√	√	√	√	-
Ryan and Han (1999)	USA *Census micro-data*	Multinomial logit	VO	√	-	-	-	√	-	-	-
Choo and Mokhtarian (2004)	USA *Attitudinal Survey*	Multinomial logit	VT	√	√	√	√	√	-	-	-
Bento *et al.* (2005)	USA *Travel Survey*	Multinomial logit	VO	√	√	-	√	√	√	√	-
Soltani (2005)	Australia *Travel Survey*	Multinomial logit	VO	√	-	-	√	√	√	-	-
Potoglou and Kanaroglou (2008a)	Canada *Internet Survey*	Multinomial logit	VO	√	-	√	√	√	√	-	-
Potoglou and Susilo (2008)	USA Netherlands Japan *Travel Survey*	Multinomial logit	VO	√	√	-	√	√	-	-	-
Potoglou (2008)	Canada *Internet Survey*	Multinomial logit	VT	√	√	-	√	√	-	-	-
Zegras (2010)	Chile *OD Survey*	Multinomial logit	VO	√	-	-	-	√	√	-	-
Caulfield (2012)	Ireland *Census Data*	Multinomial logit	VO	√	√	-	√	√	√	-	-
Wong (2013)	Macao *Travel Survey*	Multinomial logit	VO	√	-	-	√	√	-	-	-
McCarthy and Tay (1998)	USA *Consumer Survey*	Nested logit	VT	√	-	-	√	√	-	√	-
Kermanshah and Ghazi (2001)	Iran *Travel Survey*	Nested logit	VO	√	-	√	√	√	-	-	-
Mohammadian and Miller (2002)	Canada *Retrospective Survey*	Nested logit	VT	√	√	√	-	-	-	√	-
Mohammadian and Miller (2003a)	Canada	Nested logit	VT	√	√	√	-	-	-	√	-

	Retrospective Survey										
Cao *et al.* (2006)	USA *Attitudinal Survey*	Nested logit	VT	√	√	-	√	√	-	-	-
Guo (2013)	USA *Travel Survey*	Nested logit	VO	√	-	√	-	√	√	-	-
Potoglou (2008)	Canada *Internet Survey*	Random parameters logit	VT	√	√	-	√	√	-	-	-
Shay and Khattak (2011)	USA *Travel Survey*	Poisson regression	VO	√	-	-	-	√	-	-	-
Shay and Khattak (2005)	USA *Travel Survey*	Negative binomial regression	VO	√	-	-	-	√	-	-	-
Shay and Khattak (2007)	USA *Travel Survey*	Negative binomial regression	VO	√	-	-	-	√	√	-	-
Karlaftis and Golias (2002)	Greece *Roadside Interviews*	Poisson-lognormal model	VO	√	-	-	√	√	√	√	-
Beck *et al.* (2013)	Australia *Interviewer Assisted Online Survey*	Latent class multinomial logit	VT	-	-	-	-	-	-	√	-
Mohammadian and Miller (2002)	Canada *Retrospective Survey*	Artificial Neural Network	VT	√	√	√	-	-	-	√	-
Endogenous Static (29)											
Dissanayake and Morikawa (2002)	Thailand *Travel Survey*	Nested logit	VO	√	√	√	√	-	-	-	√
Salon (2009)	USA *Travel Survey*	Multinomial logit	VO	√	-	-	√	√	√	√	√
Weinberger and Goetzke (2010)	USA *Census Micro-Data*	Multinomial probit	VO	√	√	-	√	√	-	-	√
Bhat and Guo (2007)	USA *Travel Survey*	Mixed multidimensional choice modeling	VO	√	-	-	-	√	√	-	√
Yamamoto (2009)	Japan Malaysia *Trip Survey*	Trivariate binary probit	VO	√	-	-	-	√	√	-	√
Pinjari *et al.* (2011)	USA *Travel Survey*	Mixed multidimensional choice modeling	VO	√	√	-	√	√	√	√	√

Konduri *et al.* (2011)	USA *Travel Survey*	Probit-based joint discrete continuous model	VT	√	√	-	√	√	-	-	√
Anastasopoulos *et al.* (2012)	Greece *Travel Survey*	Random parameters bivariate ordered probit	VO	√	√	√	-	√	√	-	√
Paleti *et al.* (2013a)	USA *Travel Survey*	Mixed multidimensional choice modeling	VO	√	-	-	√	√	-	-	√
Paleti *et al.* (2013c)	USA *Travel Survey*	Bivariate multinomial probit	VO	√	-	√	√	√	-	-	√
Bhat and Sen (2006)	USA *Travel Survey*	Mixed multiple discrete continuous extreme value model	VT & VU	√	-	-	√	√	-	√	√
Ahn *et al.* (2008)	South Korea *Face-to-face Interview*	Mixed multiple discrete continuous extreme value model	VO & VU	-	-	-	-	-	-	√	√
Bhat *et al.* (2009)	USA *Travel Survey*	Joint nested multiple discrete continuous extreme value model	VT & VU	√	√	-	√	√	√	√	√
Vyas *et al.* (2012)	USA *Vehicle Survey*	Joint nested multiple discrete continuous extreme value model	VT & VU	√	√	-	√	√	-	-	√
Spissu *et al.* (2009)	USA *Travel Survey*	Copula based joint multinomial discrete-continuous model	VT & VU	√	√	-	√	√	√	-	√
Fang (2008)	USA *Travel Survey*	Bayesian multivariate ordered probit and tobit model	VO & VU	√	√	-	√	√	-	-	√
Brownstone and Fang (2009)	USA *Travel Survey*	Bayesian multivariate ordered probit and tobit	VO & VU	√	√	-	√	√	√	-	√

Schimek (1996)	USA *Travel Survey*	Two-equation system simultaneous equation model	VO & VU	√	√	-	√	√	√	-	√
Chen *et al.* (2008)	USA *Travel Survey*	Two-equation system simultaneous equation model	VO & VU	√	√	-	√	√	√	√	√
Bhat and Koppelman (1993)	Netherlands *Travel Survey*	Endogenous switching simultaneous equation model	VO	√	√	-	-	√	-	-	√
Golob *et al.* (1996)	USA *Telephone Survey & SP Survey*	Cross-sectional structural equation model	VT & VU	√	√	√	√	-	-	√	√
Golob *et al.* (1997)	USA *Telephone Survey & SP Survey*	Cross-sectional structural equation model	VT & VU	√	√	√	√	-	-	√	√
Giuliano and Dargay (2006)	USA Great Britain *Travel Survey*	Cross-sectional structural equation model	VO & VU	√	√	√	√	√	√	-	√
Cao *et al.* (2007a)	USA *Attitudinal Survey*	Structural equation model	VO	√	√	-	√	-	-	-	√
Gao *et al.* (2008)	USA Census Tract Data	Cross-sectional structural equation model	VO	√	√	√	-	-	-	-	√
Senbil *et al.* (2009)	Japan Malaysia *Travel Survey*	Cross-sectional structural equation model	VO	√	√	-	-	√	√	√	√
van Acker and Witlox (2010)	Belgium *Travel Behaviour Survey*	Cross-sectional structural equation model	VO & VU	√	√	√	-	√	-	-	√
de Abreu e Silva *et al.* (2012)	Canada *OD Survey*	Cross-sectional structural equation model	VO & VU	√	√	√	√	√	-	-	√
Aditjandra *et al.* (2012)	Great Britain *Quasi-longitudinal Data*	Quasi-longitudinal structural equation model	VO	√	√	-	-	√	-	-	√

Exogenous Dynamic (15)											
Prillwitz *et al.* (2006)	Germanay *Panel Waves*	Binary probit	VO	√	√	√	√	√	-	-	-
Yamamoto (2008)	Japan *Panel Survey*	Multinomial logit	VTR	√	-	-	√	√	√	-	-
Pendyala *et al.* (1995)	Netherlands *Mobility Panel Survey*	Ordered probit	VO	√	-	-	-	√	√	-	-
Matas and Raymond (2008)	Spain *Pseudo-Panel*	Ordered probit	VO	√	√	-	√	√	√	-	-
de Jong (1996)	Netherlands *Vehicle Panel Survey*	Single hazard duration model	VOD	√	√	-	-	-	-	√	√
Yamamoto and Kitamura (2000)	USA *Panel Survey*	Single hazard duration model	VOD	√	√	√	-	-	-	-	√
Gilbert (1992)	USA *Panel Survey*	Competing hazards duration model	VTR	√	√	√	√	-	-	-	-
Yamamoto *et el.* (1999)	USA *Panel Survey*	Competing hazards duration model	VTR	√	-	√	√	-	-	-	-
Mohammadian and Rashidi (2007)	Canada *Retrospective Survey*	Competing hazards duration model	VTR	√	√	√	√	√	-	√	√
Yamamoto (2008)	France *Panel Survey*	Competing hazards duration model	VTR	√	-	-	√	√	√	-	-
Kitamura and Bunch (1990)	Netherlands *Mobility Panel Survey*	Random effects ordered probit	VO	√	√	-	√	-	√	-	√
Nobile *et al.* (1997)	Netherlands *Mobility Panel Survey*	Random effects multinomial probit model	VO	√	-	-	√	√	-	-	√
Mohammadian and Miller (2003b)	Canada *Retrospective Survey*	Mixed parameter logit	VTR & VT	√	-	-	-	-	-	√	√
Bjorner and Leth-Petersen (2007)	Denmark *Panel Survey*	Random effects multinomial logit model	VO	√	√	√	-	√	-	√	√
Woldeamanuel *et al.* (2009)	Germanay *Panel Survey*	Random effects regression	VO	√	-	-	-	√	√	√	√
Nolan (2010)	Ireland *Panel Survey*	Random effects binary probit	VO	√	√	-	√	√	-	-	√

Endogenous Dynamic (6)											
Eluru *et al.* (2010a)	USA *Travel Survey*	Copula based joint GEV-based logit-regression model	VT & VU	√	-	-	√	√	√	-	√
Paleti *et al.* (2011)	USA *Vehicle Survey*	Copula based joint GEV-based logit-regression model	VT & VU	√	√	-	√	√	-	√	√
Paleti *et al.* (2013b)	USA *Travel Survey*	Multinomial probit model	VT	√	√	√	√	√	-	-	√
Golob (1990)	Netherlands *Mobility Panel Survey*	Longitudinal structural equation model	VO	√	-	-	√	√	-	-	√
Kitamura (2009)	Netherlands *Mobility Panel Survey*	Three equation simultaneous equation model	VO	√	√	-	√	√	√	-	√
Rashidi and Mohammadian (2011)	USA *Travel Panel Survey*	Hazard based Simultaneous equation model	VTR	√	√	-	-	√	√	-	√

Note: VO = vehicle ownership; VT = vehicle type; VU = vehicle use; VOD = vehicle ownership duration; VTR = vehicle transaction; OD = origin-destination

Table 2.2 Decision Matrix for Vehicle Ownership Model Selection

Vehicle Demand	Suggested Model			
	Exogenous Static	Endogenous Static	Exogenous Dynamic	Endogenous Dynamic
Vehicle count				
No heterogeneity	Generalized ordered logit	Multidimensional choice modeling	-	Simultaneous equation model
Heterogeneity	Latent class multinomial logit	Mixed multidimensional choice modeling	Mixed generalized ordered logit	-
Vehicle count and use	-	Multiple discrete continuous extreme value model	-	-
Vehicle type				
No heterogeneity	Multinomial logit	-	-	Multinomial probit
Heterogeneity	Mixed multinomial logit	-	-	-
Vehicle type and use	-	Copula based joint multinomial discrete continuous model	-	Copula based joint GEV-based logit-regression model
Vehicle ownership duration	-	-	Duration model (single hazard)	-
Vehicle transaction	-	-	Duration model (competing hazards)	Hazard based Simultaneous equation model
Vehicle transaction and type	-	-	Mixed parameter logit	-

CHAPTER 3 ANALYSIS OF THE EVOLUTION OF TRANSPORTATION EXPENDITURE: A LONG TERM PERSPECTIVE

3.1 Background

The interest among consumer researchers and marketers on the subject of household expenditure patterns and spending habits dates back to the middle of 19th century (Ferdous et al., 2010). However, researchers began to explore transportation related expenditure of households a century later – in the mid 1950's (see Houthakker, 1957). In recent years, the topic has seen a revival of interest among travel behavior researchers; the examination of household budgeting is particularly relevant at this time because of the increasing recognition of the challenges associated with global climate change and inequity in developed countries.

It is an established fact that transportation sector is one of the major contributors of greenhouse gas (GHG) emissions and hence, with the increasing recognition of global climate change issues, several countries are considering wide ranging policy measures to reduce the quantity of GHG emissions from this sector. Towards that end, a comprehensive understanding of household budgetary allocations through quantitative analysis will allow transportation professionals to simulate the positive and negative consequences of proposed policies targeting GHG reductions. For example, a framework to model households' response to policies such as gasoline tax or electric vehicle subsidy requires an understanding of how households adjust their monetary expenditures to maintain their mobility levels in response to these policies. Further, quantitative frameworks developed can also be employed to study the potential equity/distributional implications of private transportation usage penalty for vulnerable population segments. In fact, there is evidence indicating that a blanket increase in gas prices as a measure of reducing GHG emissions might adversely affect the lower income groups (Ferdous et al., 2010).

3.1.1 Modelling Budgetary Allocations

The factors affecting household budgetary allocations include household composition, employment status, household location, household evolution, and global socio-economic factors (such as economic, technological and cultural factors). Accommodating for the impact of current household characteristics in the budgetary allocation process is possible through cross-sectional databases. However, analysis using cross-sectional data can only provide a snapshot of the decision processes and not their evolution. To address this limitation, one could study household expenditure patterns over time employing longitudinal databases that track the expenditure patterns of the same households across multiple years. Unfortunately, longitudinal data suffers from respondent fatigue and retention problem (Hanly and Dargay, 2000; Yee and Niemeier, 1996). An alternative to the longitudinal data collection process is to pool multiple year cross-sectional databases of household expenditure – an approach gaining wide applicability in travel behavior literature recently (see Sanko, 2013; Dargay, 2002; Dargay and Vythoulkas, 1999). Though the multiple waves are not compiled based on the same set of households, they still provide us an opportunity to examine the impact of changing economic, social, and cultural trends on household expenses and thus provide additional policy-relevant information. Moreover, the pooled database enables us to examine if the impact of exogenous variables have evolved overtime. For example, the impact of an additional employed individual in the household on vehicle purchase expenditure might have changed from 1995 through 2005. Such analysis can only be conducted through pooled cross-sectional databases.

3.1.2 Contribution and Organization of the Chapter

The current research contributes to transportation literature by developing an econometric model of household budgetary allocations with a focus on transportation budget. Specifically, our research aims at investigating the factors affecting expenditure of households and its evolution in Canada using the public-use micro-data extracted from the Survey of Household Spending (SHS) for the years 1997 – 2009.

The proposed econometric modeling approach is built on the multiple discrete continuous extreme value model (MDCEV) framework. The MDCEV framework recognizes that households choose to allocate budgets to multiple alternatives simultaneously. Further, to incorporate the effect of observed and unobserved temporal effects, we specifically consider two versions of the MDCEV model – the scaled MDCEV (SMDCEV) and the mixed MDCEV (MMDCEV) model (see Sobhani et al., 2014). The two variants differ in the way they incorporate the influence of unobserved attributes within the decision process. We estimate both models and employ data fit comparison metrics to determine the appropriate model structure. Additionally, a policy analysis exercise is conducted to illustrate the applicability of the proposed modeling framework.

The rest of the chapter is organized as follows. Section 3.2 provides a detailed literature review on transportation expenditure. The formulations of the econometric model frameworks used in the analysis is presented in Section 3.3. In Section 3.4, the data source for the empirical analysis and sample formation procedures are described. This section also contains some descriptive statistics of the sample used for model development. The empirical analysis and the policy simulation results are presented in Section 3.5. Finally, Section 3.6 concludes the chapter.

3.2 Literature Review

In this section, we provide a summary of the literature that examined, directly or indirectly, transportation expenditure patterns of households. The term transport expenditure usually refers to outlays of time and money (Kuhnimhof and Gringmuth, 2009). In fact, time and money costs are considered as the two fundamental constraints of travel. Earlier literature on transportation money budgets can be broadly classified into two categories: (1) studies that focus on transportation expenditure in conjunction with household expenditures for other commodities and services, and (2) studies that examine transportation expenditure exclusively[4]. Transportation expenditure typically examined in these literature includes the following dimensions: vehicle acquisition costs, gasoline costs, vehicle insurance costs, vehicle operation and maintenance costs, public transportation costs, non-motorized transportation costs, intercity travel costs, and recreational vehicle related costs.

Among the studies that have examined transport expenditure in the context of various other household budgetary decisions, Choo et al. (2007) investigated whether the relationship between transportation and telecommunication is substitutive, complementary, or neither. They found that most transportation expenses are highly income elastic, whereas communication expenditures are generally less elastic than the transportation ones. The authors argue that vehicle ownership imposes substantial costs on economically disadvantaged groups, thereby limiting other consumption/expenditure opportunities. In another study, using cluster analysis techniques, Sanchez et al. (2006) analyzed the combined transportation and housing expenditure trade-offs that low-income working households make and reported that these expenditures cannot be considered in isolation. Very recently, Ferdous et al. (2010) reported that overall transportation

[4] There has been recent research exploring transport and time budget allocation in a unified framework (see Konduri et al., 2011 and Anas, 2007 for such literature).

expenditure allocation of households in USA is primarily affected by household socio-demographics. The authors found that households residing in urban areas allocate higher proportion of their income to housing as well as public transportation. They also conducted a sensitivity analysis to explore how households adjust their consumption patterns with rising fuel price. Their policy analysis results indicated that in the short run, adjustments are made in savings rates, food consumption, and vehicle purchase expenses while in the long term major shifts occur in housing and utility expenditures.

The second category of studies concentrate solely on examining the transportation related spending of households. Petrol or gasoline outlays constitute the biggest portion of the overall transportation expenses of households and thus, it can be modeled as a proxy to vehicle use. The expenditure functions developed can be used to analyze both the effectiveness and redistributive effects of price-based energy consumption reduction policies, such as petrol taxation (Asensio et al., 2003a; Oladosu, 2003). The authors concluded that household income, socio-demographics, residential location, and vehicle fleet attributes are the most significant factors in the petrol expenditure allocation process of households. In another study, Asensio et al. (2003b) analyzed the redistributive effect of urban public transport subsidies in Spain considering that the subsidies provided to the transit sector is directly related to the fare expenses incurred by households.

Rather than focusing on only one transport outlay category, Thakuriah and Liao (2005) explored the variation of a range of household vehicle ownership expenditures while controlling for socio-economic variables, demographics, lifecycle, and geographic region of residence in the country. According to their findings, vehicle owning households would spend, on average, 18 cents on vehicles for every additional dollar in monetary expenditure. Similar research was conducted by Nolan (2003) using Irish Household Budget Survey micro-data considering three transport expenditure categories: gasoline cost and conveyance cost of bus and taxi. In a later study, Thakuriah and Liao (2006) investigated the relationship between transportation expenditures (termed as mobility investments) and ability to pay (measured by income). They found that increased income leads to increased overall transportation expenditures and vice versa – presumably because mobility investments fetch accessibility benefits which in turn contribute to higher income. In the most recent work, Thakuriah and Mallon-Keita (2014) analyzed how transportation expenses changed for US households from pre-recession (2005-2006) to recession periods (2007-2009) and noted a decline in the auto-related spending of households during the economic downturn.

3.3 Econometric Framework

3.3.1 Basic Structure and Traditional MDCEV Model

It is reasonable to expect that the money allocated to different expenditure categories depends on the marginal utility that households derive from spending in those categories. Let us consider that there are *K* different expenditure categories that a household can potentially allocate its money to. If $\boldsymbol{x_k}$ represents the allotted non-negative amount of the total budget to each expenditure category *k* (including savings), the total utility derived from such allocation can be expressed in the following additive non-linear functional form (Bhat, 2008):

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\}; \ \psi_k > 0, \alpha_k \leq 1, \gamma_k > 0 \qquad (3.1)$$

where $U(\boldsymbol{x})$ is a quasi-concave, increasing, and continuously differentiable function with respect to the expenditure quantity (Kx1) - vector *x* ($x_k \geq 0$ for all k alternatives), γ_k and α_k are parameters associated with alternative *k*. ψ_k represents the baseline marginal utility. Through this term, the effect of observed and unobserved alternative attributes, decision-maker attributes, and the choice environment attributes may be introduced as $\psi_k = exp(\beta' z_k + \varepsilon_k)$, where z_k represents the vector of exogenous variables and ε_k captures the idiosyncratic characteristics that affect the baseline utility. γ_k enables corner solutions while simultaneously influencing satiation and α_k influences satiation only. Note that the above utility function is formulated considering absence of outside goods (goods that is always consumed).

If, however, an outside goods is present, the utility function can be modified as follows using the same notational preliminaries:

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1} \exp(\varepsilon_1) \{(x_1 + \gamma_1)^{\alpha_1}\} + \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k} exp(\beta' z_k + \varepsilon_k) \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \qquad (3.2)$$

In the above formula, we need $\gamma_1 \leq 0$, while $\gamma_k > 0$ for k > 1. Also, we need $(x_1 + \gamma_1) > 0$. The magnitude of γ_1 may be interpreted as the required lower bound (or a "subsistence value") for consumption of the outside goods. In the above baseline parameter expression, the term ε_1 is an idiosyncratic term assumed to be identically and independently standard type I extreme-value distributed across households, as well as independent of the terms in the baseline parameter expression for other alternatives (inside goods).

It is very challenging to identify γ_k and α_k simultaneously in empirical applications for the outside and inside goods (see Bhat, 2008 and Bhat and Eluru, 2010 for an elaborate discussion on the issue). Usually, the analyst can choose to estimate satiation using either γ_k or α_k, since these two parameters have similar role in terms of allowing for satiation. Depending on the chosen parameter structure for estimation, different utility structures can be estimated and the selection of the most appropriate form is based on statistical considerations.

If only γ_k parameters are estimated the utility simplifies to γ-profile

$$U(\boldsymbol{x}) = \exp(\varepsilon_1)\, ln\{(x_1 + \gamma_1)\} + \sum_{k=2}^{K} \gamma_k exp(\beta' z_k + \varepsilon_k) \left\{ ln\left(\frac{x_k}{\gamma_k} + 1\right)\right\} \tag{3.3}$$

Similarly, if only α_k parameter are estimated, the corresponding utility expression collapses to α-profile.

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1} \exp(\varepsilon_1)\, \{x_1{}^{\alpha_1}\} + \sum_{k=2}^{K} \frac{1}{\alpha_k} exp(\beta' z_k + \varepsilon_k)\{(x_k + 1)^{\alpha_k} - 1\} \tag{3.4}$$

Let V_k be the alternative utility. The expressions for V_k for γ-profile and α-profile utility forms are as below:

$$V_k = \beta' z_k - ln\left(\frac{x_k^*}{\gamma_k} + 1\right) - ln\, p_k (k \geq 2);\; V_1 = -ln(x_1 + \gamma_1) \tag{3.5}$$

$$V_k = \beta' z_k + (\alpha_k - 1)\, ln(x_k^* + 1) - ln\, p_k (k \geq 2);\; V_1 = (\alpha_1 - 1)\, ln(x_1^*) \tag{3.6}$$

We would assume (following Bhat, 2005 and Bhat, 2008) that ε_k's are independently and identically distributed across alternatives with a scale parameter of σ. Given the values of the alternative utilities for the two profiles, the probability expression for the expenditure allocation to the first M of the K goods (M $\geq$1) is:

$$P(e_1^*, e_2^*, e_3^*, \dots, e_M^*, 0,0,0, \dots, 0) = \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} C_i\right] \left[\sum_{i=1}^{M} \frac{1}{C_i}\right] \left[\frac{\prod_{i=1}^{M} exp^{\frac{V_i}{\sigma}}}{\left(\sum_{k=1}^{K} exp^{\frac{V_k}{\sigma}}\right)^M}\right] (M-1)! \quad (3.7)$$

where, $C_i = \frac{1-\alpha_i}{e_i^* + \gamma_i p_i}$.

In the traditional MDCEV model, the scale parameter σ is set to 1 for normalization.

3.3.2 Scaled MDCEV Model

In our context, due to the inherent differences across the expenditure databases across years and different economic conditions, we can estimate the scale parameter provided we normalize σ for one year. The σ is parameterized as $\exp(\delta y)$ where y is the vector of year indicator variables as well as the annual economic indicators and δ is the corresponding coefficient vector to be estimated. The δ parameters are significant when they are different from 0 as that would imply that the scale parameter will be different from 1. The same expression in Equation 3.7 is adopted with the appropriate σ for probability and likelihood computations.

3.3.3 Mixed MDCEV Model

The mixed MDCEV model accommodates unobserved heterogeneity in the effect of exogenous variables (random coefficients structure) and correlations across alternatives (error correlations structure). The baseline parameter expression for the inside alternatives in Equation (3.2) can be expressed as follows:

$$\psi_k = exp\{(\beta_k{}' + \alpha_k{}')\, z_k + \eta' w_k + \xi_k\} \quad (3.8)$$

In the above equation, β' and α' are column vector of parameters, where β' represents the mean effect and α' represents household level disturbance of the coefficient. The term $\eta' w_k$ constitutes the mechanism to generate household level correlation across unobserved utility components of the alternatives. In this component, w_k is specified to be a column vector of dimension *H* with each row representing a group *h* (*h=1, 2, ..., H*) of alternatives sharing common household-specific unobserved components and the vector η' may be specified as a *H*-dimensional realization from a multivariate normally distributed random vector η, $\eta \sim N(0, \Omega)$. As before, the component, ξ_k is assumed to be independently and identically Gumbel distributed across households. For complete formulation of likelihood for the MMDCEV model see, Bhat and Eluru (2010).

The parameters of MMDCEV model are estimated using maximum simulated likelihood (MSL) procedure. Specifically, scrambled Halton sequence is used to draw realizations from the population normal distribution. In this research, the stability of the parameter estimates were tested using varying number of Halton draws per observation for the specifications considered, and the results were found to be stable with 100 draws.

3.4 Empirical Data

3.4.1 Data Source

The primary data source used in this analysis is the Survey of Household Spending (SHS). This is an annual cross-sectional survey conducted by Statistics Canada since 1997 (Milligan, 2008). The survey primarily collects detailed information on household and family expenditures and spending habits in Canada (every year in the 10 provinces and usually every alternate year in the territories) on a wide variety of goods and services (see, Statistics Canada website http://www23.statcan.gc.ca/ for details on survey, sampling, and administration procedure). The SHS also collects information on individual and household socio-economic and demographic attributes, dwelling characteristics (such as type, age and tenure) and information on household equipment (such as appliances, electronics and communications equipment, and vehicles). For our study, the SHS data is augmented with several annual economic indicators such as inflation rate, unemployment rate, gross domestic product (GDP), and wage rate to accommodate for the impact of these indicators on the budgetary allocation process. For example, during recession, households are likely to reduce spending on discretionary commodities (such as new vehicle purchase). In this research, we employed public-use micro-data extracted from the SHS for the years 1997-2009.

3.4.2 Dependent Variable Compilation

The reported expenditure categories were reorganized to create the following twenty alternatives:

1. *Food:* costs incurred from purchase of food and non-alcoholic beverages from grocery stores as well as from restaurants

2. *Shelter:* rent, regular mortgage payments, condominium charges, property taxes, and home-owners' insurance premiums
3. *Secondary accommodation:* expenditure for owned vacation home and lodging while away from home (overnight or longer)
4. *Utilities:* water and sewage charges, electricity, natural gas and other fuel (*such as* propane and wood for barbeques), telephone, cellular, internet, and postal service costs
5. *Alcohol and tobacco products:* total expenditure for all tobacco products and smokers' supplies, alcoholic beverages prepared at home as well as purchased and consumed in restaurants and bars
6. *Clothing:* expenditure on purchasing clothes and clothing services (laundry and dry cleaning)
7. *Personal care:* personal care supplies, equipment, and services
8. *Household maintenance and operation:* expenses for household furniture and décor wares (such as rugs, curtains), supplies, services, accessories, and household maintenance and operation equipments
9. *Entertainment and recreation:* cost of home entertainment/sports/hobby equipment and associated services, admission fares to movies and live events, club membership fees, and recreational trip expenses
10. *Education:* costs of books, education supplies, and tuition fees
11. *Health care:* hospital expenses, cost of health care supplies and goods, prescription medicines and pharmaceutical products, eye and dental-care goods and services, health insurance premiums, and other medical services
12. *Business services and welfare activities:* total expenditure on financial services, union and professional dues, and charitable contributions
13. *Automobile acquisition:* net purchase price paid for automobiles after deducting any trade-in allowance or separate sales and costs for renting and leasing vehicles
14. *Recreational vehicle:* purchase/rent and operation of recreational vehicles
15. *Gasoline costs:* gasoline and other fuel expenses for owned and leased vehicles
16. *Vehicle insurance costs:* total public and private vehicle insurance premiums paid for owned and leased automobiles

17. *Vehicle operation and maintenance:* expenses accrued from maintenance and repair operations, garage rent and parking fee, and purchase of accessories
18. *Public transportation:* local transit and commuter transportation costs
19. *Non-motorized transport:* purchase cost of bikes, parts and accessories as well as maintenance and repair costs
20. *Intercity travel:* fare of airplane, train and highway bus travel

We retained the transportation related expenditure categories as disaggregate as possible. In addition to these alternatives, a savings alternative was created by subtracting the total annual expenditure from the total gross income. If savings were negative, then the savings variable was coded as zero. Finally, from the survey database, for each survey year, 1,000 data records were randomly sampled providing us with 13,000 observations in our pooled dataset.

3.4.3 Descriptive Analysis

A general overview of the socio-economic characteristics of the sample used in this study is presented in Table 3.1. It provides a perspective for the magnitudes and distribution of the key variables utilized in the empirical model. Some salient features of the data are as follows: one-quarter of the households are comprised of a single person, while another quarter consisted of childless couples. More than 55 percent of the households have at least one full time earner and more than 40 percent have at least one person employed part-time. Our sample has roughly equal share of households belonging to each income category and more than 60 percent listed paid employment or self-employment as their major source of income. The distribution of the automobile ownership levels in the sample indicates that more than three-quarters of households owned at least one private automobile. Majority (66.7%) of the households owned their dwelling and lived in a dwelling that is single-detached. Most of the households are located in urban areas with high population density.

Table 3.2 provides a descriptive snapshot of the 21 expenditure categories modeled in the study. The first column represents the average spending of households across the entire sample and the figures within the parenthesis values represent the percentages of household income allocated to these expenditure categories. It can be observed that all of the households in the sample spend some non-zero amount of money in the food category which accounts for approximately 11 percent of their income. As expected, housing is the highest expenditure incurred (accounting for

11.9 percent of household income) while education and secondary residence being the smallest expenditures. Moreover, household maintenance and operation, utilities, entertainment and recreation also form a substantive portion of household expenses. We also observe that 90 percent or more households incur expenditures in each of the clothing, personal care, health care, and business and welfare activities categories.

In our study, we considered eight different transportation related expenditure categories including vehicle purchase/rent/lease, recreational vehicle purchase/rent and operation, gasoline and motor fuels, vehicle insurance, vehicle operation and maintenance, public transportation, non-motorized transport, and intercity travel. These categories combined account for 13.9 percent of household income. More than one-third of the sample households allocate their resources to acquiring (purchase/rent/lease) personal automobiles while about one-quarter of the households spend money on recreational vehicle acquisition and maintenance. Of the households reporting non-zero monetary expenditures: about 85 percent spend money on fuel and motor oils while more than 70 percent incur insurance related costs. Interestingly, a sizeable number of households spend money on public transportation (more than 50 percent). On the other hand, a very small proportion of the households allocate resources on purchasing and maintaining non-motorized transports (approximately 17 percent). Expenses on intercity travel are incurred by about one-third of the households.

3.5 Empirical Analysis

3.5.1 Variables Specification

Several types of variables were considered in the model that we developed for examining expenditure allocation in each of the twenty one outlay category as well as the household savings category. The choice of these independent variables was guided by prior research on expenditure patterns. The independent variables can be broadly classified into three categories: (1) household socio-demographics, (2) residential location characteristics, and (3) temporal variables. The socio-demographic variables that were employed in our analysis included presence of children of different age groups, presence of young members (18-24 years of age), number of full- and part-time working adults, household income (gross) and its type (paid income from employment), vehicle fleet size, vehicle availability per adult, tenure type, dwelling type and family type. The residential location variables considered are: urban/rural location, population center density and region specific dummies to examine the degree of influence exerted by the region of residence on household expenditures. The regional dummies used are: Alberta (AB), British Columbia (BC), Ontario (ON), and Quebec (QC). In terms of temporal variables, we included year specific indicator variables to study time based trends in household expenditure allocation patterns. The year effects were considered in the systematic utility as well as the unobserved component of the utility. To further explain the differences in unobserved component due to temporal changes we compiled data on annual economic indicators such as inflation rate, unemployment rate, gross domestic product, and wage rate for Canada from 1997-2009 (see, http://www.statcan.gc.ca/ and http://databank.worldbank.org/data/ for details). Finally, interactions between exogenous variables from the above three categories were also considered (such as urban*regional dummies). The final specification was based on a systematic process of removing statistically insignificant variables and combining variables when their effects were not significantly different.

3.5.2 Model Specification

The model estimation results are presented in Table 3.3 (the *t*-stats are presented in parentheses). A positive (negative) coefficient for a certain variable-category combination means that an increase in the explanatory variable increases (decreases) the likelihood of budget being allocated to that expenditure category relative to the base expenditure categories. A blank entry corresponding to the effect of variable indicates no statistically significant effect at the 95 percent significance level for the variable on the choice process.

3.5.3 Model Fit Measures

The model estimation process began with the estimation of the traditional MDCEV model. Next, the scaled and mixed MDCEV models were estimated. Note that both of these two models are generalized versions of the standard MDCEV model. After extensive specification testing, the final log-likelihood values at convergence of the MDCEV, SMDCEV, and MMDCEV models were found as: -1624779, -1599533 and -1624493, respectively while the log-likelihood value for the MDCEV model with only the constants and satiation parameters is -1643408. The improvement in the data fit clearly demonstrates the superiority of the SMDCEV model over its other counterparts. The Log-likelihood Ratio (LR) test comparison and Bayesian Information Criterion

(BIC) between the SMDCEV model and the other models yields test statistics that rejects that hypothesis that all the models are similar at any reasonable level of significance. Hence, in the subsequent sections, we discuss the results of the SMDCEV model only. The exogenous variables effects are discussed by variable group, followed by constant terms and scale component results.

3.5.4 Empirical Results

3.5.4.1 Household Socio-Demographics

It has long been recognized that the presence of children affects the household budgetary allocation (Browning, 1992). Children from different age groups have different needs and requirements and thus households must allocate budgets accordingly. In our study, we found that the presence of toddlers (less than or equal to 4 years of age) contributed to higher apportioning of income to housing, clothing, personal care, household operations and maintenance, automobile acquisitions, and non-motorized transport costs. On the other hand, lower proportions of income are allocated to education, health care, and public transportation categories. Increase in expenses in the automobile acquisition category might be explained by increased travel needs with the presence of toddlers. Similar resource allocation patterns were observed for the presence of young children (5-17 years of age) variable. However, households with young children spent more on education as opposed to households with toddlers while spending less on alcohol and tobacco purchase, vehicle insurance, intercity travel, and savings. This result is not surprising, since households with more young children may sacrifice their liquor expenditures for other expenditures, in particular child care and education (Soberon-Ferrer and Dardis, 1991). When young members (18-24) are present, households spend more on both automobile acquisition and public transportation. They also apportion a higher income share to intercity travel. At 18, teenagers are allowed to drive alone and thus, households might allot more resources to acquire extra cars to allow them to drive independently (Prillwitz et al., 2006).

We considered four variables representing different life-cycle stages of households. These are: single person, couples only and, couples with a child, and other households (comprised of couple household with relatives or unrelated persons and lone parents). Compared to other households, single person households allocate higher proportion of their income to housing, secondary accommodations, alcohol and tobacco products, health care, business service and welfare activities, vehicle operation, intercity travels, and savings. The biggest proportional difference between couples only and other life cycle groups lay in their spending on transportation categories. For instance, couples only households spend more in acquiring private automobiles as well as recreational vehicles, and also on gasoline, but less in public transportation and non-motorized transport options. While high spending is enabled by high income (whether it is a "one-earner" or "dual-earner" impact that needs to be investigated), other characteristics could also provide a partial explanation for relatively high spending on transport and recreation by couple only households. For example, being childless might allow these households to participate more in activities such as recreation, and greater use of private means of transport to travel to these activities. Contrastingly, a household comprising of couples with a child tend to allocate lesser amount of resources to virtually all the expenditure categories considered.

Income share allocated to alcohol and tobacco purchases, clothing, personal care, and vehicle purchases tend to increase with increasing number of both full and part-time working adults. Interestingly, households with higher number of part-time workers tend to incur more transportation expenses including public transportation, intercity travel, and recreational vehicle purchases and maintenance. A similar result in the context of vehicle purchases and transportation expenses were reported by Thakuriah and Liao (2005). However, households with multiple full-time workers allocate more resources to housing and savings while households with higher number of part-time workers spend more on education and less on savings. Overall, the results are indicative of the variation in household expenditure patterns that occurs due to varying employment type of its members.

As expected, household income was found to be one of the influential factors affecting household's decision regarding budget allocation. Compared to low income (income < 30K) households, households with medium (income 30-70K) and high income (income >70K) spend higher proportion of their income on clothing, personal care, entertainment and recreation, health care, business service and welfare activities, and secondary accommodations. In the transportation expenditure categories, these households spend more on acquiring automobiles, gasoline, purchase and maintenance of recreational vehicles, as well as intercity travel. Reduction in public transit budget of medium income groups might be due to their being more inclined to procure personal automobiles rather than using public transit for their travels. These findings are consistent with the results reported in the existing travel behaviour literature. For example, high household income has often been reported to increase the probability of owning multiple cars (Bhat and Pulugurta, 1998; Karlaftis and Golias, 2002; Matas and Raymond, 2008). In addition to the actual amount, income type was also found to significantly affect household budgets. Compared to households living on investment income or government transfer payments, households with paid employment (wages and salaries) as their major source of income allotted more to basic necessities, gasoline expenses as well as entertainment and recreation. The results are expected because increased paid income means higher spending capability.

To capture the budget allocation patterns of multicar households, we created three household types based on their vehicle ownership levels. These are: single vehicle households, households with two cars, and households with three or more cars. All of the household types exhibited interestingly similar spending behaviour. More specifically, all of them tend to spend lesser proportion of their resources to vehicle acquisitions; the finding is in contrast to the results reported by Ferdous et al. (2010). Presumably, these households are somewhat disinclined to increase their existing vehicle fleet size and hence, they are allocating less resources to vehicle purchases, which is a promising finding. As expected, these households allocate higher proportions of income to pay for gasoline, vehicle insurance and vehicle maintenance costs. We also observed that two-vehicle owning households spend less on public transportation and more on non-motorized transportation options.

Homeowners tended to shift their spending habit from housing and alcohol and tobacco products, and direct relatively more of their monetary income towards utilities, household operation and maintenance, health care, business and welfare activities, and secondary accommodations. All of the findings are consistent with expectations and corroborate the outcomes of previous research (for example, Hong et al., 2005; Paulin, 1995). These households also spend less proportion of their earnings on vehicle purchases, public transportation, and non-motorized vehicles. On the other hand, they accrue more expenses on vehicle insurance and recreational vehicles. Similar to homeowners, single-detached households have reduced share of income apportioned towards housing and public transportation while spending more on utilities, household operation, gasoline, and recreational vehicles. Markedly different yet expected expenditure patterns were captured for households residing in apartments. Apartment residents are mostly renters (82 percent), therefore, in contrast to homeowners and single-detached housing dwellers, they do not have to pay for utilities and maintenance costs for the entire establishment by themselves, thus perhaps they allocate reduced resources in these categories. Apartment dwellers used a larger proportion of their income on public transport and intercity travel.

3.5.4.2 Residential Location Characteristics

It is evidenced in consumer literature that households living in urban areas have different needs compared to the needs of those in rural areas and, therefore, exhibit different spending patterns. In the current research context too, we captured the differences in the way that urban and rural consumers allocate their expenditures budgets. For instance, compared to their rural counterparts, households located in urban areas allocated larger share of their income on housing, clothing, personal care, entertainment and recreation, education, vehicle operation, public transportation, and intercity travels. These findings might be attributed to higher availability of consumer goods (such as education, personal care services, internet, entertainment) and to higher rents and mortgage payments in urban areas (Fousekis and Lazaridis, 2001). Reduced gasoline expenses, recreational vehicle costs and savings were also observed for these households. Reduced gasoline expenditure by urban households may be reflecting the lower distances travelled to shop and to work. According to Ferdous et al. (2010), these results are reflective of the typical "urban effect". In addition to location, population density of the area where the household is located also affected households' resource allocation decisions.

As mentioned earlier, four regional dummies were used in our model estimation and several regional differences are noted from the analysis results. Intuitively, the differences are attributable to the regional differences in housing prices, income levels, and overall prices of goods and services across the provinces analysed (see, Ferdous et al., 2010 for similar interpretations). Several interaction terms between urban and regional dummies were attempted in our model. Not accounting for these regional differences might create bias in other exogenous variable impacts.

3.5.4.3 Constant Terms

The constant variables do not have any substantive interpretations but simply capture the generic tendencies to spend in each category. Note that the baseline preference constants are introduced with the *food* category as the base category. As can be observed from the results table, all baseline preference constants without exception are negative, indicating overall reduced spending of household budget on all other expenditure categories relative to *food*. This result is consistent with expectations because all of the households in the sample spend non-zero amount in the *food* category.

As discussed earlier in the methodology section, the translation parameters (γ_k) capture the extent of decrease in marginal utility across different expenditure goods. That is, for all expenditure categories except *food*, a higher value of γ implies higher spending and less satiation. There is no γ term for *food* category because it is always consumed by all households. All of the translation parameters are statistically significant at any reasonable level of significance, thereby implying that there are clear satiation effects in household resource allocation. It is found that the γ value is the highest for *savings* and *new/used automobile purchase* alternatives, indicating that households are likely to allocate a large proportion of their budget to savings and to acquiring a vehicle, if they spend any money in these categories. On the other hand, the lowest values are observed for *personal care, public transportation,* and *non-motorized transport* categories, suggesting that the lowest proportion of money is allocated to these categories and satiation is reached very quickly for most households in these categories.

Within the set of constant parameters, the impact of year indicator was examined. The year indicator variables account for overall difference to baseline utility across the years as well as the intrinsic preferences in the dataset. If unaccounted, these effects might manifest as differences variable impacts erroneously. We found that during recession years, there was a decline in expenditure related to education as well as alcohol and tobacco products. On the other hand, utilities, personal care and health care expenditure increased. Interestingly, expenditures on gasoline also increased. Plausible reasons might be non-cessation of driving compounded by the high fuel prices during this period (see, Thakuriah and Mallon-Keita, 2014 for similar results in US).

In the current empirical context, 1997 was considered as the base year and scale coefficients were estimated for the rest twelve years (1998-2009) as well as the inflation rate economic indicator. All the parameters were highly significant indicating that there is indeed variation in the unobserved factors across the years compared to year 1997. The variation was the highest for 2003, 2000 and 2008, respectively. Fiebig et al. (2010) reported that scale heterogeneity is more important in more complex choice context. In our case, the results might be manifesting the effect of the recession periods – allocation of resources during economic downfall becomes a more complex task, thus giving rise to such variation.

3.5.5 <u>Policy Simulation Results</u>

In this section, we present the results of several policy simulations. Specifically, we assessed the impact of four different scenarios on household expenditure patterns. The scenarios considered are: (1) Zero gasoline expenditure, (2) A 15 percent increase in gasoline expenditure, (3) A 15 percent increase in health care expenses, and (4) A 10 percent reduction in savings. The policy simulations consider two possible time frames – long and short term. In the short term, households are unlikely to alter housing, utilities, education, health care, vehicle purchase, vehicle insurance, and recreational vehicle purchase/maintenance. Hence, these alternatives were assumed to be unaffected by the changes in expenditure. In the long term all alternatives are assumed to be affected by the changes in expenditure. The predicted changes in household expenditure occurring for different scenario compared to the base case are presented in Table 3.4.

Some very interesting patterns of expenditure adjustment could be discerned from these results. When households incur zero expenditure on gasoline, in the short term, they tend to allocate the extra resources towards non-motorized transportation options, intercity travel, and public transportation. Zero gasoline expenditure might be considered as a proxy for capturing the effect of all households turning their vehicle fleet into electric. These findings are thus perhaps the reflection of households becoming more environmentally conscious. In the long term, with all vehicles turning electric, households tend to spend more on purchasing recreational vehicles followed by non-motorized vehicles. For 15 percent increase in gasoline expenses, in the short term, the savings category takes the largest hit, as is expected. Due to increased gasoline expenditure, people might be relying more on public transit and non-motorized transportation for their travel and hence, we see higher spending in these categories. In the long term, as expected, there is reduction in budget allotments for most of the expenditure categories, the largest being for the automobile purchase category. As explained by Ferdous et al. (2010), households might either be postponing the purchase of more vehicles or buying cheaper automobiles in the wake of rising gasoline expenses. As opposed to short term effects, public transport expenditure is curtailed.

Interestingly, we also observe that an increase in health expenditure, in the short term, leads to a decrease in gasoline expenses, presumably increased health expenditure might lead to a cessation in driving and increased use of public transit as well as non-motorized transport. In the long term, however, adjustment is made in all the other expenditure categories. The largest reductions are observed for secondary accommodations, education, recreational vehicle purchase, and business and welfare service categories. Finally, we also see that due to a 10 percent decrease in savings, in the short term, all alternatives have increased spending outlays – particularly the discretionary alternatives. In the long term, the expenditures in the recreational home purchases and alcohol and tobacco products increase. Moreover, a sizeable decrease in non-motorized transport alternative is also observed with decrease in savings. Overall, the policy simulation exercise illustrates the applicability of the proposed model.

3.6 Summary and Conclusions

Travel behaviour literature is replete with studies focused on household travel time budget because of its possible use in forecasts of travel levels as well as modal shares. On the other hand, despite transport budgetary allocation being inextricably linked with a whole gamut of travel behavior choice processes both long and short term, it has received much lesser attention from the research community. In fact, given the strong influence on travel patterns, it would be useful to consider monetary allocation decisions as a precursor to modeling travel demand processes. The research presented in the chapter endeavours to bridge the gap in the literature by developing an econometric model of household budgetary allocations with a particular focus on transportation budget. Specifically, we aim to investigate the factors affecting expenditure of households and its evolution in Canada using the public-use micro-data extracted from the Survey of Household Spending (SHS) for the years 1997 - 2009. In terms of methodology, we adopt multiple discrete continuous extreme value model (MDCEV) framework and utilize its two variants − scaled MDCEV (SMDCEV) and mixed MDCEV (MMDCEV) models − these models simultaneously accommodate for the influence of observed and unobserved attributes on the budget allocation decisions of households across multiple time points. However, results of the data fit comparison metrics proved the superiority of the scaled model over its other counterparts.

Broadly, the model results indicated that a host of household socio-economic and demographic attributes along with the residential location characteristics affect the apportioning of income to various expenditure categories and savings. We also observed a relatively stable transportation spending behaviour over time; only the gasoline expenses displayed an increase during the recession times. Additionally, a policy analysis exercise is conducted to augment the model findings and illustrate the applicability of the proposed modeling framework. The exercise allows us to better understand how households adjust their consumption patterns with changes in scenarios. It was interesting to see that with increase in gasoline prices, people tend to allocate more resources towards public transit and non-motorized transports. The proposed framework can be employed to examine the impact of policy measures, environmental taxes, and electric vehicle subsidies on household expenditures in general and transport expenditures in particular.

3.7 Link between Chapter 3 and Chapter 4

In this chapter, household budgetary allocation decisions are investigated with particular focus on transportation budget. Recognizing the fact that households choose to allocate their limited budget to multiple expenditure categories simultaneously, a multiple discrete framework is formulated. The analysis enabled us to determine which factors significantly impacted (increased/decreased) budget allocation towards different transportation categories including vehicle purchase in comparison with other essential and discretionary household commodities.

The analysis in this chapter sheds light on the money allocation process of households for vehicle acquisition. In the vehicle decision process hierarchy, after allocating budget for vehicle acquisition, households decide on their vehicle fleet size, i.e. how many vehicles to purchase. Therefore, in the next chapter, we move towards enhancing our understanding of the factors that influence household vehicle ownership decisions. To accommodate for the potential variation in the impacts of the exogenous variables on the fleet size decisions, we use latent segmentation based frameworks to analyse the decision process.

Table 3.1 Sample Characteristics

Variables	No of Households (%)
Household Demographics	
Family Life-cycle Stage	
Single person	3276 (25.2)
Couples only	3494 (26.9)
Couples with a child	3889 (29.9)
Others	2341 (18.0)
Number of Full-time Earners	
0	5825 (44.8)
1	4788 (36.8)
≥ 2	2387 (18.4)
Number of Part-time Earners	
0	7280 (56.0)
1	4015 (30.9)
≥ 2	1705 (13.1)
Household Income	
Low income (≤ 30K)	4084 (31.4)
Medium income (30K-60K)	4168 (32.1)
High income (> 60K)	4748 (36.5)
Income Type	
Paid income	7976 (61.4)
Others	5024 (38.6)
Number of Cars	
0	2748 (21.1)
1	5657 (43.5)
2	3575 (27.5)
≥3	1020 (7.9)
Tenure Type	
Owned	8662 (66.7)
Rented	3929 (30.2)
Others	409 (3.1)
Dwelling Type	
Single detached	8233 (63.3)
Apartment	2677 (20.6)
Others	2090 (16.1)
Residential Location Attributes	
Population Centre	
Urban	10189 (78.4)
Rural	2811 (21.6)
Population Centre Size	
Low density	2811 (21.6)
Medium density (<100K)	2895 (22.3)
High density (≥100K)	7294 (56.1)
Province	
Alberta	1389 (10.7)
British Columbia	1578 (12.1)
Ontario	1832 (14.1)
Quebec	1905 (14.7)
Others	6296 (48.4)
Sample size	13000

Table 3.2 Summary Profile of Expenditure Categories and Savings

Expenditure Category	Average Spending across Entire Sample (CAD $/yr) (%)	Across Non-zero Observations	
		Average Spending (CAD $/yr)	No. of Households (%)
Food	6347.42 (10.9)	6347.42	13000 (100.0)
Shelter	6937.14 (11.9)	7050.49	12791 (98.4)
Secondary Accommodation	680.73 (1.2)	1420.46	6230 (47.9)
Utilities	3253.97 (5.6)	3265.52	12954 (99.6)
Alcohol and Tobacco Products	1375.60 (2.4)	1646.67	10860 (83.5)
Clothing	2317.64 (4.0)	2340.87	12871 (99.0)
Personal Care	866.73 (1.5)	870.08	12950 (99.6)
Household Maintenance and Operation	3581.60 (6.2)	3586.56	12982 (99.9)
Entertainment and Recreation	3039.77 (5.2)	3081.97	12822 (98.6)
Education	765.67 (1.3)	1914.91	5198 (40.0)
Health care	1568.65 (2.7)	1607.47	12686 (97.6)
Business Services and Welfare Activities	1859.63 (3.2)	1964.19	12308 (94.7)
Automobile Purchase/Rent/Lease	3313.32 (5.7)	9620.98	4477 (34.4)
Recreational Vehicle	642.15 (1.1)	2785.44	2997 (23.1)
Gasoline and Motor Fuels	1740.98 (3.0)	2077.73	10893 (83.8)
Vehicle Insurance	866.96 (1.5)	1210.98	9307 (71.6)
Vehicle Operation and Maintenance	570.24 (0.9)	758.38	9775 (75.2)
Public Transportation	207.26 (0.4)	385.97	6981 (53.7)
Non-motorized Transport	46.50 (0.1)	286.87	2107 (16.2)
Intercity travel	428.68 (0.7)	1381.11	4035 (31.0)
Savings	6430.45 (11.1)	11677.04	7159 (55.1)

Table 3.3 Estimation Results (N=13000)

Variables	HOU	SECH	UTL	ATP	CL	PC	HHMO	ENT	ED	HL	BSWA	AUTO	RECV	GAS	VEHI	VOP	PT	NMT	INTT	SAV
Household Socio-demographics																				
Children present (≤ 4yr)	0.248	---	---	---	0.064	0.300	0.301	---	-0.105	-0.097	---	0.078	---	---	---	---	-0.106	0.080	---	---
	(16.01)	---	---	---	(4.19)	(19.51)	(19.67)	---	(-5.41)	(-6.14)	---	(3.62)	---	---	---	---	(-5.36)	(2.94)	---	---
Children present (5-17 yr)	0.076	---	---	-0.086	0.175	---	0.051	0.103	0.497	-0.074	---	---	---	---	-0.181	---	-0.123	0.378	-0.090	-0.220
	(5.99)	---	---	(-6.49)	(14.00)	---	(4.18)	(9.46)	(26.81)	(-5.23)	---	---	---	---	(-15.23)	---	(-7.76)	(15.69)	(-4.40)	(-15.10)
Youth present (18-24 yr)	---	0.032	---	0.044	0.075	0.051	---	0.059	0.274	-0.082	-0.065	0.094	---	---	---	---	0.082	-0.135	---	-0.141
	---	(1.87)	---	(2.86)	(5.33)	(3.67)	---	(4.59)	(15.11)	(-5.64)	(-4.86)	(4.36)	---	---	---	---	(4.77)	(-4.99)	---	(-7.61)
Family Life Cycle Stage (Base: Other Family Types)																				
Single person	0.307	0.311	---	0.153	---	-0.023	---	---	-0.108	0.178	0.372	---	---	---	---	0.199	---	---	0.317	0.273
	(24.86)	(15.61)	---	(9.44)	---	(-1.82)	---	---	(-3.56)	(9.60)	(24.31)	---	---	---	---	(11.14)	---	---	(14.80)	(16.98)
Couple only	---	0.193	-0.080	---	---	---	---	---	-0.377	0.109	0.120	0.139	0.102	0.042	---	---	-0.253	-0.225	---	---
	---	(12.82)	(-6.94)	---	---	---	---	---	(-13.94)	(7.09)	(8.34)	(7.78)	(5.26)	(3.74)	---	---	(-15.58)	(-6.34)	---	---
Couple with one child	-0.153	---	-0.166	-0.141	-0.107	-0.107	-0.115	---	-0.046	---	-0.148	---	---	---	---	-0.145	-0.245	---	-0.227	---
	(-11.48)	---	(-13.60)	(-10.08)	(-7.61)	(-8.49)	(-8.84)	---	(-2.69)	---	(-10.43)	---	---	---	---	(-11.69)	(-14.01)	---	(-10.91)	---
No of full-time workers	0.059	---	---	0.051	0.115	0.056	---	---	---	---	---	0.050	---	---	---	---	---	---	---	0.034
	(7.62)	---	---	(5.67)	(13.71)	(6.63)	---	---	---	---	---	(4.57)	---	---	---	---	---	---	---	(2.88)
No of part--time workers	---	---	---	0.051	0.096	0.032	---	---	0.071	---	---	0.027	0.028	---	---	---	0.044	---	0.057	-0.054
	---	---	---	(6.59)	(13.26)	(4.47)	---	---	(8.30)	---	---	(2.54)	(2.64)	---	---	---	(5.21)	---	(5.87)	(-5.51)
Household Income (Base: Low Income (<30K))																				
Medium income (30-60K)	---	0.145	---	-0.024	0.100	0.023	---	0.117	---	0.053	0.094	0.087	0.059	0.051	---	---	-0.037	---	---	---
	---	(7.41)	---	(-1.73)	(7.75)	(1.84)	---	(9.18)	---	(5.26)	(7.09)	(5.44)	(3.15)	(4.85)	---	---	(-2.77)	---	---	---
High income (>70K)	---	0.262	-0.073	-0.088	0.268	0.113	0.152	0.323	---	---	0.215	---	---	---	-0.160	-0.031	---	---	0.114	0.079
	---	(12.89)	(-6.62)	(-5.13)	(16.85)	(7.33)	(13.56)	(21.99)	---	---	(13.71)	---	---	---	(-13.24)	(-2.54)	---	---	(6.47)	(5.26)
Employment Type (Base: Investment, Government Transfer and Other Types)																				
Paid income	0.027	---	---	---	---	---	---	0.045	---	-0.180	-0.091	---	---	0.040	---	---	---	---	---	-0.098
	(2.40)	---	---	---	---	---	---	(4.13)	---	(-17.94)	(-8.28)	---	---	(3.71)	---	---	---	---	---	(-6.18)
Vehicle Fleet Portfolio (Base: HH w/0 Car)																				
HH w/1 car	---	0.045	0.021	-0.104	---	---	---	---	---	0.098	0.080	-0.273	---	1.112	0.700	0.840	-0.356	---	---	0.030
	---	(3.41)	(2.28)	(-4.74)	---	---	---	---	---	(4.70)	(6.05)	(-9.51)	---	(43.81)	(28.24)	(30.91)	(-23.49)	---	---	(2.56)
HH w/2 cars	---	---	---	-0.148	---	---	---	---	---	0.116	0.099	-0.385	---	1.238	0.633	0.893	-0.466	0.047	---	---
	---	---	---	(-4.63)	---	---	---	---	---	(3.86)	(6.18)	(-9.73)	---	(39.34)	(20.81)	(24.76)	(-25.38)	(2.11)	---	---

Variables	HOU	SECH	UTL	ATP	CL	PC	HHMO	ENT	ED	HL	BSWA	AUTO	RECV	GAS	VEHI	VOP	PT	NMT	INTT	SAV
HH ≥ 3 cars	---	---	---	-0.195	---	---	---	---	-0.061	0.174	0.171	-0.277	---	1.395	0.558	0.936	-0.546	---	---	---
	---	---	---	(-4.39)	---	---	---	---	(-2.44)	(4.22)	(7.79)	(-5.08)	---	(35.27)	(14.17)	(20.15)	(-20.18)	---	---	---
Cars/Adult	---	0.141	0.062	0.157	0.018	---	0.046	0.094	0.050	-0.038	---	0.277	0.215	0.145	0.174	0.087	---	---	-0.062	-0.123
	---	(9.62)	(5.86)	(6.68)	(1.69)	---	(4.38)	(8.90)	(2.77)	(-1.68)	---	(9.25)	(11.20)	(7.94)	(9.21)	(3.77)	---	---	(-3.61)	(-9.11)
Home owner	-0.190	0.040	0.312	-0.155	---	---	0.090	---	---	0.086	0.072	-0.088	0.096	---	0.026	---	-0.198	-0.229	---	---
	(-14.55)	(2.61)	(23.65)	(-11.69)	---	---	(7.06)	---	---	(7.81)	(6.52)	(-4.98)	(3.45)	---	(2.05)	---	(-12.29)	(-9.88)	---	---
Dwelling Type (Base: Semi-detached, Terrace, Duplex and Other Types)																				
Single detached	-0.068	---	0.133	---	---	---	0.075	---	---	---	---	---	0.057	0.055	---	---	-0.103	---	---	---
	(-5.11)	---	(9.99)	---	---	---	(5.71)	---	---	---	---	---	(2.08)	(4.87)	---	---	(-6.05)	---	---	---
Apartment	0.178	---	-0.247	-0.034	0.132	---	-0.117	---	---	---	---	---	-0.107	---	---	---	0.112	---	0.160	---
	(11.02)	---	(-15.94)	(-2.25)	(10.57)	---	(-7.47)	---	---	---	---	---	(-2.48)	---	---	---	(5.94)	---	(7.59)	---
Residential Location Characteristics																				
Urban	0.317	---	---	---	0.049	0.076	---	0.184	0.140	---	---	---	-0.463	-0.139	---	0.071	0.391	---	0.205	-0.113
	(23.84)	---	---	---	(4.05)	(6.16)	---	(15.69)	(7.32)	---	---	---	(-21.61)	(-11.57)	---	(5.44)	(19.42)	---	(8.50)	(-7.55)
Population Centre Density (Base: Low Density)																				
Medium density	-0.143	0.049	---	0.037	-0.032	-0.047	---	-0.041	-0.131	-0.051	---	---	0.302	---	---	-0.082	-0.225	0.092	-0.033	0.042
	(-11.97)	(3.34)	---	(3.14)	(-2.75)	(-4.03)	---	(-3.59)	(-7.26)	(-4.54)	---	---	(13.49)	---	---	(-6.32)	(-14.82)	(3.79)	(-1.75)	(2.88)
Province (Base: Other Provinces and Territories)																				
Alberta	0.076	---	---	---	---	---	---	0.083	0.092	0.166	---	---	---	-0.100	0.167	---	0.097	---	0.048	---
	(5.04)	---	---	---	---	---	---	(5.68)	(4.21)	(11.11)	---	---	---	(-6.50)	(4.69)	---	(5.25)	---	(2.04)	---
British Columbia	0.117	---	-0.162	---	-0.112	-0.130	---	---	---	---	-0.095	-0.099	-0.119	-0.139	-0.452	---	0.427	---	0.273	-0.090
	(7.79)	---	(-11.18)	---	(-7.68)	(-8.85)	---	---	---	---	(-6.32)	(-4.31)	(-3.96)	(-8.95)	(-23.46)	---	(7.96)	---	(4.40)	(-4.78)
Ontario	0.324	-0.116	---	---	---	---	---	---	---	-0.177	---	---	-0.198	---	---	0.066	---	---	-0.090	---
	(8.12)	(-6.30)	---	---	---	---	---	---	---	(-13.39)	---	---	(-6.61)	---	---	(4.53)	---	---	(-3.98)	---
Quebec	---	-0.237	-0.336	0.031	-0.077	-0.201	-0.122	-0.134	0.046	---	-0.232	---	-0.076	---	---	---	-0.153	0.082	-0.445	-0.070
	---	(-12.01)	(-11.50)	(2.19)	(-5.69)	(-6.57)	(-9.05)	(-10.04)	(2.30)	---	(-16.96)	---	(-2.82)	---	---	---	(-8.51)	(2.96)	(-15.67)	(-4.15)
Temporal Variables (Base: 1997)																				
1998	---	---	---	---	---	---	---	---	---	---	0.050	---	---	-0.137	---	---	---	---	---	0.131
	---	---	---	---	---	---	---	---	---	---	(3.10)	---	---	(-7.83)	---	---	---	---	---	(6.59)
1999	---	-0.044	---	---	---	---	---	---	-0.043	---	---	---	---	-0.124	---	---	---	---	---	---
	---	(-1.91)	---	---	---	---	---	---	(-1.80)	---	---	---	---	(-7.00)	---	---	---	---	---	---
2000	---	---	---	---	---	---	---	---	---	0.047	---	---	---	---	0.092	---	---	---	0.090	---
	---	---	---	---	---	---	---	---	---	(2.83)	---	---	---	---	(5.33)	---	---	---	(3.45)	---

Variables	HOU	SECH	UTL	ATP	CL	PC	HHMO	ENT	ED	HL	BSWA	AUTO	RECV	GAS	VEHI	VOP	PT	NMT	INTT	SAV
2001	---	---	---	---	---	0.179	---	---	---	0.043	---	-0.066	---	-0.104	---	---	---	---	---	0.072
	---	---	---	---	---	(11.08)	---	---	---	(2.53)	---	(-2.42)	---	(-5.91)	---	---	---	---	---	(3.56)
2002	---	-0.063	---	---	---	---	---	---	---	0.091	0.029	---	---	---	---	---	---	---	-0.086	---
	---	(2.71)	---	---	---	---	---	---	---	(5.27)	(1.72)	---	---	---	---	---	---	---	(-2.85)	---
2003	---	-0.109	---	---	---	---	---	---	---	0.055	---	---	---	---	---	---	---	---	---	---
	---	(-4.45)	---	---	---	---	---	---	---	(3.13)	---	---	---	---	---	---	---	---	---	---
2004	---	-0.129	---	---	---	---	---	---	---	0.091	---	---	-0.095	---	---	---	---	---	-0.071	---
	---	(-5.31)	---	---	---		---	---	---	(5.18)	---	---	(-2.73)	---	---	---	---	---	(-2.46)	---
2005	---	-0.126	---	---	---	0.179	---	---	---	---	-0.046	---	---	---	---	---	---	---	-0.189	---
	---	(-5.28)	---	---	---	(10.51)	---	---	---	---	(-2.68)	---	---	---	---	---	---	---	(-3.72)	---
2006	---	-0.107	---	---	---	---	---	---	---	---	---	---	-0.201	---	---	---	---	---	---	---
	---	(-4.50)	---	---	---	---	---	---	---	---	---	---	(-5.43)	---	---	---	---	---	---	---
Recession years (2007-2009)	---	---	0.144	-0.023	---	0.236	---	---	-0.199	0.094	---	---	---	0.088	---	---	---	---	---	---
	---	---	(13.06)	(-1.99)	---	(20.54)	---	---	(-10.94)	(7.60)	---	---	---	(7.03)	---	---	---	---	---	---
Interaction Terms																				
QC*Urban	---	---	0.155	---	---	0.139	---	---	---	0.057	---	---	---	---	---	---	---	---	---	---
	---	---	(4.94)	---	---	(4.18)	---	---	---	(3.89)	---	---	---	---	---	---	---	---	---	---
ON*Urban	-0.164	---	---	---	---	---	---	---	-0.130	---	---	---	---	---	---	---	---	---	---	---
	(-3.91)	---	---	---	---	---	---	---	(-5.72)	---	---	---	---	---	---	---	---	---	---	---
AB*Urban	---	---	---	---	---	---	-0.054	---	---	---	-0.058	---	---	---	0.071	---	---	---	---	---
	---	---	---	---	---	---	(-3.40)	---	---	---	(-3.61)	---	---	---	(1.84)	---	---	---	---	---
BC*Urban	---	-0.135	---	-0.057	---	---	-0.120	---	0.044	---	---	---	---	---	---	---	-0.245	---	-0.260	---
	---	(-6.50)	---	(-3.48)	---	---	(-7.78)	---	(1.91)	---	---	---	---	---	---	---	(-4.38)	---	(-3.96)	---
Constant Terms and Satiation Parameters																				
Constants	-7.731	-8.948	-6.694	-8.014	-7.404	-7.211	-7.292	-7.521	-9.003	-7.711	-7.960	-8.912	-9.122	-8.596	-8.758	-8.966	-8.169	-9.300	-9.050	-8.392
	(-386.8)	(-347.7)	(-291.5)	(-412.3)	(-442.5)	(-387.6)	(-388.6)	(-474.5)	(-306.3)	(-374.7)	(-448.3)	(-351.9)	-(229.9)	(-367.9)	(-420.1)	(-372.7)	(-296.0)	(-336.2)	(-323.0)	(-419.6)
γ-parameters	7.804	7.063	6.031	7.019	6.077	5.092	6.562	6.421	6.972	6.449	6.354	9.560	7.568	6.428	6.865	6.025	5.707	5.740	7.384	9.898
	(416.7)	(390.2)	(240.3)	(464.9)	(294.6)	(230.7)	(294.9)	(323.7)	(308.9)	(406.6)	(370.0)	(370.1)	(262.4)	(336.8)	(475.0)	(381.3)	(327.2)	(194.7)	(338.0)	(437.2)
Log-likelihood at convergence = -1599533																				

HOU = Shelter; SECH = Secondary Accommodation; UTL = Utilities; ATP = Alcohol and Tobacco Product; CL = Clothing; PC = Personal Care; HHMO = Household Maintenance and Operation; ENT = Entertainment and Recreation; ED = Education; HL = Health Care; BSWA = Business Services and Welfare Activities; AUTO = Automobile Acquisition; RECV = Recreational Vehicle; GAS = Gasoline Costs; VEHI = Vehicle Insurance Costs; VOP = Vehicle Operation and Maintenance; PT = Public Transportation; NMT = Non-motorized Transport; INTT = Intercity Travel; SAV = Savings.

Table 3.4 Policy Simulation Results

enditure gories	Gas Expenditure = 0		Gas Expenditure Increased (15%)		Health Expenditure Increased (15%)		Savings Reduced (10%)	
	Short Term % Difference	Long Term % Difference	Short Term % Difference	Long Term % Difference	Short Term % Difference	Long Term % Difference	Short Term % Difference	Long Term Differen
	5.37	2.76	0.48	-0.08	0.27	-0.58	1.98	1.44
U	---	3.17	---	-0.24	---	-0.92	---	1.44
	---	3.25	---	-0.03	---	-0.73	---	1.29
B	15.38	4.9	6.93	0.13	5.87	-0.91	13.59	4.03
	7.85	3.37	2.01	-0.02	1.57	-0.8	4.5	1.4
R	10.38	3.38	4.47	0.03	3.76	-0.51	8.35	1.85
M	8.58	3.67	2.48	0.13	1.75	-0.83	4.26	1.57
	10.08	3.14	4.04	-0.26	3.88	-0.6	7.94	1.44
	---	6.92	---	1.15	---	-1.57	---	2.86
T	---	3.85	---	-0.21	---	---	---	2.19
V	9.44	4.17	2.6	0.17	1.69	-1.35	6.74	2.7
H	---	6.86	---	-1.72	---	-0.23	---	0.42
S	---	---	---	---	-0.52	-0.83	3	2.55
	---	4.43	---	-0.78	---	-0.91	---	1.15
P	12.51	5.51	3.32	0.21	2.23	-1.09	7.85	3.32
	19.78	4.39	10.71	-0.26	10	-0.98	16.16	2.47
CH	16.96	7.02	5.79	0.43	4.01	-1.6	15.79	7.23
	29.84	7.4	16.8	0.65	14.28	-0.97	18.19	-0.31
CV	---	12.07	---	1.4	---	-1.41	---	16.13
T	47.05	10.39	29.91	1.45	25.16	-1.06	19.06	-5.92
V	10.26	7.56	-1.48	0.21	-2.62	-1.95	---	---

FD = Food; HOU = Shelter; SECH = Secondary Accommodation; UTL = Utilities; ATP = Alcohol and Tobacco Product; CL = Clothing; PC = Personal Care; HHMO = Household Maintenance and Operation; ENT = Entertainment and Recreation; ED = Education; HL = Health Care; BSWA = Business Services and Welfare Activities; AUTO = Automobile Acquisition; RECV = Recreational Vehicle; GAS = Gasoline Costs; VEHI = Vehicle Insurance Costs; VOP = Vehicle Operation and Maintenance; PT = Public Transportation; NMT = Non-motorized Transport; INTT = Intercity Travel; SAV = Savings

.

CHAPTER 4 ADDRESSING POPULATION HETEROGENEITY IN DISCRETE CHOICE MODELS

4.1 Introduction

Literature on vehicle ownership is vast and expansive. However, it is surprising that there are very few vehicle ownership studies in the context of Canadian urban regions (Potoglou and Kanaroglou, 2008a; Roorda et al., 2000). The most recent study, conducted for the city of Hamilton, Ontario, Canada, was based on an internet survey that considered respondents who were employees of either City of Hamilton or McMaster University. The dataset employed in the analysis does not reflect the overall vehicle ownership preference of urban residents in Hamilton. The first objective of our study in this chapter is to address this limitation. We propose to estimate a vehicle ownership model using data from an entire metropolitan area, specifically the Quebec City region. The second objective of our study is to investigate the potential existence of population heterogeneity in the context of vehicle ownership. Towards this end, we propose the application of the latent class version of the ordered and unordered response models. Specifically, we estimate latent segmentation based ordered logit (LSOL) and latent segmentation based multinomial logit (LSMNL) models. Finally, we also undertake a comparison exercise of the latent class models with their traditional counterparts in the choice context examined.

4.1.1 Contribution and Organization of the Chapter

In this chapter, we apply a new modeling framework called the latent segmentation approach to assign households probabilistically to different segments based on a host of socio-demographic and land-use attributes (Bhat, 1997; Eluru et al., 2012a). That is, the formulated model allows us to partition households into segments and estimate the influence of exogenous variables on vehicle ownership decisions separately within each segment. Of course, the results from the analysis need to be examined carefully by the analyst to ensure that the outputs are not just statistical manifestations but are based on intuition and past evidence from literature. The proposed approach addresses two concerns: (1) ensures that the parameters are estimated employing the full sample for each segment while employing all data points for model estimation and (2) provides valuable insights on how the exogenous variables affect segmentation. This is the first implementation of latent segmentation framework for both ordered and unordered response models in the context of vehicle fleet size decision of households.

The remainder of the chapter is organized in the following order. Section 4.2 discusses the model structure and estimation procedure. Section 4.3, describes the main data sources and the sample formation procedure. Empirical results are presented and discussed in Section 4.4. Model validation outcomes and elasticity effects are also included in the same section. Finally, we summarize the major findings of the research in Section 4.5.

4.2 Econometric Framework

The latent class approach recognizes that households can be probabilistically assigned to different behaviorally similar segments as a function of observed attributes (Bhat, 1997; Srinivasan et al., 2009). Since the segments are unobserved to the analyst, they are termed as latent or endogenous. Within each segment, separate vehicle ownership models predict household choice behavior. Let us consider S homogenous segments of households (the optimal number of S is to be determined). We need to determine how to assign the households probabilistically to the segments for the segmentation model. The utility for assigning a household q (1,2,..Q) to segment s is defined as:

$$U_{qs}^{*} = \beta_s' z_q + \xi_{qs} \tag{4.1}$$

$\boldsymbol{z_q}$ is a (M x 1) column vector of attributes that influences the propensity of belonging to segment s, $\boldsymbol{\beta_s'}$ is a corresponding (M x 1) column vector of coefficients, and ξ_{qs} is an idiosyncratic random error term assumed to be identically and independently Type 1 Extreme Value distributed across households q and segment s. Then the probability that household q belongs to segment s is given as:

$$P_{qs} = \frac{\exp(\beta_s' z_q)}{\sum_s \exp(\beta_s' z_q)} \tag{4.2}$$

Within the latent segmentation approach, the probability of household q choosing auto ownership level k is given as:

$$P_q(k) = \sum_{s=1}^{S} (P_q(k) \mid s)(P_{qs}) \tag{4.3}$$

where $P_q(k)\ |s$ represents the probability of household q choosing auto ownership level k within the segment s. Note that the choice construct of car ownership considered to compute $P_q(k)\ |s$ may be either the ordered or unordered response mechanism.

Now, if we consider the car ownership levels of households (k) to be ordered,

$$y_{qs}^* = \alpha_s' \, x_q + \varepsilon_{qs}\,, \qquad y_q = k \qquad if\ \psi_{s_{k-1}} < y_{qs}^* < \psi_{s_k} \tag{4.4}$$

where y_{qs}^* is the latent propensity of household q conditional on q belonging to segment s. y_{qs}^* is mapped to the ownership level y_q by the ψ thresholds ($\psi_{s_0} = -\infty$ and $\psi_{s_k} = \infty$) in the usual ordered-response fashion. x_q is a (L x 1) column vector of attributes that influences the propensity associated with car ownership. α is a corresponding (L x 1) column vector of coefficients, and ε_{qs} is an idiosyncratic random error term assumed to be identically and independently standard logistic distributed across households q. The probability that household q chooses car ownership level k is given by:

$$P_q(k)|s = \Lambda\big(\psi_{s_k} - \alpha_s' x_q\big) - \Lambda\big(\psi_{s_{k-1}} - \alpha_s' x_q\big) \tag{4.5}$$

where $\Lambda(.)$ represents the standard logistic cumulative distribution function (cdf).

If we consider the car ownership levels (k) to be unordered, we can employ the usual random utility based multinomial logit (MNL) structure. Equation (4.6) represents the utility U_{qk} that household q associates with car ownership level k if that household belongs to segment s.

$$U_{qk}|\ s = \alpha_s' \, x_q + \varepsilon_{qk} \tag{4.6}$$

x_q is a (L x 1) column vector of attributes that influences the propensity associated with car ownership. α is a corresponding (L x 1)-column vector of coefficients and ε_{qk} is an idiosyncratic random error term assumed to be identically and independently Type 1 extreme value distributed across households q. The probability that household q chooses car ownership level k is given as:

$$P_q(k)\ |\ s = \frac{\exp(\alpha_s' x_q)}{\sum_k \exp(\ \alpha_s' x_q)} \tag{4.7}$$

The log-likelihood function for the entire dataset with appropriate $P_q(k)\ |\ s$ for ordered and unordered regimes is provided below:

$$L = \textstyle\sum_{q=1}^{Q} \log(P_q\big(k_q^*\big)), \tag{4.8}$$

where k_q^* represents the ownership level chosen by household q.

The model estimation process begins with a model considering two segments. The final number of segments is determined by adding one segment at a time until further addition does not enhance intuitive interpretation and data fit (Tang and Mokhtarian, 2009; Eluru et al., 2012a). The evaluation of the model fit in terms of the appropriate number of segments is based on the Bayesian Information Criterion (BIC)[5]. Estimation of the model is terminated when the increase in the number of segments results in an increase in the BIC value. Finally, the number of segments corresponding to the lowest value of BIC is considered the appropriate number of segments. The decision regarding the optimal number of classes should be taken considering the significance of the number of parameters and the interpretability as well as parsimony of the model (Beckman and Golias, 2008; Bujosa et al., 2010). The model estimates provide the segment characteristics, the segment specific discrete choice model estimates and number of segments.

4.3 Data

The proposed latent segmentation models are estimated using data derived from the Origin-Destination (O-D) surveys of Quebec City for the year 2001. The Quebec City database contained a total of 27,822 household data. After removing inconsistent and missing/miscoded values, we were left with 26,362 usable household records. From this, we randomly sampled 5,218 records for estimation and 1,326 records for model validation purpose.

Car ownership levels in the dataset were classified as no car, one car, two cars, and three or more cars. The dependent variable was truncated at three because the number of households with more than three automobiles was relatively small in the dataset. Table 4.1 provides a summary of the characteristics of the sample used in this study. The distribution of auto ownership levels in the estimation sample indicate that the number of two or more cars owning households is noticeably higher (42.7%) in Quebec City. From the descriptive analysis, we can observe that about 37 percent of the households have two or more full-time workers, about 9 percent have one or more part-time workers, and about 70 percent have two or more license holders. About three-

[5] The BIC for a given empirical model is equal to [– 2 (LL) + K ln (Q)], where (LL) is the log likelihood value at convergence, K is the number of parameters, and Q is the number of observations. BIC is found to be the most consistent Information Criterion (IC) for correctly identifying the appropriate number of segments in latent segmentation models (for more details, see Nylund et al., 2007; Roeder et al., 1999).

quarters of the households respectively have no children and no retirees, and more than two-thirds have no students.

4.4 Empirical Analysis

4.4.1 Variables Considered

The variables considered in our analysis can be broadly categorized into two categories: (1) household socio-demographic characteristics and (2) land use patterns. The demographic variables included number of children, number of employed adults (full-time and part-time), presence of executives, number of retirees, number of students, number of transit pass holders, number of household members, and number of licensed drivers.

In order to assess the impact of different land use characteristics on car ownership, three indicators were used: residential density, entropy index (EI_j) representing land use mix, and transit accessibility (A_j). For all calculations involving residential density, only residential land use area was used. The entropy index, EI_j is defined as: $EI_j = - \sum_k \frac{[p_k ln\, p_k]}{\ln(K)}$, where: p_k is the proportion of the developed land in the *k*th land use type. In our study, five (*K*=5) land use types were considered including residential, commercial, industrial, institutional[6] and park facilities. The value of this index varies between zero and one (since the measure was normalized by $\ln(K)$, zero (no mix) corresponds to a homogenous area characterized by single land use type and one to a perfectly heterogeneous mix). This index has been used in numerous studies for measuring land use mix (Chu, 2002; Kockelman, 1997; Potoglou and Kanaroglou, 2008a; Miranda-Moreno et al., 2011). The transit accessibility indicator takes into account the number of bus lines in the vicinity of the household, distance (in km) from the household to the closest bus stop of each of these lines (d_{ij}), and the average daily headway for each of these lines ($\bar{h}_i$). The formula for transit accessibility is: $Aj = \sum_{i=1}^{n} \frac{1}{d_{ij} * \bar{h}_i}$. This means that as the bus-stop distances and/or headways increase, the transit accessibility of household's decreases (Miranda-Moreno et al., 2011). On the other hand, a stop being closer or a smaller headway would mean a larger contribution to transit accessibility. This

[6] Institutional land use refers to land uses that cater to community's social and educational needs (schools, town hall, police station) while park facilities refer to land used for recreational or entertainment purposes.

variable was used as a proxy for the level-of-service (LOS) measure of the local public transit system.

The final specification was based on a systematic process of removing statistically insignificant variables (in our analysis we considered 90 percent significance level) and combining variables when their effects were not significantly different. The specification process was also guided by prior research, intuitiveness and parsimony considerations.

4.4.2 Model Specification and Performance Evaluation

In this research, we considered three different model specifications from both ordered and unordered choice mechanism. From the ordered category we estimated: (1) traditional ordered logit (OL) model, (2) latent segmentation based ordered logit model with two segments (LSOL II), and (3) latent segmentation based ordered logit model with three segments (LSOL III). From the unordered category we estimated: (1) traditional multinomial logit (MNL) model, (2) latent segmentation based multinomial logit model with two segments (LSMNL II), and (3) latent segmentation based multinomial logit model with three segments (LSMNL III).

Prior to discussing the model results, we compare the performance of the OL, LSOL II, and LSOL III models as well as the MNL, LSMNL II, and LSMNL III models. These models are not nested within each other. Hence, for evaluating their performance, we employ the Bayesian Information Criterion (BIC) measure. The model with the lowest value of BIC is preferred. The BIC values for the final specifications of the OL, LSOL II and LSOL III, MNL, LSMNL II and LSMNL III models are 7398, 7298, 9063, 7469, 7219 and 10334, respectively. These test statistics clearly prove that the specifications with two segments (LSOL II and LSMNL II) outperform all the other models within their respective regimes. Moreover, if more than two classes are included in the model, the third group represent only a small portion of the total households and thus does not yield any interpretable segment characteristics. Additionally, the LSMNL II has the lowest BIC value indicating that it fits the data better than the LSOL II model. These results provide strong evidence in favour of our hypothesis that car ownership of households can be better investigated through segmentation of households. From here on, we restrict ourselves to the discussion of only the LSOL II and LSMNL II models. The results for the traditional models are presented in Appendix A.

4.4.3 Behavioral Interpretation

Prior to discussing the impact of various coefficients on segmentation and car ownership, it is important to discuss the overall segmentation characteristics. The model estimations can be used to generate information regarding: (1) percentage household share across the two segments and (2) overall car ownership level shares within each segment. These estimates are provided in Table 4.2. Strikingly, we notice that the various measures computed for the LSOL II and LSMNL II exhibit very similar trends. In fact, the similarity across the ordered and unordered models confirms the presence of segmentation in the sample population.

In the two models, the likelihood of households being assigned to segment 1 is substantially higher than the likelihood of being assigned to segment 2. Further, the car ownership probabilities for households, conditional on their belonging to a particular segment, indicate that the two segments exhibit very distinct car ownership profiles. The households allocated to segment 1 are less likely to own zero cars (only 7% or 8%) whereas the households assigned to segment 2 are less likely to own 3 or more cars (only 2%). We also estimated the mean values of the segmentation variables within each segment to characterize and explain each segment more intuitively (Table 4.2, see Bhat, 1997 for details on computing these means). Based on the differences in the mean values of the segmentation variables, we can observe that the variables transit accessibility and transit pass holders offer the most substantial differences across the two segments. Hence, we employ these two variables to characterize our segments: segment 1 as households with reduced transit access and usage – transit independent (TI) and segment 2 as households with high transit access and usage – transit friendly (TF).

4.4.4 Estimation Results

4.4.4.1 Latent Segmentation Component

The LSOL II and LSMNL II model estimation results, for the segmentation component and the car ownership components for the two segments are presented in Table 4.3. In the following discussion, we discuss the variable effects on car ownership for the LSOL II and LSMNL II model simultaneously.

The latent segmentation component determines the probability that a household is assigned to one of the two segments identified. In our empirical analysis, Segment-1 is chosen to be the base and the coefficients presented in the table correspond to the propensity for being a part of the Segment-2. The constant term clearly indicates a larger likelihood for households being part of Segment-1. We found that the segment share is influenced by socio-demographic characteristics of household as well as land-use patterns. The attributes include: transit accessibility, entropy index, number of transit pass holders, number of household members and if any employed member of the household holds an executive position.

For all segmentation variables, both systems offer similar behavior. An increase in transit accessibility is likely to increase the probability that the household is part of Segment-2. With increase in the land use mix, represented by the entropy index, the likelihood of assigning the households to Segment-2 increases. Higher values of the entropy index imply that household members have the option to easily access many activities and amenities by walking or biking in addition to riding transit, thereby minimizing their need to procure and use cars (Cervero and Kockelman, 1997; Hess and Ong, 2002). Again, the higher the number of transit-pass holders in a household, the higher is the likelihood for assigning the household to Segment-2. Interestingly, households with two or more than two members were also more likely to be part of Segment-2. As expected, increased presence of executive job holders increases the chance that households would be assigned to Segment-1.

4.4.4.2 Car Ownership Component: Segment-1

Households with more employed adults (both full-time and part-time) and persons with driving license were associated with higher levels of car ownership; an indicator that these households have greater mobility needs (Kim and Kim, 2004; Potoglou and Kanaroglou, 2008a). The effect of full-time working adults is greater than that of part-time working adults. This is expected since full-time working adults have greater time-constraints and daily commitments, hence greater needs for personal vehicles. Gradually increasing alternative specific coefficients of full-time working adults and license holders in the LSMNL II imply that their effect on household's utility is higher as levels of car ownership increases.

Interestingly, number of children was associated with reduced likelihood of owning multiple cars. The result might seem counterintuitive at first glance. However, the negative effect of increased number of children on car ownership could be explained by the increased living expenses (food, clothing, and housing) that might curtail the amount of financial resources available for expenditures on acquiring and maintaining cars (Bhat and Koppelman, 1993; Soltani, 2005). The negative coefficients are gradually increasing, meaning that households associate greater disutility to multiple vehicle ownership levels with increase in the number of children. Similar to number of children, number of students also had a significant negative impact on car ownership. It is expected because households with more students would have increased budget constraints and hence, would be less inclined to own cars. Moreover, students may share their activities with friends and other household members that might further reduce the need for owning multiple cars (Vovsha et al., 2003). The result of the LSMNL II indicates that the likelihood of owning three or more cars decreases with increase in number of students in households. We also found that increase in number of retirees was associated with increased likelihood of owning multiple cars. Please note that the variable was significant in the LSMNL II model only. The finding is probably indicating that households with more retired persons are in a financially healthy situation (Matas and Raymond, 2008). Further, it is possible that these individuals prefer car mode for participating in activities.

The only land use variable that affected car ownership in this segment was residential density. As expected, the results indicated that as the residential density increases, the likelihood of households owning more cars decreases. The effect was found significant in both latent segmentation models. The signs of the coefficients as well as their magnitudes in the model show the expected trend (gradual increase in the disutility with increasing car ownership levels in the LSMNL II model). Households in denser areas tend to have fewer cars presumably due to lower car level-of-service (LOS) resulting from congestion, parking space constraints leading to escalated parking cost (Bhat and Koppelman, 1993) as well as more frequent and easily accessible public transport services (Hess and Ong, 2002). The lower speed in the dense residential zones might also be another deterrent to increased car ownership (Karlaftis and Golias, 2002).

4.4.4.3 <u>*Car Ownership Component: Segment-2*</u>

With increase in the number of employed adults (both full-time and part-time) in households, the likelihood of owning multiple cars increases (same as Segment-1). The results of the LSMNL II model show that higher number of workers increases the probability of households owning one or two cars (relative to zero or 3+ cars). This is expected since households in this segment have better transit accessibility and improved land use mix which might be obviating the need for purchasing and using more vehicles.

Similar to Segment-1, number of licensed drivers emerged as another important factor affecting car demand in Segment-2 as well. The number of licensed drivers was used as a surrogate for potential drivers in the household. The increase in potential drivers is more likely to increase the car ownership level of households. It is interesting to note that the contribution of licensed drivers reduces for the 3 or more car ownership category for the MNL system. The result indicates that the increase in the utility for households is not the same for car ownership levels of 3 and higher. Number of children has a negative impact on car ownership decision of households in LS (same as Segment-1). Interestingly, households with higher number of students had higher likelihood of owning more vehicles in both models. With increase in number of retirees, households in Segment-2 have a higher likelihood of purchasing multiple cars. From the LSMNL II estimates, it is found that increased number of retirees increased the probability of households owning two cars. This might be explained by the fact that retirees, who presumably have the time flexibility to take frequent leisure trips, are more likely to be dependent on cars for their mobility needs due to old age.

Overall, we see that the results for the LSOL II and LSMNL II models offer very similar interpretations. The difference in the mathematical framework and the differences in the formulation of the two frameworks can lead to the minor differences we observe. The results clearly underscore the importance of considering population heterogeneity through latent class models in the context of car ownership. Further, we also tend to observe that the additional flexibility of the MNL regime allows the LSMNL II model to better explain the dependent variable.

4.4.5 Validation Results

To ensure that the statistical results obtained were not a manifestation of over fitting to data, we evaluate the performance of the estimated models on a hold-out validation sample (1326 household data). This subsample of data was set aside during model estimation. Our validation analysis is conducted for both LSOL II and LSMNL II models.

To undertake the validation exercise, we employ the final parameters of the models to predict alternative probabilities for the households in the hold-out sample. To evaluate the performance, both aggregate and disaggregate measures of fit were computed. At the aggregate level, we compared the predicted[7] and actual auto ownership level shares and computed the root mean square error (RMSE) as well as the mean absolute percentage error (MAPE) of the predicted shares. At the disaggregate-level, we computed the predictive log-likelihood by calculating the log-likelihood for the predicted probabilities of the sample (Eluru et al., 2008). The results are reported in Table 4.4. The MAPE statistic shows that the LSMNL II model performs better than the LSOL II model in the overall. The predictive performance of the LSMNL II model is also superior compared to that of the LSOL II model based on the predictive log-likelihood value. Hence, there is enough evidence to suggest that LSMNL II performs slightly better in the empirical analysis compared to its ordered counterpart.

4.4.6 Elasticity Effects

The exogenous variable coefficients do not directly provide the magnitude of impacts of variables on the probability of car ownership levels. For better understanding the impacts of exogenous factors, we compute the relevant elasticities for changes in selected variables. The calculation results are presented in Table 4.5. For the analysis, we selected three socio-demographic variables (number of employed adults, number of children and number of transit pass holders) and two land use attributes (transit accessibility and residential density). Note that the elasticity effects were computed for the OL, LSOL II, MNL and LSMNL II models.

The results illustrate that both full-time working adults and part-time working adults increase household car ownership levels. However, as expected full-time working adults had greater impact on increasing vehicle ownership levels (2 or more) compared to the part-time working adults. The impact of change in number of children demonstrates the likelihood of vehicle fleet size reduction with similar impacts in magnitude in all the models. The reduction in fleet size observed in the elasticity analysis, while counterintuitive, is consistent with the coefficients of that variable in the models and is similar across all models; in particular, with respect to the large percentage increase in zero-car households it should be kept in mind that the base proportion of

[7] The aggregated predicted probabilities of car ownership outcome k of households belonging to a particular segment s can be calculated using the following equation: $\frac{\sum_q P_{qs} \times [P_q(k)|s]}{Q}$ and the overall predicted share is obtained by summing these probabilities over all segments.

those households is not very large (10%). It might be useful to investigate this result further in future analysis. Increase in number of transit pass holders resulted in a decrease in car ownership levels. The decreasing effect was more pronounced for 3 or more car ownership level. We can also see from the table that increase in transit accessibility and residential density reduces the probability of household's owning 2 or more cars. However, between the two attributes, residential density has a greater impact on car ownership levels than transit accessibility. The computation exercise provides an illustration of the applicability of the proposed framework for policy analysis.

4.5 Summary and Conclusions

There has been substantial interest in the transportation and planning literature on examining the factors that influence household car ownership levels. The topic is of great interest to policy makers given the growing focus on global warming, public health, and sustainable development issues. Two alternative model structures: ordered and unordered, have been extensively applied in the empirical studies to examine the underlying choice process for household's auto ownership preferences. These studies assume that the influence of exogenous variables remain the same for the entire population, although it is possible that the exogenous variable effects might vary across the population. The current research proposes the use of latent class modeling approach in the context of vehicle ownership. In latent class model, segment membership is probabilistically determined as a function of the socio-demographic and land use attributes of households. The approach accommodates heterogeneity within the systematic component as opposed to heterogeneity within the unobserved component captured in the simulation based mixed model approaches.

In our study, we estimate latent segmentation based ordered logit (LSOL) and latent segmentation based multinomial logit (LSMNL) models of car ownership using the data from Quebec City region of Canada. Using several goodness of fit criteria we conclude that the model specification with two-segments offered the best data fit. In both models, the probability of belonging to any segment was a function of land use characteristics (transit accessibility and entropy index) and household demographics (number of transit pass holders, presence of more than two household members and executive job position holding by any household member). Based on the differences in the mean values of the segmentation variables, we characterized our segments: Segment-1 as transit independent (TI) and Segment-2 as transit friendly (TF). In

Segment-1, higher number of employed adults and license holders increase the propensity of more cars, while increased number of children and students reduce the propensity. For Segment-2, in addition to number of employed adults and license holders, number of retirees was associated with increased car ownership of households. Similar to Segment-1, number of children has a negative impact on household's decision to own higher number of cars.

For a better understanding of the impacts of exogenous factors, we computed the relevant elasticities for changes in selected variables (number of employed adults, number of children, transit pass holders, transit accessibility and residential density). The elasticity effects indicated that both full-time working adults and part-time working adults increase household car ownership levels. On the other hand, increase in number of children, transit pass holders, transit accessibility and residential density, reduced the probability of multiple vehicle ownership. Between the two land use attributes, residential density was found to have a greater impact on car ownership levels.

4.6 Link between Chapter 4 and Chapter 5

In this chapter, in addition to identifying factors affecting vehicle fleet size decision, we also assessed the relative performance of the LSOL and LSMNL models with their traditional counterparts using several measures of fit. The results clearly offered evidence in favour of the hypothesis that car ownership can be better examined through segmentation of households. Moreover, between the two latent class models, the unordered choice mechanism appears to perform slightly better than the ordered response mechanism. But the evaluation results were not conclusive by any means. To investigate further into the matter, we extended the comparison exercise between the latent class models in the next chapter by estimating mathematically equivalent ordered and unordered models. Specifically, from the ordered regime latent class version of the generalized ordered logit model (LSGOL) as proposed in Yasmin et al. (2014) was estimated and compared with latent class multinomial logit (LSMNL).

Table 4.1 Summary Statistics of Variables

Variables	Frequency	%
Car Ownership Levels of Households		
0 Car	562	10.8
1 Car	2427	46.5
2 Cars	1886	36.1
≥ 3 Cars	344	6.6
Household Demographics		
Number of Full-time Employed Adults		
0	1510	28.9
1	1795	34.4
≥ 2	1913	36.7
Number of Part-time Employed Adults		
0	4763	91.3
1	437	8.4
≥ 2	18	0.3
Number of License Holders		
0	342	6.6
1	1231	23.6
≥ 2	3645	69.8
Number of Children		
0	3791	72.7
1	638	12.2
≥ 2	789	15.1
Number of Students		
0	4234	81.1
1	755	14.5
≥ 2	229	4.4
Number of Retirees		
0	3722	71.3
1	919	17.6
≥ 2	577	11.1
Sample size, N	5218	100

Table 4.2 Segment Characteristics and Mean Values of Segmentation Variables (N = 5218)

		Latent OL		Latent MNL	
		Segment-1	**Segment-2**	**Segment-1**	**Segment-2**
Household share		0.71	0.29	0.80	0.20
Car ownership within each segment					
0 Car		0.08	0.18	0.07	0.26
1 Car		0.41	0.57	0.44	0.53
2 Cars		0.39	0.22	0.41	0.18
≥ 3 Cars		0.11	0.02	0.08	0.02
Mean Values of Demographic and Land Use Variables in Each Segment					
	Overall Market	**Segment-1**	**Segment-2**	**Segment-1**	**Segment-2**
Transit Accessibility	317.23	275.9	418.5	281.97	454.90
Entropy Index	0.37	0.35	0.43	0.35	0.45
Number of Transit Pass Holders	0.19	0.03	0.60	0.08	0.61
Two Persons	0.38	0.42	0.27	-	-
More than Two Persons	0.41	0.31	0.63	0.37	0.54
Executive Job Holder	0.12	0.12	0.12	0.12	0.09

Table 4.3 Parameter Estimates (N=5218)

Variables	Latent OL				Latent MNL			
	Segment-1		Segment-2		Segment-1		Segment-2	
	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat
Segmentation Component								
Constant	-	-	-4.2017	-6.716	-	-	-3.0933	-11.487
Land Use Variables								
Transit Accessibility	-	-	0.0013	3.796	-	-	0.0011	4.835
Entropy Index	-	-	2.2129	3.792	-	-	1.5455	3.455
Household Demographics								
Number of Transit Pass Holders	-	-	3.0622	7.330	-	-	1.642	9.256
Number of Household Members (Base: Single person)								
Two persons			0.9768	2.132			-	-
More than two persons	-	-	2.4409	4.619	-	-	0.7454	3.341
Executive Job Holder	-	-	-0.5191	-1.949	-	-	-0.6368	-2.517
Car Ownership Component								
Constants/Thresholds								
Threshold 1/Constant 1	1.4568	6.782	1.6718	5.379	-16.3443	-2.242	-3.3287	-5.891
Threshold 2/Constant 2	7.3418	30.792	6.1667	15.706	-23.1409	-3.150	-5.8831	-8.523
Threshold 3/Constant 3	11.7227	39.31	10.0706	18.382	-31.3775	-4.191	-4.4137	-4.228
Land Use Variables								
Residential Density	-0.0012	-4.488	-0.0022	-6.181				
1 Car					-0.0032	-2.271	-0.0022	-3.545
2 Cars					-0.0048	-3.304	-0.0039	-4.643
≥ 3 Cars					-0.0067	-4.314	-0.0039	-4.643
Household Demographics								
Number of Full-time Employed Adults	0.8524	11.455	0.9945	7.327				
1 Car					7.6091	1.853	1.0878	4.979
2 Cars					8.5137	2.068	1.8853	6.858
≥ 3 Cars					8.8236	2.136	-	-
Number of Part-time Employed Adults	0.5636	3.395	0.6311	3.121				
0 Car					-0.6486	-3.678	-0.7045	-2.13
1 Car					-0.6486	-3.678	-0.7045	-2.13
2 Cars					-	-	-	-

≥ 3 Cars					-	-	-	-
Number of License Holders	3.6924	30.201	1.6565	13.74				
1 Car					19.2163	2.588	2.6001	10.636
2 Cars					22.4946	3.019	2.6001	10.636
≥ 3 Cars					25.1929	3.359	2.2289	3.885
Number of Children	-3.7143	-26.59	-1.1172	-7.615				
0 Car					-	-	2.0193	7.069
1 Car					-10.6869	-2.587	-	-
2 Cars					-13.9118	-3.350	-	-
≥ 3 Cars					-16.8293	-3.994	-	-
Number of Students	-0.3415	-2.733	0.2271	1.786				
1 Car					-	-	-0.6363	-4.373
2 Cars					-	-	-	-
≥ 3 Cars					-0.6073	-3.439	-	-
Number of Retirees	-	-	1.0505	6.187				
0 Car					-5.5836	-1.919	-	-
1 Car					-	-	-	-
2 Cars					-	-	1.2481	4.183
≥ 3 Cars					-	-	-	-
Log-likelihood at zero	-7233.68				-7233.68			
Log-likelihood at sample shares	-5964.82				-5964.82			
Log-likelihood at convergence	-3568.07				-3481.08			
Log-likelihood at convergence of traditional OL and MNL	-3647.70				-3619.15			
Log-likelihood at convergence of traditional OL and MNL with interaction terms	-3641.60				-3607.67			

Note: - denotes variables which are not significant. Also, the coefficient estimates across different alternatives are constrained to be same when the effects are not significantly different.

Table 4.4 Measures of Fit in the Validation Sample (N = 1326)

DISAGGREGATE MEASURES OF FIT		
Summary Statistic	**LSOL II**	**LSMNL II**
Log-likelihood at zero	-1838.23	-1838.23
Log-likelihood at sample shares	-1489.29	-1489.29
Predictive log-likelihood	-877.91	-852.71
Number of observations	1326	1326
Number of parameters estimated	19	30
Predictive adjusted likelihood ratio index	0.398	0.407

AGGREGATE MEASURES OF FIT			
Car Ownership Levels/ Measures of fit	**Actual Shares**	**Predictions**	
		LSOL II	**LSMNL II**
0 car	10.1	10.7	10.9
1car	46.7	45.8	45.1
2 cars	37.3	36.4	37.3
≥ 3 cars	6.0	6.9	6.7
RMSE	-	0.83	0.88
MAPE	-	6.30	5.75

Table 4.5 Elasticity Effects[8] of Important Variables

Models	Car ownership levels	Number of Full-time Employed Adults	Number of Part-time Employed Adults	Number of Children	Number of Transit Pass Holders	Transit Accessibility	Residential Density
		Variables Considered					
OL	0 Car	-24.31	-14.72	170.26	60.48	4.12	3.19
	1 Car	-10.35	-5.96	20.79	15.78	0.41	0.39
	2 Cars	13.26	8.09	-63.16	-29.61	-1.41	-1.14
	≥ 3 Cars	39.26	21.35	-79.44	-47.77	-1.91	-1.70
LSOL II	0 Car	-32.96	-22.46	187.82	44.39	1.88	6.86
	1 Car	-14.11	-9.17	19.15	17.30	0.36	0.50
	2 Cars	19.07	13.16	-67.35	-25.17	-0.76	-2.26
	≥ 3 Cars	49.15	29.62	-77.61	-57.40	-1.46	-2.54
MNL	0 Car	-32.58	-1.38	174.08	68.29	4.13	2.11
	1 Car	-14.81	-15.92	11.21	13.89	0.50	0.71
	2 Cars	23.55	16.50	-51.46	-27.30	-1.77	-1.25
	≥ 3 Cars	28.52	24.05	-81.36	-59.85	-0.60	-1.64
LSMNL II	0 Car	-24.20	-2.77	156.31	49.65	3.41	4.31
	1 Car	-16.54	-16.29	23.52	5.41	-0.11	1.47
	2 Cars	25.05	20.23	-62.01	-17.19	-0.73	-2.32
	≥ 3 Cars	18.34	7.74	-82.27	-25.43	-0.86	-4.69

[8] For the ordinal variables (number of employed adults, number of children and number of transit pass holders), the variable was increased by one unit and for the continuous variables (transit accessibility and residential density), the value was increased by 25 percent and the resulting percentage change in probability was calculated. The elasticity effects represent percentage change in the share of the dependent variable for a unit increase (increased by 1 for ordinal variables and 25% for continuous variables) in the independent variable.

CHAPTER 5 APPROPRIATE FRAMEWORK OF ANALYSIS WHILE ADDRESSING POPULATION HETEROGENEITY IN DISCRETE CHOICE MODELS

5.1 Introduction

Discrete choice models and their variants are employed extensively for analyzing decision processes in various fields including transportation, marketing, social sciences, bio-statistics and epidemiology. These models serve as a useful tool for understanding the impact of exogenous variables on the decision process by quantifying their impact on the dependent variable. The quantitative estimates provide an important tool for policy analysis. The number and form of the parameters is determined by the exact econometric modeling structure employed. To obtain accurate estimates of the impact of exogenous variables it is important that the appropriate model structure for the dataset under consideration is employed. Our proposed research is geared towards understanding the impact of population heterogeneity in ordered discrete variables and investigate the appropriate framework (ordered or unordered) of analysis while addressing population heterogeneity in the context of vehicle ownership.

5.1.1 Contribution and Organization of the Chapter

In our research, we propose a comparison of the latent class version of the ordered and unordered response models. We investigate the potential existence of population heterogeneity in the context of vehicle ownership in Greater Montreal Area (GMA). Towards this end, we estimate latent segmentation based ordered logit (both traditional ordered and generalized ordered response model) and latent segmentation based multinomial logit models. The analysis is undertaken in two steps. First, we estimate the latent class models using the vehicle ownership database for GMA. Second, we employ the estimates to predict car ownership levels for a hold out validation sample. The comparison exercise provides insights on population heterogeneity in terms of vehicle ownership choice while also providing insights on applicability of latent class ordered and unordered models for examining vehicle ownership.

In summary, this chapter makes several contributions. *First,* we enhance our understanding of the appropriate framework to study ordered discrete variables in the presence of population heterogeneity. *Second,* we propose a novel way of identifying important explanatory variables, though qualitatively, for better understanding the class-specific choice models. The exogenous variable coefficients of the class specific choice models only provide information regarding the directionality and magnitude of impacts of the variables. For instance, among others, number of full-time workers might have a positive impact while number of bus stops might have a negative impact on the vehicle ownership levels for all the identified classes. A natural question then arises is, in what preference order do these variables affect household's decision process? If there is an order and if we can identify it, it would enable us to make better inferences about the identified classes and their preferences/choices made. *Finally,* we contribute to the travel behavior literature on Canadian urban regions by investigating vehicle ownership decision processes in the Greater Montreal region. Towards that end, we use and examine a rich set of explanatory variables, including household socio-demographics, land use and transit measures.

The remainder of the chapter is structured in the following order. Section 5.2 describes the mathematical structure details of the latent segmentation based ordered and unordered models. In Section 5.3, the empirical context of the research is presented including brief highlights of the study area and the empirical findings of earlier literature. Information regarding data source, sample formation, and sample descriptive statistics are provided in Section 5.4. Section 5.5 contains a detailed discussion of empirical results. Validation measures are also presented in the same section. Finally, we summarize the major findings of the research in Section 5.6.

5.2 Econometric Framework

Let us consider S homogenous segments of decision makers (the optimal number of S is to be determined). We need to determine how to assign the decision makers probabilistically to the segments for the segmentation model. The most commonly used structure corresponds to the multinomial logit structure (see Bhat, 1997; Greene and Hensher, 2003; Eluru et al., 2012a). The utility for assigning a decision maker *h (1,2,..H)* to segment s ($s = 1,2, \dots S$) is defined as:

$$U_{hs}^{*} = \beta_s' z_h + \xi_{hs} \tag{5.1}$$

z_h is a (M x 1) column vector of attributes that influences the propensity of belonging to segment s, β_s' is a corresponding (M x 1) column vector of coefficients and ξ_{hs} is an idiosyncratic random error term assumed to be identically and independently Type 1 Extreme Value distributed across household h and segment s. The probability that household h belongs to segment s is given as:

$$P_{hs} = \frac{\exp(\beta_s' z_h)}{\sum_s \exp(\beta_s' z_h)} \tag{5.2}$$

Within the latent segmentation approach, the unconditional probability of decision makers h choosing alternative level k is given as:

$$P_h(k) = \sum_{s=1}^{S} (P_h(k) \mid s)(P_{hs}) \tag{5.3}$$

where $P_h(k)|s$ represents the conditional probability of household h choosing auto ownership level k within the segment s.

Now, if we consider the dependent variable (k) to be ordered, we can analyze the fleet size decision using a GOL model. In the GOL model:

$$y_{hs}^* = \alpha_s' x_h + \varepsilon_{hs}, \qquad y_h = k \qquad if\ \psi_{s_{k-1}} < y_{hs}^* < \psi_{s_k} \tag{5.4}$$

where y_{hs}^* is the latent propensity of decision maker h, conditional on h belonging to segment s. y_{hs}^* is mapped to the dependent variable y_h by the ψ thresholds ($\psi_{s_0} = -\infty$ and $\psi_{s_k} = +\infty$) in the usual ordered-response fashion $\left(-\infty < \psi_{s_1} < \psi_{s_2} < \ \dots\dots\dots < \psi_{s_{k-1}} < +\infty\right)$. x_h is a (L x 1) column vector of attributes that influences the propensity associated with vehicle ownership. α is a corresponding (L x 1) column vector of coefficients and ε_{hs} is an idiosyncratic random error term assumed to be identically and independently standard logistic distributed across households h. To maintain the ordering conditions and allow the thresholds to vary across households within each segment, Eluru et al. (2008) propose the following non-linear parameterization of the thresholds as a function of exogenous variables:

$$\psi_{s_k} = \psi_{s_{k-1}} + exp(\delta_{sk} Z_{sh}) \tag{5.5}$$

where δ_{sk} is a segment-specific and vehicle fleet level-specific row vector of parameters to be estimated and Z_{sh} is a corresponding column vector of segment-specific exogenous variables. Given the above set-up, the conditional probability that household h chooses car ownership level k is given by:

$$P_h(k)|k = \Lambda\left(\psi_{s_k} - \alpha_s' x_h\right) - \Lambda\left(\psi_{s_{k-1}} - \alpha_s' x_h\right) \tag{5.6}$$

where $\Lambda(.)$ represents the standard logistic cumulative distribution function (cdf).

On the other hand, if we consider dependent variable (k) to be unordered, we can employ the usual random utility based MNL structure. Equation (7) represents the utility U_{hks} that household h associates with car ownership level k if that household belongs to segment s

$$U_{hks} = \alpha_s' \, x_h + \varepsilon_{hks} \tag{5.7}$$

x_h is a (L x 1) column vector of attributes that influences the propensity associated with vehicle ownership. α is a corresponding (L x 1)-column vector of coefficients and ε_{hks} is an idiosyncratic random error term assumed to be identically and independently Type-1 extreme value distributed across decision maker h. Then the conditional probability that decision maker h chooses dependent variable level k is given as:

$$P_h(k)|s = \frac{\exp(\alpha_s' x_h)}{\sum_k \exp(\alpha_s' x_h)} \tag{5.8}$$

The log-likelihood function for the entire dataset with appropriate $P_h(k)|s$ for ordered and unordered regimes is provided below:

$$L = \sum_{h=1}^{N} log \left[\sum_{s=1}^{S} (P_h(k)|s) \times (P_{hs}) \right] \tag{5.9}$$

It is worthwhile to mention here that the estimation of latent segmentation based models using quasi-Newton routines can be computationally unstable (see Bhat, 1997 and Sobhani et al., 2013 for a discussion). The estimation of such models requires employing good starting values for the estimation procedure. Hence, for our analysis, the log-likelihood function and its corresponding gradient function were coded in the Gauss matrix programming language. The coding of the gradient function ensures the reduction in instability associated with such an estimation process.

5.2.1 Post Estimation Qualitative Analysis

In this research, to further explain the relationships observed in the latent segmentation based discrete models, we propose an importance ranking of variables affecting the dependent variables in each segment. The ranking of variables will allow us to see what variables affect the choice process and how these variables might differ across segments in the population. To compute the ranking, within each segment, we compute the contribution of each variable (summed across multiple alternatives if the variable is present in multiple alternatives) relative to the total utility across all alternatives. The ratio computed is averaged across the entire population. The process is repeated for all variables and a ranking is generated in the descending order. The approach is similar to one that was employed by Walker and Li (2007). Note that the ranking approach is qualitative and intended to understanding the segment characteristics and the rating results are a function of the underlying data being used for the model development and hence, cannot be generalized.

5.3 Empirical Context

5.3.1 Study Area

Montreal is the second largest Census Metropolitan Area (CMA) in Canada characterized by a diverse urban form and a unique heterogeneous multimodal transportation system comprising of metro, commuter train, and an extensive bus service. The urban region size with its land use mix, public transportation system, and active transportation infrastructure and culture makes Montreal an ideal subject to study vehicle fleet size decisions of households. Moreover, since the city has a diverse socio-demographic composition, investigation of the possible existence of observed population heterogeneity and its sources in the context of vehicle ownership decision are useful.

5.3.2 Earlier Work

To help articulate the position of the current research, we present a brief summary of the most important characteristics of earlier research efforts investigating vehicle fleet size decision of households. From an empirical standpoint, studies have found significant relationships between vehicle ownership and variables such as household socio-demographic characteristics, residential location attributes, built environment variables, and vehicle attributes. It is well established in the literature that among the different socio-demographic characteristics, household income dominates the choice process of household auto ownership. As expected, affluent households, irrespective of country and region, are always found to have a stronger preference towards higher number of private cars compared to middle-and low-income families (Karlaftis and Golias, 2002; Cao et al., 2007b; Li et al., 2010). In addition, families residing in their own homes are likely to own multiple private vehicles compared to families that rented or lived in apartments (Li et al., 2010; Zegras, 2010). The other two important demographic factors found in literature are presence (and number) of employed adults and license holders. Household vehicle fleet size tend to increase with the increase in number in these demographic categories (Bhat and Pulugurta, 1998; Ryan and Han, 1999; Chu, 2002; Karlaftis and Golias, 2002). In short, the empirical analyses show that the impact of socio-demographics on vehicle ownership is pretty well-defined. In terms of the impact of the build environment and land use variables, however, researchers so far have found mixed and inconclusive results. Some have found that increased population and residential density, and transit accessibility negatively impact number of automobiles owned by households while others have reported minimal or no association at all (see Bhat and Guo, 2007). The reason might be that the relationship between built environment attribute and vehicle ownership is complex and cannot be disentangled easily. Among others, only few researchers have examined the effect of non-traditional variables such as, change in residential location or moving on vehicle ownership (Yamamoto, 2008; Weinberger and Goetzke, 2010) or attempted to capture the common unobserved factors influencing car ownership and other decision processes (Bhat et al., 2013).

5.4 Data

5.4.1 Data Source

The primary data source used in the current analysis is the 2008 cross-sectional Origin-Destination (O-D) survey of Greater Montreal Area (GMA). The O-D surveys are conducted every five year and they are the primary source of information on individual mobility patterns in the Montreal region. *Agence Metropolitaine de Transport (AMT)* of Quebec provided us the survey data where a total of 66,124 household records were available. This data only contained socio-demographic information of the surveyed households. Hence, we decided to augment it using secondary data sources. The secondary sources used were: (1) census 2006 data summary files, and (2) GIS layers of land use, road network, and transit facilities. Using these sources, several demographic, land use, and built environment variables at the census tract (CT) level were extracted, aggregated accordingly, and appended to the O-D database for exploring them in the model specification. The final random sample for analysis comprised approximately 7,000 households, of which 4,913 data records were used for estimation and 1,972 data records were set aside for model validation. This was done to reduce the data compilation burden using ArcGIS.

5.4.2 Descriptive Statistics

Car ownership levels were classified as no car, one car, two cars, and three or more cars. Table 5.1 provides a summary of the characteristics of the sample used in this study. The distribution of auto ownership levels in the estimation sample is as follows: 20.3% of the households were carless, 43.3% owned one car, 28.7% owned two cars, and 7.8% of the households had a fleet size in excess of two cars. Moreover, the auto ownership descriptive analysis indicated an average ownership of 1.27 vehicles per household. Figure 5.1 presents the spatial distribution of the household by vehicle ownership levels in the GMA at the Traffic Analysis Zone (TAZ) level. The color coding changes from light to dark grey as the proportion of household increases by ownership level. Interestingly, as the number of cars in the households increases, the outer TAZ areas become darker and inner neighborhoods become pretty light or almost white. This in other words provides evidence that households with larger number of cars live in farther suburbs while households in the central areas have fewer cars. Some other salient characteristics of the sample are: majority of the households (60.8%) reside in medium income census tracts, two-thirds have at least one male adult (67%), one full-time employed member (64.6%), about three-quarters have at least one part-time worker, nearly one-third of the households have at least one child and one-third of the households have at least one retiree.

5.5 Empirical Analysis

5.5.1 Variables Considered

5.5.1.1 Segmentation Component

In the current research, we considered segmentation based only on land use variables although the proposed modeling approach theoretically can accommodate classification of segments based on the universal set of variables. This is done so for two specific reasons. First, while it is plausible to consider all variables in the latent segmentation consideration, the estimation of latent segmentation models with the entire variables set is likely to result in convergence challenges as well as difficulty in interpreting the results (see Sobhani et al., 2013 and Eluru et al., 2012a for discussions on challenges associated with latent segmentation models). Second, there has been much interest amongst the travel behavior researchers in disentangling the effects of built environments on vehicle ownership (see Bhat and Eluru, 2009; Eluru et al., 2010a; de Abreu e Silva et al., 2012). Research on this issue is divided into two major streams. One stream of studies treat land use characteristics as purely exogenous factors in models of vehicle ownership. However, doing so might result in self-selection bias. The other stream of studies overcome the bias by capturing the common unobserved heterogeneity between land use and vehicle ownership (Bhat and Guo, 2007; Pinjari et al., 2011). However, the modeling approach require extensive simulation. On the other hand, latent class models offer an elegant and tractable way to capture both unobserved and observed taste variation (Srinivasan et al., 2009). Therefore, in our study, we segment households based on observable land use attributes so that those with similar land use preferences are grouped together while implicitly accounting for the self-selection bias.

5.5.1.2 <u>Vehicle Ownership Component</u>

In the current study, a comprehensive set of exogenous attributes were considered to examine vehicle ownership levels of households. The independent variables can be broadly classified into three categories: (1) household socio-demographic characteristics, (2) land use characteristics, and (3) transit accessibility measures. *Household socio-demographic variables* that were employed in our analysis included number of employed members (full-time and part-time), number of male and female adults, number of adults belonging to different age groups, presence of children of different ages, number of retirees, and number of students. The following *land use variables* (at the census tract level where the household is located) were considered in our study: number of single and semi-detached as well as rented households, number of employed adults, high-school and trade certificate holders, and number of driver, passenger, transit, and walk commuters. Additionally, total area (in hectare) of the census tract, land use mix or entropy index (EI), population density, and median income of households in the census tract based on residential location were also included. The entropy index, EI_j is defined as: $EI_j = -\sum_k \frac{[p_k \ln p_k]}{\ln(K)}$, where: p_k is the proportion of the developed land in the kth land use type. In our study, five (K = 5) land use types were considered including residential, commercial, industrial, government and institutional[9], and park

[9] Institutional land use refers to land uses that cater to community's social and educational needs (schools, town hall, police station) while park facilities refer to land used for recreational or entertainment purposes.

facilities. The value of this index ranges from zero to one (since the measure was normalized by $\ln(K)$, zero (no mix) corresponds to a homogenous area characterized by single land use type and one to a perfectly heterogeneous mix). This index has been used in numerous studies for measuring land use mix (Chu, 2002; Kockelman, 1997; Potoglou and Kanaroglou, 2008a). We also created another variable called older suburbs if the majority of the households in the census tract were constructed pre-1946 (dwellings built before 1946 identifies the historical center of a city). Further, we introduced location specific (borough indicators) variables to examine the degree of influence exerted by the area of residence on household car ownership levels. These variables are expected to capture attributes of household's activity travel environment as well as the utility/disutility of automobile maintenance and operation (such as parking costs) in particular areas. *The transit accessibility measures* considered, as a proxy for ease of transit accessibility and level of service of alternative modes, (within 600m buffer[10] of household residential location) were: number of bus, commuter rail, and metro stops as well as length of bus line (km), length of commuter rail line (km), and length of metro line (km).

The final specification was based on a systematic process of removing statistically insignificant variables (in our analysis we considered 90 percent significance level) variables and combining variables when their effects were not significantly different. The specification process was also guided by prior research, intuitiveness and parsimony considerations.

5.5.2 Model Specification and Performance Evaluation

In this research, we considered six different model specifications from ordered choice mechanism and three different model specifications from the unordered choice mechanism. From the ordered category we estimated: (1) traditional ordered logit (OL) model, (2) latent segmentation based ordered logit model with two segments (LSOLII), (3) latent segmentation based ordered logit model with three segments (LSOLIII), (4) generalized ordered logit (GOL) model, (5) latent segmentation based generalized ordered logit model with two segments (LSGOLII), and (6) latent segmentation based generalized ordered logit model with three segments (LSGOLIII). From the unordered category we estimated: (1) traditional multinomial logit (MNL) model, (2) latent

[10] Buffers were established around household geocoded locations with 600m radius. In earlier literature, the acceptable walking distance to transit stops and stations is often assumed to be 400m (Larsen et al., 2010). Hence, we employed a slightly larger buffer than the 400m to allow for the low-density developments in Canadian cities that might require people to walk further to reach transit stations from their households.

segmentation based multinomial logit model with two segments (LSMNLII), and (3) latent segmentation based multinomial logit model with three segments (LSMNLIII). The performance of the developed models was evaluated in four steps. In the first step, appropriate latent segmentation scheme for the OL, GOL, and MNL models was identified. Next, the latent class ordered (OL and GOL) and unordered models, obtained from the first step, were compared with their traditional counterparts (OL, GOL, and MNL). Then, the performance of the two latent class models from the ordered regime (OL and GOL) was evaluated. Finally, an in-depth comparison of the latent class versions of the ordered and unordered response models was carried out.

The model estimation process began with a model considering two segments. The final number of segments was determined by adding one segment at a time until further addition did not enhance intuitive interpretation and data fit (Tang and Mokhtarian, 2009; Eluru et al., 2012a). The evaluation of the model fit in terms of the appropriate number of segments was based on the Bayesian Information Criterion (BIC)[11]. Estimation of the model was terminated when the increase in the number of segments resulted in an increase in the BIC value. Finally, the number of segments corresponding to the lowest value of BIC was considered as the appropriate number of segments. It should be noted that the decision regarding the optimal number of classes should be taken considering the significance of the number of parameters and the interpretability as well as parsimony of the model (Beckman and Golias, 2008; Bujosa et al., 2010).

The BIC (number of parameters estimated) values for the LSOL model with two and three segments were, 8995.61 (43) and 9002.78 (52), respectively. The BIC (number of parameters estimated) values for the LSGOL model with two and three segments were, 8962.49 (51) and 8964.08 (57), respectively. The calculated BIC (number of parameters estimated) values for the LSMNL model with two and three segments were, 9282.41 (68) and 9315.77 (78), respectively. Therefore, we selected two segments as the appropriate number of segments for both LSOL, LSGOL, and LSMNL models. The traditional and latent class models are not nested within each other. Hence, for evaluating their performance, we employed the BIC measure and chose the model with the lowest BIC value. The BIC (number of parameters estimated) values for the final specifications of the OL, LSOL II, GOL, LSGOLII, MNL, and LSMNLII models were: 9251.91

[11] The BIC for a given empirical model is equal to [– 2 (LL) + K ln (Q)], where (LL) is the log likelihood value at convergence, K is the number of parameters, and Q is the number of observations. BIC is found to be the most consistent Information Criterion (IC) for correctly identifying the appropriate number of segments in latent segmentation models (for more details, see Nylund et al., 2007; Lee and Timmermans, 2007; Roeder et al., 1999).

(27), 8995.61 (43), 9056.90 (46), 8962.49 (51), 9392.50 (63) and 9282.41 (68), respectively. These test statistics clearly prove that the specifications with two segments (LSOLII, LSGOLII, and LSMNLII) outperform all the other models within their respective regimes. Moreover, the LSGOL II has the lowest BIC value between the two latent class ordered models indicating that it fits the data better than the LSOLII model. And between the ordered and the unordered model, we can clearly see that the LSGOLII model outperforms the LSMNLII model by a large margin. Hence, from here on, we restrict ourselves to the discussion of only LSGOLII model.

5.5.3 Behavioural Interpretation

Prior to discussing the impact of various coefficients on segmentation and car ownership, it is important to discuss the overall segmentation characteristics. The segment membership model provides information as to who is likely to be in each segment, whereas the segment specific choice models provide information on how each segment behaves. Thus, the model estimations can be used to generate: (1) percentage of household share across the two segments and (2) overall car ownership level shares within each segment. These estimates are provided in Table 5.2. Interestingly, it is observed that the likelihood of households being assigned to any of the two segments is almost equal. Further, the car ownership probabilities for households, conditional on their belonging to a particular segment, indicate that the two segments exhibit very distinct car ownership profiles. The households allocated to Segment-1 are less likely to own zero cars (only 13%) whereas the households assigned to Segment-2 are less likely to own 3 or more cars (only 2%).

In order to characterize the segments better and more intuitively, we calculated the mean values of the segmentation variables (Table 5.2, see Bhat,1997 for details on computing these means) and created an "importance rating" of variables listed for each specific class. Overall, Segment-2 is characterized by higher population density, lower land area, and increased number of rented households in the census tract. The variable means clearly indicate that second segment corresponds to households from well-connected neighborhoods that are less reliant on automobile usage compared to the households from Segment-1.

The importance rating of variables generated as described in the Econometric Framework section are presented in Table 5.3. From the results we observe that household socio-demographics are more important in Segment-1. The vehicle ownership decision of these households is more likely to be governed by the monetary solvency, individual preference (age and gender), and mobility needs of their members. On the other hand, households in Segment-2 are likely to emphasize more on surrounding land use attributes while making their decision regarding vehicle fleet size. Increased number of employed adults in the census tract where the household resides appears as the most important variable followed by the number of commuters by automobile mode. Household socio-demographics are the other important attributes for this segment. Based on these segment specific car ownership shares, variable means, and importance rating, we can characterize Segment-1 as low density high car ownership segment (LHS) and Segment-2 as high density low car ownership segment (HLS).

5.5.4 Estimation Results

5.5.4.1 Latent Segmentation Component

The LSGOLII model estimation results, for the segmentation component and the car ownership components for the two segments for Greater Montreal Area (GMA) are presented in Table 5.4. The latent segmentation component determines the probability that a household is assigned to one of the two segments identified. In our empirical analysis, Segment-1 is chosen to be the base and the coefficients presented in the table correspond to the propensity for being a part of Segment-2. The constant term does not have any substantive interpretation. The land use variables found to significantly impact the probability of class membership in the LSGOLII model are: whether the household is located outside of Montreal Island, the number of rented households in the CT, whether the household is located in the older suburbs, population density, and area of the census tract where the household is located.

If households are located outside of Montreal Island, they are more likely to be assigned to Segment-1 which is the high car ownership segment. As is the case for the vast majority of North American Metropolitan areas, Greater Montreal urban areas extend well beyond its urban core – the Island of Montreal. In fact, it is reported that 80 percent of the urban area is located outside the central city (Canada CM, 2010). These are regions that are characterized by low population density, high auto accessibility, and reduced transit accessibility. Hence, it is likely that households residing in these areas will tend to own more cars (Zegras, 2010).

When households reside in a CT with increased number of rented households (which are generally multi-family housing unit), they are more likely to be part of Segment-2 which is the low car ownership segment. The result might be explained in light of the fact that in census tracts with more rental units, the average income may be lower indicating lower purchasing power (Dieleman and Everaers, 1994), which may in turn translates to potentially reduced car ownership levels. It is also reported in literature that census tracts with higher percentage of rented dwelling units tend to have higher job accessibility and households residing in areas with high job accessibility are less auto oriented (Gao et al., 2008). In addition, rental dwellings are more often built as compact medium or high-rise developments as opposed to spacious owner-occupied dwellings. Moreover, in the case of rented dwellings, garaging vehicles might prove to be difficult and inconvenient, due to space constraint as well as high parking costs. These issues are likely to play a key role in the vehicle ownership decision of households (Guo, 2013). Thus the observed result might be reflecting the effect of housing tenure decision as well as parking supply on household's fleet composition decision.

Households living in older suburbs have higher probability of being assigned to low-car ownership segment (Segment-2). These areas, built before World war II, are characterized by medium land-use mix, higher population and building density, grid street-design, and also have pedestrian facilities in the form of sidewalks, paths, and crosswalks allowing daily requirements to be met either by walking or taking transit (Badland and Schofield, 2005; Moudon et al., 1997). Hence, households living in these neighbourhoods are less geared towards car ownership and use.

The positive coefficient for population density indicated that with increase in the population density the likelihood of the households being part of low car ownership (Segment-2) increases. The result is intuitive and conforms to the findings of previous studies (see Chen et al., 2008; Li et al., 2010; Schimek, 1996; Dargay and Hanly, 2007). Areas that are dense in population often have clustered opportunities for both work and non-work related activities. High concentration of population and/or activities, in turn, allows for easily accessible, reliable and efficient transit service (Giuliano, 2004). Therefore, households might be less inclined to own and use automobile and more inclined towards walking or using transit (Reilly and Landis, 2002) in areas with increased population density. We also observed that the bigger the census tract, in terms of area (hectares), the higher is the probability that the households would belong to Segment-1. Census tract area could be taken as a proxy for the number and spatial distribution of opportunities available (Giulianno and Dargay, 2006; García-Palomares, 2010). Households living in large census tracts might need to rely on auto for reaching different services for fulfilling their needs.

5.5.4.2 Car Ownership Component

In the following discussion, we discuss the variable effects on car ownership for the LSGOL II.

Segment-1

For obvious reasons, presence of children in households significantly affects their fleet size decision. In particular, surprisingly though, we found that households with children between 5 to 9 years have a lower propensity of possessing multiple vehicles. On the other hand, with the presence of teenaged children (15-19 years of age), households tend to own multiple automobiles. At 18, teenagers are allowed to drive alone and thus, households might acquire extra cars to allow them to drive independently (Prevedouros and Schofer, 1992; Prillwitz et al., 2006). The negative effect of this variable on the threshold separating two and more than two cars categories indicates an increased likelihood of household's owning more than two cars.

Households with more employed adults (both full-time and part-time) were associated with higher levels of car ownership; an indicator that these households have greater mobility needs (Kim and Kim, 2004; Potoglou and Kanaroglou, 2008a). As anticipated, the effect of full-time working adults is greater than that of part-time working adults. This is understandable since full-time working adults have greater time-constraints and daily commitments, hence greater needs for personal vehicles (Bhat and Pulugurta, 1998; Dargay and Hanly, 2007). With increase in the number of full-time workers, LSGOLII model indicated that two vehicles may be the upper bound on automobile ownership choice for the households. Increase in the number of male, female, young, and middle-aged adults in the household, increases the probability of high car ownership of households. The positive effect of number of male adults on the threshold separating one and two cars categories indicates a higher probability of household limiting their fleet size to one vehicle.

It is evident from previous literature that income is one of the most influential factors affecting household's decision regarding their vehicle fleet size (Chu, 2002). In our analysis, household income was unavailable to us. However, to address the unavailability we employed census tract median income as a proxy measure for the relative affluence of households. From our analysis results, we found that households living in medium and high income census tracts have a stronger preference towards owning more cars. The result is in close agreement with the findings of the previous literature (Ma and Srinivasan, 2010; Karlaftis and Golias, 2002; Li et al., 2010).

Number of transit commuters was negatively associated with the vehicle ownership levels of households, intuitively suggesting that localities where public transit is more used, the likelihood of vehicle purchase will be lower. It might be due to two reasons: the CT is located in the urban core where the dwelling density and land use mix is higher (Pinjari et al., 2007) and it has better transit accessibility and frequency of service (Legrain et al., 2015; Ewing and Cervero, 2010, Moniruzzaman and Paez, 2012). From our analysis, we also observed that number of high-school certificate holders in the census tract positively impact household's vehicle fleet choice. On the other hand, number of trade certificate holders impact the decision negatively. The positive sign of the coefficient of number of single and semi-detached households in the first threshold indicates higher probability of owning one car by households. In our analysis, in addition to the above land-use measures we considered a host of borough variables. For Segment-1, only the Ville-Marie borough was found to be significant. Interestingly, although the borough represents medium to high dense neighbourhood around the downtown region with good transit accessibility in general, households located in this borough exhibited an increased propensity towards multiple vehicle ownership. The result needs to be further investigated.

The results corresponding to transit accessibility measures highlight the important role of public transit in Montreal. Increase in the number of metro stops as well as the length of bus line within the household buffer zone negatively impacted household's propensity to own multiple cars. The result lends support to the idea that increased transit access and high quality of transit service can significantly reduce the number of automobiles owned by households (Ryan and Han, 1999; Bento et al., 2005; Kim and Kim, 2004; Cullinane, 2002).

Segment-2

Presence of toddlers (children 0-4 years) is associated with an increased propensity to own multiple vehicles. The effect of presence of children 5 to 9 years was only found significant in the first threshold in the LSGOLII model indicating a higher probability of owning a single vehicle when children of this age range is present. Households might enjoy the extra flexibility that personal automobiles offer in terms of traveling with children (for example, dropping children off to daycare, school and/or participate in wide variety of leisure activities) and hence are more inclined towards owning more vehicles (Nolan, 2010).

With increase in the number of employed adults (both full-time and part-time) in households, the likelihood of owning multiple cars increases (same as Segment-1). The positive impact of part-time worker variable on the threshold demarcating single and two vehicles suggests an increased probability of single vehicle ownership. Similar to Segment-1, number of male and female adults have a positive association with fleet size decision of households. Interestingly, households with higher number of students had higher likelihood of owning more vehicles as well. With increase in number of retirees, households in Segment-2 have a higher likelihood of purchasing multiple cars. The result might be explained in light of the fact that retirees, who presumably live alone and have the time flexibility to take frequent leisure trips, are more likely to be dependent on cars for their mobility needs due to old age. However, increased number of young adults decreased the probability of high car ownership.

Our results indicated that households tend to own more vehicles when they live in a CT with increased number of driver commuters. There are two plausible explanations: these census tracts have low population density and are not well served by transit; also, the accessibility at the job locations via non-auto mode for these commuters are poor, thereby increasing the likelihood of owning more cars (Chen et al., 2008). Similar increasing impact of number of commuters as passengers on household vehicle fleet size was also observed. Intuitively, households residing in a census tract with more walk commuters tend to own less cars. Unlike segment-1, number of high-school degree certificate holders has a negative impact on household vehicle ownership level. In our analysis, total employed persons in the census tract positively impacted the second threshold that separated two and more than two cars ownership categories. The result suggest that two vehicles may be the upper bound on automobile ownership choice for the households when they are located in neighbourhood with increased number of employed persons. Westmount was the only borough variable that was found significant for Segment-2. Households in these boroughs had higher propensity of owning more vehicles. Given that it is one of the affluent on-island suburbs, the result makes intuitive sense.

The only transit attribute found significant was number of metro stations and length of bus lines. As expected, these variables have a negative effect on the probability of a household to own more vehicles. As the number of metro stops and bus coverage in terms of length increases at the place of residence, it is apparent that better transit service is available (Caulfield, 2012), hence obviating the need to have large fleet size.

Overall, we saw that the results for the LSGOLII and LSMNLII models offered very similar interpretations. The difference in the mathematical framework and the differences in the formulation of the two frameworks can lead to the minor differences we observe. However, the results from both models clearly underscore the importance of considering population heterogeneity through latent class models in the context of car ownership.

5.5.5 Validation Results

A validation experiment is also carried out in order to ensure that the statistical results obtained above are not a manifestation of over fitting to data. The validation analysis is conducted for the LSGOLII and LSMNLII models using a subsample of data (1972 records) that was set aside. Both aggregate and disaggregate measures of fit were computed. The results are reported in Table 5.5.

At the disaggregate-level, we computed the predictive log-likelihood which is computed by calculating the log-likelihood for the predicted probabilities of the sample (Eluru et al., 2008). In terms of disaggregate validation measures, the LSGOLII model consistently outperforms the LSMNLII model. At the aggregate level, we compare the predicted[12] and actual auto ownership level shares and compute the root mean square error (RMSE) as well as the mean absolute percentage error (MAPE) of the predicted shares. We compute these measures for set of full validation sample as well as specific sub-samples within that validation population – presence of children (0-4 years, 5-9 years, and 15-19 years), median income of census tract (medium and high), and full-time worker. Both LSGOLII and LSMNLII performed well at the aggregate level. However, the RMSE and MAPE values indicated that the predictive performance of the LSGOLII model is far superior to that of LSMNLII model for both full and sub-samples. Hence, there is enough evidence to suggest that LSGOLII performs significantly better in the empirical analysis compared to its unordered counterpart.

[12]The aggregated predicted probabilities of car ownership outcome k of households belonging to a particular segment s can be calculated using the following equation: $\frac{\sum_q P_{qs} \times [P_q(k)|s]}{Q}$ and the overall predicted share is obtained by summing these probabilities over s.

5.6 Summary and Conclusions

Latent class modeling is an elegant and tractable approach that accommodates for population heterogeneity relevant to the choice at hand (in our case, vehicle ownership). The current research examined alternative approaches (ordered and unordered) for incorporating population heterogeneity in the context of vehicle ownership. Towards that end, we estimated latent segmentation based ordered logit (both traditional ordered and generalized ordered response model) and latent segmentation based multinomial logit models using 2008 Origin-Destination (O-D) survey data of the Montreal region, Canada. The performance of the alternative frameworks was examined in the context of model estimation and validation (at the aggregate and disaggregate level) by using a host of comparison metrics. The results from the exercise illustrated the superiority of the generalized ordered framework in comparison with its unordered counterpart in modeling vehicle ownership decisions of households.

Other empirical findings of our study can be summarized as follows. The results support the hypotheses that there is preference heterogeneity and that the heterogeneity can be explained in part by the observable land use attributes – thereby implicitly capturing the residential self-selection bias in the vehicle fleet size decisions. The novel element added in this research is the identification procedure of the important variables affecting the class specific choice models. The preference order of the variables help us in better understanding the population segments identified. For example, in our analysis, model results revealed the existence of two population segments – households allocated to segment 1 are less likely to own zero cars (only 13%) whereas the households assigned to segment 2 are less likely to own 3 or more cars (only 2%). From the importance rating, we found that vehicle ownership decision of households in segment 1 is more likely to be governed by the monetary solvency, individual preference (age and gender), and mobility needs of their members. On the other hand, households in Segment-2 put more emphasis on surrounding land use attributes while making their decision regarding vehicle fleet choice. Based on these segment specific car ownership shares, variable means, and importance rating, we can characterize segment 1 as low density high car ownership segment (LHS) and segment 2 as high density low car ownership segment (HLS). The current research can be extended in several directions. Beginning with the context studied here, we can extend the analysis by allowing the choice-model coefficients to be randomly distributed. While this will bring additional computational complexity, but it would be an interesting exercise.

5.7 Link between Chapter 5 and Chapter 6

In this chapter, an extensive empirical comparison of alternative model frameworks in the context of vehicle ownership in the presence of population heterogeneity is conducted. The performance evaluation results (aggregate and disaggregate) provided evidence that latent generalized ordered logit framework is a promising tool to examine ordered variables such as vehicle ownership while accommodating for population heterogeneity.

In the previous two chapters, we used cross-sectional data for our analysis. As a result, only a snapshot of the fleet profile of households was captured, and not the evolution of the decision process. In Chapter 6, we move our analysis of vehicle fleet size further in another direction using an innovative cross-sectional data compilation technique. We stitch cross-sectional data from 1998, 2003, and 2008 Origin-Destination (OD) survey of Greater Montreal Area (GMA) and analyze it using variants of the generalized ordered logit (GOL) framework to incorporate the variances across different time points adequately.

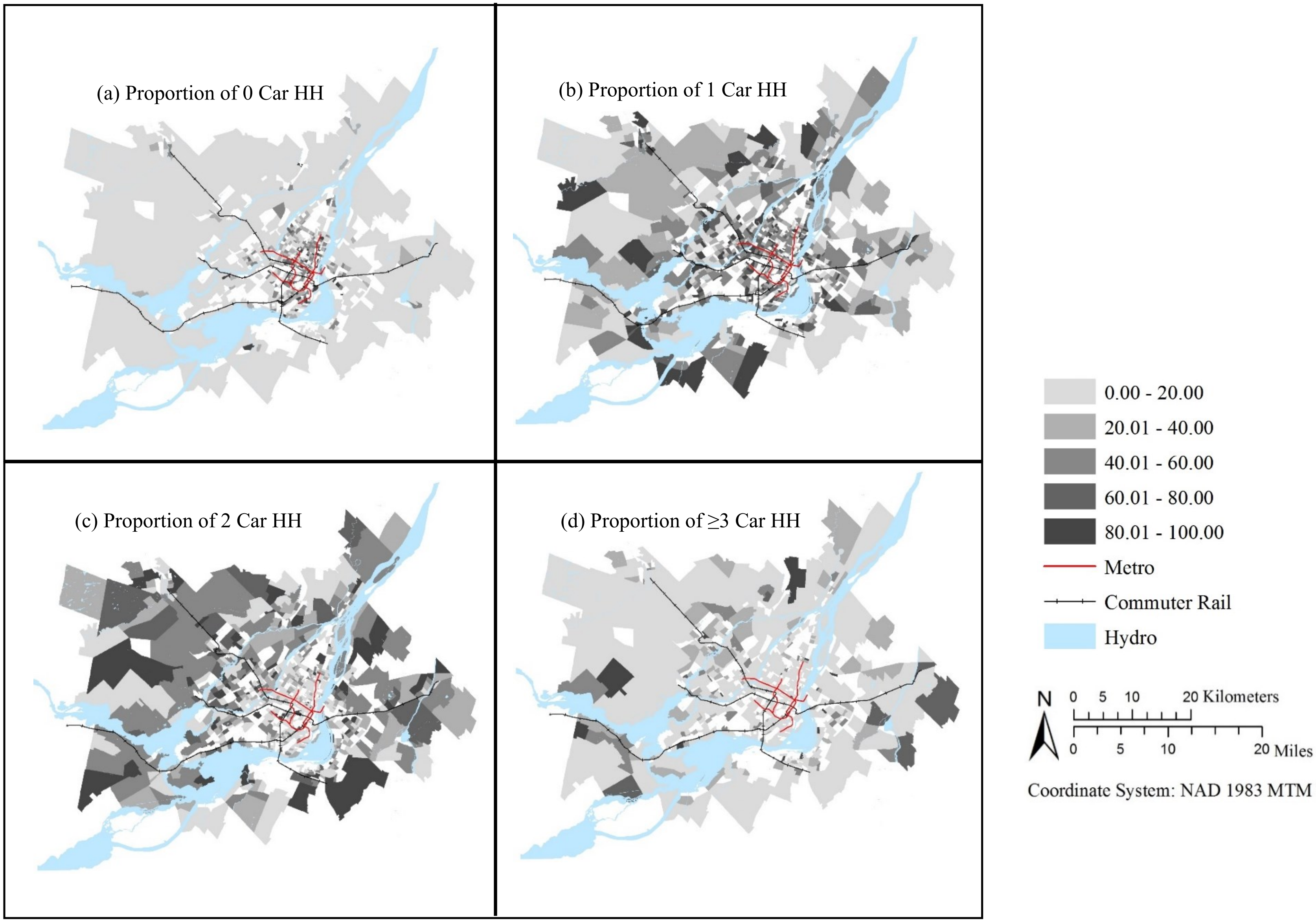


Figure 5.1 Proportion of Households (HH) in Different Vehicle Ownership Categories in TAZs

Table 5.1 Summary Statistics of Variables*

Variables	Frequency	%
Car Ownership Levels of Households		
0 Car	997	20.3
1 Car	2126	43.3
2 Cars	1409	28.7
≥ 3 Cars	381	7.8
Household Socio-demographics		
Number of adult males		
0	1199	24.4
1	3294	67.0
≥ 2	420	8.6
Number of full-time workers		
0	1740	35.4
1	1674	34.1
≥ 2	1499	30.5
Number of part-time workers		
0	4429	74.9
1	489	11.8
≥ 2	554	13.3
Number of children		
0	3357	68.3
1	675	13.7
≥ 2	881	18.0
Number of retirees		
0	3219	65.5
1	1094	22.3
≥ 2	600	12.2
Land Use Measures		
Income (CT level)		
Low (Less than 40K)	375	7.7
Medium (40K –80K)	2989	60.8
High (	1549	31.5
Sample size	4913	100
*The numbers in the table represent the percentage distribution of households in the sample for the OD years		

Table 5.2 Segment Characteristics and Mean Values of Segmentation Variables

Components		LSGOLII	
		Segment-1	Segment-2
Household share		0.51	0.49
Car ownership within each segment			
0 Car		0.13	0.25
1 Car		0.39	0.50
2 Cars		0.36	0.23
≥ 3 Cars		0.12	0.02
Mean Values of Land Use Variables in Each Segment			
	Overall Market Share	**Segment-1**	**Segment-2**
Outside Island of Montreal	0.47	0.70	0.23
Population Density	44.59	29.38	60.65
Total Area (1000's Ha)	0.74	1.24	0.21
Ln (# of rented HH in CT)	6.43	6.00	6.87
Older Suburb	0.06	0.01	0.11

Table 5.3 Importance Rating of Variables in the Segment Specific Choice Model

Rank	Variables in Segment 1	Variables in Segment 2
1	Ln (high school certificate holders)	Ln (number of employed persons)
2	Number of adult males	Ln (drive commuters)
3	Full time workers	Ln (passenger commuters)
4	Ln (trade certificate holders)	Ln (high school certificate holders)
5	Ln (transit commuters)	Ln (walk commuters)
6	Number of female adults	Full time workers
7	Ln (single and semi-detached households)	Number of male adults
8	Medium income (CT)	Number of female adults
9	Number of middle aged adults	Part time workers
10	Ln (bus length in the buffer)	Number of retirees
11	High income (CT)	Number of bus stops
12	Number of young adults	Number of young adults
13	Presence of children (15-19 years)	Number of students
14	Number of metro stops	Presence of children (5-9 years)
15	Ville-Marie	Presence of children (<4 years)
16	Part time workers	Westmount
17	Presence of children (5-9 years)	–

Table 5.4 LSGOLII Estimation Results (N=4913)

Segmentation Component				
Variables	**Segment 1**		**Segment 2**	
	Estimate	**t-stat**	**Estimate**	**t-stat**
Constant	–	–	-3.503	-4.260
Land use characteristics				
Household (HH) Location				
Outside Island of Montreal	–	–	-0.973	-4.361
Census Tract (CT) Characteristics				
Population Density	–	–	0.009	2.361
Total Area (1000's Ha)	–	–	-0.346	-2.737
Ln (# of rented HH in CT)	–	–	0.555	4.532
Older Suburb	–	–	1.638	2.160

Vehicle Ownership Components						
Variables	**Latent Propensity**	**Threshold 1**	**Threshold 2**	**Latent Propensity**	**Threshold 1**	**Threshold 2**
Constant	4.295 (5.401)	1.183 (10.602)	1.066 (14.598)	3.008 (3.423)	1.715 (7.844)	-2.848 (-1.270)
Household socio-demographic characteristics						
Presence of Children (0-4 years)	–	–	–	1.079 (5.300)	–	–
Presence of Children (5-9 years)	-0.417 (-2.196)	–	–	–	0.223 (2.797)	–
Presence of Children (15-19 years)	0.600 (2.742)	–	-0.290 (-3.441)	–	–	–
Number of Male Adults	2.868 (9.186)	0.184 (3.811)	–	0.813 (5.887)	–	–
Number of Female Adults	1.108 (7.677)	–	–	0.551 (3.838)	–	–
Full-time Workers	1.685 (5.582)	0.143 (2.929)	0.183 (4.308)	1.272 (10.927)	–	–

Part-time Workers	0.573 (3.188)	–	–	1.010 (4.953)	0.132 (1.812)	
Number of Young Adults	0.877 (5.775)	–	–	-0.693 (-7.067)	–	–
Number of Middle Aged Adults	0.718 (5.740)	–	–	0.349 (3.309)	–	–
Number of Students	–	–	–	0.449 (5.702)	–	–
Number of Retirees	–	–	–	0.825 (5.707)	–	–
Land use characteristics						
Ln (# of driver commuters in CT)	–	–	–	0.940 (6.061)	–	–
Ln (# of passenger commuters in CT	–	–	–	–	-0.117 (-2.431)	–
Ln (# of transit commuters in CT)	-0.307 (-3.710)	–	–	–	–	–
Ln (# of walk commuters in CT)	–	–	–	-0.293 (-3.222)	–	–
Ln (# of high school degree holders in CT)	1.146 (4.477)	–	–	-0.567 (-3.709)	–	–
Ln (# of trade certificate holders in CT)	-0.449 (-2.361)	–	–	–	–	–
Ln(# of single, semi-detached HH in CT)	–	0.033 (2.405)	–	–	–	–
Ln (Total employed adults in CT)	–	–	–	–	–	0.520 (1.753)
Median Income (Base: Low Income)						
Medium income (40-80K)	–	-0.149 (-2.605)	-0.120 (-1.849)	–	–	–
High income (Above 80K)	1.475 (4.960)	–	–	1.244 (5.934)	–	–
Boroughs (Base: Other Boroughs)						
Ville-Marie	2.254 (2.782)	–	–	0.332 (2.886)	–	–
Westmount	–	–	–	1.819 (3.452)	–	–
Transit attributes						

Number of bus stops	–	–	–	-0.010 (-2.964)	–	–
Number of metro Stations	-0.355 (-3.375)	–	–	–	–	–
Ln (Bus line length)	-0.051 (-2.637)	–	–	–	–	–

Notes:
"---"denotes the variable is insignificant at the 10% level
Threshold 1 is the threshold between one and two cars; Threshold 2 is the threshold between two and more than two cars
HH = household; CT = census tract

Table 5.5 Measures of Fit in Validation Sample (N=1972)

Disaggregate Level		
Summary statistic	**LSGOLII Predictions**	**LSMNLII Predictions**
Number of observations	1972	1972
Number of parameters	51	68
Log-likelihood at zero	-2733.77	-2733.77
Log-likelihood at sample shares	-2458.45	-2458.45
Predictive Log-likelihood	-1771.99	-1787.09
Predictive adjusted likelihood ratio index	0.26	0.25

Aggregate Level				
Vehicle Ownership Levels		**Actual shares**	**LSGOLII Predictions**	**LSMNLII Predictions**
0 Car		20.13	20.73	20.81
1 Car		43.97	44.34	44.27
2 Cars		27.54	27.10	27.32
≥ 3 Cars		8.37	7.83	7.59
RMSE		–	0.50	0.54
MAPE		–	2.97	3.54
Presence of Children (0-4 years)	0 Car	5.83	7.83	12.34
	1 Car	41.26	40.54	41.44
	2 Cars	44.66	44.10	38.82
	≥ 3 Cars	8.25	7.52	7.39
	RMSE	–	2.39	9.06
	MAPE	–	11.56	33.97
	Predictive Log-likelihood	–	-183.52	-192.75

Presence of Children (5-9 years)	0 Car	4.15	6.83	6.76
	1 Car	40.41	44.99	43.80
	2 Cars	43.52	41.95	42.35
	≥ 3 Cars	11.92	6.24	7.09
	RMSE	–	7.65	6.33
	MAPE	–	31.84	28.69
	Predictive Log-likelihood	–	-171.92	-176.68
Presence of Children (15-19 years)	0 Car	6.98	7.10	8.21
	1 Car	30.23	32.08	30.69
	2 Cars	37.21	36.61	38.99
	≥ 3 Cars	25.58	24.21	22.11
	RMSE	–	3.07	5.31
	MAPE	–	3.70	9.38
	Predictive Log-likelihood	–	-269.51	-279.56
Medium Income (40-80K)	0 Car	23.73	24.99	24.79
	1 Car	48.85	48.12	48.10
	2 Cars	21.44	21.29	21.77
	≥ 3 Cars	5.97	5.59	5.33
	RMSE	–	9.24	9.08
	MAPE	–	3.46	4.56
	Predictive Log-likelihood	–	-1102.07	-1115.30
High Income (Above 80K)	0 Car	4.83	5.12	5.25
	1 Car	35.94	37.28	37.54
	2 Cars	44.26	43.53	43.41
	≥ 3 Cars	14.98	14.08	13.80
	RMSE	–	5.39	6.62

	MAPE	–	4.38	5.78
	Predictive Log-likelihood	–	-532.23	-538.19
No Full-time worker	0 Car	36.07	36.78	37.38
	1 Car	51.11	50.81	50.21
	2 Cars	10.58	10.84	10.69
	≥ 3 Cars	2.23	1.57	1.72
	RMSE	–	3.76	6.03
	MAPE	–	8.62	7.32
	Predictive Log-likelihood	–	-597.96	-597.26

CHAPTER 6 ANALYSIS OF VEHICLE OWNERSHIP EVOLUTION USING PSEUDO PANEL ANALYSIS

6.1 Introduction

Longitudinal studies have a respected place in travel behavior research. Truth be told, such studies are still a rarity primarily due to data unavailability. The current study is motivated from the need to address this longitudinal data availability challenge. We take motivation from the fact that we might not have an abundance of longitudinal data but fortunately, large collections of cross-sectional data that are compiled over many years (for instance, the Origin Destination (O-D) survey data for Montreal or Quebec City or Sherbrooke which are collected every five year) do exist and efficient usage of this huge data resource combined with application of appropriate econometric modeling technique can be a possible solution to the issue.

Specifically, we intend to develop vehicle ownership frameworks employing cross sectional databases compiled over multiple time points and thereby investigate the factors affecting vehicle ownership and its evolutions in recent years in the Greater Montreal Area (GMA). Towards that end, three origin-destination (O-D) surveys from years 1998, 2003 and 2008 are utilized. The study approach is built on the generalized ordered logit (GOL) framework. Further, to incorporate the effect of observed and unobserved temporal effects, we specifically consider two versions of the GOL model – the mixed GOL model and the scaled GOL model. The two variants differ in the way they incorporate the influence of unobserved attributes within the decision process. We estimate both models and employ data fit comparison metrics to determine the appropriate model structure. The model specification is undertaken so as to shed light on how the changes to Montreal region across the study years and boroughs have affected household vehicle ownership.

6.1.1 <u>Contribution and Organization of the Chapter</u>

All the studies employing OR models ignore the potential impact of unobserved time specific attributes on the decision process. The studies that explore these unobserved effects (Dargay and Vythoulkas, 1999; Dargay, 2002; Nobile et al., 1997) employ either linear regression frameworks or multinomial probit (MNP) models. The applicability of linear regression and unordered approaches to study vehicle ownership is arguable as the vehicle ownership variable is an ordinal discrete variable. A more appropriate framework to examine this variable would be the OR framework. However, one important limitation of the OR models is that they constrain the impact of the exogenous variables to be monotonic for all alternatives. To overcome this issue, researchers have resorted to the unordered response (UR) models that allow the impact of exogenous variables to vary across car ownership levels (Bhat and Pulugurta, 1998; Potoglou and Kanaroglou, 2008a; Potoglou and Susilo, 2008). However, the increased flexibility from the UR models is obtained at the cost of neglecting the inherent ordering of the car ownership levels. The recently proposed GOL model relaxes the monotonic effect of exogenous variables of the traditional OR models while still recognizing the inherent ordered nature of the variable (Eluru et al., 2008). In fact, recent evidence comparing the performance of GOL model with its unordered counterparts (such as MNL, nested logit (NL), ordered generalized extreme value (OGEV), and mixed multinomial logit (MMNL)) has established the GOL model as an appropriate framework to study ordered variables (see Eluru, 2013; Yasmin and Eluru, 2013). Hence, in our study, we employ the GOL framework to study car ownership. To elaborate, we contribute to literature by employing two variants - the scaled GOL model (SGOL) and mixed GOL (MGOL) model.

Further, we study car ownership evolution in Montreal region using a comprehensive set of exogenous variables with a particular focus on land use and urban form characteristics. We also incorporate the impact of temporal changes to borough location on the choice process. As mentioned earlier, in addition to the observed attributes, the study also considers the impact of unobserved attributes on the decision process. In summary, the current study contributes to literature in two ways. First, *methodologically*, the study employs an approach to stitch together multiple year cross-sectional datasets to generate a rich pooled dataset that will allow us to study the evolution of vehicle ownership. Second, *empirically*, the study contributes to vehicle ownership literature by estimating the GOL models using a rich set of exogenous variables including household socio-demographics, transit accessibility measures, land use characteristics and observed and unobserved effects of the year of data collection (and their interaction with other observed variables).

The rest of the chapter is organized as follows. Section 6.2 provides detailed formulations of the econometric model frameworks used in the analysis. In Section 6.3, the data source for the empirical analysis and sample formation procedures are described. This section also contains some descriptive statistics of the sample used for model development. The empirical analysis and the policy simulation results are presented in Section 6.4. Section 6.5 concludes the chapter.

6.2 Econometric Framework

In this section, we briefly provide the details of the econometric framework of the models considered for examining vehicle ownership level evolution of households. For convenience, first we introduce the traditional ordered logit (OL) model, then discuss about the generalized ordered logit model (GOL), scaled generalized ordered logit model (SGOL), and finally present the mixed version of the generalized ordered logit (MGOL) model.

If we consider the car ownership levels of households (k) to be ordered,

$$y_q^* = \alpha' x_q + \varepsilon_q , \qquad y_q = k \qquad if\ \psi_{k-1} < y_q^* < \psi_k \tag{6.1}$$

where y_q^* is the latent car owning propensity of household q. y_q^* is mapped to the vehicle ownership level y_q by the ψ thresholds ($\psi_0 = -\infty$ and $\psi_k = \infty$) in the usual ordered-response fashion. x_q is a column vector of attributes (not including a constant) that influences the propensity associated with car ownership. α' is a corresponding column vector of coefficients and ε_q is an idiosyncratic random error term assumed to be identically and independently standard logistic distributed across households q. The probability that household q chooses car ownership level k is given by:

$$P_q(k) = \Lambda\left(\psi_k - \alpha' x_q\right) - \Lambda\left(\psi_{k-1} - \alpha' x_q\right) \tag{6.2}$$

where $\Lambda(.)$represents the standard logistic cumulative distribution function (cdf).

GOL is a flexible form of the traditional OL model that relaxes the restriction of constant threshold across population. The GOL model represents the threshold parameters as a function of exogenous variables (Srinivasan, 2002; Eluru et al., 2008). In order to ensure the ordering of observed discrete vehicle ownership levels $\left(-\infty < \psi_{q,1} < \psi_{q,2} < \cdots \dots \dots < \psi_{q,k-1} < +\infty\right)$, we employ the following parametric form as employed by Eluru et al. (2008):

$$\psi_{q,k} = \psi_{q,k-1} + exp(\gamma_{qk} + \delta'_{qk} Z_{qk}) \tag{6.3}$$

where, Z_{qk} is a set of explanatory variables associated with the k^{th} threshold (excluding a constant), δ'_{qk} is a vector of parameters to be estimated and γ_{qk} is a parameter associated with car ownership levels of households (k). The remaining structure and probability expressions are similar to the OL model. For identification, we need to restrict one of the δ'_k vectors to zero.

For both OL and GOL model, the probability expression of Equation 6.2, is derived by assuming that the variance in propensity over different car ownership levels across years is unity.

However, we can introduce a *scale parameter (λ)*, which would scale the coefficients to reflect the variance of the unobserved portion of the utility for each time point. The probability expression can then be written as:

$$P_q(k) = \Lambda\left[\frac{(\psi_k - \alpha' x_q)}{\lambda}\right] - \Lambda\left[\frac{(\psi_{k-1} - \alpha' x_q)}{\lambda}\right] \tag{6.4}$$

where λ is the parameter of interest and is equal to $exp\ (\sigma x_i)$ and x_i is the time elapsed variable. This yields the SGOL model. If the σ parameters are not significantly different from 0, the expression in equation (6.4) collapses to the expression in Equation (6.2) yielding either the OL or GOL model depending on the threshold characterization.

The mixed GOL accommodates unobserved heterogeneity in the effect of exogenous variables on household car ownership levels in both the latent car owning propensity function and the threshold functions (Srinivasan, 2002; Eluru et al., 2008). The equation system for MGOL model can be expressed as:

$$y_q^* = (\alpha' + \beta')x_q + \varepsilon_q \tag{6.5}$$

$$\psi_{q,k} = \psi_{q,k-1} + exp[(\delta'_{qk} + \theta'_{qk})Z_{qk}] \tag{6.6}$$

We assume that $\beta'$and θ'_{qk} are independent realizations from normal distribution for this study. The proposed approach takes the form of a random coefficients GOL model thus allowing us to capture the influence of year specific error correlation through elements of x_q and Z_{qk}. This approach is analogous to splitting the error term (ε_q) into multiple error components (analogous to error components mixed logit model). The parameters to be estimated in the MGOL model are the mean and covariance matrix of the distributions of $\beta'$and θ'_{qk}. In this study, we use the Halton sequence (200 Halton draws) to evaluate the multidimensional integrals (see Eluru et al., 2008 for a similar estimation process). In our analysis, x_q vector includes the year elapsed allowing us to estimate observed and unobserved variations with respect to time.

6.3 Data

The proposed models are estimated using data derived from the cross-sectional Origin-Destination (O-D) surveys of Greater Montreal Area (GMA) for the years 1998, 2003 and 2008. These surveys are conducted every five years and are the primary source of information on individual mobility patterns in the Montreal region. The survey data, provided by *Agence Metropolitaine de Transport* (AMT) of Quebec, was at the trip level. For the current research, data from each O-D year was aggregated at the household level which yielded three datasets with 67,225, 58, 962 and 68,132 household level data, respectively. From this database, for each year, 4,000 data records were randomly sampled. These three samples were pooled together to obtain a sample of 12,000 records for model analysis.

Car ownership levels were classified as no car, one car, two cars, and three or more cars. The dependent variable was truncated at three because the number of households with more than three automobiles was relatively small in the dataset. Table 6.1 provides a summary of the characteristics of selected socio-demographic and land use variables used in this study. The distribution of auto ownership levels by year (1998-2008) in the estimation samples indicate that in each of the three survey years, percentage of households owning one car accounted for the largest share. We can also see that proportion of zero car owning households increased somewhat in 2008 compared to 1998. On the other hand, a slight decrease could be observed in the proportions of households owning single and two cars. Interestingly, there is a noticeable increase in the number of households owning more than two cars in 2008 (7.5%). Some other salient characteristics of the sample are: in 1998, one-half of the households belonged to low income census tracts, but in recent years, more households were residing in medium and high-income census tracts. Over the years, about two-thirds of the households had at least one full time employed adult and zero students, more than 10 percent had at least one part-time employed person and more than 50 percent had two or more license holders. As expected in a North American city, there is a gradual increase in the number of retirees in the households.

6.4 Empirical Analysis

6.4.1 <u>Variables Considered</u>

In the current study, a comprehensive set of exogenous attributes were considered to study vehicle ownership levels. The independent variables can be broadly classified into four categories: (1) household socio-demographic characteristics, (2) transit accessibility measures (3) land use characteristics, and (4) temporal variables. *Household socio-demographic* variables that were employed in our analysis included number of employed adults (full-time and part-time), number of males, average age of the household members, presence of children of different ages, number of retirees, number of students and number of licensed drivers. The *transit accessibility measures* considered, as a proxy for ease of transit accessibility and level of service of alternative modes, (within 600m buffer[13] of household residential location) were: bus stops, commuter rail stops, metro stops, length of bus line (km), length of commuter rail line (km), and length of metro line. In order to assess the impact of different *land use characteristics* on car ownership, the following land use variables were considered in our study: residential, commercial, government and institutional, resource and industrial, park and recreational, open and water area. Moreover, average distance of work location from the households, population density, and the median income of households in the census tract (CT) based on residential location were also included. Further, we introduced location specific (borough indicators) variables to examine the degree of influence exerted by the area of residence on household car ownership levels. These variables are expected to capture attributes of household's activity travel environment as well as the utility/disutility of automobile maintenance and operation in particular areas. In terms of *temporal variables*, we introduced a variable called "time elapsed from 1998" which is the time difference between the most recent O-D survey years (2008 and 2003) from the base survey year (1998). Both linear and polynomial effects of the time elapsed variable were tested. Moreover, interaction of exogenous variables with the time elapsed variable (linear and polynomial) were utilized to control for time varying variable effects. As a result, it would be possible to apply the developed models for future year scenarios. The final specification was based on a systematic process of removing statistically insignificant variables and combining variables when their effects were not significantly different. The specification process was also guided by prior research, intuitiveness and parsimony considerations.

6.4.2 <u>Estimation Results</u>

In this research, we considered three different model specifications of the GOL model. These are: (1) GOL (2) SGOL, and (3) MGOL. As explained earlier, all of these models are generalized

[13] Buffers were established around household geocoded locations with 600m radius. In earlier literature, the acceptable walking distance to transit stops and stations is often assumed to be 400m (Larsen et al., 2010). Hence, we employed a slightly larger buffer than the 400m to allow for the low-density developments in Canadian cities that might require people to walk further to reach transit stations from their households.

versions of the standard OL model. After extensive specification testing, the final log-likelihood values (number of parameters) at convergence of the GOL, SGOL and MGOL models were found as: -8647.92 (49), -8646.05 (50) and -8556.61 (53), respectively. The performance of the models was tested using Log-likelihood Ratio test, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) measures. The AIC (BIC) values for the final specifications of the GOL, SGOL and MGOL models are 17393.84 (17756.08), 17392.10 (17761.73), and 17219.22 (17611.04), respectively. The improvement in the data fit clearly demonstrates the superiority of the MGOL model over its other counterparts. Hence, in the following sections, we discuss results of the MGOL model only.

The model estimation results are presented in Table 6.2. Note that there are three columns in the table. The first column corresponds to the car ownership propensity, the second column corresponds to the first threshold that demarcates the one and two car ownership categories and the third column corresponds to the second threshold that demarcates the two and more than two car ownership categories. In the following presentation, we discuss both variable effects and unobserved heterogeneity effects on the latent car ownership propensity and the two thresholds. The effect of each category of variables on the thresholds provides a sense of how the probability of car ownership in specific ownership categories is affected.

6.4.2.1 Constants

The constant variables do not have any substantive interpretation. Within the set of constant parameters, the impact of the time elapsed variable was examined. The effect of the variable was found significant for both propensity and the second threshold that separates two car ownership level from three or more cars ownership level. The effects indicate that households in recent times are more likely to have an increased fleet size. The findings confirm our observations of an increase in households with at least two cars in the data.

6.4.2.2 Household Demographics

Increased number of male household members increases the likelihood of multiple car ownership of households and the gender effect is found to be highly significant. For obvious reasons, presence of children in households significantly affects their fleet size decision. In particular, we found that households with children between 5 to 9 years have a higher propensity of possessing multiple vehicles presumably owing to the increased travel needs, such as, chauffeuring them to and from daycare and/or school. Presence of young children (aged between 10 to 14 years) in the household also have similar positive effect. The result is intuitively understandable since children of this age have diversified activity requirements and are mostly dependent on the adult householders for their mobility which might result in additional vehicle purchase. The presence of teenaged children (15-19 years of age) do not have a direct effect on propensity, however, a positive impact of the interaction term between the presence of 15-19 year old children and elapsed time was observed in our analysis. Moreover, the effect of the variable on the threshold indicates increased likelihood of single vehicle ownership. A plausible reason for the smaller fleet size might be that teens of this age can travel by themselves, unaccompanied by an adult or peer and are soon to move out of the house.

Our results underscore the increased latent propensity of owning multiple vehicles by middle aged households (average age of householders 30 to 60 years). The effect of this variable is also significant for the threshold demarcating two and more than two vehicles. The negative sign of the coefficient in the threshold indicates higher likelihood of owning more than two vehicles. As expected, households with more number of full time employed adults are more likely to have higher levels of vehicle ownership; an indicator that these households have greater mobility needs complemented by enhanced buying capability (Kim and Kim, 2004; Bhat and Pulugurta, 1998; Potoglou and Kanaroglou, 2008a). Interestingly, we also observe that with elapsed time, the impact of full time workers on vehicle ownership levels is reducing. The result is quite encouraging for policy makers highlighting that in the recent years, growing environmental consciousness and increased inclination towards using transit might actually be contributing towards lowering vehicle ownership levels. Similar to full time workers, increase in the number of part time workers also increases household's propensity to own multiple cars. The latent propensity is found to be normally distributed with a mean of 0.3719 and standard deviation of 0.6510, suggesting that in 28.43% of the households, an increase in part-time worker has a positive impact on car ownership. With increase in number of retirees, households have a higher likelihood of purchasing more cars. Retirees live primarily in single-person households (Nobis, 2007) and hence, they are more likely to be dependent on cars for their mobility needs.

The negative impact of number of students on the propensity indicates that households with higher number of students are less inclined to own several cars. It is expected because households with more students would have increased budget constraints and hence, would be less inclined to own cars. Moreover, students may share their activities with friends and other household members that might further reduce the need for owning multiple cars (Vovsha et al., 2003). The results associated with the number of licensed drivers (surrogate for potential drivers in the household) reflect the anticipated higher probability of households owning multiple cars. The effect of the variable on the thresholds is quite interesting. The variable exhibits significant impact on both the thresholds. It is very hard to establish the exact impact of these threshold parameters as their impact is quite non-linear and is household specific. The GOL model with its flexibility in allowing for such variations across the households provides a better fit to the observed vehicle ownership profiles. We also found that when immobile persons are present, households become less likely to own higher number of cars.

6.4.2.3 Transit Accessibility Measures

The results corresponding to transit accessibility measures highlight the important role of public transit in Montreal. Increase in the number of bus stops as well as bus and metro line length within the household buffer zone negatively impact household's propensity to own cars. The result lends support to the concept that increased transit access and high quality of transit service can significantly reduce the number of automobiles owned by households (Ryan and Han, 1999; Bento et al., 2005; Kim and Kim, 2004; Cullinane, 2002). Of particular interest are the effects of number of bus stops and metro line length. The impact of number of bus stops on fleet size is normally distributed with a mean of -0.0324 and standard deviation of 0.0473. The effect of metro line on vehicle ownership propensity is also normally distributed with a mean of -0.2939 and standard deviation of 0.6368. It suggests that the impact of number of bus stops and metro line varies substantially across the various parts of the urban region. The distribution measures indicate that for approximately 25% of households number of bus stops have a reduced propensity for vehicle ownership while the metro variable has reducing effect for 32% of households.

6.4.2.4 Land Use Measures

It is evident from previous literature that income is one of the most influential factors affecting household's decision regarding their vehicle fleet size. In our analysis, household income was unavailable to us. However, to address the unavailability we employed census tract median income as a proxy measure for the affluence of households. From our analysis results, we find that households living in medium income areas have a stronger preference to have more cars. The result is in close agreement with the findings of previous literature (Karlaftis and Golias, 2002; Li et al., 2010). Interestingly, we also observe that with elapsed time, the impact of living in medium income census tract on vehicle ownership levels is reducing. Location of households in highly advantaged areas does not have any effect on the vehicle ownership propensity, however, its impacts on threshold parameterization are relatively complex. It has a negative impact on the threshold between one and two cars and a positive impact on the threshold between two and more than two cars. From the results of the interaction of high income with time elapsed variable, we observe that with time, high income households are becoming more inclined towards owning two cars and less inclined to have a fleet size of more than two cars.

As expected, when distance between household and work location increases, households have a higher likelihood of owning multiple vehicles and the effect is getting stronger with passing time. This is perhaps the consequence of the fact that when home and work locations are far apart, car ownership becomes a necessity since driving appears to be the only convenient and reliable mode to reach work destination. Our results indicate that households in census tract areas with increased commercial as well as government and institutional land use are less likely to have multiple cars. When households are located in such areas with increased heterogeneous land use mix, their members have the option to easily access many activities and amenities by walking or biking in addition to riding transit, thereby minimizing their need to procure and use cars (Cervero and Kockelman, 1997; Hess and Ong, 2002).

In our analysis, in addition to the above land-use measures we considered a host of borough variables. Of these variables, some regions exhibited distinct car ownership profiles across the years. These include Ville-Marie (VM), Cote-des-Neiges (CDN), and Plateau-Mont-Royal (PMR). These boroughs represent medium to high dense neighbourhoods around the downtown region with good transit accessibility in general. We find that the impact of all three of the borough dummies on vehicle owning propensity of households is negative and significant, indicating that households in these areas tend to have lower automobile ownership. The interaction effects of the VM and CDN boroughs with the time elapsed variable showed similar in magnitude positive impacts. It is suggesting that the trend of reduced propensity is diminishing with passing time. Interestingly, VM borough also has a negative impact on the second threshold meaning an increased tendency of households to own more than two cars which tends to increase in recent years. These two results involving VM borough suggest that the vehicle ownership is likely to be in the extremes in the region (either 0 or ≥3). The local agencies of these boroughs need to investigate the reasons for this dramatic change. The impact of CDN and PMR boroughs are normally distributed suggesting the presence of unobserved factors influencing the vehicle fleet size decision of households living in these areas. More specifically, the distribution measures indicate that for approximately 23.5% of households located in CDN borough have a reduced propensity for vehicle ownership whereas living in PMR borough has reducing effect for 33% of households. Given that PMR borough has emerged as one of the most environmentally conscious neighbourhoods in Montreal, the results are not surprising. In fact, the borough policies (such as parking cost mechanisms, altering traffic flow patterns) serve as a case study for policy makers interested in reducing vehicle ownership.

6.4.3 Elasticity Effects and Policy Analysis

The exogenous variable coefficients do not directly provide the magnitude of impacts of variables on the probability of each car ownership levels. Moreover, the impacts of coefficients of the MGOL framework might not be readily interpretable due to the interactions between propensity and thresholds. Hence, to provide a better understanding of the impacts of exogenous factors, we compute two measures: (1) the aggregate level elasticity effects and (2) disaggregate level changes in vehicle ownership levels.

6.4.3.1 Aggregate Elasticity Effects

The elasticity computation results are presented in Table 3. Following observations can be made based on the elasticity results. First, the results illustrate that possession of license, employed status (full-time and/or part-time), and location of the household in the Ville-Marie borough and economically advantaged (medium and/or high income census tracts) areas are the most important variables resulting in higher household car ownership levels. Second, in terms of vehicle ownership reduction, presence of teenaged children (15-19 yeas), increased number of students and location of household in CDN borough contribute significantly. Third, of the three transit accessibility measures, number of bus stops and length of bus lines have a greater impact on reducing vehicle ownership levels. Finally, we observe that socio-demographic variables are likely to have more significant impact on vehicle ownership levels compared to the impact of transit and land use attributes.

6.4.3.2 <u>Disaggregate Level Changes</u>

In this section, we focus on the borough level variables (VM and PMR) to illustrate the variation in vehicle ownership probabilities across the years. Towards this purpose, we consider synthetic households (SH1 – SH4) with certain attributes and generate the probability profiles by changing the attributes for the household.

The first household (SH1) is a two person household located in low income area comprised of a young male and a young female adult who are students and do not possess a driving license. For this type of household, the probability of being carless is the highest in 1998 and 2003 (ranging from 64-71%) which is expected (see (a) in Figure 1 and 2). Interestingly, the probability drops to 46% in 2008. The probability of zero car ownership for PMR borough highlights the increase of such households whereas for the VM borough the trend is reversed particularly for 2008.

The second household (SH2) is similar to HH1, except that the male householder is a full-time worker and holds a driving license. Also, a toddler (0-4 years of age) is present in the household. The status of the female member was unchanged. As we can see, with employment and driver license, the probability of zero car ownership drops down drastically. For such households we see that VM borough has larger probability for one car in 1998 and 2003 (see (b) in Figure 1). However, for 2008, these households have higher likelihood of owning two cars. On the other hand, for the PMR region, the most likely outcome for the household is to own one car (see (b) in Figure 2).

The third household (SH3) is formed by changing the employment status of the female member into a part-time worker with a driving license from HH2. Also, the household resides in a medium income census tract area. In VM borough, the vehicle ownership shares vary substantially for the household across the three years (see (c) in Figure 1). In the PMR borough, the probability plots indicate that for all years, the probability of owning two cars is the highest (60-65%) (see (c) in Figure 1).

The fourth and the final synthetic household (SH4) was formed by changing the employment status of the female adult of HH3 into full time worker as well as changing their age from young to middle age. Also, the child member was considered to be between 5-9 years. For

VM borough, the household is more likely to own three or more cars in 1998 and 2003 while two cars in 2008 (see (d) in Figure 1). In PMR borough, the household fleet is more likely to be composed of either two or more than two cars (see (d) in Figure 2).

6.5 Summary and Conclusions

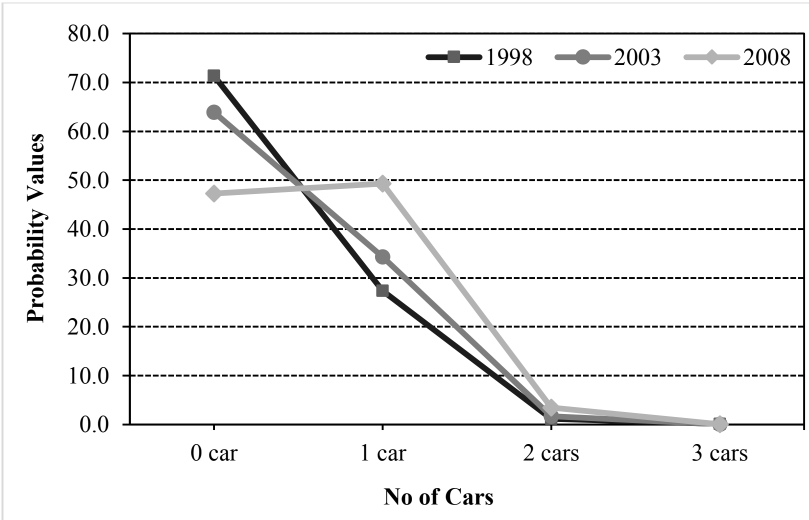
The current study examines vehicle ownership evolution in the Greater Montreal Area (GMA), Canada using cross sectional databases compiled over multiple time points. The study approach is built on the GOL framework that relaxes the restrictive assumption of the traditional OL model. Further, to incorporate the effect of observed and unobserved temporal effects, we consider two variants of the GOL model – the mixed GOL model and the scaled GOL model. After extensive specification testing, we found that the MGOL performed better than its other counterparts. The empirical model specification was based on a rich set of exogenous variables including household socio-demographics, transit accessibility measures, land use characteristics, and temporal factors. Further, observed and unobserved effects of the elapsed time from the base year (1998) of data collection (and their interaction with other observed variables) are explicitly considered in our analysis enabling us to examine trends in variable impacts across the years.

In accordance with the existing literature, socio-demographic variables were found to be an important predictor of automobile ownership of households. Our results also confirmed that the impact of some socio-demographic variables varied with time. For instance, we observed that in recent years, the impact of full time workers on vehicle ownership levels has been reducing. The result is quite encouraging for policy makers highlighting that in the recent years, growing environmental consciousness and increased inclination towards using transit might actually be contributing to lower vehicle ownership levels. In fact, the results corresponding to transit accessibility measures highlighted the important role of public transit in Montreal. The number of bus stops, and increase in bus and metro line length within the household buffer zone negatively impacted household's propensity to own cars. Since households tended to own more cars when they lived farther away from the work location, focusing on establishing good network connections between place of residence and place of work might reduce reliance on cars for day-to-day commute.

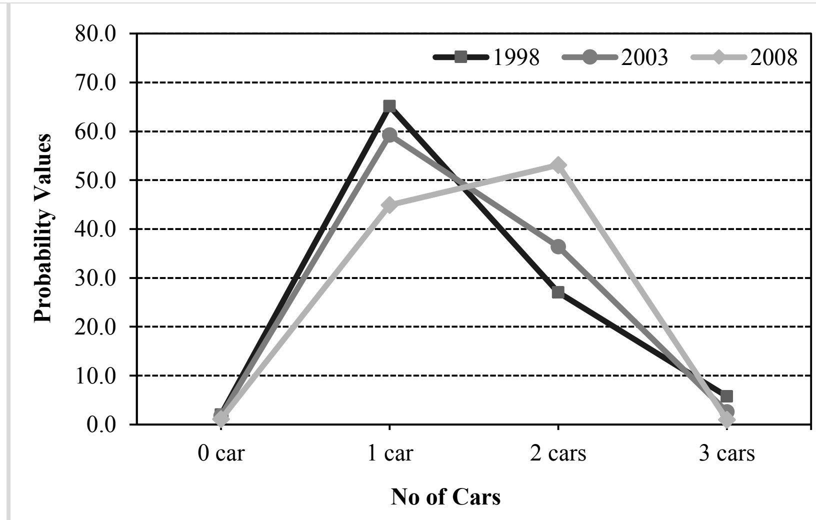
In our analysis, the boroughs which exhibited significant impact on car ownership include Ville-Marie, Cote-des-Neiges, and Plateau-Mont-Royal. Specifically, Ville–Marie borough transitioned from a negative propensity for car ownership towards a positive car ownership propensity from 1998 to 2008. The local agencies of this borough need to investigate the reasons for this dramatic change in such dense neighbourhood. In fact, they might need to review the borough policies such as parking cost mechanisms and/or altering traffic flow patterns with congestion pricing or implementation of more one way streets. In fact, combining different policies with information and advertising campaigns that promote more sustainable transport choices can help to bring about behavioral change and discourage unnecessary car use and in the long run, the ownership of multiple cars.

6.6 Link between Chapter 6 and Chapter 7

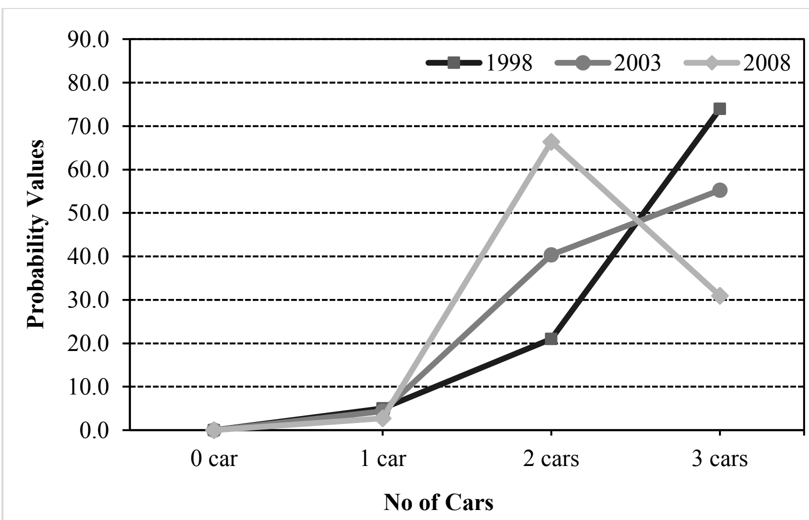
In the previous chapters, our research focus were on the first two dimensions of the vehicle decision process – vehicle acquisition budget allocation and vehicle fleet size. Vehicle usage is another equally important part of the decision hierarchy, particularly the short term usage decisions. Since, short term vehicle usage decisions have direct implications for the fuel consumed and emissions generated. Therefore, in the next chapter, we analyze vehicle type choice preferences of households along with three other choice dimensions using data from the Quebec City Travel and Activity Panel Survey (QCTAPS).



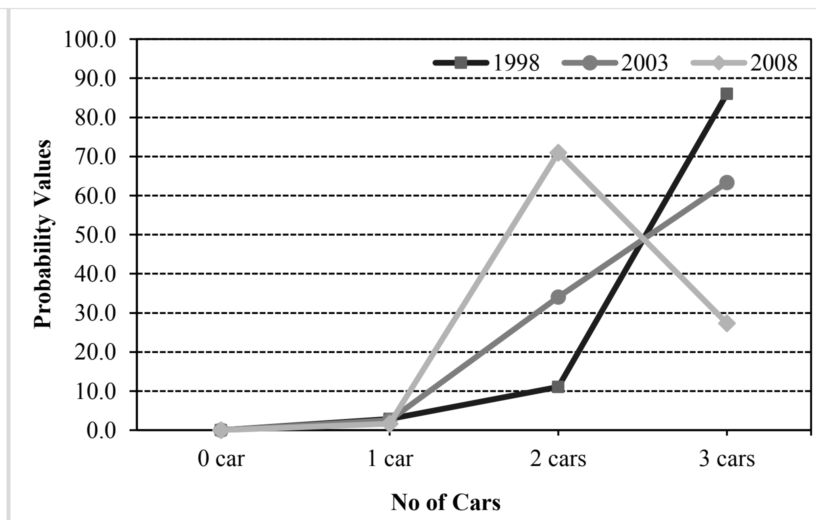
(a) SH1



(b) SH2

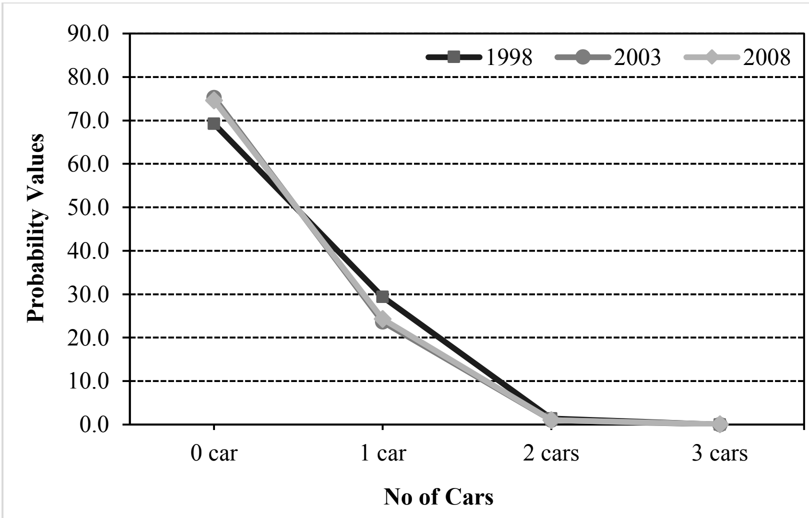


(c) SH3

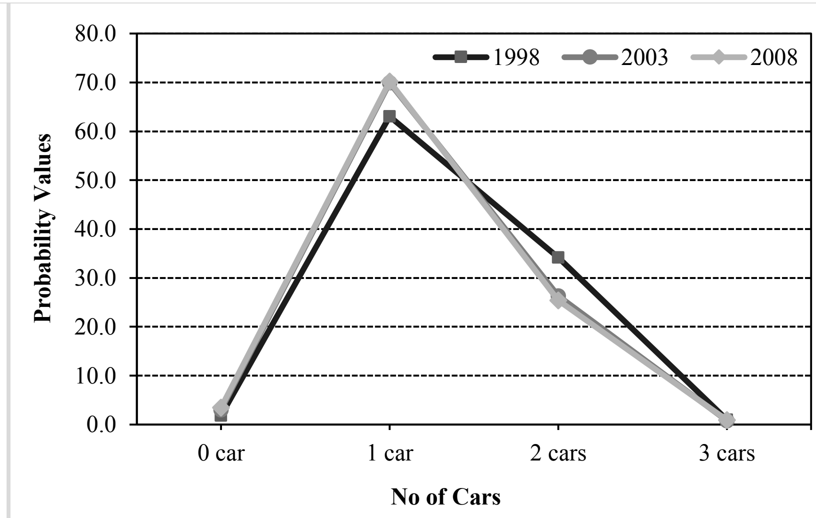


(d) SH4

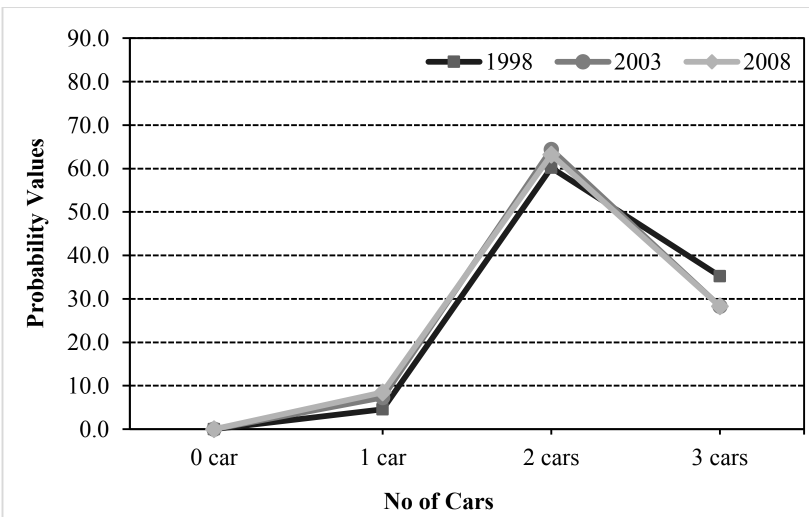
Figure 6.1 Evolution of Car Ownership Levels across Years for Artificial Households in Ville-Marie Borough



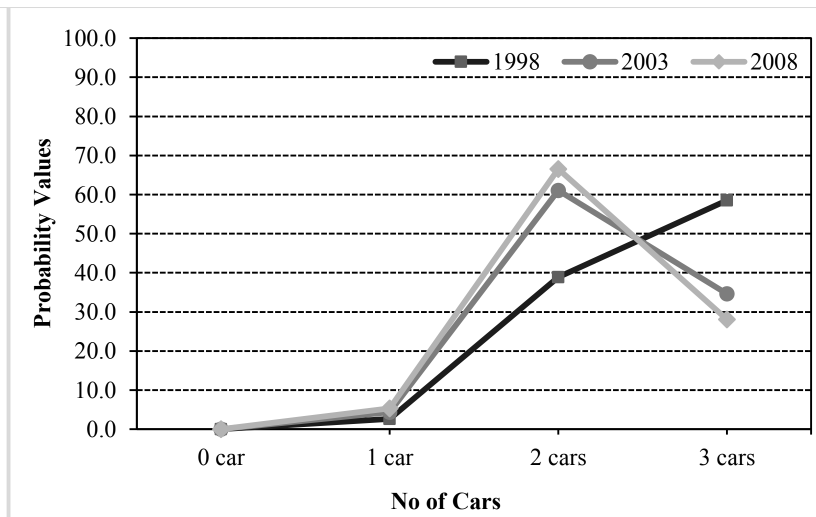
(a) SH1



(b) SH2



(c) SH3



(d) SH4

Figure 6.2 Evolution of Car Ownership Levels across Years for Artificial Households in Plateau-Mont-Royal Borough

Table 6.1 Summary Statistics of Socio-demographic and Land Use Variables

Variables	OD Years[*]		
	1998	2003	2008
Car Ownership Levels of Households			
0 Car	19.5	19.0	21.1
1 Car	45.2	44.5	42.8
2 Cars	29.7	30.1	28.6
≥ 3 Cars	5.7	6.5	7.5
Household Socio-Demographics			
No of Males			
0	29.5	34.4	36.1
1	33.6	33.3	33.1
≥ 2	36.9	32.3	30.8
No of Middle Aged Adults			
0	59.5	56.9	51.4
1	22.2	23.1	26.5
≥ 2	18.3	20.0	22.1
Number of Full-time Employed Adults			
0	31.6	32.6	36.2
1	38.5	37.9	33.6
≥ 2	29.9	29.5	30.2
Number of Part-time Employed Adults			
0	88.4	89.4	89.8
1	10.8	10.0	9.5
≥ 2	0.8	0.6	0.7
Number of License Holders			
0	11.7	11.6	13.4
1	33.4	33.5	32.6
≥ 2	54.9	54.9	54.0
Number of Students			
0	62.9	64.7	68.0
1	18.4	18.2	16.0
≥ 2	18.7	17.1	16.0
Number of Retirees			
0	75.2	72.7	64.3
1	15.4	18.1	23.2
≥ 2	9.4	9.2	12.5
Land Use Measures			
Income *(CT level)*			
Low (Less than 40K)	51.2	40.6	33.9
Medium (40K – 80K)	47.5	54.1	57.5
High (Above 80K)	1.3	5.3	8.6
Sample size	4000	4000	4000

[*] *The numbers in the table represent the percentage distribution of households in the sample for the OD years*

Table 6.2 MGOL Estimation Results (N=12000)

Variables	Latent Propensity		Threshold between One and Two Cars, ψ_2		Threshold between Two and Three or More Cars, ψ_3	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant	2.5475	16.708	1.2326	44.900	1.4529	17.385
Time Elapsed	0.0526	3.385	---	---	-0.0241	-3.955
Household Socio-Demographics						
No of Males	0.2263	6.220	---	---	---	---
Presence of Children						
5-9 years	0.3549	3.688	---	---	---	---
10-14 years	0.8062	4.266	0.0625	2.258	---	---
15-19 years	---	---	0.0430	2.747	---	---
15-19 years * Time elapsed	0.0310	2.025	---	---	---	---
Middle Aged Adults	0.1052	1.864	---	---	-0.1063	-3.481
Full-time Working Adults	0.4974	8.099	---	---	---	---
Full-time Working Adults* Time elapsed	-0.0385	-3.093	-0.0053	-2.978	0.0144	4.745
Part-time Working Adults						
Mean	0.3719	4.909	---	---	---	---
Standard Deviation	0.6510	4.101	---	---	---	---
No of Retirees	0.4411	5.856	0.0389	2.885	---	---
No of Seniors	---	---	0.0497	4.927	---	---
No of Students	-0.2903	-5.349	---	---	---	---
No of License Holders	4.0030	30.213	0.2921	20.756	-0.0965	-2.701
Presence of Immobile Persons	-0.2844	-5.395	---	---	---	---
Transit Accessibility Measures						
No of Bus Stops						
Mean	-0.0324	-8.015	---	---	---	---
Standard Deviation	0.0473	6.782	---	---	---	---
Length of Bus Lines (km)	-0.0063	-3.219	---	---	---	---
Length of Metro Lines (km)						

Mean	-0.2940	-5.640	---	---	---	---
Standard Deviation	0.6368	6.905	---	---	---	---
Land Use Measures						
Income (Base: Low Income)						
Medium Income (40K-80K)	0.5425	6.813	---	---	---	---
Medium Income * Time elapsed	-0.0326	-2.355	---	---	---	---
High Income (Above 80K)	---	---	-0.2849	-4.895	0.3460	2.721
High Income * Time elapsed	---	---	0.0255	3.670	-0.0426	-2.599
Ln (Distance to work)	0.0812	2.953	---	---	0.0475	3.844
Distance to work*Time Elapsed	0.0010	2.231	---	---	---	---
Type of Land Use						
Commercial (KM^2)	-1.9289	-3.950	---	---	---	---
Government and Institutional (KM^2)	-1.5299	-4.261	---	---	---	---
Population Density* Time elapsed	-0.1047	-6.149	---	---	---	---
Boroughs						
Ville-Marie	-1.0289	-2.984	---	---	-0.6569	-2.054
Ville-Marie * Time Elapsed	0.1293	2.569			0.0935	2.380
Cote-des-Neiges	---	---	---	---	---	---
Mean	-1.1942	-3.982	---	---	---	---
Standard Deviation	1.6522	5.145	---	---	---	---
Cote-des-Neiges * Time Elapsed	0.1233	3.219	---	---	---	---
Plateau-Mont-Royal						
Mean	-0.9257	-3.700	---	---	---	---
Standard Deviation	2.1003	5.686	---	---	---	---
Log-likelihood at sample shares, LL (c)	-14641.984					
Log-likelihood at convergence, LL (β)	-8556.612					
Number of observations	12000					

Table 6.3 Elasticity Effects

Variables	0 Car	1 Car	2 Cars	≥ 3 Cars
Household Demographics				
No of Males	-5.750	-3.039	5.214	14.171
Presence of Children				
5-9 years	-8.966	-4.866	8.291	22.222
10-14 years	-19.700	-1.644	11.001	21.620
15-19 years	0.000	6.686	-5.780	-18.140
Middle Aged Household	-2.745	-1.389	-2.544	28.463
Full-time Working Adults	-12.273	-6.790	10.888	33.282
Part-time Working Adults	-8.070	-5.437	6.075	32.973
No of Retirees	-10.953	-0.033	5.394	9.456
No of Students	7.825	3.710	-7.209	-16.020
No of License Holders	-74.474	-17.800	28.701	213.700
Presence of Immobile Persons	7.576	3.704	-6.962	-16.313
Transit Accessibility Measures				
No of Bus Stops	0.995	0.270	-0.757	-1.425
Length of Bus Lines (km)	0.568	0.166	-0.447	-0.827
Length of Metro Lines (km)	0.983	-0.441	-0.111	0.401
Land Use Measures				
Income				
Medium Income (40K-80K)	-14.250	-7.469	13.864	30.578
High Income (Above 80K)	0.000	-35.925	46.742	28.600
Distance to work	-0.203	-0.228	0.691	-0.937
Land Use Type				
Commercial	0.520	0.045	-0.317	-0.467
Government and Institutional	4.063	0.723	-2.761	-4.797
Boroughs				
Ville-Marie	30.205	11.730	-48.560	46.054
Plateau-Mont-Royal	34.489	1.449	-26.348	2.572
Cote-des-Neiges	40.656	6.617	-31.392	-27.454

CHAPTER 7 A JOINT ECONOMETRIC ANALYSIS OF ACTIVITY FLEXIBILITY, VEHICLE TYPE CHOICE AND PRIMARY DRIVER SELECTION

7.1 Introduction

Flexibility indicators help us gain a better understanding of the process of prioritization of activities that precede the actual execution of activities/trips. In other words, it helps us to understand "how activity patterns are derived" (Doherty, 2006). While it might appear that considering spatial and temporal flexibility might not be readily a choice – it can be perceived as a broad grouping of various activities in space and time. The current effort builds on earlier research that has suggested the use of "more salient features of activity" such as measures of spatial and temporal flexibility rather than being limited to the use of generic activity types where the flexibility of each activity is typically assumed to be fixed (for example, work is routine and shopping is impulsive) (Doherty, 2006). The current study explores interconnectedness of the flexibility of activities in space and time with the short term vehicle type choice and primary driver allocation.

7.1.1 Earlier Work

Of the four activity travel choices under consideration, vehicle type choice has received significant attention. Broadly, vehicle type choice studies can be classified into two major categories: (1) long-term and (2) short-term. The relevant studies for our study are the short term studies. For example, Konduri et al. (2011) and Paleti et al. (2012) have explicitly modeled vehicle type choice in tour-based models. Both of these two studies used mixed multidimensional choice model systems to better understand the complex relationship between different tour attributes (e.g. tour length, tour complexity) and the type of vehicle used to undertake the tour by individuals in a household. At the activity level, Faghih-Imani et al. (2014) applied mixed multiple discrete continuous extreme value (MMDCEV) framework to examine daily vehicle type and usage decisions while incorporating the influence of activity type and accompaniment type choices.

Research efforts concerning the effect of perceived flexibility of activities are comparatively fewer in number. Recently, researchers examined how the trips and activities are considered and adopted for execution, i.e. individual's perception of activity attribute and its impact on activity scheduling. For instance, Mohammadian and Doherty (2005) reported that temporally and spatially flexible activities are more likely to be impulsive or near-impulsive since they need less time to plan. In a later study (Mohammadian and Doherty, 2006), the authors modeled the duration of time between planning and execution of pre-planned activities using the same dataset. The findings of these two studies suggested that in addition to conventional activity and individual attributes, flexibility/fixity of activities plays an important role in the choice of activity-planning sequence. Based on their findings, the authors alluded to a possible interdependency between spatio-temporal flexibility and activity-travel attributes.

Individual's perception of spatial and temporal flexibility of activity was investigated by Miranda-Moreno and Lee-Gosselin (2008) and Lee-Gosselin and Miranda-Moreno (2009) using data from Quebec City, Canada (same dataset explored in our current study). In the first study, they explored the activity travel patterns of baby-boomers to find out whether they lived lives that are highly routine or flexible. In the latter study, they examined the impact of information and communication technologies (ICT), on the frequency of different temporally and spatially flexible categories of the executed out-of-home activities. They reported that access to mobile phones was associated with the propensity to pre-arrange activities both in time and in space, while internet was significantly and negatively associated with the number of habitual activities, again in time and in space.

Finally, the primary driver choice has received attention more recently in travel behavior literature. Households acquire different vehicles to satisfy various transportation needs while accommodating for preferences of the household members. In multiple vehicle households, individuals routinely face vehicle type decisions for activity participation. For instance, Kitamura et al. (2001) reported that male primary users are more likely to use pickup trucks, and younger people are more likely to use sports cars, SUVs, and pickup trucks. People with college degrees or long-distance commuters are more likely to use four-door sedans. A decade later, Vyas et al. (2012) conducted another study using vehicle survey data from California. The authors found that middle aged, senior and female drivers prefer SUVs and that workers and female drivers have an inclination to drive newer cars.

7.1.2 <u>Methods</u>

In the literature, various approaches have been proposed to accommodate for inter-dependency in the behavioral process. One of the simplest approaches employed is to ignore these interdependencies and apply a sequential approach to modeling multiple choice dimensions. The approach is intuitive and easy to employ in practice. However, in this approach not only do we neglect interdependencies between choices, but there is also the question of which sequence to be employed (Pendyala and Bhat, 2004; Kuppam and Pendyala, 2001; Goulias et al., 1990). The exact sequence of the choice processes has significant implications for policy analysis. More recently, a latent segmentation approach that simultaneously allows for different causal structures has been proposed (Chakour and Eluru, 2014). However, the approach is applicable to choice contexts with a small number of dependent variables. An alternative approach accommodates for the interdependency between multiple choices by tying together the unobserved components of the various choices using appropriate distributional assumptions yielding a multivariate joint choice model framework. The approach, while mathematically appealing, requires extensive simulation for model estimation (Bhat and Guo, 2007; Pinjari et al., 2011; Paleti et al., 2013a; Paleti et al., 2013b; Paleti et al., 2013c).

A third approach involves considering the multiple choice processes as a package of decisions made simultaneously. In this approach, every alternative from each choice is coupled with alternatives from other choices to yield a set of combination alternatives. The exact number of combination alternatives generated is obtained by computing the product of number of alternatives for all choice processes (Faghih-Imani et al., 2013; Salon, 2009; Eluru et al., 2010b). The approach, while resulting in an explosion of the number of alternatives, accommodates the dependencies between choices through the systematic component. Further, the methodology employed to study the influence of exogenous factors is usually based on traditional modeling approaches – thus making it a more appealing framework for practice and policy analysis.

7.1.3 Contribution and Organization of the Chapter

Our current research attempt falls within the last category of methodology efforts. Specifically, the current research contributes to our understanding of short term vehicle usage decisions by examining four activity travel choice processes: spatial flexibility of the activity, temporal flexibility of the activity, activity vehicle choice (characterized as vehicle type for auto users and other for non-auto users), and primary driver (for auto users). We consider these four choices as a packaged choice. To model the choice dimensions, we adopt a panel mixed multinomial logit (MMNL) model that accounts for the intrinsic unobserved taste preferences across multiple records for each individual from the longitudinal survey. The data used in the research is drawn from a panel survey conducted in Quebec City, Canada from 2003 – 2006.

The remainder of the chapter is structured in the following order. Section 7.2 describes the data source and the choice set formation procedure. Some descriptive statistics of the sample is also presented in the same section. In Section 7.3, the econometric framework adopted for the analysis is discussed in detail. Empirical results are presented and discussed in Section 7.4. Finally, we summarize the major findings of the research in Section 7.5.

7.2 Data

The primary data used in the current analysis were collected using a longitudinal panel survey of households in the Quebec City region of Canada. The survey, titled "Quebec City Travel and Activity Panel Survey (QCTAPS)", is comprised of three waves, about one year apart for a given household and was carried out from 2003 through 2006. This section of the research first describes the survey instrument with primary focus on the elements relevant to this analysis, and subsequently presents a descriptive analysis of the data sample used for model formulation.

7.2.1 Survey Instrument

The QCTAPS employed a multi-instrument package known as OPFAST (Observed and Perceived Flexibility of Activities in Space and Time) to investigate the decision processes employed by individuals and households to organize their activities in space and time. Specifically, the survey attempts to investigate respondent's perceptions of temporal and spatial flexibility in the organization of their activities. Part of the instrument was an executed activity/travel diary that covered seven consecutive days in wave 1 and two days in the second and third waves. For our analysis, an individual is the unit of analysis for the panel data where repetition of observations of the same individual are accommodated. Information reported in the travel diaries was validated and augmented by a home interview following the diary week, including the geographical location of each activity. A total of 250 households took part in the survey and a high retention rate of 67% was observed from wave 1 to wave 3.

A unique feature of the survey was that respondents were trained to classify every activity that they executed, in or out of the home, according to whether they were ‘‘routine" (or habitual), ‘‘planned" (pre-arranged) or ‘‘impulsive" (or spontaneous) in time and space – using a trichotomy suggested by Garling et al. (1998). The distinction between planned and impulsive is that, for the latter ‘‘one hour in advance, I did not know [that] (temporal dimension) [where] (spatial dimension), I was going to do the activity” (see Lee-Gosselin and Miranda-Moreno, 2009 for more detailed classification of activities by their degree of spatial and temporal spontaneity, with examples). The multi-instrument package OPFAST is described in more detail in Lee-Gosselin (2005).

7.2.2 Choice Set Formation and Descriptive Statistics

The following steps were followed for creating the choice set for our analysis. First, from the activity file, the out-of-home activities were separated out. Second, the several dimensions of analysis were characterized. Perceived temporal flexibility and spatial flexibility of activity is categorized as: (1) Routine, (2) Planned, and (3) Impulsive. The vehicle type alternatives are classified as: (1) Compact sedan, (2) Large sedan, (3) Van and Minivan, (4) Sport Utility Vehicle (SUV), (5) Pick-up and Trucks and (6) Other vehicles including walking, biking, and transit – these vehicle types are available to every individual for any out-of-home activity. The choice set from which households make their choices is defined by the available alternatives in the data set. Hence, the vehicle type dimensions are appropriately matched with the household vehicle ownership information (i.e. if a household does not own a SUV, the individual will not have alternatives corresponding to SUV available to him/her for the activity). For the purpose of our analysis, we considered as many drivers as there were adults in the household and assigned them with numbers for identification. In the dataset, a maximum of four adults are present, hence, the driver dimension comprised a maximum of four alternatives. Third, the MNL model component alternatives are formed as combinations of three perceived temporal flexibility alternatives with the three perceived spatial flexibility options, six travel vehicle type choice alternatives and four driver options. Overall, these categories resulted in a total of 216 discrete alternatives (3*3*6*4). Of course, the reader would recognize that across different individuals the number of alternative available will change based on vehicle fleet available and number of adults in the house.

The database contained a total of 46,730 activity records comprising of both in-home and out-of-home activities. Of these, 14,579 activities were conducted out-of-home. After removing inconsistent and missing/miscoded values, we were left with 8,098 usable out-of-home activity records of 234 households and 378 individuals. Of these households, only 8.1 percent were carless, more than 50 percent owned one private vehicle, approximately 30 percent owned two vehicles, and 6.4 percent owned three or more vehicles. Moreover, the driver count indicated that 49.6 percent of them had one driver, 46.2 percent had two drivers, 3.8 percent had 3 drivers and 0.4 percent had 4 drivers.

Table 7.1 provides descriptive statistics for this sample of out-of-home activities. Several interesting features can be observed from the Table. For instance, across all activities, the percentage of temporally routine activity undertaking (36.3%) is slightly lower than temporally planned activity execution (37.7%). In terms of vehicle type choice, as expected, compact sedans have the largest share (41.7%) while other vehicles alternative has a reasonable share (27.0%). Across temporal flexibility and vehicle type combination, compact sedan for planned activity is the most prevalent combination. It is interesting to note that, among activities undertaken using SUV, planned activities are common. Not so surprisingly, the largest share for other vehicle alternative is for temporally routine activities. Across all activities, spatially routine activities are the most common.

Similar to the temporal flexibility, the most common combination for spatial flexibility is the routine and compact sedan. In the case of activities pursued by SUV, spatially routine activities are preferred. Again, similar to temporal flexibility, the largest share for other vehicle alternatives is for spatially routine activities. Also, observed from the descriptive analysis is a clear trend of distinct proportion of vehicle type usage by temporal and spatial flexibility. The trend is particularly strong for spatial flexibility. For instance, for spatially planned activities, the vehicle type chosen ranges from 19.7 percent (for other) through 44.3 percent (pickups/trucks). We also observed (between the activity flexibility types themselves) that 18 percent of the spatially routine activities are temporally impulsive. At the same time, only 3.6 percent of activities are temporally routine but spatially impulsive highlighting the complex interaction between spatial and temporal flexibility.

The distribution of temporal flexibility, spatial flexibility and vehicle type dimensions by selected individual socio-demographic characteristics is presented in Table 7.2. Several observations can be noted from the Table. First, males perform more temporally and spatially routine activities while women engage themselves more in planned activities. For vehicle choice, compact sedan is the preferred alternative for all individuals. Second, temporally fixed activities are conducted more by middle aged persons; seniors prefer planned and impulsive activities both spatially and temporally. It is interesting to see the involvement of young persons in spatially routine activities. If a personal vehicle is used in the activity execution, young people mostly use compact sedans; seniors use large sedan. Third, as expected, university degree holders prefer routine activities and they use more compact sedans relative to non-university degree holders.

7.3 Econometric Framework

In this analysis, we use a panel mixed multinomial logit (MMNL) model formulation. Let *i* be the index for the discrete choice combination of activity flexibility (temporal and spatial), activity vehicle type choice and primary driver choice ($i = 1, 2, \ldots I$). With this notation, the random utility formulation takes the following familiar form:

$$U_{qit} = \beta_i{}' x_{qt} + (\mu_{qi} + \eta_{qit}) \tag{7.1}$$

In the above equation, U_{qit} represents the utility obtained by the q^{th} individual in choosing the i^{th} alternative at the t^{th} choice occasion. x_{qt} is a vector of attributes influencing the choice framework at the t^{th} choice occasion. $\beta_i{}'$ is a corresponding vector of mean coefficients. μ_{qi} and η_{qit} form the complete error term. The first component μ_{qi} is a vector of normal random terms with zero mean (tied to the number of individuals in the dataset) representing the error components while the second term η_{qit} is an idiosyncratic error term assumed to be identically and independently Type-1 extreme value distributed (tied to the number of activity records in the dataset). According to the utility maximization principle, an individual q will choose the alternative that offers the highest utility. The unconditional probability expression for choosing alternative i across a series of activities for individual q is given by:

$$P_i = \int \prod_{t=1}^{T} \frac{\exp\ [\beta_i{}' x_{qt} + \mu_{qi}]}{\sum_{j \epsilon C} \exp\ [\beta_i{}' x_{qt} + \mu_{qi}]} d\boldsymbol{F}(\mu_{qi}) d\mu_{qi} \tag{7.2}$$

where *C* represents the choice set for individual q and *T* represents the number of records per individual. The log-likelihood function is constructed based on the above probability expression and maximum simulated likelihood (MSL) estimation is employed to estimate β' parameters. For this particular study, we use a quasi-Monte Carlo (QMC) approach with 150 draws for the MSL estimation (see Bhat, 2001 for more details).

7.4 Empirical Analysis

7.4.1 Variable Specification

Several categories of exogenous variables were considered in the model including individual and household socio-demographics, household residential location characteristics, activity attributes, and contextual variables. The *individual socio-demographics* considered are: gender, age, education, and cellular phone usage. The *household and residential attributes* considered include household income, dwelling type, family type and location of household. The *residential location and type variables* capture attributes of a household's activity–travel environment. Three types of *activity attributes* were considered in our analysis: activity location, activity type and accompaniment type. In terms of *contextual variables*, we included season and day of week. The choice of these independent variables was guided by prior research on activity-based modeling and also constrained by data availability. Note that in the current context, we do not have any alternative specific variables for drivers since the driver alternatives are unlabeled and characterized by driver attributes. Moreover, it is not possible to evaluate the effect of household and residential location characteristics on primary driver selection directly.

The final variable specification was based on a systematic process of removing statistically insignificant variables (in our analysis we considered 90 percent significance level), combining and constraining variables when their effects were not significantly different. Estimating all potential exogenous variable effects for all of the alternatives (up to 216) would result in a cumbersome and likely inefficient model specification. Hence, in the current research, variable effects are considered across the four dimensions. This allows capturing majority of the exogenous variable impacts while retaining a fairly parsimonious model specification (see Faghih-Imani et al., 2013 for a similar analysis).

7.4.2 Estimation Results

The model estimation process began with the estimation of the traditional MNL model. Next, the panel mixed MNL model was estimated. After extensive specification testing, the final log-likelihood values at convergence of the MNL and mixed MNL models were found as: -20363.23 and -19990.96, respectively. The improvement in the data fit clearly demonstrates the superiority of the mixed MNL model over its traditional counterpart. The Log-likelihood Ratio (LR) test comparison between the MMNL and MNL model yields a test statistic value that rejects that hypothesis that all the models are similar at any reasonable level of significance. Hence, in the subsequent sections, we discuss about the results of the MMNL model only.

The final specification results of the joint model are presented in Table 7.3 (the *t*-stats are presented in parentheses). A positive (negative) coefficient for a certain variable-category combination means that an increase in the explanatory variable increases (decreases) the likelihood of that alternative being chosen relative to the base alternative. A blank entry corresponding to the effect of variable indicates no statistically significant effect of the variable on the choice processes. In the following sections, we discuss the effects of variables by variable category.

7.4.2.1 Constants

The constant term clearly indicates that there is a greater probability of temporally pre-planned activities being pursued. Spatially impulsive activities are the least likely to be chosen as evidenced by the high negative constant relative to other flexibility indicators. Among the vehicle type themselves, SUVs and sedans (both compact and large) are the most likely vehicle type choice for out-of-home activity participation, if they are available. On the other hand, vans/mini-vans and pick-ups/trucks are the least likely vehicle type choice.

Within the set of constant parameters, the impact of wave indicator was examined. Specifically, these indicators are expected to capture the across wave variations. The effect of the wave dummy variable was found significant for both types of activity flexibility and vehicle type choice. We observed that individuals in wave 1 were more inclined towards performing routine and pre-planned activities. The negative coefficient for spatial flexibility indicates that individuals were less likely to take part in pre-planned or impulsive activities. Compact sedan has higher propensity of being chosen compared to large sedan, vans/minivans, and SUVs.

7.4.2.2 Individual Socio-demographics

The parameter estimates for individual demographic characteristics underscore their importance on daily activity travel decisions. We find that females are unlikely to drive sedans and vans/minivans. In terms of the role of primary driver, we find that females are more likely to be assigned to the responsibility relative to men. The result might be explained in light of the fact that compared to men, women are in charge of taking care of kids and they pursue more household chores which might require them to be the primary driver of the household.

We introduced age as dummy variables since they provided the best model fit. Our analysis results regarding young aged people supports the notion that they are more likely to perform temporally impulsive activities (Mohammadian and Doherty, 2005). People of this age also tend to be disinclined to use compact sedans and vans/minivans for activity engagements. The choice of vehicles of young people might be driven by their preference for environmentally friendly alternatives (such as transit and active forms of transportation) or personal view towards the vehicles - i.e. they might think sedans and vans as “boring” and hence, they would rather drive the stylish SUVs or rugged pick-up trucks. Also, young individuals are likely to be designated primary drivers relative to middle aged individuals. This is plausible because these individuals are likely to be living alone and do not share their car with anyone. Seniors are found to be indifferent towards any type of activity flexibility indicators which is understandable since people of this age are generally free from fixed employment and have the liberty to pursue activities at their will without any time and space constraint. With respect to vehicle type choice dimension, seniors have a lower preference for compact sedans and SUVs for out of home activities. Similar to young individuals, seniors are likely to be designated primary drivers in their household.

Turning to education effects, household members with university degree and greater level of education prefer vans/minivans as their vehicle choice. Individuals’ education levels are certainly correlated to their occupations and income (Choo and Mokhtarian, 2004). As such, they are more inclined towards routinized life and driving large vehicles such as vans/minivans. For the same reasons, they are also more likely to be the primary driver of the household. Contrastingly, individuals with diploma degrees tend to engage more in temporally impulsive activities and less in spatially pre-planned activities. Large sedans, vans/minivans, and pick-ups/trucks are their preferred choice of vehicle for activity participation. Non-usage of cell phones is significantly and negatively associated with temporally and spatially pre-planned and impulsive activities and these individuals are disinclined to use large sedan for performing their out-of-home activities.

7.4.2.3 Household Socio-demographics

Among household demographics, several behaviourally intuitive yet interesting findings were observed. For instance, individuals residing in single-detached dwellings are more likely to undertake temporally pre-planned as well as impulsive and spatially pre-planned activities. For travel, they tend to prefer sedans. The single detached dwelling housing stock is predominantly owner-occupied and located in low density areas. Preference for luxury cars of this type of households is noted in the literature (Kitamura et al., 2001). Moreover, residents of apartment are more likely to use compact sedans for their activity participation. Both medium and high income is negatively associated with temporal flexibility meaning that individuals belonging to these households are less likely to take part in planned or impulsive activities, presumably reflecting their time constraints resulting from job commitment issues. Members from medium income households are more disinclined towards spatially pre-planned activities. Similar vehicle type choice preference is observed between the members of these two types of households. Individuals from medium income group are more likely to opt for large sedans and vans/minivans for their activity participation while members of affluent households prefer both vans/minivans and SUVs (Kitamura et al., 2001; Bhat and Sen, 2006)

An individual from a household with children is disinclined to engage in temporally pre-planned or impulsive activities presumably owing to the responsibility of tending to the child/children; individuals from these households might have decreased ability or desire to take part in activities which are planned in a short period of time. They are more likely to choose vans/minivans for their activity execution. The results are intuitively understandable - for chauffeuring kids they use vans/minivans since these vehicles are more spacious, safe and comfortable for travel with children (Bhat and Sen, 2006). Very interestingly, individuals belonging to a childless household are also unlikely to pursue non-routine activities. They tend to avoid large sedans for travel.

7.4.2.4 <u>Household Residential Location Attributes</u>

With regards to residential location attributes, we considered the following categories: peripheral areas, old suburbs, new suburbs and Central Business Districts (CBD). The categories were created applying a k-means cluster analysis using population density, land use mix and transit accessibility indices. The peripheral areas have the lowest values for all three indices. Old suburbs have medium land use mix and population density, and are served by the main transit lines. New suburbs are characterized by low to medium density, land use mix and transit accessibility. CBD represents mostly downtown core and central neighborhoods, with the highest values for all three indices.

The modeling results show that individuals living in peripheral areas have a higher propensity of getting involved in temporally and spatially impulsive activities. Peripheral areas have the lowest land use mix, transit accessibility and population densities, and are very auto oriented neighborhoods thus allowing for impulsive activity participation (temporal and spatial). As is expected, these individuals also tend to prefer large sedans, vans/minivans, and SUVs for their traveling purposes. Persons living in CBDs are also more likely to conduct temporally and spatially impulsive activities. On the other hand, these individuals are unlikely to employ sedans and SUVs highlighting an overall preference for non-auto oriented travel. People living in older suburbs are more inclined towards spatially impulsive activities and disinclined towards using any types of sedans.

7.4.2.5 *Contextual Variables*

Walk/bike/transit appears to be the preferred vehicle type choice in summer season. This is intuitive, since the weather is more conducive to undertaking activities by walking or biking or taking public transit than winter. During winter, individuals tend to be less inclined towards temporally impulsive activities. Preference for routine activities during winter might be explained in light of the unfavorable weather for activities outside home during this season. People prefer not to choose vans/minivans during winter. During winter maintaining larger cars such as vans is expensive (increased heating leading to increased gas cost) and difficult (more snow cleaning, parking difficulty) in Canada and this might be deterring individuals from using these vehicles for their out of home travel purposes. During weekends, people are free and relaxed and hence, they tend to undertake more pre-planned and impulsive activities – a result which is intuitively understandable. Sedans and vans/minivans are the preferred vehicle choices for weekend activities. It appears that people also allocate Fridays for spatially impulsive activities and prefer to use compact sedans for these activities, presumably due to possible congestion on Fridays.

7.4.2.6 *Activity Attributes*

Activities undertaken in peripheral and CBD areas tend to be of temporally impulsive in nature. In case of CBDs, the activities are also spatially flexible. As expected, contrasting choice of vehicles is observed between these two activity locations. In peripheral areas people prefer larger vehicles such as vans/minivans and SUVs whereas in CBDs people tend to undertake activities by walk/bike/transit. The results are in line with expectations. Unlike peripheral areas, these neighborhoods have diverse land use mix and increased number of activity centers. Moreover, these areas are also known for their "pedestrian oriented" urban form and parking restrictions which might be deterring individuals from using vehicles and opting for walk/bike/transit mode instead. Individuals tend not to use compact sedans and/or pick-ups/trucks for undertaking activities in older suburbs.

Activities that involve basic need (e.g. meals) are either routine or impulsive in time and less likely to be pre-planned (see Mohammadian and Doherty, 2005 for similar results). On the other hand, the location is more likely to be pre-planned or selected impulsively. It might be indicating that dining out or doing grocery might be a temporally routine/impulsive activity for individuals but the location (restaurant or superstore) for these activities might be either pre-planned or impulsive. For activities involving basic needs, people tend to prefer walk/bike/transit (except pick-up trucks). Among all activity types, work/school is considered to be fixed or mandatory in time and space and our results conform to this expectation. In terms of vehicle choice, pick-ups/trucks are likely to be chosen (please note that if pickup trucks are available in the household they are potentially work related vehicle purchases). As expected, our results suggest that both shopping and social/recreational activities are more likely to be impulsively undertaken by individuals. For shopping activity types, individuals tend not to use sedans presumably due to the fact that these activities are usually conducted in groups and thus people might prefer larger vehicle. Activities conducted alone tends to be of routine nature and pursued using walk/bike/transit indicating that usage of vehicle is deemed required with increased number of accompanying people (except for pickup trucks which could be related to small share of pickup trucks).

7.4.2.7 <u>*Error Components*</u>

The final model specification included two error components, confirming the presence of common unobserved attributes among joint choice alternatives. Specifically, the dimensions that exhibited strong correlations include: spatial flexibility (pre-planned) and temporal flexibility (pre-planned). The error component parameters provide important insights regarding the sensitivity of joint choice alternatives sharing the same temporal and spatial flexibility.

7.5 Summary and Conclusions

Given the overwhelming contribution of private vehicles towards GHG emissions, it is not surprising that travel behavior researchers have examined vehicle fleet choices (number, type and usage) extensively. Traditionally, vehicle fleet decisions are examined as a long term choice with annual usage metrics. Only recently, travel behavior models have started examining vehicle usage decisions (type and mileage) as a short-term decision in the context of activity travel analysis. Our study contributes to the growing literature on short term vehicle usage decisions by examining four activity travel choice processes: spatial flexibility of the activity, temporal flexibility of the activity, activity vehicle type, and primary driver (for auto users). The four choice dimensions considered in this research are of significance for policy making and urban transportation planning purposes.

The data used in the study is drawn from a longitudinal panel survey of households in the Quebec City region of Canada. The survey comprised of three waves, about one year apart for a household and carried out from 2003 through 2006. The survey attempts to investigate respondent's perceptions of temporal and spatial flexibility in the organization of their activities. In terms of methodology, a panel mixed multinomial logit model (MMNL) is applied to account for the intrinsic unobserved taste preferences across individuals from the longitudinal survey.

The analysis results revealed that several individual and household socio-demographic characteristics, residential location and activity attributes as well as contextual variables influence the packaged choice of temporal flexibility, spatial flexibility, vehicle type choice and primary driver selection. For example, young individuals' temporally impulsive activity proneness was confirmed. Individuals belonging to both medium and high income households are less likely to take part in temporally pre-planned or impulsive activities. We also observed that compared to the central business districts, individuals living in the peripheral areas are more likely to conduct impulsive activities and use larger automobiles. Moreover, we also identify the presence of common unobserved attributes among the joint choice alternatives.

Table 7.1 Distribution of Perceived Temporal and Spatial Flexibility by Vehicle Type

Dimensions	Vehicle Type						Within Vehicle Type Choice (%)
	Compact Sedan	*Large Sedan*	*Van/Minivan*	*SUV*	*Pick-ups/Trucks*	*Walk/Bike/Transit*	
Perceived Temporal Flexibility							
Routine	1179 (34.9%)	475 (33.6%)	178 (31.8%)	124 (30.7%)	48 (30.4%)	938 (42.9%)	2942 (36.3%)
Planned	1345 (39.8%)	606 (42.9%)	235 (42.0%)	185 (45.8%)	67 (42.4%)	617 (28.2%)	3055 (37.7%)
Impulsive	852 (25.2%)	332 (23.5%)	146 (26.1%)	95 (23.5%)	43 (27.2%)	633 (28.9%)	2101 (25.9%)
Perceived Spatial Flexibility							
Routine	1971 (58.4%)	796 (56.3%)	319 (57.1%)	205 (50.7%)	68 (43.0%)	1343 (61.4%)	4702 (58.1%)
Planned	928 (27.5%)	428 (30.3%)	150 (26.8%)	141 (34.9%)	70 (44.3%)	430 (19.7%)	2147 (26.5%)
Impulsive	477 (14.1%)	189 (13.4%)	90 (16.1%)	58 (14.4%)	20 (12.7%)	415 (19.0%)	1249 (15.4%)
Within Temporal and Spatial Flexibility (%)	3376 (41.7%)	1413 (17.4%)	559 (6.9%)	404 (5.0%)	158 (2.0%)	2188 (27.0%)	8098 (100.0%)

Table 7.2 Distribution of Individual Characteristics across Temporal Flexibility, Spatial Flexibility and Vehicle Type

Dimensions	Gender		Age			Education		Total (%)
	Male	*Female*	*Young (<= 30)*	*Middle Aged (31 - 60)*	*Senior (> 60)*	*University*	*Others*	
Perceived Temporal Flexibility								
Routine	1589 (38.9%)	1353 (33.8%)	557 (37.4%)	2085 (38.8%)	300 (24.3%)	1126 (39.0%)	1816 (34.8%)	2942 (36.3%)
Planned	1486 (36.3%)	1569 (39.1%)	514 (34.5%)	1980 (36.8%)	561 (45.5%)	1093 (37.9%)	1962 (37.6%)	3055 (37.7%)
Impulsive	1015 (24.8%)	1086 (27.1%)	419 (28.1%)	1309 (24.4%)	373 (30.2%)	665 (23.1%)	1436 (27.5%)	2101 (25.9%)
Perceived Spatial Flexibility								
Routine	2409 (58.9%)	2293 (57.2%)	891 (59.8%)	3182 (59.2%)	629 (51.0%)	1690 (58.6%)	3012 (57.8%)	4702 (58.1%)
Planned	1061 (25.9%)	1086 (27.1%)	357 (24.0%)	1408 (26.2%)	382 (31.0%)	777 (26.9%)	1370 (26.3%)	2147 (26.5%)
Impulsive	620 (15.2%)	629 (15.7%)	242 (16.2%)	784 (14.6%)	223 (18.1%)	417 (14.5%)	832 (16.0%)	1249 (15.4%)
Vehicle Type								
Compact Sedan	1555 (38.0%)	1821 (45.4%)	709 (47.6%)	2152 (40.0%)	515 (41.7%)	2403 (46.1%)	973 (33.7%)	3376 (41.7%)
Large Sedan	804 (19.7%)	609 (15.2%)	47 (3.2%)	950 (17.7%)	416 (33.7%)	898 (17.2%)	515 (17.9%)	1413 (17.4%)
Van/Minivan	291 (7.1%)	268 (6.7%)	32 (2.1%)	504 (9.4%)	23 (1.9%)	337 (6.5%)	222 (7.7%)	559 (6.9%)
SUV	248 (6.1%)	156 (3.9%)	37 (2.5%)	304 (5.7%)	63 (5.1%)	170 (3.3%)	234 (8.1%)	404 (5.0%)
Pick-ups/Trucks	151 (3.7%)	7 (0.2%)	1 (0.1%)	133 (2.5%)	24 (1.9%)	149 (2.9%)	9 (0.3%)	158 (2.0%)
Walk/Bike/Transit	1041 (25.5%)	1147 (28.6%)	664 (44.6%)	1331 (24.8%)	193 (15.6%)	1257 (24.1%)	931 (32.3%)	2188 (27.0%)
Total (%)	4090 (50.5%)	4008 (49.5%)	1490 (18.4%)	5374 (66.4%)	1234 (15.2%)	5214 (64.4%)	2884 (35.6%)	8098 (100.0%)

Table 7.3 Estimation Results (N = 378 individuals and 8098 activity records)

Variables	Temporal Flexibility *(Base: Routine)*		Spatial Flexibility *(Base: Routine)*		Vehicle Type *(Base: Walk, bike and transit)*					Primary Driver
	Planned	**Impulsive**	**Planned**	**Impulsive**	**Compact Sedan**	**Large Sedan**	**Van/ Minivan**	**Sport Utility Vehicle (SUV)**	**Pick-up/ Trucks**	
Constants	1.283	0.151	-0.100	-1.222	2.517	0.690	-2.584	1.657	-2.894	---
	(8.550)	(0.957)	(-0.602)	(-8.403)	(11.982)	(2.102)	(-4.415)	(4.181)	(-8.392)	---
Wave1	---	-0.324	-0.203	-0.528	0.957	0.363	0.461	0.461	-0.459	---
	---	(-4.608)	(-2.767)	(-6.519)	(11.528)	(3.117)	(3.464)	(3.464)	(-1.683)	---
Individual Characteristics										
Female	---	---	---	---	-1.071	-0.666	-0.325	---	---	1.433
	---	---	---	---	(-10.941)	(-5.570)	(-2.018)	---	---	(18.094)
Age (Base: Middle aged (31-60))										
Young (Age ≤30)	---	0.415	---	---	-0.889	---	-0.889	---	---	0.862
	---	(4.309)	---	---	(-6.689)	---	(-6.689)	---	---	(7.491)
Senior (Age >60)	---	---	---	---	-1.482	---	---	-1.667	---	2.159
	---	---	---	---	(-9.142)	---	---	(-5.515)	---	(14.969)
Education Level (Base: Other degree)										
University Degree	---	---	---	---	-2.857	---	1.009	-1.530	---	2.102
	---	---	---	---	(-12.125)	---	(2.290)	(-6.441)	---	(12.182)
Diploma Degree	---	0.258	---	---	-0.851	1.415	2.646	---	2.094	---
	---	(3.733)	---	---	(-5.040)	(7.465)	(6.499)	---	(6.409)	---
Don't Use Cell Phone	-0.240	-0.461	-0.283	-0.462	---	-0.630	---	---	---	---
	(-2.632)	(-6.423)	(-3.109)	(-6.413)	---	(-6.052)	---	---	---	---
Household Socio-demographics										
Dwelling Type (Base: Others)										
Detached House	0.349	0.381	0.314	---	0.385	0.531	---	-1.306	---	---
	(3.115)	(4.602)	(2.636)	---	(3.831)	(3.983)	---	(-4.309)	---	---
Apartment	---	---	---	---	0.331	---	---	---	---	---
	---	---	---	---	(2.681)	---	---	---	---	---
Income (Base: Low Income (< 20K))										
Medium Income (20K-60K)	-0.564	-0.564	-0.420	---	---	0.448	2.620	---	---	---
	(-5.297)	(-5.297)	(-3.851)	---	---	(3.623)	(6.378)	---	---	---
High Income (> 60K)	-0.528	-0.528	---	---	---	---	1.034	1.371	---	---
	(-4.161)	(-4.161)	---	---	---	---	(2.452)	(5.369)	---	---
Family structure (Base: Single Adult)										
Couples with Children	-0.539	-0.539	-0.392	-0.363	---	-0.381	0.528	-0.596	1.051	---
	(-5.274)	(-5.274)	(-2.555)	(-3.954)	---	(-1.933)	(2.622)	(-2.383)	(4.080)	---
Couples without Children	-0.340	-0.464	-0.298	-0.233	---	-0.807	---	---	---	---

	(-2.732)	(-4.875)	(-2.091)	(-2.557)	---	(-3.997)	---	---	---	---
Household Residential Attributes (Base: New Suburbs)										
Peripheral Areas	---	0.197	---	0.190	---	0.384	0.384	1.671	---	---
	---	(2.255)	---	(2.008)	---	(3.297)	(3.297)	(6.711)	---	---
Central Business District	---	0.263	---	0.236	-0.741	-0.623	---	-0.948	---	---
	---	(2.589)	---	(2.217)	(-6.497)	(-2.905)	---	(-3.721)	---	---
Old Suburbs	---	---	---	0.150	-0.651	-1.043	---	---	---	---
	---	---	---	(1.669)	(-7.430)	(-7.606)	---	---	---	---
Contextual Variables										
Season (Base: Spring and Fall)										
Summer	---	---	---	---	-0.210	-0.545	-0.990	-0.805	---	---
	---	---	---	---	(-2.894)	(-4.965)	(-5.998)	(-4.202)	---	---
Winter	---	-0.307	---	---	---	---	-1.401	---	---	---
	---	(-2.949)	---	---	---	---	(-4.796)	---	---	---
Day of Week										
Weekend	0.922	0.922	0.569	0.314	0.826	0.773	0.489	---	---	---
	(11.279)	(11.279)	(7.954)	(3.754)	(8.571)	(5.932)	(2.602)	---	---	---
Friday	---	---	---	0.172	0.364	---	---	---	---	---
	---	---	---	(1.845)	(3.989)	---	---	---	---	---
Activity Attributes										
Activity Location (Base: New Suburbs)										
Peripheral Areas	---	0.214	---	---	---	---	0.529	0.962	---	---
	---	(2.129)	---	---	---	---	(2.426)	(4.131)	---	---
Central Business District	0.241	0.380	0.325	0.460	-1.091	-0.869	-0.493	-0.493	---	---
	(3.146)	(4.501)	(4.399)	(5.529)	(-12.881)	(-7.698)	(-3.350)	(-3.350)	---	---
Old Suburbs	---	---	---	---	-0.256	---	---	---	-0.445	---
	---	---	---	---	(-2.961)	---	---	---	(-1.662)	---
Activity Type (Base: Other Activities)										
Basic Needs	-0.918	---	0.247	1.584	-1.469	-1.246	-1.580	-1.313	---	---
	(-10.162)	---	(2.519)	(12.814)	(-11.123)	(-7.791)	(-6.260)	(-5.027)	---	---
Work/School	-1.600	-2.299	-0.765	-1.482	-0.845	-0.351	-1.139	---	1.471	---
	(-19.683)	(-18.780)	(-9.166)	(-8.420)	(-7.348)	(-2.749)	(5.999)	---	(6.139)	---
Shopping	0.953	2.214	0.428	1.708	-0.293	---	---	---	---	---
	(9.206)	(20.935)	(5.514)	(15.163)	(-2.525)	---	---	---	---	---
Social/Recreational	---	0.790	---	0.865	-1.810	-1.660	-1.968	-1.729	---	---
	---	(10.128)	---	(7.568)	(-15.421)	(-12.148)	(-9.353)	(-7.918)	---	---
Accompaniment Type										
Alone	-0.453	-0.142	-0.585	-0.585	-0.513	-0.998	-0.420	-0.912	---	---
	(-6.381)	(-1.899)	(-10.348)	(-10.348)	(-6.316)	(-8.584)	(-2.556)	(-5.420)	---	---
Error Components	0.799	---	0.821	---	---	---	---	---	---	---

	(17.249)	---	(15.738)	---	---	---	---	---	---	---
Log-likelihood at convergence						-19992.50				

Note: --- denotes insignificant variables. Also, the coefficient estimates across different alternatives are constrained to be same when the effects are not significantly different.

CHAPTER 8 CONCLUSIONS AND RECOMMENDATIONS

8.1 Introduction

The decisions to own, operate, and maintain personal vehicles (in other words, allocating resources for transportation) are the most fundamental transportation related financial decisions that households make. Therefore, this dissertation aimed at developing original methodologies and providing important empirical evidences on household's travel behavior choice processes across disparate temporal scales, i.e. long term, medium term, and short term. As part of long term decision, budget allocation decisions of Canadian households were investigated by proposing an original methodology that allowed to understand how household transportation expenditures have evolved in Canada in comparison with a wide assortment of other essential and discretionary commodities, goods, and services that households incur expenses on. The medium term choice of vehicle fleet size of households is explored next by developing and validating state-of-the-art econometric models. Finally, in terms of short term decision, the vehicle usage pattern is investigated. Toward that end, a joint econometric framework of vehicle type choice, activity flexibility (temporal and spatial), and primary driver choice is developed. The current dissertation develops and presents novel econometric modeling approaches to address the following underexplored aspects related to: (1) transport expenditure and its evolution, (2) population heterogeneity, (3) appropriateness of econometric frameworks (ordered vs unordered), (4) multiple time point data pooling (pseudo panel analysis), and (5) short term vehicle usage. In doing so, it contributes methodologically by developing advanced discrete choice models while simultaneously contributing empirically by considering a wide range of exogenous variables in developing the models.

The rest of the chapter is organized as follows. Sections 8.2 through 8.6 discuss the findings and contributions – methodological and empirical, of the dissertation. Section 8.7 concludes the dissertation by documenting the limitations of the current research and presenting the directions for future research.

8.2 Analysis of the Evolution of Transportation Expenditure: A Long Term Perspective

The primary focus of the research effort presented in Chapter 3 is to investigate the household budgetary allocation decisions with a particular focus on transportation budget. Specifically, the aim was to present a methodology to identify the factors affecting expenditure patterns of households and its evolution in Canada. Recognizing the fact that households choose to allocate their limited budget to multiple expenditure categories simultaneously, a multiple discrete framework is formulated using the public-use micro-data extracted from the Survey of Household Spending (SHS) for the years 1997 – 2009. Since, 13 years of data were pooled together to form the estimation dataset, two variants of the multiple discrete continuous extreme value (MDCEV) model – scaled MDCEV (SMDCEV) and mixed MDCEV (MMDCEV) model were utilized for accommodating the effect of observed and unobserved temporal effects. After extensive specification testing, data fit comparison metrics were employed to identify the best model structure. Important findings from the research undertaking include:

- Transportation related spending of households in Canada remained relatively stable over the span of 13 years of study period (12-16% of the total budget).
- The best model fit was obtained for the scaled MDCEV model confirming variation in the unobserved factors across the study years compared to 1997. The variation was more pronounced for the recession year of 2008.
- In terms of the model results, household socio-economic and demographic attributes along with residential location characteristics were found to impact household expenditure significantly.
- The policy simulation exercise revealed that with increase in gasoline prices, households tend to spend more on public transit and non-motorized transportation options in the short run while in the long run, less amount of budget is allocated to vehicle purchasing.

8.2.1 Implications and Suggested Policy Measures

The research reinstated the need and importance of simultaneous exploration of different expenditure categories including transportation to get a more holistic picture of household budget allocation patterns. Moreover, the simultaneous examination also helps glean more information about potential trade-offs amongst different outlays of money. For instance, in our study, we found that households residing in single detached dwelling located in a medium-density urban area spend less on housing as well as public transportation, while spending more on gasoline. On the other hand, apartment dwellers in high density urban areas allocate higher proportion of their income on housing and public transportation while spending less on gasoline. The result might be suggesting that transportation related benefits of high density urban areas might be associated with increasing housing cost (Palm et al., 2014) and thus have intriguing implications for the "smart growth" policies intended towards reducing household vehicle ownership and usage. It can be argued that the policy of densification and diversification of metropolitan areas need to be complemented with strategies to reduce housing expenses. If gas price is lowered, it's the low density dwellers who are usually less inclined to use transit would be the beneficiaries in terms of reduced transportation cost – which would incentivize suburban living.

8.3 Addressing Population Heterogeneity in Discrete Choice Models

In Chapter 4, we proposed an econometric approach that relaxes the population homogeneity assumption and accommodates systematic heterogeneity based on observed attributes in the data in the context of vehicle ownership. Using the Origin-Destination (OD) survey data of Quebec City for the year 2001, a total of six models, three from the ordered regime and three from the unordered regime were estimated. Specifically, the models estimated from the ordered regime were: (1) traditional ordered logit (OL), (2) latent segmentation based ordered logit model with two segments (LSOL II), and (3) latent segmentation based ordered logit model with three segments (LSOL III). On the other hand, models estimated from the unordered framework were: (1) traditional multinomial logit (MNL), (2) latent segmentation based multinomial logit model with two segments (LSMNL II), and (3) latent segmentation based multinomial logit model with three segments (LSMNL III). The performance of the developed latent class models in comparison with their traditional counterparts were evaluated using Bayesian information criterion (BIC). The comparison result indicated that latent models outperformed the traditional models by a large margin offering evidence in favour of the hypothesis that car ownership can be better examined through segmentation of households. Other important findings from the study include:

- The latent class models outperformed their traditional counterparts by a large margin providing evidence regarding the presence of segmentation in the population with respect to vehicle ownership decisions. This is the most important finding that emerged from this research study.

- Households were assigned probabilistically to two segments – transit independent (TI) and transit friendly (TF) based on a host of socio-demographic (number of household members and job type) and land use variables (transit accessibility, number of transit pass holders, and entropy index).
- From the elasticity effects, it was observed that residential density had greater negative impact on household's fleet size.
- The predictive performance evaluation exercise between the two latent class models on a validation sample did not yield a conclusive result. However, the unordered choice mechanism appeared to perform slightly better than the ordered response mechanism.

8.3.1 Implications and Suggested Policy Measures

The findings of this chapter have important implications, both from a methodological standpoint and a policy perspective. From modeling perspective, it is apparent that travel demand modeling tools can be benefitted from incorporating the proposed modeling frameworks which are capable of accommodating population heterogeneity in the data.

From the policy aspect, increase in the number of full and part-time workers appeared as the two most important demographic variables and residential density appeared as the most important land use measure influencing households' decision to change their fleet size. Presumably, with increasing working members, there will be conflict over the usage of vehicles for commuting. Hence, the decision to purchase additional vehicle might be taken in an effort to minimize these conflicts. Once purchased, households get dependent on private automobiles for their daily travel. That being said, there is a merit in adopting policy measures oriented towards discouraging households from purchasing additional vehicles by providing better quality transportation alternatives (such as reliable public transit, infrastructure for active transportation, along with bike and car sharing services) to the automobiles (Roorda et al., 2009). In addition, other strategies such as increasing vehicle registration fees (Chin and Smith, 1997) coupled with making locations closer to central neighborhoods more desirable to families through advertisement (campaigns as that of "A Life Near Everything" by the City of Montreal) (Manaugh et al., 2010) are likely to have a greater effect on reducing vehicle ownership. As Macfarlane et al. (2015) puts it – "the best policy agenda is likely to be a multi-faceted approach, of which densification may be only one part".

8.4 Appropriate Framework of Analysis while Addressing Population Heterogeneity in Discrete Choice Models

An empirical comparison of the latent class versions of the ordered and unordered models in the context of household vehicle ownership levels was presented in Chapter 5. The potential existence of population heterogeneity was investigated through the developed models while the comparison exercise provided guideline regarding the appropriate framework (ordered or unordered) for analyzing vehicle fleet size decisions. The alternative modeling approaches considered included: from the ordered regime – ordered logit (OL), generalized ordered logit (GOL), latent segmentation based ordered logit with two segments (LSOLII), latent segmentation based ordered logit with three segments (LSOLIII), latent segmentation based generalized ordered logit with two segments (LSGOLII), and latent segmentation based generalized ordered logit with two segments (LSGOLIII) and from the unordered regime – multinomial logit (MNL), latent segmentation based multinomial logit with two segments (LSMNLII), and latent segmentation based multinomial logit with three segments (LSMNLIII). The empirical analysis was conducted using Origin-Destination (O-D) survey of Greater Montreal Area (GMA) for the year 2008. The performance of the model frameworks was evaluated in the context of both model estimation and validation (aggregate and disaggregate level). Important findings from the empirical analysis include:

- Empirical results indicate the presence of two distinct population segments based on the observed land use attributes: low density high car ownership and high density low car ownership segments.
- The choice of vehicle fleet in segment-1 is more influenced by household socio-demographics while surrounding land use characteristics tend to govern the ownership decision of the households belonging to segment-2 more. The finding has intriguing policy implications.
- It was evident from the estimation results that LSGOLII model outperformed its traditional counterpart from the ordered regime as well as its unordered counterpart by a large margin.

- The validation exercise revealed that LSGOLII demonstrated superior prediction performance both at the aggregate and disaggregate levels. In addition, superior prediction for various sub-samples at the aggregate and disaggregate level was also noted.

8.4.1 Implications

In summary, the comparison exercise supports the hypothesis that LSGOL is a promising ordered response framework for accommodating population heterogeneity and for relaxing the fixed threshold assumption in the context of ordered variables.

8.5 Analysis of Vehicle Ownership Evolution Using Pseudo Panel Analysis

The major focus of Chapter 6 was to propose a workaround approach to the longitudinal data availability issue for studying the evolution of vehicle ownership decision process. To achieve the objective, first, we stitched together multiple year cross-sectional datasets to form a pseudo panel from Origin-Destination (O-D) surveys of Greater Montreal Area (GMA) for the years 1998, 2003, and 2008. Afterwards, we utilized this pooled dataset to undertake vehicle fleet size analysis. Towards that end, we estimated four different models from the ordered regime: (1) ordered logit (OL), (2) generalized ordered logit (GOL), (3) scaled generalized ordered logit (SGOL), and (4) mixed generalized ordered logit (MGOL) model while employing a comprehensive set of exogenous variables. The comparison exercise, based on information criterion metrics, highlighted the superiority of the MGOL models in terms of data fit compared to its other ordered outcome counterparts. Important results from the empirical analysis include:

- From the analysis result it was observed that households in Montreal in recent times are more likely to have increased vehicle fleet size.
- Both socio-demographics and land-use variable impacts are changing with time. The most striking is the reduction in the impact of the full-time workers on increased vehicle ownership. It might be a manifestation of the environmental consciousness of households in recent years.
- Moreover, variation due to unobserved factors are captured for part-time working adults, number of bus stops, and length of metro lines.

- In terms of the effect of location of households, we found that some neighborhoods exhibited distinct car ownership temporal dynamics over the years.

8.5.1 Implications and Suggested Policy Measures

This study demonstrated a simple yet efficient data usage method − pooling cross-sectional data from multiple time points to examine vehicle ownership and its evolution. However, usefulness of the proposed method is not limited to only vehicle ownership analysis. In fact, in the absence of panel data, travel demand and behaviour analysis could benefit from such application of the multiple year cross sectional databases and improve the forecasting capability of the developed models. A possible reduction in data collection burden, in terms of cost, is also achievable through the use of pooled data.

In addition to increased sample size, combining cross-sectional datasets offers the potential to gain insight into the temporal change in the impact of the exogenous variables conveying useful information for policy makers. For example, based on the result of the full-time working adults, policy makers can ponder upon strategies targeted towards behavioral modification through awareness. The psychological and behavioral strategies can range from providing specific information on public transport, travel campaigns, and travel education – make people aware of the consequences of their travel habits (Wotton, 1999; Fujii and Taniguchi, 2005). In fact, recent research by Gaker et al. (2011) has shown that educated people are willing to adjust their transport behavior to reduce carbon dioxide emissions. So, it can be a potential policy avenue to pursue. Other potential softer policies include encouraging workers to telework and/or teleconference or car share or make multi-modal commute trips by increasing transit accessibility at the work location. In fact, it is found in literature that employment density at work is more important than population density at residential location (Zhang, 2004; Chen et al., 2008). Therefore, accessibility of work locations in combination with that of residential locations should be considered to be able to achieve higher non-auto mode shares (Pinjari et al., 2011; Frank et al., 2008; Maat and Timmermans, 2009). In addition, integration of the already owned private automobiles with the public transport system through park-and-ride or park-and-pool or kiss-and-ride schemes can also be considered (Santos et al., 2010; Turnbull et al., 2004). In fact, researchers have argued that efforts to attract auto drivers to park-and-ride options will be more successful than attracting them to other modes (Habib, 2014). These schemes can induce households not to increase their use existing fleet. The underlying reason for these policies being, if the need for using cars can be reduced by other attractive alternatives, this might eventually result in reduced vehicle acquisition by households.

8.6 A Joint Econometric Analysis of Activity Flexibility, Vehicle Type Choice and Primary Driver Selection

A unified model system of activity travel choices consistent with the microeconomic utility maximization theory of behaviour was proposed in Chapter 7. The activity travel dimensions analyzed in this chapter include activity vehicle type choice, activity flexibility (temporal and spatial), and primary driver choice. The interconnectedness of the choice dimensions was examined using panel mixed multinomial model (MMNL). The data for this research was extracted from the Quebec City Travel and Activity Panel Survey (QCTAPS). Since, the dataset used in the analysis is a quite unique and rare-to-obtain longitudinal dataset (more so in the Canadian context), this study offers valuable insight into the daily activity travel patterns in a large urban context in Canada. Important findings and policy suggestions from the empirical analysis include:

- The findings reported lends credence to the notion of packaged nature of activity-travel choices, warranting simultaneous modeling of various choice dimensions in a unifying framework.
- The interdependency between the joint choice alternatives is further confirmed by the significant error component parameters.
- From the model estimation results, a clear gender related difference in vehicle type choice for activity participation is identified.
- Moreover, a pronounced location effect on activity flexibility and vehicle type choice was also captured from the analysis. It was found that individuals undertook spatially and temporally flexible activities more if either their residence or the activity location is located in the peripheral or CBD areas. However, peripheral areas were associated with the use of larger automobiles while the CBDs were associated with non-auto oriented travel modes.

8.6.1 <u>Implications and Suggested Policy Measures</u>

The model developed can be incorporated in an activity based forecasting system of travel and emissions. The analysis conducted and the model developed in this chapter is the front-end of actual GHG emission analysis. That is, once the vehicle type choice in the short term is determined in an activity based model, it can be embedded within an emission modeling module to obtain actual emissions estimates.

As suggested in previous works, increasing land use mix and population densification will increase accessibility to amenities which in turn might encourage changes in the spatial organization of activities (even travel reduction in the longer term). For example, increased concentrations of employment opportunities in the peripheral locations might be one policy to consider (Manaugh et al., 2010). Application of "road-diet"– constriction of vehicular roadway space, both at the intersection and mid-block locations, in favor of crosswalks and/or bike lanes might encourage a mode shift from private motor vehicles to more sustainable modes such as walk/bike/transit for activity participation (Buliung and Kanaroglou, 2006). Roadway infrastructural changes could be complemented by providing bike-and-ride facilities (Bachand-Marleau et al., 2011). Additionally, reduction in parking spaces or increase in the parking cost at or in the vicinity of the workplace or shopping centres might deter individuals from using larger vehicles for these activity purposes. Altogether, these policies might induce individuals not to use their private vehicles and thereby redistribute traffic resulting in reduced congestion and emission. The implementation of strategies could be effective in treated neighbourhoods; however, spillover effects to other neighborhoods could be generated.

8.7 Directions for Future Research

The preceding sections of this chapter discuss the contributions of the current dissertation in examining transportation expenditure and vehicle ownership (number and type) decision of households. Admittedly, the research efforts presented in this dissertation are not without limitations. In the following sections, we discuss the limitations of the research efforts and discuss possible extensions for the future.

8.7.1 **Limitations of the Current Research**

Comprehensive set of exogenous variables including household socio-demographic attributes, land use characteristics, and transit accessibility measures were utilized for the model development purpose. However, information on psychological factors such as individuals' lifestyles, attitudes, and preferences was not available to us (unfortunately, most travel surveys do not collect such information). If we could incorporate these factors along with the aforementioned factors, it would enhance the findings and provide greater behavioral realism. For example, several previous studies (Wu et al., 1999; Schwanen and Mokhtarian, 2005; Cao et al., 2006, Choo and Mokhtarian, 2004) have shown that people's attitudes, lifestyle preferences, and values undoubtedly play a significant role in shaping their travel choices. Another aspect related to data is that we have used cross-sectional travel survey datasets of either a single year or multiple years for our empirical analysis (except for the study presented in Chapter 7). The dataset used for the short term vehicle usage study in Chapter 7 was a unique longitudinal dataset. Nevertheless, it had a relatively small sample size which limited our capability to fully investigate the impact of time on the choice processes (temporal flexibility, spatial flexibility, vehicle type, and primary driver). Additional research on larger data samples would be useful in confirming our findings and accommodating more temporal trends. The land use variables used in the models throughout the dissertation was obtained at the traffic analysis zone (TAZ) and/or census tract (CT) levels. Admittedly, higher spatial resolution on these variables could potentially improve the results. For instance, particularly helpful additional variables include parking price or availability, crime rates, and school quality information, all by census tract.

8.7.2 **Research Extensions**

The current research endeavored to contribute to the travel behavior literature both methodologically and empirically by addressing five specific questions as discussed before. In the following section, the potential future extensions of the research are identified.

8.7.2.1 Residential Self-selection

Residential self-selection, otherwise known as residential sorting, is a phenomenon where households with a proclivity towards certain lifestyle and travel pattern may deliberately choose to live in neighborhoods conducive to their travel abilities, needs, and mobility orientations (Bhat and Guo, 2007; Pinjari et al., 2007; Wang et al., 2011). That is, residential location choice is endogenous to vehicle ownership and/or type choice. However, the vehicle ownership models developed in this dissertation do not explicitly account for the residential self-selection effects, although some form of self-selection is captured through the latent segmentation models presented in Chapter 4 and 5 – the latent segments identified could be a reflection of households' locational preferences. Therefore, the works of chapter 5 and 6 can be extended in two possible directions. First, by including attitudes as exogenous variables in the models if and when they can be observed (see Schwanen and Mokhtarian, 2005; Cao et al., 2006). Second, it would be interesting to examine whether endogeneity exists, if so, how much, by considering the vehicle fleet choice and residential location choice as occurring simultaneously, even after accounting for heterogeneity through latent segmentation models.

8.7.2.2 Intra-household Interactions

Traditionally, in the extant travel behavior literature, each household has been regarded as a single utility-maximizing agent. However, households are composed of multiple individuals who play different roles in the household and interact in many ways in the decision-making. Therefore, choice behavior of households such as decisions regarding monetary budget allocations to various necessary and discretionary expenditure categories, vehicle fleet size and vehicle type choice for activity-travel is influenced by the varying travel needs and preferences, and/or opinions by their members. The modeling frameworks proposed in the current dissertation can be extended to explore these complex intra-household interactions (interaction among the household members) and the impact of such interactions on household decision making process as well as the outcome. In the context of contemporary times, where female members of households have increased labor force participation rate and increased rate of owning driver's licenses, it would be interesting to investigate the variations in model performance and interpretations between single household-level utility based approach (as used in the current dissertation) and a multiple individual-level utility approach.

8.7.2.3 Data Fusion

Data fusion is one avenue of research which holds considerable promise in the context of travel demand modeling in light of the rising costs and associated difficulties of conducting traditional travel surveys (Stopher and Greaves, 2007). In the current dissertation, the benefits of cross-sectional data pooling was demonstrated. However, the exercise was limited to pooling of multiple year data extracted from the traditional consumer expenditure and O-D survey. It is only reasonable to say that the travel survey data needs to be supplemented by integrating it with other readily available compatible data sources if we want to understand the full continuum of household travel behavior. For example, activities and travel are inevitably linked with household expenditure. Hence, in order to understand the full spectrum of implications of fuel price change it is important not only to understand the activity travel patterns but also the budget allocation patterns of households simultaneously and this can be achieved through data fusion. Of course, integration of disparate data sources is a complex process and would require application of appropriate statistical data-stitching techniques. Additionally, fusing and supplementing traditional survey data with travel data collected through Global Positioning System (GPS) and smartphone sensors could be the next big step towards enhancing data quality to a great extent (Abdulazim et al., 2013). The econometric modeling frameworks proposed in the current dissertation can be further extended to develop larger integrated model systems using the fused data source. Another approach for enhancing the model estimations might be the integration of exogenous factors from multiple sources such as targeted marketing (TM) data, credit reporting farm data, and vehicle registration databases. Further, the study presented in Chapter 5 may be enhanced by incorporating trip length or duration and activity type into the model system. Of course that would require the formulation of a discrete-continuous modeling framework. Further, it might be beneficial to compare our categorization with models that employ activity type categorization while using spatial and temporal flexibility as independent variables.

8.7.2.4 Transferability of Models

The empirical models developed in this dissertation are based on data from different metropolitan regions of Canada. Therefore, it is not possible to generalize the results obtained from the models for other urban regions or countries. An important avenue of research would be to examine the transferability of the current model applications to other geographical contexts. The usage of wide variety of data would enable us to gain better understanding on the influence of wide variety of socio-demographics and built environments on various choice dimensions and thereby, help in informed policy formulation. In addition, the vehicle ownership models developed in the dissertation could be used as input in regional models for calculating fuel consumption rates and greenhouse gas (GHG) emissions.

APPENDIX A

Table A.1 Traditional Ordered Logit Model (OL) Estimates with ***Transit Accessibility*** Interactions (N = 5218)

Variables	Estimate	t-stat
Thresholds		
Threshold 1	0.2538	2.122
Threshold 2	5.2263	33.044
Threshold 3	9.0811	46.899
Land Use Variables		
Residential Density	-0.0075	-3.909
Transit Accessibility	-0.0085	-6.014
Entropy Index	-1.0490	-5.523
Household Demographics		
Number of Children	-2.4389	-30.272
Number of Full-time Employed Adults	0.6943	12.914
Number of Part-time Employed Adults	0.6115	4.227
Number of Students	-0.2255	-2.963
Executive Job Holder	0.4975	5.018
Number of License Holders	2.5471	36.298
Number of Transit Pass Holders	-1.0615	-14.12
Two-person Household	0.2689	3.439
Interactions		
Number of Part-time Employed Adults*Transit Accessibility	-0.0061	-1.724
Number of Retirees*Transit Accessibility	0.0042	3.557
Log-likelihood at Convergence	-3641.60	
Log-likelihood at Sample Shares	-5964.82	

Table A.2 Latent Segmentation based Ordered Logit Model (OL) Estimates with only *Transit Accessibility* as the Segmentation Variable (N = 5218)

Variables	Segment-1		Segment-2	
	Estimate	t-stat	Estimate	t-stat
Segmentation Component				
Constant	-	-	-1.6363	-10.172
Transit Accessibility	-	-	0.0134	6.039
Car Ownership Component				
Thresholds				
Threshold 1	0.6958	2.724	0.6583	0.982
Threshold 2	6.0852	26.558	8.276	9.222
Threshold 3	10.8423	37.554	14.1706	6.41
Household Demographics	-	-	-	-
Number of Full-time Employed Adults	0.8859	11.468	-	-
Number of Part-time Employed Adults	0.6297	3.883	-	-
Number of Children	-3.3168	-26.74	-2.1336	-7.471
Number of License Holders	3.2494	30.945	3.0808	12.241
Number of Students	-0.1860	-1.835	-0.6084	-2.769
Two-person Household	0.3073	2.443	-	-
Number of Transit Pass Holders	-1.0387	-10.887	-1.716	-6.9
Executive Job Holder	0.7622	5.009	-	-
Two-person Household	-	-	-0.6333	-1.62
Land Use Variables				
Residential Density	-0.0099	-3.566	-0.0131	-2.059
Entropy Index	-1.4137	-5.011	-1.5671	-1.935
Log-likelihood at Convergence	-3568.22			
Log-likelihood at Sample Shares	-5964.82			

Table A.3 Traditional Multinomial Logit Model (MNL) Estimates with *Transit Accessibility* Interactions (N = 5218)

Variables	Estimate	t-stat
Constants		
Constant 1	-1.8042	-5.916
Constant 2	-5.952	-16.707
Constant 3	-11.7669	-23.932
Land Use Variables		
Residential Density		
1 Car	-0.0062	-1.862
2 Cars	-0.0135	-3.316
≥ 3 Cars	-0.0163	-2.490
Entropy Index		
1 Car	-1.5132	-3.368
2 Cars	-2.2993	-4.618
≥ 3 Cars	-3.6049	-5.648
Household Demographics		
Number of Children		
1 Car	-3.9217	-11.205
2 Cars	-5.7777	-15.711
≥ 3 Cars	-7.7994	-19.737
Number of Part-time Employed Adults		
1 Car	-	-
2 Cars	0.5244	3.969
≥ 3 Cars	0.7539	3.378
Number of Full-time Employed Adults		
1 Car	1.0900	6.735
2 Cars	1.7981	10.131
≥ 3 Cars	2.0468	9.801
Executive Job Holder		
1 Car	-	-
2 Cars	0.3804	3.277
≥ 3 Cars	0.8097	4.318
Number of Retirees		
1 Car	0.3561	1.849
2 Cars	0.3864	1.805
≥ 3 Cars	0.5452	2.124
Number of License Holders		
1 Car	4.2629	18.613
2 Cars	6.2085	25.023
≥ 3 Cars	7.9213	29.222
Number of Transit Pass Holders		
1 Car	-1.4514	-10.137
2 Cars	-2.3630	-14.219
≥ 3 Cars	-3.4770	-14.691
Interactions		
Number of License Holders*Transit Accessibility		
1 Car	-0.0124	-5.636
2 Cars	-0.0163	-6.893
≥ 3 Cars	-0.0173	-6.528
Number of Children*Transit Accessibility		
1 Car	0.0139	2.958
2 Cars	0.0172	3.360
≥ 3 Cars	0.0202	2.966

Number of Retirees*Transit Accessibility		
1 Car	5.5952	2.009
2 Cars	9.4581	2.891
≥ 3 Cars	9.4581	2.891
Log-likelihood at Convergence	-3607.67	
Log-likelihood at Sample Shares	-5964.82	

Table A.4 Latent Segmentation based Multinomial Logit Model (OL) Estimates with only *Transit Accessibility* as Segmentation Variable (N = 5218)

Variables	Segment-1		Segment-2	
	Estimate	t-stat	Estimate	t-stat
Segmentation Component				
Constant	-	-	-2.4578	-15.551
Transit Accessibility	-	-	0.0124	6.071
Car Ownership Component				
Constants				
Constant 1	-4.8171	-4.571	-2.3356	-3.216
Constant 2	-10.5113	-9.547	-7.3323	-2.714
Constant 3	-17.5067	-14.762	-4.0809	-5.186
Land Use Variables				
Residential Density				
1 Car	-0.0155	-2.192	-	-
2 Cars	-0.0272	-3.551	-	-
≥ 3 Cars	-0.0337	-3.567	-	-
Entropy Index				
1 Car	-2.7973	-2.17	-3.4785	-3.352
2 Cars	-4.0389	-3.066	-	-
≥ 3 Cars	-5.6967	-4.103	-	-
Household Demographics				
Number of Children				
1 Car	-8.964	-6.828	-	-
2 Cars	-11.8039	-8.796	-	-
≥ 3 Cars	-14.0273	-10.36	-	-
Number of Part-time Employed Adults				
1 Car	-	-	-	-
2 Cars	0.6543	3.807	-	-
≥ 3 Cars	1.055	3.906	-	-
Number of Full-time Employed Adults				
1 Car	2.0097	5.179	1.342	4.336
2 Cars	2.9361	7.331	-	-
≥ 3 Cars	3.3985	8.066	-	-
Executive Job Holder				
1 Car	-	-	-	-
2 Cars	0.6729	3.68	-	-
≥ 3 Cars	1.2247	4.977	-	-
Number of Retirees				
0 Car	-	-	1.2385	3.886
1 Car	1.5333	3.11	-	-
2 Cars	1.676	3.336	-	-
≥ 3 Cars	1.9844	3.764	-	-
Number of License Holders				
1 Car	8.627	6.888	1.7672	6.314
2 Cars	11.4739	8.987	1.9901	1.945
≥ 3 Cars	13.4881	10.482	1.095	2.114
Number of Transit Pass Holders				
1 Car	-2.1339	-7.158	-	-
2 Cars	-3.1675	-9.926	-	-
≥ 3 Cars	-4.4106	-11.789	-	-
Log-likelihood at Convergence	-3482.08			
Log-likelihood at Sample Shares	-5964.82			

Table A.5 Traditional Ordered Logit Model (OL) Estimates with All Variables (N = 5218)

Variables	**Estimate**	**t-stat**
Thresholds		
Threshold 1	0.2107	1.768
Threshold 2	5.1967	32.64
Threshold 3	9.0296	46.626
Land Use Variables		
Transit Accessibility	-0.0073	-5.599
Entropy Index	-1.0536	-5.555
Residential Density	-0.0073	-3.809
Household Demographics		
Number of Transit Pass Holders	-1.0727	-14.238
Number of Household Members		
Two persons	0.4634	4.063
More than two persons	0.3516	2.171
Number of Children	-2.4805	-31.269
Number of Full-time Employed Adults	0.5816	12.001
Number of Part-time Employed Adults	0.3369	3.205
Number of Students	-0.3068	-4.027
Number of License Holders	2.5195	33.326
Executive Job Holder	0.4887	4.942
Log-likelihood at Convergence	-3647.70	
Log-likelihood at Sample Shares	-5964.82	

Table A.6 Traditional Multinomial Logit Model (MNL) Estimates with All Variables (N = 5218)

Variables	Estimate	t-stat
Constants		
Constant 1	-1.4137	-5.109
Constant 2	-5.3681	-16.354
Constant 3	-11.1327	-23.854
Land Use Variables		
Residential Density		
1 Car	-0.0060	-1.748
2 Cars	-0.0128	-3.083
≥ 3 Cars	-0.0174	-2.769
Transit Accessibility		
1 Car	-0.0089	-3.738
2 Cars	-0.0159	-5.591
≥ 3 Cars	-0.0159	-5.591
Entropy Index		
1 Car	-1.4165	-3.186
2 Cars	-2.1731	-4.399
≥ 3 Cars	-3.5333	-5.63
Household Demographics		
Number of Children		
1 Car	-3.1663	-13.176
2 Cars	-4.9127	-18.926
≥ 3 Cars	-6.8824	-23.876
Number of Part-time Employed Adults		
1 Car	-	-
2 Cars	0.5231	3.968
≥ 3 Cars	0.7516	3.374
Number of Full-time Employed Adults		
1 Car	1.0842	6.728
2 Cars	1.7922	10.133
≥ 3 Cars	2.0430	9.812
Executive Job Holder		
1 Car	-	-
2 Cars	0.3786	3.272
≥ 3 Cars	0.8035	4.295
Number of Retirees		
1 Car	0.6376	4.621
2 Cars	0.7832	4.934
≥ 3 Cars	0.9313	4.373
Number of License Holders		
1 Car	3.6048	19.696
2 Cars	5.4324	26.646
≥ 3 Cars	7.1236	31.016
Number of Transit Pass Holders		
1 Car	-1.4623	-10.089
2 Cars	-2.3874	-14.251
≥ 3 Cars	-3.5138	-14.812
Log-likelihood at Convergence	-3619.15	
Log-likelihood at Sample Shares	-5964.82	

REFERENCES

Abdulazim, T., Abdelgawad, H., Habib, K. M. N., & Abdulhai, B. (2013). Using smartphones and sensor technologies to automate the collection of travel data. *Transportation Research Record: Journal of the Transportation Research Board, 2383*, 44-52.

Aditjandra, P. T., Cao, X., & Mulley, C. (2012). Understanding neighborhood design impact on travel behaviour: An application of structural equations model to a British metropolitan data. *Transportation Research Part A: Policy and Practice, 46*(1), 22-32.

Adjemian, M. K., Cynthia Lin, C. Y., & Williams, J. (2010). Estimating spatial interdependence in automobile type choice with survey data. *Transportation Research Part A: Policy and Practice, 44*(9), 661-675.

Ahn, J., Jeong, G., & Kim, Y. (2008). A forecast of household ownership and use of alternative fuel vehicles: A multiple discrete-continuous choice approach. *Energy Economics, 30*(5), 2091-2104.

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics, 24*(1-2), 3-61.

Anas, A. (2007). A unified theory of consumption, travel and trip chaining. *Journal of Urban Economics, 62*(2), 162-186.

Anastasopoulos, P., Karlaftis, M., Haddock, J., & Mannering, F. (2012). Household automobile and motorcycle ownership analyzed with random parameters bivariate ordered probit model. *Transportation Research Record: Journal of the Transportation Research Board, 2279*, 12-20.

Asensio, J., Matas, A., & Raymond, J.-L. (2003a). Petrol expenditure and redistributive effects of its taxation in Spain. *Transportation Research Part A: Policy and Practice, 37*(1), 49-69.

Asensio, J., Matas, A., & Raymond, J.-L. (2003b). Redistributive effects of subsidies to urban public transport in Spain. *Transport Reviews, 23*(4), 433-452.

Axsen, J., & Kurani, K. S. (2012). Interpersonal influence within car buyers' social networks: applying five perspectives to plug-in hybrid vehicle drivers. *Environment and Planning A, 44*(5), 1047-1065.

Bachand-Marleau, J., Larsen, J., & El-Geneidy, A. M. (2011). Much-anticipated marriage of cycling and transit: How will it work? *Transportation Research Record: Journal of the Transportation Research Board, 2247*, 109-117.

Badland, H., & Schofield, G. (2005). Transport, urban design, and physical activity: an evidence-based update. *Transportation Research Part D: Transport and Environment, 10*(3), 177-196.

Başar, G., & Bhat, C. (2004). A parameterized consideration set model for airport choice: An application to the San Francisco Bay Area. *Transportation Research Part B: Methodological, 38*(10), 889-904.

Beck, M. J., Rose, J. M., & Hensher, D. A. (2013). Environmental attitudes and emissions charging: An example of policy implications for vehicle choice. *Transportation Research Part A: Policy and Practice, 50*, 171-182.

Beckman, J., & Goulias, K. (2008). Immigration, residential location, car ownership, and commuting behavior: a multivariate latent class analysis from California. *Transportation, 35*(5), 655-671.

Bento, A. M., Cropper, M. L., Mobarak, A. M., & Vinha, K. (2005). The effects of urban spatial structure on travel demand in the United States. *Review of Economics and Statistics, 87*(3), 466-478.

Bhat, C. R. (1997). An endogenous segmentation mode choice model with an application to intercity travel. *Transportation science, 31*(1), 34-48.

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

Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological, 35*(7), 677-693.

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

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

Bhat, C. R. (2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B: Methodological, 45*(7), 923-939.

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

Bhat, C. R., & Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B: Methodological, 43*(7), 749-765.

Bhat, C. R., & Eluru, N. (2010). The Multiple discrete-continuous extreme value (mdcev) model: formulation and applications. In S. Hess & A. Daly (Eds.), *Choice Modelling: The State-of-the-Art and the State-of-Practice, Proceedings from the inaugural International Choice Modelling Conference* (pp. 71-100): Emerald Group Publishing Limited.

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

Bhat, C. R., & Koppelman, F. S. (1993). An endogenous switching simultaneous equation system of employment, income, and car ownership. *Transportation Research Part A: Policy and Practice, 27*(6), 447-459.

Bhat, C. R., & Pinjari, A. R. (2008). Duration modeling. In D. A. Hensher & K. J. Button (Eds.), *Handbook of transport modelling* (2nd ed., pp. 91-112). New York: Pergamon.

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

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

Bhat, C. R., Eluru, N., & Copperman, R. B. (2008). Flexible model structures for discrete choice analysis. In D. A. Hensher & K. Button (Eds.), *Handbook of transport modelling* (2nd ed., Vol. 1, pp. 75-104). New York: Pergamon.

Bhat, C. R., Paleti, R., Pendyala, R. M., Lorenzini, K., & Konduri, K. C. (2013). Accommodating immigration status and self-selection effects in a joint model of household auto ownership and residential location choice. *Transportation Research Record: Journal of the Transportation Research Board, 2382*, 142-150.

Bhat, C. R., Sen, S., & Eluru, N. (2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B: Methodological, 43*(1), 1-18.

Bhat, C., & Lockwood, A. (2004). On distinguishing between physically active and physically passive episodes and between travel and activity episodes: an analysis of weekend recreational participation in the San Francisco Bay area. *Transportation Research Part A: Policy and Practice, 38*(8), 573-592.

Bjørner, T. B., & Leth-Petersen, S. (2007). A dynamic random effects multinomial logit model of household car ownership. *Nationaløkonomisk Tidsskrift, 145*(1), 83-100.

Browning, M. (1992). Children and household economic behavior. *Journal of Economic Literature*, 1434-1475.

Brownstone, D., & Fang, H. A. (2009). *Vehicle ownership and utilization choice model with endogenous residential density*. Paper presented at the 88th Annual Meeting of the Transportation Research Board Washington, D.C.

Brownstone, D., & Golob, T. F. (2009). The impact of residential density on vehicle usage and energy consumption. *Journal of Urban Economics, 65*(1), 91-98.

Bujosa, A., Riera, A., & Hicks, R. (2010). Combining discrete and continuous representations of preference heterogeneity: A latent class approach. *Environmental and Resource Economics, 47*(4), 477-493.

Buliung, R. N., & Kanaroglou, P. S. (2006). Urban form and household activity-travel behavior. *Growth and Change, 37*(2), 172-199.

Bunch, D. S. (2000). Automobile demand and type choice. In D. A. Hensher & K. Button (Eds.), *Handbook of transport modelling* (1st ed., Vol. 1, pp. 463-479). New York: Pergamon.

Canada, N. R. (2009). Canadian vehicle survey summary report. Retrieved March 30, 2015, from http://oee.nrcan.gc.ca/publications/statistics/cvs09/pdf/cvs09.pdf

Canada, T. (2006). The cost of urban congestion in Canada. Retrieved March 30, 2015, from http://www.adec-inc.ca/pdf/02-rapport/cong-canada-ang.pdf

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2006). Neighborhood design and vehicle type choice: Evidence from Northern California. *Transportation Research Part D: Transport and Environment, 11*(2), 133-145.

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2007a). Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. *Transportation, 34*(5), 535-556.

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2007b). Cross-sectional and quasi-panel explorations of the connection between the built environment and auto ownership. *Environment and Planning A, 39*(4), 830-847.

Castro, M., Paleti, R., & Bhat, C. R. (2013). A spatial generalized ordered response model to examine highway crash injury severity. *Accident Analysis & Prevention, 52*, 188-203.

Caulfield, B. (2012). An examination of the factors that impact upon multiple vehicle ownership: The case of Dublin, Ireland. *Transport Policy, 19*(1), 132-138.

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

Chakour, V., & Eluru, N. (2014). Analyzing commuter train user behavior: a decision framework for access mode and station choice. *Transportation, 41*(1), 211-228.

Chan, S., Miranda-Moreno, L. F., Patterson, Z., & Barla, P. (2013). *Spatial Analysis of Demand for Hybrid Electric Vehicles and Its Potential Impact on Greenhouse Gases in Montreal and Quebec City, Canada.* Paper presented at the Transportation Research Board 92nd Annual Meeting.

Chapman, L. (2007). Transport and climate change: a review. *Journal of Transport Geography, 15*(5), 354-367.

Chen, C., Gong, H., & Paaswell, R. (2008). Role of the built environment on mode choice decisions: Additional evidence on the impact of density. *Transportation, 35*(3), 285-299.

Chin, A., & Smith, P. (1997). Automobile ownership and government policy: The economics of Singapore's vehicle quota scheme. *Transportation Research Part A: Policy and Practice, 31*(2), 129-140.

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

Choo, S., Lee, T., & Mokhtarian, P. (2007). Do transportation and communications tend to be substitutes, complements, or neither?: U.S. consumer expenditures perspective, 1984-2002. *Transportation Research Record: Journal of the Transportation Research Board, 2010*, 121-132.

Chu, Y.-L. (2002). Automobile ownership analysis using ordered probit models. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 60-67.

Clark, S. D. (2007). Estimating local car ownership models. *Journal of Transport Geography, 15*(3), 184-197.

Cullinane, S. (2002). The relationship between car ownership and public transport provision: A case study of Hong Kong. *Transport Policy, 9*(1), 29-39.

Dargay, J. M. (2002). Determinants of car ownership in rural and urban areas: A pseudo-panel analysis. *Transportation Research Part E: Logistics and Transportation Review, 38*(5), 351-366.

Dargay, J. M., & Vythoulkas, P. C. (1999). Estimation of a dynamic car ownership model: A pseudo-panel approach. *Journal of Transport Economics and Policy, 33*(3), 287-302.

Dargay, J., & Hanly, M. (2007). Volatility of car ownership, commuting mode and time in the UK. *Transportation Research Part A: Policy and Practice, 41*(10), 934-948.

Daziano, R. A., & Bolduc, D. (2013). Incorporating pro-environmental preferences towards green automobile technologies through a Bayesian hybrid choice model. *Transportmetrica A: Transport Science, 9*(1), 74-106.

de Abreu e Silva, J., Morency, C., & Goulias, K. G. (2012). Using structural equations modeling to unravel the influence of land use patterns on travel behavior of workers in Montreal. *Transportation Research Part A: Policy and Practice, 46*(8), 1252-1264.

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

de Jong, G. C., & Kitamura, R. (2009). A review of household dynamic vehicle ownership models: Holdings models versus transactions models. *Transportation, 36*(6), 733-743.

de Jong, G., Fox, J., Daly, A., Pieters, M., & Smit, R. (2004). Comparison of car ownership models. *Transport Reviews, 24*(4), 379-408.

de Lapparent, M., & Cernicchiaro, G. (2012). How long to own and how much to use a car? A dynamic discrete choice model to explain holding duration and driven mileage. *Economic Modelling, 29*(5), 1737-1744.

de Lorimier, A., & El-Geneidy, A. M. (2013). Understanding the factors affecting vehicle usage and availability in carsharing networks: A case study of Communauto carsharing system from Montréal, Canada. *International Journal of Sustainable Transportation, 7*(1), 35-51.

Depaire, B., Wets, G., & Vanhoof, K. (2008). Traffic accident segmentation by means of latent class clustering. *Accident Analysis & Prevention, 40*(4), 1257-1266.

Dieleman, F. M., & Everaers, P. C. (1994). From renting to owning: life course and housing market circumstances. *Housing Studies, 9*(1), 11-25.

Dissanayake, D., & Morikawa, T. (2002). Household travel behavior in developing countries: Nested logit model of vehicle ownership, mode choice, and trip chaining. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 45-52.

Doherty, S. T. (2006). Should we abandon activity type analysis? Redefining activities by their salient attributes. *Transportation, 33*(6), 517-536.

Dupuy, G. (1999). From the 'magic circle' to 'automobile dependence': Measurements and political implications. *Transport Policy, 6*(1), 1-17.

Eluru, N. (2013). Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. *Accident Analysis & Prevention, 55*, 1-11.

Eluru, N., & Bhat, C. R. (2007). A joint econometric analysis of seat belt use and crash-related injury severity. *Accident Analysis and Prevention, 39*(5), 1037-1049.

Eluru, N., Bagheri, M., Miranda-Moreno, L. F., & Fu, L. (2012a). A latent class modeling approach for identifying vehicle driver injury severity factors at highway-railway crossings. *Accident Analysis & Prevention, 47*, 119-127.

Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention, 40*(3), 1033-1054.

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

Eluru, N., Chakour, V., & El-Geneidy, A. (2012b). Travel mode choice and transit route choice behavior in Montreal: insights from McGill University members commute patterns. *Public Transport, 4*(2), 129-149.

Eluru, N., Pinjari, A., Pendyala, R., & Bhat, C. (2010b). An econometric multi-dimensional choice model of activity-travel behavior. *Transportation Letters, 2*(4), 217-230.

Ewing, R., & Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association, 76*(3), 265-294.

Faghih-Imani, A., Ghafghazi, G., Eluru, N., & Pinjari, A. R. (2014). A multiple-discrete approach for examining vehicle type use for daily activity participation decisions. *Transportation Letters: The International Journal of Transportation Research, 6*(1), 1-13.

Fang, H. A. (2008). A discrete–continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B: Methodological, 42*(9), 736-758.

Ferdous, N., Pinjari, A., Bhat, C., & Pendyala, R. (2010). A comprehensive analysis of household transportation expenditures relative to other goods and services: an application to United States consumer expenditure data. *Transportation, 37*(3), 363-390.

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Marketing Science, 29*(3), 393-421.

Fousekis, P., & Lazaridis, P. (2001). Food expenditure patterns of the urban and the rural households in Greece: A kernel regression analysis. *Agricultural Economic Review, 2*(1), 5-19.

Frank, L., Bradley, M., Kavage, S., Chapman, J., & Lawton, T. K. (2008). Urban form, travel time, and cost relationships with tour complexity and mode choice. *Transportation, 35*(1), 37-54.

Fujii, S., & Taniguchi, A. (2005). Reducing family car-use by providing travel advice or requesting behavioral plans: An experimental analysis of travel feedback programs. *Transportation Research Part D: Transport and Environment, 10*(5), 385-393.

Gaker, D., Vautin, D., Vij, A., & Walker, J. L. (2011). The power and value of green in promoting sustainable transport behavior. *Environmental Research Letters, 6*(3), 1-10.

Gao, S., Mokhtarian, P. L., & Johnston, R. A. (2008). Exploring the connections among job accessibility, employment, income, and auto ownership using structural equation modeling. *Annals of Regional Science, 42*(2), 341-356.

García-Palomares, J. C. (2010). Urban sprawl and travel to work: the case of the metropolitan area of Madrid. *Journal of Transport Geography, 18*(2), 197-213.

Gaundry, M. J. I., & Dagenais, M. G. (1979). The dogit model. *Transportation Research Part B: Methodological, 13*(2), 105-111.

Gilbert, C. C. S. (1992). A duration model of automobile ownership. *Transportation Research Part B: Methodological, 26*(2), 97-114.

Gilbert, R. (2003). Greater Toronto area comparisons. Retrieved March 30, 2015, from http://www.richardgilbert.ca/Files/2003/Greater%20Toronto%20Area%20Comparisons.pdf

Giuliano, G. (2004). Land use impacts of transportation investments: Highway and transit. In S. Hanson & G. Giuliano (Eds.), *The geography of urban transportation* (3rd ed., pp. 237-273). New York: Guilford Press.

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

Golob, T. F. (1990). The dynamics of household travel time expenditures and car ownership decisions. *Transportation Research Part A: General, 24*(6), 443-463.

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

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

Goulias, K. G., Pendyala, R. M., & Kitamura, R. (1990). Practical method for the estimation of trip generation and trip chaining. *University of California Transportation Center*.

Greene, W. H. (2003). *Econometric analysis*. Upper Saddle River, N.J.: Prentice Hall.

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological, 37*(8), 681-698.

Guo, Z. (2013). Does residential parking supply affect household car ownership? The case of New York City. *Journal of Transport Geography, 26*(0), 18-28.

Haas, P., Makarewicz, C., Benedict, A., & Bernstein, S. (2008). Estimating transportation costs by characteristics of neighborhood and household. *Transportation Research Record: Journal of the Transportation Research Board, 2077*, 62-70.

Habib, K. (2014). Household-level commuting mode choices, car allocation and car ownership level choices of two-worker households: the case of the city of Toronto. *Transportation, 41*(3), 651-672.

Handy, S. (1993). A cycle of dependence: automobiles, accessibility, and the evolution of the transportation and retail hierarchies. *Berkeley Planning Journal, 8*, 21-43.

Handy, S., Cao, X., & Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D: Transport and Environment, 10*(6), 427-444.

Hanly, M., & Dargay, J. (2000). Car ownership in Great Britain: Panel data analysis. *Transportation Research Record: Journal of the Transportation Research Board, 1718*, 83-89.

Hensher, D. A. (1992). *Dimensions of automobile demand: a longitudinal study of household automobile ownership and use*. Amsterdam; New York; New York, N.Y., U.S.A.: North-Holland

Hensher, D. A., & Mannering, F. L. (1994). Hazard-based duration models and their application to transport analysis. *Transport Reviews, 14*(1), 63-82.

Hess, D. B., & Ong, P. (2002). Traditional neighborhoods and automobile ownership. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 35-44.

Holtzclaw, J., Clear, R., Dittmar, H., Goldstein, D., & Haas, P. (2002). Location efficiency: Neighborhood and socio-economic characteristics determine auto ownership and use - Studies in Chicago, Los Angeles and San Francisco. *Transportation Planning and Technology, 25*(1), 1-27.

Hong, G.-S., Fan, J. X., Palmer, L., & Bhargava, V. (2005). Leisure travel expenditure patterns by family life cycle stages. *Journal of Travel & Tourism Marketing, 18*(2), 15-30.

Houthakker, H. S. (1957). An international comparison of household expenditure patterns, commemorating the centenary of Engel's law. *Econometrica, Journal of the Econometric Society*, 532-551.

IPCC. (2007). Climate change 2007: Synthesis report. Retrieved March 30, 2015, from http://www.ipcc.ch/publications_and_data/ar4/syr/en/spm.html

Karlaftis, M., & Golias, J. (2002). Automobile ownership, households without automobiles, and urban traffic parameters: Are they related? *Transportation Research Record: Journal of the Transportation Research Board, 1792*, 29-35.

Kenworthy, J. (1991). The land use and transit connection in Toronto. *Australian Planner, 29*(3), 149-154.

Kenworthy, J. R., & Laube, F. B. (1999). Patterns of automobile dependence in cities: An international overview of key physical and economic dimensions with some implications for urban policy. *Transportation Research Part A: Policy and Practice, 33*(7-8), 691-723.

Kermanshah, M., & Ghazi, F. (2001). Modeling automobile ownership decisions: A disaggregate approach. *Scientia Iranica, 8*(1), 29-37.

Kim, H. S., & Kim, E. (2004). Effects of public transit on automobile ownership and use in households of the USA. *Review of Urban and Regional Development Studies, 16*(3), 245-260.

Kim, J. K., Kim, S., Ulfarsson, G. F., & Porrello, L. A. (2007). Bicyclist injury severities in bicycle-motor vehicle accidents. *Accident Analysis and Prevention, 39*(2), 238-251.

Kitamura, R. (2000). Longitudinal Methods. In D. A. Hensher & K. Button (Eds.), *Handbook of transport modelling* (1st ed., Vol. 1, pp. 113–129). New York: Pergamon.

Kitamura, R. (2009). A dynamic model system of household car ownership, trip generation, and modal split: model development and simulation experiment. *Transportation, 36*(6), 711-732.

Kitamura, R., & Bunch, D. S. (1990). *Heterogeneity and state dependence in household car ownership: a panel analysis using ordered-response probit models with error components*. Paper presented at the Eleventh International Symposium on Transportation and Traffic Theory, Yokohama, Japan.

Kitamura, R., Akiyama, T., Yamamoto, T., & Golob, T. (2001). Accessibility in a metropolis: Toward a better understanding of land use and travel. *Transportation Research Record: Journal of the Transportation Research Board, 1780*, 64-75.

Kockelman, K. (1997). Travel behavior as function of accessibility, land use mixing, and land use balance: Evidence from san francisco bay area. *Transportation Research Record: Journal of the Transportation Research Board, 1607*, 116-125.

Komanoff, C., & Pucher, J. (2003). Bicycle transport in the US: Recent trends and policies *Sustainable Transport* (pp. 450-464).

Konduri, K., Astroza, S., Sana, B., Pendyala, R., & Jara-Díaz, S. (2011). Joint analysis of time use and consumer expenditure data. *Transportation Research Record: Journal of the Transportation Research Board, 2231*, 53-60.

Konduri, K., Ye, X., Sana, B., & Pendyala, R. (2011). Joint model of vehicle type choice and tour length. *Transportation Research Record: Journal of the Transportation Research Board, 2255*, 28-37.

Koppelman, F. S., & Sethi, V. (2008). Closed-form Discrete Choice Models. In D. A. Hensher & K. J. Button (Eds.), *Handbook of Transport Modelling* (2nd ed., Vol. 1, pp. 211-227). New York: Pergamon.

Kuhnimhof, T., & Gringmuth, C. (2009). Multiday multiagent model of travel behavior with activity scheduling. *Transportation Research Record: Journal of the Transportation Research Board, 2134*, 178-185.

Kuhnimhof, T., Zumkeller, D., & Chlond, B. (2013). Who made peak car, and how? A breakdown of trends over four decades in four countries. *Transport Reviews, 33*(3), 325-342.

Kuppam, A., & Pendyala, R. (2001). A structural equations analysis of commuters' activity and travel patterns. *Transportation, 28*(1), 33-54.

Larsen, J., El-Geneidy, A., & Yasmin, F. (2010). Beyond the quarter mile: re-examining travel distances by active transportation. *Canadian Journal of Urban Research, 19*(1), 70-88.

Lee, B., & Timmermans, H. J. P. (2007). A latent class accelerated hazard model of activity episode durations. *Transportation Research Part B: Methodological, 41*(4), 426-447.

Lee-Gosselin, M. E. H. (2005). A data collection strategy for perceived and observed flexibility in the spatio-temporal organization of household activities and associated travel. In H. J. P. Timmermans (Ed.), *Progress in activity-based analysis*. Amsterdam; San Diego, CA; Oxford: Elsevier.

Lee-Gosselin, M., & Miranda-Moreno, L. F. (2009). What is different about urban activities of those with access to ICTs? Some early evidence from Québec, Canada. *Journal of Transport Geography, 17*(2), 104-114.

Legrain, A., Buliung, R., & Ahmed, M. (2015). Who, what, when, and where: revisiting the influences of transit mode share. *Transportation Research Record (forthcoming)*.

Li, J., Walker, J., Srinivasan, S., & Anderson, W. (2010). Modeling private car ownership in China. *Transportation Research Record: Journal of the Transportation Research Board, 2193*, 76-84.

Litman, T. (2002). The costs of automobile dependency and the benefits of balanced transportation. Retrieved March 30, 2015, from http://www.brocku.ca/tren/courses/tren3p18/Litman%202002%20-%20The%20Costs%20of%20Automobile%20Dependency.pdf

Litman, T. (2003). *Social inclusion as a transport planning issue in Canada.* Paper presented at the Transport and Social Exclusion G7 Comparison Seminar. London.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice, 44*(5), 291-305.

Ma, L., & Srinivasan, S. (2010). Impact of individuals' immigrant status on household auto ownership. *Transportation Research Record: Journal of the Transportation Research Board, 2156*, 36-46.

Maat, K., & Timmermans, H. J. P. (2009). Influence of the residential and work environment on car use in dual-earner households. *Transportation Research Part A: Policy and Practice, 43*(7), 654-664.

Macfarlane, G. S., Garrow, L. A., & Mokhtarian, P. L. (2015). The influences of past and present residential locations on vehicle ownership decisions. *Transportation Research Part A: Policy and Practice, 74*, 186-200.

Manaugh, K., Miranda-Moreno, L., & El-Geneidy, A. (2010). The effect of neighbourhood characteristics, accessibility, home–work location, and demographics on commuting distances. *Transportation, 37*(4), 627-646.

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

Manski, C. F., & Sherman, L. (1980). An empirical analysis of household choice among motor vehicles. *Transportation Research, Part A (General), 14A*(5-6), 349-366.

Matas, A., & Raymond, J.-L. (2008). Changes in the structure of car ownership in Spain. *Transportation Research Part A: Policy and Practice, 42*(1), 187-202.

McCarthy, P. S., & Tay, R. S. (1998). New vehicle consumption and fuel efficiency: A nested logit approach. *Transportation Research Part E: Logistics and Transportation Review, 34*(1), 39-51.

Millard-Ball, A., & Schipper, L. (2011). Are we reaching peak travel? Trends in passenger transport in eight industrialized countries. *Transport Reviews, 31*(3), 357-378.

Milligan, K. (2008). The evolution of elderly poverty in Canada. *Canadian Public Policy, 34*(4), S79-S94.

Miranda-Moreno, L. F., & Lee-Gosselin, M. (2008). A week in the life of baby boomers: How do they see the spatial–temporal organization of their activities and travel? *Transportation, 35*(5), 629-653.

Miranda-Moreno, L., Bettex, L., Zahabi, S., Kreider, T., & Barla, P. (2011). Simultaneous modeling of endogenous influence of urban form and public transit accessibility on distance traveled. *Transportation Research Record: Journal of the Transportation Research Board, 2255*, 100-109.

Mohammadian, A., & Doherty, S. (2005). Mixed logit model of activity-scheduling time horizon incorporating spatial-temporal flexibility variables. *Transportation Research Record: Journal of the Transportation Research Board, 1926*, 33-40.

Mohammadian, A., & Doherty, S. T. (2006). Modeling activity scheduling time horizon: duration of time between planning and execution of pre-planned activities. *Transportation Research Part A: Policy and Practice, 40*(6), 475-490.

Mohammadian, A., & Miller, E. (2002). Nested logit models and artificial neural networks for predicting household automobile choices: Comparison of performance. *Transportation Research Record: Journal of the Transportation Research Board, 1807*, 92-100.

Mohammadian, A., & Miller, E. (2003a). Empirical investigation of household vehicle type choice decisions. *Transportation Research Record: Journal of the Transportation Research Board, 1854*, 99-106.

Mohammadian, A., & Miller, E. (2003b). Dynamic modeling of household automobile transactions. *Transportation Research Record: Journal of the Transportation Research Board, 1831*, 98-105.

Mohammadian, A., & Rashidi, T. (2007). Modeling household vehicle transaction behavior, competing risk duration approach. *Transportation Research Record: Journal of the Transportation Research Board, 2014*, 9-16.

Mohan, D. (2002). Road safety in less-motorized environments: future concerns. *International Journal of Epidemiology, 31*(3), 527-532.

Moniruzzaman, M., & Páez, A. (2012). Accessibility to transit, by transit, and mode share: application of a logistic model with spatial filters. *Journal of Transport Geography, 24*(0), 198-205.

Montreal, C. M. d. (2010). Portrait of Greater Montreal *Metropolitan* (Vol. 1).

Moriarty, P. (2002). Household travel time and money expenditures. *Road and Transport Research, 11*(4), 14-23.

Moudon, A., Hess, P., Snyder, M., & Stanilov, K. (1997). Effects of site design on pedestrian travel in mixed-use, medium-density environments. *Transportation Research Record: Journal of the Transportation Research Board, 1578*, 48-55.

Newman, P., & Kenworthy, J. (1999). Costs of automobile dependence: Global survey of cities. *Transportation Research Record: Journal of the Transportation Research Board, 1670*, 17-26.

Nobile, A., Bhat, C. R., & Pas, E. I. (1997). A random-effects multinomial probit model of car ownership choice. In C. Gatsonis, J. Hodges, R. Kass, R. McCulloch, P. Rossi, & N. Singpurwalla (Eds.), *Case Studies in Bayesian Statistics* (Vol. 121, pp. 419-434): Springer New York.

Nobis, C. (2007). Multimodality: Facets and causes of sustainable mobility behavior. *Transportation Research Record: Journal of the Transportation Research Board, 2010*, 35-44.

Nolan, A. (2003). The determinants of urban household's transport decisions: A microeconometric study using Irish data. *International Journal of Transport Economics, 30*(1), 103-129.

Nolan, A. (2010). A dynamic analysis of household car ownership. *Transportation Research Part A: Policy and Practice, 44*(6), 446-455.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535-569.

Oladosu, G. (2003). An almost ideal demand system model of household vehicle fuel expenditure allocation in the United States. *The Energy Journal*, 1-21.

Paleti, R., Bhat, C. R., & Pendyala, R. M. (2013a). *An integrated model of residential location, work location, vehicle ownership, and commute tour characteristics.* Paper presented at the 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., USA.

Paleti, R., Bhat, C. R., Pendyala, R. M., & Goulias, K. G. (2013b). *The modeling of household vehicle type choice accommodating spatial dependence effects*. Paper presented at the 92nd Annual Meeting of the Transportation Research Board, Washington D.C., USA.

Paleti, R., Eluru, N., Bhat, C., Pendyala, R., Adler, T., & Goulias, K. (2011). design of comprehensive microsimulator of household vehicle fleet composition, utilization, and evolution. *Transportation Research Record: Journal of the Transportation Research Board, 2254*, 44-57.

Paleti, R., Pendyala, R. M., Bhat, C. R., & Konduri, K. C. (2012). *A joint tour-based model of tour complexity, passenger accompaniment, vehicle type choice, and tour length.* Paper presented at the 91st Annual Meeting of the Transportation Research Board, Washington, D.C.

Paleti, R., Pendyala, R. M., Bhat, C. R., Lorenzini, K., & Konduri, K. C. (2013c). *Accommodating immigration status and self-selection effects in a joint model of household auto ownership and residential location choice.* Paper presented at the 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., USA.

Palm, M., Gregor, B., Wang, H., & McMullen, B. S. (2014). The trade-offs between population density and households' transportation-housing costs. *Transport Policy, 36*, 160-172.

Paulin, G. D. (1995). A comparison of consumer expenditures by housing tenure. *Journal of Consumer Affairs, 29*(1), 164-198.

Pendyala, R. M., Kostyniuk, L. P., & Goulias, K. G. (1995). A repeated cross-sectional evaluation of car ownership. *Transportation, 22*(2), 165-184.

Pendyala, R., & Bhat, C. (2004). An exploration of the relationship between timing and duration of maintenance activities. *Transportation, 31*(4), 429-456.

Pinjari, A. R. (2011). Generalized extreme value (GEV)-based error structures for multiple discrete-continuous choice models. *Transportation Research Part B: Methodological, 45*(3), 474-489.

Pinjari, A. R., Pendyala, R. M., Bhat, C. R., & Waddell, P. A. (2007). Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. *Transportation, 34*(5), 557-573.

Pinjari, A., Eluru, N., Bhat, C., Pendyala, R., & Spissu, E. (2008). Joint model of choice of residential neighborhood and bicycle ownership: accounting for self-selection and unobserved heterogeneity. *Transportation Research Record: Journal of the Transportation Research Board, 2082*, 17-26.

Pinjari, A., Pendyala, R., Bhat, C., & Waddell, P. (2011). Modeling the choice continuum: An integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. *Transportation, 38*(6), 933-958.

Potoglou, D. (2008). Vehicle-type choice and neighbourhood characteristics: An empirical study of Hamilton, Canada. *Transportation Research Part D: Transport and Environment, 13*(3), 177-186.

Potoglou, D., & Kanaroglou, P. S. (2008b). Disaggregate demand analyses for conventional and alternative fueled automobiles: A review. *International Journal of Sustainable Transportation, 2*(4), 234-259.

Potoglou, D., & Kanaroglou, P. S. (2008a). Modelling car ownership in urban areas: a case study of Hamilton, Canada. *Journal of Transport Geography, 16*(1), 42-54.

Potoglou, D., & Susilo, Y. (2008). Comparison of vehicle-ownership models. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 97-105.

Prevedouros, P. D., & Schofer, J. L. (1992). Factors affecting automobile ownership and use. *Transportation Research Record: Journal of the Transportation Research Board, 1364*, 152-160.

Prillwitz, J., Harms, S., & Lanzendorf, M. (2006). Impact of life-course events on car ownership. *Transportation Research Record: Journal of the Transportation Research Board, 1985*, 71-77.

Pucher, J., & Renne, J. L. (2003). Socioeconomics of urban travel: evidence from the 2001 NHTS. *Transportation Quarterly, 57*(3), 49-77.

Rashidi, T. H., & Mohammadian, A. (2011). A dynamic hazard-based system of equations of vehicle ownership with endogenous long-term decision factors incorporating group decision making. *Journal of Transport Geography, 19*(6), 1072-1080.

Reilly, M., & Landis, J. (2002). *The Influence of Built Form and Land Use on Mode Choice*. (IURD WP 2002-4(1)).

Rodrigue, J-P, Comtois, C., & Slack, B. (2006). *The geography of transport system* (1st ed.). New York: Routledge.

Roeder, K., Lynch, K. G., & Nagin, D. S. (1999). Modeling uncertainty in latent class membership: A case study in criminology. *Journal of the American Statistical Association, 94*(447), 766-776.

Roorda, M. J., Carrasco, J. A., & Miller, E. J. (2009). An integrated model of vehicle transactions, activity scheduling and mode choice. *Transportation Research Part B: Methodological, 43*(2), 217-229.

Roorda, M., Mohammadian, A., & Miller, E. (2000). Toronto area car ownership study: A retrospective interview and its applications. *Transportation Research Record: Journal of the Transportation Research Board, 1719*, 69-76.

Roorda, M., Morency, C., & Woo, K. (2008). Two cities, two realities?: A closer look at the evolution of trip rates in Toronto and Montreal, Canada. *Transportation Research Record: Journal of the Transportation Research Board, 2082*, 156-167.

Ryan, J., & Han, G. (1999). Vehicle-Ownership model using family structure and accessibility application to Honolulu, Hawaii. *Transportation Research Record: Journal of the Transportation Research Board, 1676*, 1-10.

Salon, D. (2009). Neighborhoods, cars, and commuting in New York City: A discrete choice approach. *Transportation Research Part A: Policy and Practice, 43*(2), 180-196.

Sanchez, T. W., Makarewicz, C., Haas, P. M., & Dawkins, C. J. (2007). *Transportation Costs, Inequities, and Trade-Offs*. Paper presented at the 86th Annual Meeting of the Transportation Research Board, Washington D.C.

Sanko, N. (2013). Travel demand forecasts improved by using cross-sectional data from multiple time points. *Transportation*, 1-23.

Santos, G., Behrendt, H., & Teytelboym, A. (2010). Part II: Policy instruments for sustainable road transport. *Research in Transportation Economics, 28*(1), 46-91.

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis & Prevention, 43*(5), 1666-1676.

Schimek, P. (1996). Household motor vehicle ownership and use: How much does residential density matter? *Transportation Research Record: Journal of the Transportation Research Board, 1552*, 120-125.

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography, 13*(1), 83-99.

Selander, J., Nilsson, M. E., Bluhm, G., Rosenlund, M., Lindqvist, M., Nise, G., & Pershagen, G. (2009). Long-term exposure to road traffic noise and myocardial infarction. *Epidemiology, 20*(2), 272-279.

Sen, S. (2006). *A joint multiple discrete continuous extreme value (MDCEV) model and multinomial logit model (MNL) for examining vehicle type/vintage, make/model and usage decisions of the household.* (Ph.D Dissertation), The University of Texas at Austin, Ann Arbor.

Senbil, M., Kitamura, R., & Mohamad, J. (2009). Residential location, vehicle ownership and travel in Asia: A comparative analysis of Kei-Han-Shin and Kuala Lumpur metropolitan areas. *Transportation, 36*(3), 325-350.

Shay, E., & Khattak, A. (2005). Automobile ownership and use in neotraditional and conventional neighborhoods. *Transportation Research Record: Journal of the Transportation Research Board, 1902*, 18-25.

Shay, E., & Khattak, A. (2007). Automobiles, trips, and neighborhood type: comparing environmental measures. *Transportation Research Record: Journal of the Transportation Research Board, 2010*, 73-82.

Shay, E., & Khattak, A. J. (2011). Household travel decision chains: Residential environment, automobile ownership, trips and mode choice. *International Journal of Sustainable Transportation, 6*(2), 88-110.

Small, K. A. (1994). Approximate generalized extreme value models of discrete choice. *Journal of Econometrics, 62*(2), 351-382.

Soberon-Ferrer, H., & Dardis, R. (1991). Determinants of household expenditures for services. *Journal of Consumer Research*, 385-397.

Sobhani, A., Eluru, N., & Faghih-Imani, A. (2013). A latent segmentation based multiple discrete continuous extreme value model. *Transportation Research Part B: Methodological, 58*, 154-169.

Sobhani, A., Eluru, N., & Pinjari, A. R. (2014). *Evolution of adults' weekday time use patterns from 1992 to 2010: A Canadian perspective*. Paper presented at the 93rd Annual Meeting of the Transportation Research Board (TRB), Washington, D.C.

Soltani, A. (2005). Exploring the impacts of built environments on vehicle ownership. *Proceedings of the Eastern Asia Society for Transportation Studies, 5*, 2151-2163.

Spissu, E., Pinjari, A., Pendyala, R., & Bhat, C. (2009). A copula-based joint multinomial discrete–continuous model of vehicle type choice and miles of travel. *Transportation, 36*(4), 403-422.

Srinivasan, K. (2002). Injury severity analysis with variable and correlated thresholds: Ordered mixed logit formulation. *Transportation Research Record: Journal of the Transportation Research Board, 1784*(-1), 132-141. doi: 10.3141/1784-17

Srinivasan, K., Naidu, G., & Sutrala, T. (2009). Heterogeneous decision rule model of mode choice incorporating utility maximization and disutility minimization. *Transportation Research Record: Journal of the Transportation Research Board, 2132*, 59-68.

Stopher, P. R., & Greaves, S. P. (2007). Household travel surveys: Where are we going? *Transportation Research Part A: Policy and Practice, 41*(5), 367-381.

Swait, J., & Ben-Akiva, M. (1986). *Analysis of the effects of captivity on travel time and cost elasticities.* Paper presented at the Behavioural Research for Transport Policy: The 1985 International Conference on Travel Behavior, The Netherlands.

Tang, W. L., & Mokhtarian, P. L. (2009). Accounting for taste heterogeneity in purchase channel intention modeling: an example from northern California for book purchases. *Journal of Choice Modelling, 2*(2), 148-172.

Terza, J. V. (1985). Ordinal probit: A generalization. *Communications in Statistics-Theory and Methods, 14*(1), 1-11.

Thakuriah, P., & Liao, Y. (2005). Analysis of variations in vehicle ownership expenditures. *Transportation Research Record: Journal of the Transportation Research Board, 1926*, 1-9.

Thakuriah, P., & Liao, Y. (2006). Transportation expenditures and ability to pay: Evidence from consumer expenditure survey. *Transportation Research Record: Journal of the Transportation Research Board, 1985*, 257-265.

Thakuriah, P., & Mallon-Keita, Y. (2014). *An analysis of household transportation spending during the 2007-2009 US economic recession* Paper presented at the 93rd Annual Meeting of the Transportation Research Board, Washington D.C.

Train, K. (2009). *Discrete choice methods with simulation*. Cambridge; New York: Cambridge University Press.

Turcotte, M. (2008). Dependence on cars in urban neighbourhoods. *Canadian Social Trends, 85*, 20-30.

Turcotte, M. (2011). Commuting to work: Results of the 2010 general social survey. *Canadian Social Trends, 92*, 25-36.

Turnbull, K. F., Pratt, R. H., Evans, J. E., & Levinson, H. S. (2004). Park-and-Ride/Pool. In: *traveler response to transportation system changes.* Transportation Research Board, Washington D.C.

van Acker, V., & Witlox, F. (2010). Car ownership as a mediating variable in car travel behaviour research using a structural equation modelling approach to identify its dual relationship. *Journal of Transport Geography, 18*(1), 65-74.

Vovsha, P., Petersen, E., & Donnelly, R. (2003). Explicit modeling of joint travel by household members: Statistical evidence and applied approach. *Transportation Research Record: Journal of the Transportation Research Board, 1831*, 1-10.

Vyas, G., Paleti, R., Bhat, C., Goulias, K., Pendyala, R., Hu, H.-H., . . . Bahreinian, A. (2012). Joint vehicle holdings, by type and vintage, and primary driver assignment model with application for California. *Transportation Research Record: Journal of the Transportation Research Board, 2302*, 74-83.

Wafa, Z., Bhat, C. R., Pendyala, R. M., & Garikapati, V. M. (2015). A latent-segmentation based approach to investigating the spatial transferability of activity-travel models. *Transportation Research Record (forthcoming)*.

Walker, J. L., & Li, J. (2007). Latent lifestyle preferences and household location decisions. *Journal of Geographical Systems, 9*(1), 77-101.

Wang, D., Chai, Y., & Li, F. (2011). Built environment diversities and activity–travel behaviour variations in Beijing, China. *Journal of Transport Geography, 19*(6), 1173-1186.

Washington, S., Karlaftis, M. G., & Mannering, F. L. (2011). *Statistical and econometric methods for transportation data analysis*. Boca Raton, FL: CRC Press.

Weinberger, R., & Goetzke, F. (2010). Unpacking preference: How previous experience affects auto ownership in the United States. *Urban Studies, 47*(10), 2111-2128.

Weinberger, R. (2007). The high cost of free highways. *Idaho Law Review, 43*(2), 475-503.

Whelan, G. (2007). Modelling car ownership in Great Britain. *Transportation Research Part A, 41*(3), 205-219.

Winkelmann, R. (2008). Econometric analysis of count data. from http://public.eblib.com/EBLPublic/PublicView.do?ptiID=371726

Woldeamanuel, M., Cyganski, R., Schulz, A., & Justen, A. (2009). Variation of households' car ownership across time: application of a panel data model. *Transportation, 36*(4), 371-387.

Wong, K. I. (2013). An analysis of car and motorcycle ownership in Macao. *International Journal of Sustainable Transportation, 7*(3), 204-225.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, Mass.: MIT Press.

Wu, G., Yamamoto, T., & Kitamura, R. (1999). Vehicle ownership model that incorporates the causal structure underlying attitudes toward vehicle ownership. *Transportation Research Record: Journal of the Transportation Research Board, 1676*, 61-67.

Xie, Y., Zhao, K., & Huynh, N. (2012). Analysis of driver injury severity in rural single-vehicle crashes. *Accident Analysis & Prevention, 47*, 36-44.

Yamamoto, T. (2008). The impact of life-course events on vehicle ownership dynamics: The cases of france and Japan. *IATSS Research, 32*(2), 34-43.

Yamamoto, T. (2009). Comparative analysis of household car, motorcycle and bicycle ownership between Osaka metropolitan area, Japan and Kuala Lumpur, Malaysia. *Transportation, 36*(3), 351-366.

Yamamoto, T., & Kitamura, R. (2000). An analysis of household vehicle holding durations considering intended holding durations. *Transportation Research Part A: Policy and Practice, 34*(5), 339-351.

Yamamoto, T., Kitamura, R., & Kimura, S. (1999). Competing-risks-duration model of household vehicle transactions with indicators of changes in explanatory variables. *Transportation Research Record: Journal of the Transportation Research Board, 1676*, 116-123.

Yamamoto, T., Matsuda, T., & Kitamura, R. (1997). An analysis of household vehicle holding durations relative to intended holding durations. *Infrastructure Planning Review, 14*(14), 799-808.

Yasmin, S., & Eluru, N. (2013). Evaluating alternate discrete outcome frameworks for modeling crash injury severity. *Accident Analysis & Prevention, 59*, 506-521.

Yasmin, S., Eluru, N., Bhat, C. R., & Tay, R. (2014). A latent segmentation based generalized ordered logit model to examine factors influencing driver injury severity. *Analytic Methods in Accident Research, 1*, 23-38.

Yee, J. L., & Niemeier, D. (1996). Advantages and disadvantages: Longitudinal vs. repeated cross-section surveys. *Project Battelle, 94*, 16-22.

Zegras, C. (2010). The built environment and motor vehicle ownership and use: Evidence from Santiago de Chile. *Urban Studies, 47*(8), 1793-1817.

Zhang, M. (2004). The role of land use in travel mode choice: evidence from Boston and Hong Kong. *Journal of the American Planning Association, 70*(3), 344-360.