Demand Modelling of Cross-Regional Commuting Trips in Multimodal Networks

By

Mohamed Salah Mahmoud

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Graduate Department of Civil Engineering
University of Toronto

Demand Modelling of Cross-Regional Commuting Trips in Multimodal Networks

Mohamed Salah Mahmoud

Doctor of Philosophy

Department of Civil Engineering

University of Toronto

2015

ABSTRACT

The continual expansion of major metropolitan areas caused by the inexorable urban sprawl away from city centres has produced persistent growth in the number of cross-regional trips. This research focuses on acquiring the required understanding of cross-regional travellers' travel behaviour by quantifying the effects of the key factors that influence their travel decisions. A practice-ready framework for modelling cross-regional travel demand in the context of multimodal transportation networks is presented. The framework adopts a comprehensive mode choice model that takes into account the three main decisions faced by intermodal travellers, namely the main mode, access mode, and access station location choices. The three choices and their interactions are carefully considered in the different phases of the model development.

The modelling framework is developed over two phases. In phase I, data from a travel survey are used along with detailed information on transit stations to develop access station choice models for regional and local transit park-and-ride users. In phase II, the access station choice models are utilized within an innovative multimodal trip planner tool that is developed and integrated with the Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*). *SCRIPT* is an online

respondent-customized survey that consists of three sections, collecting revealed preference (RP) data, stated preference (SP) data, and household and personal information. Using *SCRIPT*'s RP/SP data, a set of advanced econometric joint main mode and access mode choice models are developed. The joint RP/SP models reveal meaningful insights into cross-regional commuters' mode choice behaviour. The developed models are validated and calibrated to develop the Interactive Model for Policy Analysis of Cross-Regional Travel (*IMPACT*) to be used for the prediction of corresponding changes in aggregate modal shares in response to the introduction of new policies. As such, the developed models offer a significant step forward in modelling cross-regional travel decisions by introducing a policy-sensitive travel demand modelling framework.

DEDICATION

To the soul of my beloved sister.

ACKNOWLEDGMENTS

All praise is due to Allah.

This life is but a journey with numerous stops. Here I have arrived at the stop to bring my doctoral journey to a fortunate end. I am utterly grateful to have completed my dissertation with guidance and support from my supervisors, colleagues, family, and friends. Without them, the completion of this work would not have been possible. Therefore, I am indebted to them for their technical assistance, continuous motivation, and genuine moral support.

In particular, I would like to express my deep and sincere gratitude to my supervisors: **Prof. Amer Shalaby** and **Prof. Khandker M. Nurul Habib** for their outstanding support, guidance, and valuable feedback throughout the past few years. My dear professors, I consider myself lucky to have been able to get to know you, work under your supervision, and learn from your exceptional expertise. The way you treat your students and the overall work environment in your research groups made my journey much more fruitful and enjoyable than what I would have ever expected. Thank you for this unique experience.

I would also like to thank the members of the examination committee for their valuable suggestions and comments. Thanks for spending your valuable time and energy on my PhD dissertation.

I would also like to thank my colleagues from the ITS Lab at the University of Toronto for such a wonderful learning experience. Specifically, I would like to thank Adam Weiss for providing insightful suggestions for Section 7.5.2 and Tamer Abdulazim for his technical support in the survey web design. Also, I would like to thank Metrolinx for partially funding this research.

I sincerely thank my beloved parents, sisters, and family for their continuous love and unfaltering support. Thank you for being my source of motivation and inspiration during the thick and thin.

Last but not least, I am of course also indebted to my wife for her love, support, and patience. My gratitude and appreciation to you, Sara, are beyond what words can express.

I cannot do justice in just a few words to acknowledge the support I have received for the completion of my dissertation. However, I thank everyone who has contributed in any capacity at any point during the course of my doctoral journey. Thank you all.

LIST OF PUBLICATIONS

The following chapters of this dissertation have been reproduced with modifications from my previously published and presented material:

Chapter 2 – **LITERATURE REVIEW** and Chapter 3 – **CONCEPTUAL FRAMEWORK**

Mahmoud, M. S., K. M. N. Habib, and A. Shalaby. Modelling Transit Mode Choice for Inter-regional Commuting Trips: Joint Model of Mode Choice, Access Mode Choice and Access Station Choice. Presented at the ACT Sustainable Mobility Summit, Hamilton, ON, Canada, 2012.

Mahmoud, M. S., A. Weiss, and K. M. N. Habib. Latent Captivation or Mode Culture? Investigation into Mode Choice Preference Structures in Competitive Modal Arrangements. Presented at the 94th Annual Meeting of Transportation Research Board, Washington, D.C., 2015.

Chapter 4 – **MODELLING ACCESS STATION CHOICE FOR CROSS-REGIONAL COMMUTER TRIPS**

Mahmoud, M. S., K. M. N. Habib, and A. Shalaby. Park-and-Ride Access Station Choice Model for Cross-Regional Commuter Trips: Case Study of Greater Toronto and Hamilton Area, Canada. Presented at the 93rd Annual Meeting of Transportation Research Board, Washington, D.C., 2014.

Mahmoud, M. S., K. M. N. Habib, and A. Shalaby. Park-and-Ride Access Station Choice Model for Cross-Regional Commuter Trips: Case Study of Greater Toronto and Hamilton Area, Canada. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2419*, Transportation Research Board of the National Academies, Washington, D.C., 2014, pp. 92–100.

Chapter 5 and Chapter 6 – **SURVEY of CROSS-REGIONAL INTERMODAL PASSENGER TRAVEL**

Mahmoud, M. S., K. M. N. Habib, and A. Shalaby. Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*): A Joint RP/SP Survey Integrated with a Multimodal Trip Planner Tool. Presented at the 94th Annual Meeting of Transportation Research Board, Washington, D.C., 2015.

Mahmoud, M. S., K. M. N. Habib, and A. Shalaby. Survey of Cross-Regional Intermodal Passenger Travel (SCRIPT): A Joint RP/SP Survey Integrated with a Multimodal Trip Planner Tool. Forthcoming in *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C.

Chapter 7 – **MODE CHOICE MODELLING of CROSS-REGIONAL INTERMODAL COMMUTING TRIPS**

Mahmoud, M. S., K. M. N. Habib, and A. Shalaby. Demand Modelling of Cross-Regional Intermodal Commuting Trips. Report submitted to Metrolinx – Strategic Policy and Systems Planning Department, November 2014.

Mahmoud, M. S., K. M. N. Habib, and A. Shalaby. Data Fusion of Revealed and Stated Preference Information for Joint Estimation of Mode Choice Models: A Case Study on Cross-Regional Intermodal Commuting Trips in the Greater Toronto and Hamilton Area. Forthcoming at the 14th International Conference on Travel Behaviour Research, International Association for Travel Behaviour Research (*IATBR*), London, UK, 2015.

TABLE OF CONTENTS

DEDICATION iv

ACKNOWLEDGMENTS v

LIST OF PUBLICATIONS vi

TABLE OF CONTENTS viii

LIST OF TABLES xiii

LIST OF FIGURES xv

LIST OF APPENDICES xvii

GLOSSARY xviii

CHAPTER 1 1

1 **INTRODUCTION** 1

1.1 Chapter Overview 1

1.2 Problem Statement 1

1.3 Motivation 2

1.4 Research Objectives and Methods 3

1.5 Dissertation Layout 5

CHAPTER 2 7

2 **LITERATURE REVIEW** 7

2.1 Chapter Overview 7

2.2 Cross-Regional Travel and Multimodal Trips 7

2.2.1 Characteristics of Cross-Regional Commuting Trips 7

2.2.2 Transit Modal Integration in Multimodal Networks 10

2.3 Current Practice of Cross-Regional Travel Demand Modelling in Multimodal Networks 10

2.3.1 Access Station Location Choice 11

2.3.2 Main Mode and Access Mode Choice 15

2.4 Current Practice of Discrete Choice Models 18

2.5 Current Practice in RP/SP Survey Design 23

2.5.1 The Need for RP/SP Surveys 23

2.5.2 SP Experimental Design 24

2.5.3 Customized SP Experiments and Trip Planner Tools 25

2.5.4 RP/SP Data and Mode Choice Models 26

2.6 Chapter Summary 28

CHAPTER 3 30

3 **CONCEPTUAL FRAMEWORK** 30

3.1 Chapter Overview 30

3.2 Introduction 30

3.3 Conceptual Framework of Modelling Cross-Regional Commuting Trips 31

3.4 Framework Components under Investigation 34

3.5 Chapter Summary 39

CHAPTER 4 41

4 **MODELLING ACCESS STATION CHOICE FOR CROSS-REGIONAL COMMUTER TRIPS** 41

4.1 Chapter Overview 41

4.2 Introduction 41

4.3 Study Area, Data Description, and Spatial Analysis 42

4.4 Econometric Model 47

4.5 Empirical Models 48

4.6 Results, Discussion, and Validation 50

4.7 Chapter Summary 53

CHAPTER 5 56

5 **SURVEY OF CROSS-REGIONAL INTERMODAL PASSENGER TRAVEL (*SCRIPT*)** 56

5.1 Chapter Overview 56

5.2 Introduction 56

5.3 Survey Study Area 57

5.4 Travel Mode Options and Description 59

5.5 Experimental Design 59

5.6 Survey Sample Design 61

5.6.1 Target Population 61

5.6.2 Sampling Method, Sample Size, and Sample Allocation 61

5.7 Survey Instrument Design 62

5.7.1 Survey Data Model 62

5.7.2 Multimodal Trip Planner Tool 65

5.7.3 Section A: Revealed Preference (RP) Data 67

5.7.4 Section B: Stated Preference (SP) Data 67

5.7.5 Section C: Household and Personal Information 69

5.7.6 Respondents' Feedback 73

5.8 Chapter Summary 73

CHAPTER 6 74

6 **SURVEY IMPLEMENTATION AND DATA COLLECTION** 74

6.1 Chapter Overview 74

6.2 Data Collection 74

6.3 Descriptive Data Statistics 75

6.3.1 Revealed Preference Data 75

6.3.2 Stated Preference Data 76

6.3.3 Household and Personal Information....80

6.4 Chapter Summary....85

CHAPTER 7....86

7 **MODE CHOICE MODELLING OF CROSS-REGIONAL COMMUTING TRIPS**....86

7.1 Chapter Overview....86

7.2 *SCRIPT* Data....86

7.2.1 Generating Level-of-service Attributes....86

7.2.2 Data Fusion of RP/SP Information....87

7.2.3 Estimation and Validation Datasets....87

7.3 Econometric Modelling Frameworks....88

7.3.1 Multinomial Logit Model....88

7.3.2 IIA-relaxed Models....90

7.4 Empirical MNL Models....96

7.4.1 RP-only MNL Model....97

7.4.2 SP-only MNL Model....100

7.4.3 Joint RP/SP MNL Model....103

7.4.4 Model Validation....117

7.5 Empirical IIA-Relaxed Models....120

7.5.1 Joint RP/SP NL Model....120

7.5.2 Joint RP/SP PLC Model....125

7.6 Comparison of Joint RP/SP MNL, NL, and PLC Models....129

7.7 Policy Analysis Tool and Practical Implications....132

7.8 Chapter Summary....135

CHAPTER 8 137

8 **CONCLUSIONS, RESEARCH CONTRIBUTIONS, AND FUTURE RESEARCH** 137

8.1 Chapter Overview 137

8.2 Research Summary 137

8.3 Research Contributions 142

8.4 Directions for Future Research 145

9 **REFERENCES** 147

APPENDIX A 156

LIST OF TABLES

Table 4-1 Definitions of Variables 49

Table 4-2 Regional and Local Transit Park-and-ride Access Station Choice MNL Model 51

Table 4-3 Regional Transit (GO Train) Park-and-ride Access Station Choice MNL Model 51

Table 4-4 Local Transit (TTC Subway) Park-and-ride Access Station Choice MNL Model 51

Table 5-1 Travel Modes and Description 60

Table 5-2 Mode Feasibility 66

Table 5-3 Travel Modes’ Attributes 70

Table 5-4 Travel Modes’ Attribute Levels 71

Table 6-1 *SCRIPT* Implementation Metrics 75

Table 6-2 *SCRIPT* and TTS Sample Distribution by Mode Split 75

Table 6-3 *SCRIPT* and TTS Sample Distribution by Trip Origin 75

Table 6-4 *SCRIPT* and TTS Sample Distribution by Gender 80

Table 7-1 Definitions of Variables 96

Table 7-2 RP-only MNL Model 99

Table 7-3 SP-only MNL Model 101

Table 7-4 Joint RP/SP MNL Model 106

Table 7-5 Value of Travel Time Savings 109

Table 7-6 Average Direct Elasticities for Travel Cost and In-vehicle Travel Time 111

Table 7-8 Joint RP/SP NL Model 123

Table 7-9 Joint RP/SP PLC Model 127

Table 7-10 Joint RP/SP MNL Model for NL/PLC Comparison........ 130

Table 7-11 Value of Travel Time Savings........ 131

Table 7-7 Predicted GO Transit Mode Shares 134

LIST OF FIGURES

Figure 1-1 Thesis Roadmap ... 6

Figure 3-1 Conceptual Framework ... 33

Figure 3-2 Tri-variate Joint Decision Structure ... 34

Figure 3-3 Decision Tree Structure ... 36

Figure 3-4 Framework Development ... 38

Figure 4-1 Distribution of Cross-Regional Park-and-Ride Users' Home and Work Locations ... 44

Figure 4-2 Park-and-Ride Parking Lot Capacities, Parking Cost, and Parking Demand ... 45

Figure 4-3 Park-and-Ride Catchment Areas ... 46

Figure 4-4 Relative Station Direction (Alpha αi) of Park-and-Ride Stations (Si) from Home Location (H) to Work Location (W) ... 46

Figure 4-5 Marginal Effects of Access Distance and Relative Station Direction ... 54

Figure 5-1 The Greater Toronto Hamilton Area ... 58

Figure 5-2 SCRIPT Data Model ... 64

Figure 5-3 The Integration of the Multimodal Trip Planner Tool with *SCRIPT* ... 65

Figure 5-4 Polices under Investigation ... 68

Figure 5-5 Snapshot of a Sample SP Experiment ... 72

Figure 6-1 Spatial Distribution of Cross-Regional Trip End ... 77

Figure 6-2 Cross-Regional Trip Mode Shares ... 78

Figure 6-3 Level of Confidence in SP Choice ... 79

Figure 6-4 Mode Availability 79

Figure 6-5 Age Distribution 80

Figure 6-6 Household Size Distribution 81

Figure 6-7 Vehicle Ownership Distribution 81

Figure 6-8 Bike Ownership Distribution 82

Figure 6-9 Household Yearly Income Distribution 82

Figure 6-10 Dwelling Type Distribution 83

Figure 6-11 Marital Status Distribution 83

Figure 6-12 Occupation Type Distribution 84

Figure 6-13 Driving Licence Possession Distribution 84

Figure 6-14 Transit Pass Possession Distribution 85

Figure 7-1 The Artificial Nested Structure of Joint RP/SP MNL Model 89

Figure 7-2 The Artificial Nested Structure of the Joint RP/SP NL Model 91

Figure 7-3 The Artificial Nested Structure of Joint RP/SP PLC Models 93

Figure 7-4 Marginal Effects 116

Figure 7-5 Joint RP/SP Model Validation 119

Figure 7-6 Joint RP/SP NL Empirical Model Structure 121

Figure 7-7 Snapshot of *IMPACT* 133

LIST OF APPENDICES

Appendix A: Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*) …..…………156

GLOSSARY

API	Application programming interface
CAA	Canadian Automobile Association
COVNL	Covariance nested logit
DEFF	Design effect
DOGEV	Dogit generalized extreme value model
EMME	A travel demand modelling system for urban, regional, and national transportation forecasting, developed and marketed by INRO
GAUSS ®	A matrix programming language for mathematics and statistics, developed and marketed by Aptech Systems
GEV	Generalized extreme value
GO Transit	Government of Ontario inter-regional public transit system
GTA	Greater Toronto Area
GTFS	General Transit Feed Specifications
GTHA	Grater Toronto and Hamilton Area
IIA	Independence of irrelevant alternatives
IID	Identically and independently distributed
IMPACT	Interactive Model for Policy Analysis of Cross-Regional Travel
K&R	Kiss-and-ride
MAXLIKE	A package that provides a likelihood-based approach to estimate model parameters
MNL	Multinomial logit model
MNP	Multinomial probit model
Ngene®	Software for generating experimental designs that are used in stated choice experiments for the purpose of estimating choice models, particularly of the logit type
NL	Nested logit
O/D	Origin/destination
P&R	Park-and-ride
PCL	Paired combinatorial logit

PLC	Parameterized logit captivity
R	Software environment for statistical computing and graphics
RP	Revealed preference
RP/SP	Revealed preference/stated preference
RUM	Random utility maximization
SCRIPT	Survey of Cross-Regional Intermodal Passenger Travel
SP	Stated preference
TTC	Toronto Transit Commission
TTS	Transportation Tomorrow Survey
UTMS	Urban Transportation Modelling System

CHAPTER 1

1 INTRODUCTION

1.1 Chapter Overview

This chapter introduces the research conducted for this thesis. Section 1.2 presents a discussion of the problem statement. Next, the motivation of this research is presented in Section 1.3. The following section, Section 1.4, highlights the main objectives of the dissertation and provides an overview of the research methodology. Finally, Section 1.5 presents the dissertation's layout.

1.2 Problem Statement

We live in the era of "mobility," an era in which individuals have become more mobile in terms of their ability to plan trips using various travel modes within multimodal transportation systems. Such systems provide travel options that consider various modes (walking, cycling, automobile, public transit, etc.) and connections among them. That is, they offer the promise of supporting and leveraging sustainable transportation solutions. However, the planning of multimodal transportation systems is relatively complicated. Travel modes are not perfect substitutes; rather, each mode may only be appropriate for particular users and/or uses. As such, providing the proper coordination between multiple modes within one integrated system, given the differences in their attributes and the variations among their potential users' characteristics, is a challenge.

The development of multimodal transportation systems cannot be achieved without proper planning of the levels of integration between the various travel modes. Transit modal integration is the cornerstone of this development. Transit modal integration refers to schemes that provide combinations of transit services and other motorized or non-motorized modes. That is, individuals can seamlessly interchange between different modes of travel. As such, more opportunities are created from the efficient integration of transit services with other modes by increasing users' travel options and providing more sustainable transportation choices. Therefore, improving transit modal integration is one of the promising strategies that have been under investigation by regional transit operators. However, moving forward in this direction requires appropriate planning of the contiguous transportation system/network and proper understanding of individuals' travel behaviour.

This research is focused on acquiring a thorough understanding of the travel behaviour of cross-regional travellers, many of whom have viable intermodal options, by quantifying the effects of the key factors that influence their travel decisions. Ultimately, this research develops a detailed travel demand model for cross-regional commuting trips in multimodal networks. In particular, further emphasis is dedicated to understanding cross-regional travellers' behaviour of travel mode choices in a multimodal transportation system context. The following sections present the motivation behind defining this specific focus and the objectives of conducting this research.

1.3 Motivation

The continual expansion of major metropolitan areas caused by the inexorable urban sprawl away from city centres has produced persistent growth in the number of cross-regional commuter trips (i.e. trips that cross boundaries of municipal or regional jurisdictions that have different transit operators) (DMG, 1998, 2008). As such, cross-regional transit trips are typically intermodal trips. Intermodal trips are defined as are trips that involve transfer between contiguous local transit systems, transfer between local and regional transit systems, or transfer between automobiles and local or regional transit systems.

In 2011, approximately one in three trips in the Greater Toronto and Hamilton Area (GTHA) crossed a regional boundary (DMG, 2013). As a result, a growing group of commuters experiences long travel times, especially for travel modes other than driving. Like many other North American regions, the peak-period commuting trips in the GTHA are dominated by private automobile drivers. Despite the availability of alternative and more sustainable travel modes (such as transit, carpooling, car-sharing, walking and biking), driving alone is often more attractive since it provides shorter travel times and ubiquitous accessibility while still being financially competitive.

The phenomenon of private automobile domination becomes more prominent in the case of cross-regional commuting trips. Unlike intra-regional commuting travel (trips originating from and destined to the same region), cross-regional commuters have a unique set of possible travel modes since non-motorized modes are mostly infeasible. In addition, cross-regional trips often originate in suburban areas where transit accessibility is relatively inadequate because of low population density and sparse land use. That is, non-automobile cross-regional trips may involve the use of multiple transit services or the interaction between two travel modes, which often results in delays caused by the typical lack of service coordination, not to mention the absence of fare integration. Therefore, in such cases, transit options do not provide competitive travel times and/or costs compared to automobile options, which explains the latter's dominance. This highlights the importance of dedicating special attention to the enhancement of multimodal transportation systems' current planning practices by improving transit modal integration, which, as explained earlier, requires proper understanding of individuals' travel behaviour.

As explained in Chapter 2, few studies have focused on developing demand models that are capable of explaining the behaviour of such a unique group of travellers. According to the cited literature, a comprehensive framework that encapsulates all possible decisions faced by individuals (such as mode, departure time, and route choices) in the context of cross-regional trips does not exist. Many of the previous studies on demand modelling have overlooked intermodal travellers' behaviour in general and cross-regional commuters' behaviour in particular. Most of the existing regional demand models are developed using large-scale datasets that are heavily imbalanced towards intra-regional commuters since they often represent a higher share compared to cross-regional commuters. Therefore, the results of such models are highly influenced by the behaviour/travel patterns of intra-regional commuters which are different than those of cross-regional commuters, many of whom have viable intermodal options. Therefore, an enhanced mode choice model that can be used as a core component of a conceptual travel demand framework for modelling cross-regional trips with special treatment of intermodal options is required to fill this significant gap in the literature.

1.4 Research Objectives and Methods

The overall objective of this research is to develop a practice-ready framework for modelling cross-regional travel demand in the context of multimodal transportation networks. The primary objective is to develop a comprehensive mode choice model that takes into account the three main decisions faced by intermodal travellers, namely the main mode, access mode, and access station location choices. Such models offer a significant step forward in modelling cross-regional travel decisions by providing policy-sensitive models that are capable of capturing individuals' behavioural changes in accordance with transit modal integration policy initiatives. The mode choice model along with other pre-developed model components form together a detailed modelling framework of cross-regional trips. To meet the research objectives, three main tasks are defined as follows:

1. Develop a conceptual framework for modelling the choice behaviour of cross-regional commuters.

2. Design and execute an innovative data collection tool to provide the required customized data on cross-regional commuters' travel behaviour.

3. Develop, validate, and calibrate policy-sensitive models to be used to evaluate the effectiveness of policy initiatives on cross-regional commuters' travel choices.

This research adopts state-of-the-art methods to fulfill the goals of the study. First, a fully disaggregate practice-ready travel demand model is conceptualized. The framework is developed at the individual level, providing a continuous updating process that attempts to capture observed travel behaviour. The framework consists of a series of model components; each component represents one of the travel decisions faced by cross-regional travellers. The framework produces a detailed set of discrete-choice travel decisions with special treatment of the interaction between different travel options. These decisions include individuals' expected departure time choice, joint decisions on the main travel mode and access mode choices (if any), access station location choice (if any), and driving and/or transit route choices.

The framework places more emphasis on the mode choice model component. Therefore, a comprehensive modelling structure for studying the mode choice decision of cross-regional trips is developed. Accordingly, three decisions are identified: main mode, access mode, and access station location. The development of a model structure that encompasses the three decisions requires detailed and exhaustive data on individuals' preferences. Hence, the second step is to design a data collection tool that is capable of providing such comprehensive information on cross-regional travellers' travel preferences. As such, a joint revealed preference/stated preference travel survey is designed and implemented. The survey adopts an innovative multimodal trip planner tool that is developed to generate customized travel options to each respondent including intermodal travel options such as park-and-ride.

Finally, using the collected data, policy-sensitive mode choice model are developed, validated and calibrated to develop a policy analysis tool for evaluating the effectiveness of policy initiatives on cross-regional commuters' travel choices. In addition, advanced model structures are examined to provide in-depth insights in terms of explaining cross-regional travellers' behaviour.

1.5 Dissertation Layout

Figure 1-1 shows the dissertation's layout; the highlighted components identify the key research contributions presented herein. The remainder of the dissertation (after Chapter 1) is organized as follows. Chapter 2 provides a review of the relevant studies in the literature and identifies the gaps in the relevant body of knowledge. Chapter 3 presents the conceptual framework for modelling cross-regional commuting trips, phases of framework development, and highlights elements of the framework to be investigated in this study. Chapter 4 presents an investigation of cross-regional park-and-ride commuters' access station location choice. A set of discrete choice models is developed to be used as a tool to predict park-and-ride cross-regional commuters' choice of access station location. Chapter 5 presents the Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*), an on-line data collection tool. *SCRIPT* collects data on respondents' current commuting trips as well as their stated mode choice in response to hypothetical changes in the current mode attributes. The survey features an innovative multimodal trip planner tool that generates respondent-customized travel options. Chapter 6 provides an overview of the survey implementation process and data collection procedure. The collected data is prepared to be used for empirical modelling as discussed in Chapter 7. In Chapter 7, the development of cross-regional commuters' joint main mode and access mode choice models using revealed preference (RP) and stated preference (SP) data is presented. In addition, detailed discussion on the developed models' validation and calibration to be used for policy analysis is provided. Finally, Chapter 8 provides a summary of the conducted research, research contributions, and directions for future research.

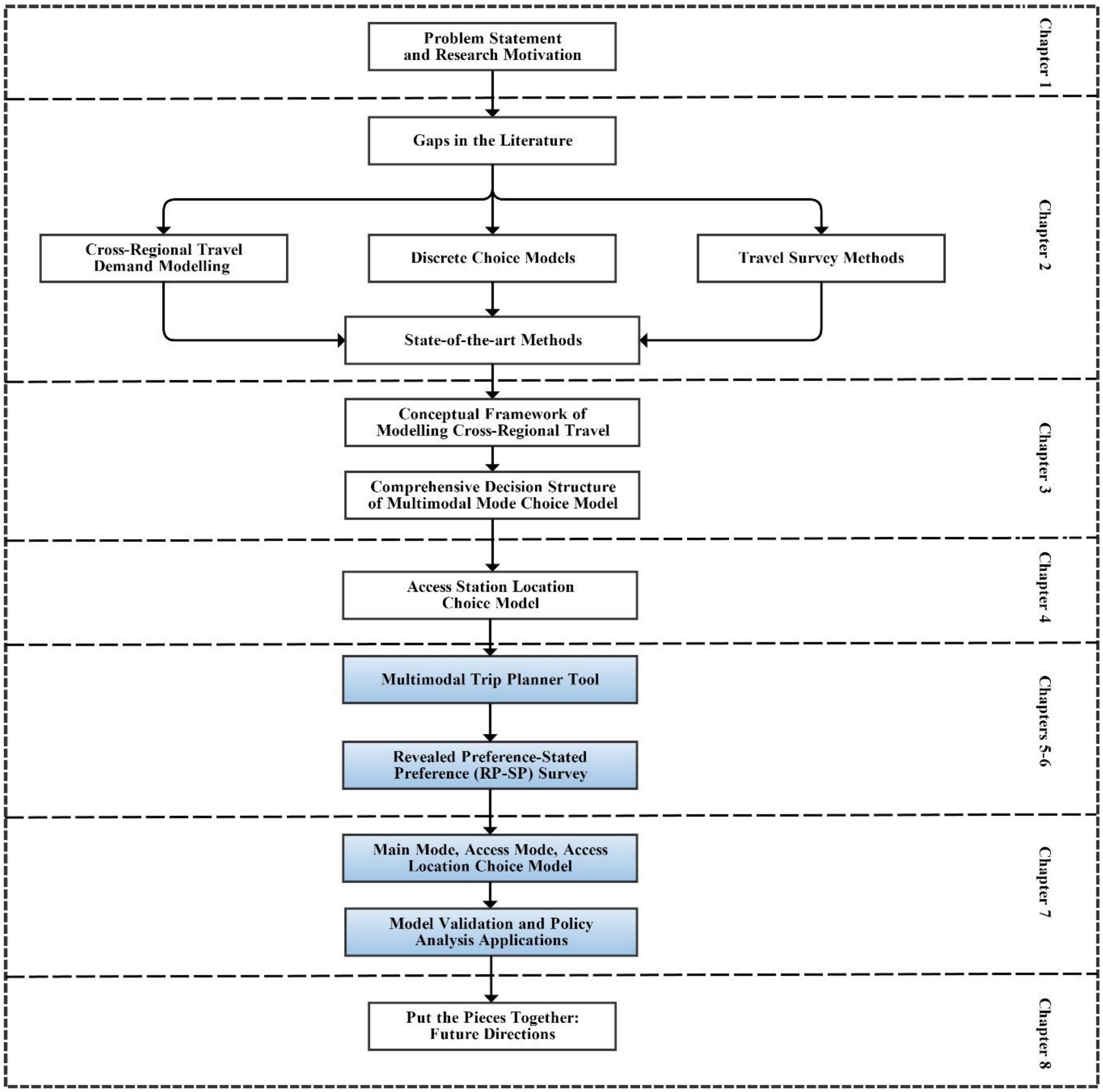


Figure 1-1 Thesis Roadmap

CHAPTER 2

2 LITERATURE REVIEW

2.1 Chapter Overview

This chapter provides a review of the relevant studies in the literature. The chapter starts with Section 2.2, which provides an overview of the characteristics of cross-regional travel within the context of multimodal transportation networks with an emphasis on transit modal integration. In Section 2.3, the current practice of cross-regional travel demand modelling is presented. Then, a review of discrete choice models used for modelling travel demand is presented in Section 2.4. Next, Section 2.5 highlights the need for stated preference (SP) data and reviews the current practice in RP/SP survey designs. Finally, Section 2.6 provides a summary of this chapter and identifies the gaps in the reviewed literature.

2.2 Cross-Regional Travel and Multimodal Trips

2.2.1 Characteristics of Cross-Regional Commuting Trips

Cross-regional trips are trips that cross boundaries of municipal or regional jurisdictions that have different transit operators. Such trips are growing in number steadily, producing persistent pressure on regions' transportation systems (DMG, 1998, 2008). In 2011, approximately one in three trips in the Greater Toronto and Hamilton Area (GTHA) crossed a regional boundary (DMG, 2013). However, the literature on cross-regional travel is evidently limited; few research efforts have studied long-distance trips, which share some similarities with cross-regional trips. Long-distance commuting trips are often classified based on a predefined travel distance or time threshold. Typically, based on their definitions, the travel time of a one-way long-distance commuting trip ranges from 30 to 45 minutes (Sandow and Westin, 2010; Van Ham and Hooimeijer, 2009). Other studies have focused on weekly commuting with travel distances of up to 200 Km (Öhman and Lindgren, 2003; Zhang and Chen, 2011). However, such generic definitions may not represent long-distance commuting trips in regions similar to the GTHA. In addition, these definitions do not appropriately take into account the interaction between different travel modes across regional boundaries and how it might affect individuals' choice behaviour. That is, models developed to explain long-distance trip patterns may not specifically represent the travel behaviour of cross-regional and/or intermodal commuters.

One study emphasized the substantial share of long-distance commuting trips compared to the total travel demand and confirmed significant growth over time (Sandow and Westin, 2010). Other studies have highlighted the importance of dedicating special research efforts to investigate long-distance and inter-regional commuting trips, claiming that the use of typical travel demand models is inadequate to explain the travel behaviour of this special market segment (Lee, 1996). Based on the results of a set of logistic regression models, another study emphasized that long-distance commuters have distinct characteristics relative to other types of commuters (Van Ham and Hooimeijer, 2009). For instance, commuters with higher income and education levels are more likely to commute longer distances. Similarly, another study investigated the socioeconomic and land use attributes that influence long-distance commuters' choices (Titheridge and Hall, 2006). The results showed that one of the driving motives for long-distance commutes is the lack of job opportunities in the vicinity of their residential areas. These results are consistent with the findings of other studies on long-distance travel (Outwater et al., 2010; Sandow and Westin, 2010).

Other studies have focused on long-distance commuters' travel-related choices. The associated decisions of long-distance commuters may involve complex choices that can be either long-term decisions, such as the choice of work and/or home location choices, or short-term decisions, such as mode, route, and departure time choices. These decisions are influenced by various factors related to individuals' socioeconomic and demographic characteristics, household interactions, land use attributes, relative level-of-service characteristics of competing travel modes, and current travel demand management policies. Previous studies have shown that such factors can strongly affect long-distance travellers' travel choices; for instance, home owners are less likely to change home locations to avoid long-distance commutes (Titheridge and Hall, 2006; Van Ham and Hooimeijer, 2009).

Long-distance trips are often mixed with intercity trips that are, in most cases, non-commuting trips, so the trade-offs between travel mode and route choices are perceived differently (Lu and Marsh, 2011). In addition, service frequency and accessibility are expected to significantly affect individuals' choices for commuting trips more than for intercity trips. This can be supported by the fact that users are less concerned with modal and spatial accessibility for non-commuting trips, which are less frequent trips compared to daily commuting trips. In addition, the effect of access and/or egress mode choices is less significant for intercity trips, as the access/egress travel time and cost proportions are relatively minor compared to the total trip travel time and cost.

In light of the above, research efforts investigating long-distance commuting trips with the objective of understanding travellers' behaviour in choosing their modes of transport have been limited. Many of the studies cited above have not explicitly investigated the effects of intermodal interactions and travel demand policies on long-distance trips. Therefore, applying traditional demand models for long-distance trips may not adequately represent the travel behaviour of long-distance commuters. Despite sharing similar characteristics, there are major differences between cross-regional and long-distance trips. The complexity of cross-regional trips stems from the multimodal nature of their long distance travel, which might not be the case for typical long-distance trips.

To investigate the effect of the different travel demand management and land use policies on their intermodal travel decisions, the distinct characteristics of cross-regional commuters must be well identified. Unlike goods' shipping, the common practice of most passenger transportation service providers lacks proper service integration since local services are typically administered by different jurisdiction agencies (Goetz and Vowles, 2000; Graham, 2000). For instance, a non-automobile cross-regional trip (in its simplest form) can be broken down into three components: access, main, and egress trips. Both access and egress trips can be conducted in either motorized (automobile or transit) or non-motorized (walk or bike) modes.

In the GTHA, a typical cross-regional trip may involve automobile access from the trip origin to a regional commuter rail station, followed by a commuter rail trip to a central station, and finally a local transit trip to the trip destination. That is, for such typical intermodal cross-regional trips, the total trip time and cost are hardly known *a priori* to the travellers, different fare systems might apply, and mode transfers are not necessarily coordinated. This lack of coordination among different modes limits the efficiency of the entire transportation system. As mentioned above, a few studies have attempted to define long-distance and/or cross-regional commuters' characteristics; however, none of them has investigated such unique travel characteristics within the multimodal urban context. Therefore, dedicated in-depth research efforts are required to fill this gap in the literature of cross-regional travel demand modelling.

2.2.2 Transit Modal Integration in Multimodal Networks

The integration and/or interaction of different travel modes within the context of multimodal networks play a major role in defining the feasibility of cross-regional intermodal trips and therefore cross-regional travellers' choices. Transit modal integration, in an urban transportation system context, refers to schemes that provide combinations of transit services and other motorized or non-motorized modes. This type of modal integration, the need for it, and the issues related to its implementation can be clearly recognized when studying cross-regional commuting trips, which often require multiple transfers between different modes/transit services.

Improving transit modal integration is one of the promising strategies that have been under investigation by regional transit operators (Metrolinx, 2008). Such improvements can be achieved at the transit integration level (e.g. between local transit and regional transit) by providing synchronized transfers, easier access to major stations in areas where transit accessibility is inadequate, advanced fare integration schemes *via* smart cards, and the integration of station areas with supporting land uses (e.g. transit oriented developments). Similarly, improvements at the integration level between transit and automobiles can be achieved through a system of connected mobility hubs where parking facilities are available for park-and-ride or passenger drop-off/pick-up areas (i.e. kiss-and-ride) to provide seamless transfers between different modes of travel. Evaluating the effectiveness of such initiatives and supporting policies requires a proper understanding of individuals' travel choices, especially for a target population with unique travel characteristics, such as cross-regional commuters.

2.3 Current Practice of Cross-Regional Travel Demand Modelling in Multimodal Networks

As mentioned previously, many of the previous studies on demand modelling overlooked intermodal travellers' behaviour in the context of cross-regional travel. Most of the regional demand models are developed using large-scale datasets that are heavily imbalanced towards intra-regional commuters since they often represent a higher share compared to cross-regional commuters. Therefore, the results of such models are highly influenced by the behaviour/travel patterns of intra-regional commuters which are different than those of cross-regional commuters as explained in the previous sections. That is, there is a need to direct research efforts towards investigating the behaviour of this significant segment of the population (i.e. cross-regional commuters).

Previous research has defined the decision structure of intermodal cross-regional trips by three choices: main mode, access mode, and access station location. These studies attempted to model similar decision structures using four different approaches: incorporating accessibility measures within traditional transit mode split models, inherently modelling the access mode and main mode choices separately, jointly modelling main and access mode choices, and jointly modelling access mode and station location choices (Beimborn et al., 2003; Cervero and Kockelman, 1997; Debrezion et al., 2007, 2009; Fan et al., 1993; Korf and Demetsky, 1981; Kumar and Gur, 1982; Miller, 2007; Park et al., 2014; Polydoropoulou and Ben-Akiva, 2001; Roorda et al., 2006; Sobieniak et al., 1979; Vijayakumar et al., 2011; Wen et al., 2012). To develop models that are capable of capturing individuals' current and possible future changes in behaviour, the interactions among the three decisions should be considered within the decision structure. In the following sections, a review of previous research efforts on main mode, access mode, and access station location choices is presented.

2.3.1 Access Station Location Choice

Access station location choice refers to individuals' decision on the station location that they choose to access a transit service. In most cases, individuals are assumed to use the station closest to their trip origin. This might be a fair assumption in the case of non-motorized access to transit. However, in the case of automobile access, the choice of access station location becomes challenging. In the context of cross-regional trips, which often originate in suburban areas where transit accessibility is significantly inadequate, automobile access has shown a significant increase in transit access mode shares, especially to access regional transit services. Automobile access to regional transit services has become popular in big cities. In the Montreal region, 38% of the commuter rail users access the system via their private automobiles (Vijayakumar et al., 2011). Similarly, in the Greater Toronto and Hamilton Area (GTHA), 81% of GO Train commuters' station access is done by automobiles, including park-and-ride (P&R), kiss-and-ride (K&R), and carpool (Metrolinx, 2011).

Among the different means of modal integration, the use of private cars for transit access is the most flexible form of modal integration (Cairns, 1998; Metrolinx, 2008). Park-and-ride facilities have become an important component of urban transportation systems (Bos et al., 2005; Cairns, 1998; Chakour and Eluru, 2014; Forsey et al., 2013; García and Marín, 2002; Holguín-Veras et al., 2012). They act as a key connection between the road network and the transit system, playing an important role in facilitating cross-modal integration and providing access to transit for users who may not otherwise consider transit as a travel alternative. Well-designed park-and-ride facilities along with optimally utilized transit facilities can efficiently increase the capacity of the entire transportation systems (García and Marín, 2002). From users' perspectives, opportunities created by the efficient integration of transit services with other modes (e.g. private car) increase users' travel options and provide more sustainable transportation choices (Bos et al., 2005; Foote, 2000; Holguín-Veras et al., 2012; Liao et al., 2010).

Transit service providers tend to provide park-and-ride and kiss-and-ride facilities or expand existing parking capacities at locations that could attract more users to the system. Along with increasing transit system supply levels, travel demand management policies help control existing driving demand and make transit a competitive alternative to private car use. For instance, several cities have increased parking costs within the city centres to attract more users to the park-and-ride travel option. Other policies/strategies, such as providing reserved parking, priority parking for carpooling passengers, and bicycle parking may also attract more users. However, evidence of such effects cannot be established without proper investigation of current and forecasted travel demand, which requires further investigation to fill this gap.

Access station choice, a key determinant in park-and-ride commuters' travel decisions, is not isolated from the choice of park-and-ride as a travel mode. Commuters choose the access station in accordance with both main mode (transit) and access mode (automobile) choices. Many factors may affect this choice, including traveller characteristics, station attributes, station orientation relative to individuals' home and work locations, surrounding land use, and quality of the contiguous transportation network. In other words, the question of which station to choose to access the transit system is an internal/endogenous question posed by transit users that choose the automobile as their access mode. Previous studies have shown that higher transit service frequencies along with better station amenities/facilities and network connectivity make stations more attractive as access points (Debrezion et al., 2007).

A limited number of explicit park-and-ride access station choice models can be found in the literature. Kastrenakes (1988) studied rail station choice in New Jersey. The developed multinomial logit model incorporated variables including station access time, frequency of service at the boarding station, station location relative to home location, and generalized trip cost from access stations to final destinations. However, other variables, such as parking availability and parking fees, showed counterintuitive signs and were dropped from the final model. With more focus on the effect of parking attractiveness on station choice, Wardman and Whelan (1999) studied rail station choice for inter-urban trips in London, UK. The study concluded that parking availability and better passenger-related facilities are important factors that attract more travellers to a specific station.

Fan et al. (1993) studied the access mode and access station choices of commuter rail and subway morning peak period trips in the Greater Toronto Area (GTA). A multinomial logit model was developed to model access station choice. A choice set that consists of the closest five stations was defined for subway park-and-ride users. However, for commuter rail park-and-ride users, the two closest stations on the two closest lines defined their choice set. Although this study has provided a solid foundation for future research in access station location choice modelling, it suffered from a few limitations. The two models did not incorporate station-specific attributes such as parking prices and station amenities. In a later study, Roorda et al. (2006) studied the inclusion of minor modes of transport (e.g. drive access to commuter rail and subway) in a tour-based mode choice model. To account for the commuter rail access station choice, a multinomial logit model was estimated considering a choice set of the two closest stations on the two closest commuter rail lines. Similarly, a multinomial logit model was estimated to predict individuals' subway access station choices with drive access. The developed models consider variables such as access time, parking cost, parking capacity, and commuter rail frequency. However, similar to the study by Fan et al. (1993), the developed models were estimated using level-of-service attributes at the aggregate (zonal) level rather than the spatially disaggregate (individual) level. The regional travel demand model of the GTA applies similar model structures to determine access station choice for subway and commuter rail park-and-ride users (Miller, 2007).

Debrezion et al. (2007) developed a multinomial logit model to study passengers' choice of departure railway station. The model showed that the effect of service frequency is relatively small compared to the effect of the distance between the home and station locations. However, no accessibility measures or station-specific attributes were included in the analysis. In addition, the study was limited to aggregate choices made by travellers at the household postal code area level without considering the trip purpose or destination.

Station access location is often modelled conditionally on access mode choice in a nested model structure with access mode choice in the upper level and station choice in the lower level. Fan et al. (1993), developed a nested logit model with access mode choice at the upper level and access rail station choice at the lower level. Similarly, Mukundan (1991) developed a nested logit model to study the access mode and station choices for Metro rail trips in Washington, D.C. The access station choice set was defined, based on predetermined modal impedance functions, as the two best access stations for the walk mode and the six best access stations for all the other modes. In a later study, Debrezion et al. (2009) developed a joint model of access mode and railway station choice. The estimated nested logit model was capable of jointly capturing both access mode and station location choice preferences. The provision of parking spaces and bicycle racks showed a significant positive effect on drive and bike access mode choices, respectively. However, the two studies (Debrezion et al., 2009; Mukundan, 1991) did not focus primarily on examining station choice for commuters who access the system using their private cars. Further, the sample frame used for the two studies incorporated a significant percentage of intercity trips, which, as explained previously, have different characteristics than long-distance or cross-regional commuting trips.

Vijayakumar et al. (2011) conducted a study to understand the variables that affect driving distance to suburban rail stations and demand for them. A multivariate regression model was developed to measure how individuals' socioeconomic attributes and station characteristics affect driving access to stations. The results showed that additional parking spots at park-and-ride locations and higher train frequencies attract more users to drive further to rail stations. In addition, better street connectivity to a rail station, as a measure of station accessibility, contributes to the total demand served by the station. Recently, Chakour and Eluru (2014) studied commuter rail users' mode and station choice behaviour using data from an on-board survey. The study focused on relaxing the hierarchical nature of the nesting choice decision structure using a latent segmentation approach in which access mode and station choice are tested to determine individuals' choice sequence. The trip, socioeconomic, and level-of-service attributes as well as built environment factors were used to study commuters' behaviour. The results showed that the developed latent segmentation model has higher explanatory power than the traditional nested logit models.

In light of the above, the access location choice of transit station is not a straightforward decision, especially for automobile access; in fact, there are many factors that affect individuals' choice decisions, such as access distance, the location of stations relative to home and work locations, station amenities, etc. In the GTHA, more than 30% of cross-regional commuters who access rail transit by car choose a station that is not the closest one to their home locations (DMG, 2013). Therefore, a better understanding of individuals' transit access station is needed to enhance the forecasting capabilities of travel demand models at the station level. Such models can be useful when defining factors that affect individuals' choice of access stations as well as stations' catchment areas. From the transit service provider's point of view, this can be beneficial for station expansion and/or service improvement planning. This stresses the importance of studying the choice of access stations for cross-regional commuters. As such, by filling this gap in the knowledge, high-fidelity operational demand models that are tailored for cross-regional travel, with special treatment of intermodal travel options, can be developed.

2.3.2 Main Mode and Access Mode Choice

Mode choice or mode split models represent the third step of the traditional four-stage Urban Transportation Modelling System (UTMS). Such models address the issue of selecting the most favourable travel mode alternative for each individual to conduct his/her travel plans. Mode choice models are considered by many researchers as the most important classical models in the area of transportation planning (Ortúzar and Willumsen, 2011). This section provides a literature review on travelers' mode choice and access mode choice within the context of cross-regional trips in multimodal transportation networks. The next section provides details of the state-of-the-art methods in discrete choice models that are used for modelling such choices.

The complexity of multimodal transportation planning and modelling stems partly from dissimilarities among the different modes in terms of their availability and level-of-service attributes. They are not equally substitutes; rather, each can be suitable and/or available for specific travellers or trip purposes (Fan et al., 1993; Litman, 2011). For intermodal trips, transit access mode choice is an important component in defining individuals' travel choices. Individuals may have options for access modes such as non-motorized modes (walking or biking), transit feeder services, or automobiles. In addition, travel modes within the same category (e.g. local transit and regional transit as public transit modes) have some similarities, but each of them has its own unique characteristics. This becomes clearer when studying the choice of access mode choices to different transit modes. For instance, studies on commuting mode choice in big cities have revealed that regional transit with local transit access, regional transit with walk access, local transit park-and-ride, and regional transit park-and-ride are often independent modal options. In the GTHA, it has been empirically proven that these four modal options are independent travel alternatives (Habib, 2013; Habib et al., 2009; Habib and Weiss, 2014).

In a multimodal transport system, different modes compete with and/or complement one another in a seamless, integrated fashion to provide mobility and accessibility for all trips, including intermodal trips. As such, travellers define a feasible choice set of possible travel mode alternatives before choosing one of them. According to the random utility maximization (RUM) concept adopted by McFadden (1973) and assuming that individuals are rational, the selection of a travel mode alternative is made such that it maximizes individuals' benefits. This approach assumes that individuals assign weights to each of the attributes that contribute to a satisfaction measure (a utility function), based on which they choose a travel mode alternative that maximizes their level of satisfaction. Apart from captivity constraints and/or household interactions, individuals typically tend to minimize the total generalized travel cost, including travel time and gas cost or transit fare, while maintaining an acceptable level of comfort and service convenience.

In general, the literature on access mode choice for transit services is rich. Tsamboulas et al. (1992) used an in-station survey to develop a multinomial logit model for access mode choice. Based on a population segmentation by trip purpose, they found that individuals' trip purpose has a significant effect on the access mode choice. For commuting trips, the availability and characteristics of travel modes for accessing and egressing transit stations play a critical role in transit use and ridership. Most of the related previous studies have focused largely on the access mode choice than on the egress mode choice because fewer modes are typically available for the latter. Sobieniak et al. (1979) and Korf and Demetsky (1981) have primarily studied and provided conceptual frameworks for modelling access mode choices of intercity and rapid transit stations using discrete choice models, respectively. Travel cost, access time, and waiting time came on top of the significant attributes, along with other relevant individual socioeconomic factors and trip characteristics that affect their decisions. Sobieniak et al. (1979) showed that considerations of passengers' convenience level significantly affect their access mode choice.

Previous research efforts have shown the dominance of non-motorized access modes, mainly walking, to access local transit services (Bergman et al., 2011; Cervero and Kockelman, 1997; Givoni and Rietveld, 2007; Park et al., 2014). However, many of the previous studies have not explicitly studied cross-regional commuters' behaviour of access mode choice and how they differ from commuter urban transit riders or non-commuter intercity travellers. Unlike local transit services within transit-friendly urban networks or dense city centres, regional transit services have wider station spacing; thus, users are more likely to use non-motorized modes to access such services (Wells, 1997). Therefore, automobile access to suburban transit services has become more popular in larger cities (Holguín-Veras et al., 2012; Meek et al., 2009). As such, special attention has been dedicated to studying and analyzing regional transit commuters' access mode choices, particularly automobile access such as park-and-ride (Li et al., 2007; Tsang et al., 2005; Washbrook et al., 2006).

Park-and-ride is often treated under the umbrella of automobile access to transit (Bos et al., 2004; Chakour and Eluru, 2014; Holguín-Veras et al., 2012). However, as discussed earlier, park-and-ride should be treated as a separate mode (Habib and Weiss, 2014). Kumar and Gur (1982) used a sequence of simple logit models to model the access mode choice given the use of transit as the main travel mode. The main mode choices included automobile, transit, rail, and express bus, while transit access mode choices included walk, bus, auto driver, and auto passenger. However, Fan et al. (1993), in a later study, showed that their approach was not fully consistent with the random utility choice theory. Alternatively, Fan et al. (1993) studied access mode choice and access station location choice for commuter rail users in the GTA. A nested logit model with access mode choice at the upper level and access rail station choice at the lower level was developed. As such, the choice of the main travel mode was not investigated in their study. In addition, parking price as well as other station-level factors were not included in the developed models. Li et al. (2007) applied a network equilibrium formulation to model commuters' park-and-ride mode choice in a multimodal network context considering three modes: auto, transit with walk access, and park-and-ride. The study showed that traditional demand models often misestimate park-and-ride demand. The study concluded that introducing park-and-ride facilities may affect the overall system positively or negatively depending on several factors, including parking charges and offered spaces. Therefore, capturing the effects of station-level attributes and the relative station locations with respect to individuals' home/work locations is essential when studying access mode choice in general and park-and-ride users' behaviour in particular.

2.4 Current Practice of Discrete Choice Models

As described above, cross-regional intermodal trips involve three main choices: main mode, access mode, and access station location choices. Typically, discrete choice models are used for modelling such choices. That is, individuals select an option from a finite subset of a universal set of alternatives (Ortúzar and Willumsen, 2011). Assuming that individuals are rational, they select the option that maximizes their benefits. This approach has been adopted within the conventional four-stage model for decades. Traditionally, the mode choice components of such frameworks do not account for the explicit interactions between the three choices of main mode, access mode, and access station location. As such, the main mode and access mode choices are not often explicitly considered for trips that involve the use of transit with a combination of other travel modes. Similarly, the choice of access station location is not considered in many of the traditionally developed models; rather, individuals are assumed to choose the closest station to their trip origins (Lu, 2003). As a result, limited policy measures targeted at integrating transit with other modes at the regional level can be analyzed. Alternatively, the conventional four-stage aggregate modelling approach can be reformed to a series of discrete choice models in a disaggregate fashion with enhanced mode choice models that are capable of capturing the interactions between different decisions. Therefore, choice decisions can be formulated as sequential or simultaneously joint discrete choice models. In its simplest form, a series of discrete choice models is used to represent a finite set of alternatives that are available to the individuals, such as making a trip, destination location, mode choice, and route choice, by which this decision structure mimics the four stages of the UTMS: trip generation, trip distribution, mode choice, and route choice.

An overview of the development of discrete choice models is presented in Ben-Akiva and Lerman (1985), Khan (2007), Manski and McFadden (1981), and McFadden (1973, 1984). Various decision rules have been discussed in the literature, among which the utility concept is most commonly used in the context of discrete models. Traditionally, the random utility maximization (RUM) concept adopted by McFadden (1973) has been used to model travel mode choices assuming that utilities are random to the modeller and that the choice is deterministic from the decision maker's perspective. Based on individual-specific weights assigned to travel mode alternatives' attributes, individuals choose one alternative that maximizes their utility with respect to other possible alternatives. In other words, the probability of choosing a particular alternative is proportional to the difference between its estimated utility and the estimated utility of other available alternatives (Koppelman and Bhat, 2006; McFadden, 1973; Meyer and Miller, 2001). In general, discrete choice models postulate that "*the probability of an individual choosing a given option is an auction of their socioeconomic characteristics and the relative attractiveness of the option*" (Ortúzar and Willumsen, 2011). Those characteristics and attributes allow researchers to derive the utility formulation of each alternative, which is composed of observed and random components. Typically, the observed utility is formulated as a linear function in which the endogenous variables represent both individuals' and alternatives' characteristics. The unobserved random utility component is represented as a non-systematic error term. This error term can be assumed to follow different distributions that are defined based on the assumed correlation between random residuals. Consequently, various types and mathematical formulations of discrete choice models have been developed.

The logit model family is the most widely used structure to develop discrete choice models; therefore, it is considered the core of consumer behavioural analysis and choice modelling. The multinomial logit (MNL) model is the simplest logit formulation since the parameter estimation process can be done in a tractable closed-form manner. Despite being the most commonly used discrete choice model, the MNL model formulation suffers from limitations such as the independence from irrelevant alternatives (IIA) property, which is the constant ratio of any two alternatives' choice probability irrespective of the existence of a third alternative. In other words, the individuals' decision in choosing one of two alternatives is independent of other alternatives (Koppelman and Bhat, 2006). Such a limitation results from the assumption of the random error terms that are identically and independently distributed (IID) following the double exponential (Gumbel Type I extreme value) distribution with a homogeneous matrix of variance-covariance across all alternatives. In other words, the error terms are assumed to be uncorrelated and have the same variance across all alternatives. Hence, the difference between the error terms is distributed logistic. To overcome the MNL model limitations, other structures have been developed, such as the multinomial probit models (MNP), generalized extreme value (GEV) models, and mixed logit models (McFadden, 1986).

The multinomial probit (MNP) model structure results from the assumption that the error term distribution is normally distributed. The mathematical complexity of the MNP model formulation limits its practical use in terms of parameter estimation and interpretation. In contrast, generalized extreme value (GEV) models are developed based on a generalization of the extreme value distribution that allows for capturing the correlation across error terms of a set of specified alternatives. A nested logit (NL) model (one of the most common applications of the GEV model family) structure accommodates different degrees of similarity between subsets of alternatives and relaxes the IIA assumption of the MNL model. The NL model structure assumes an equal covariance of alternatives within each nest (subset of similar alternatives) and zero correlation across other nests. Similar to the NL model structure, Bhat (1997) and Koppelman and Wen (2000) introduced a closed form of the covariance nested logit (COVNL) and paired combinatorial logit (PCL) models, respectively, which allow for heterogeneity across alternatives by introducing different covariance among the nested alternatives. Alternatively, in addition to the IID distributed error term, an additional random error component can be added to the capture the correlation and taste of variation across individuals and alternatives. Adding this additional random error term results in the mixed logit model formulation. One of the early applications of mixed logit models that allows for several random coefficients simultaneously is that of Revelt and Train (1998) and Bhat (1998). Several other applications of the mixed logit models have been used in the area of discrete choice models. A detailed review of the different model formulations and their practical uses can be found in Khan (2007).

Most of the aforementioned model formulations do not explicitly consider captivity constraints. The definition of users' captivity, modal availability, or the omission of specific modes from individuals' choice sets greatly affect the accuracy of studies of mode choice behaviour (Stopher, 1980; Tardiff, 1976). This situation is more prominent in the case of cross-regional trips where commuters have a large number of travel mode alternatives including intermodal travel options in their universal choice set. However, not all modes within the pre-defined universal choice set are considered by individuals. One case of alternative omission is the dependent or captivate case, whereby the decision maker rejects all alternatives but one, thereby guaranteeing the selection of a single alternative within the universal choice set. In many choice-modelling applications, some or all aspects of this dependency are often ignored completely. Typical choice models that follow the multinomial logit formulation often apply rules to define modal availability (i.e. if the decision maker does not have a driver's license, then driving is not an available alternative to them), thereby capturing what has been previously defined as "forced captivity" (Jacques et al., 2013). This form of forced dependency can be determined using logical rules and is therefore relatively easy to encapsulate within most model specifications.

Conversely, individuals may have no desire to consider some of the alternatives that may be readily available to them and therefore are implicitly dependent on other modes of travel. This implicit dependency may be a result of individuals' preconceptions of specific alternatives or their travel inertia towards these modes. For instance, perceived higher mobility with a car or viewing bicycling to work as an efficient way to exercise may implicitly cause a commuter to be attracted to one mode over another. Therefore, even if a mode may become more attractive from a level-of-service perspective, the decision maker might not consider the rational trade-off between modes (i.e. based on their level-of-service attributes) due to unobserved reasons. Unfortunately, it is difficult to capture directly such implicit decisions regarding which alternatives to be considered within individuals' choice sets. Therefore, modellers often treat the perception of modal dependency as a latent choice. Despite these modelling challenges, understanding how latent/implicit dependency affects commuting mode choice is highly relevant for policy- and decision-makers. A recent study by Chouros and Dellaert (2010) presented a mode choice model considering travellers' inertia and the learning-based lock-in effects. The results suggested that travellers' inertia towards a specific mode may be harder to 'break' compared to conventional mode choice models' forecasted mode-switch. This stems from the notion that dependent users will inherently not consider alternative modes even if they become more attractive from a level-of-service perspective. This has implications for determining policies to encourage mode-switching behaviour, as discussed in Habib and Weiss (2014).

Modal dependence and choice set generation are concepts of significant interest and development within the field of discrete choice modelling. Manski (1977) introduced the problem of choice set generation for decision-makers. However, a limited discussion on the concept of latent modal dependence was presented; this discussion postulated that the latent choice set can be indirectly inferred through sufficient details and personalized information regarding the decision-maker and associated choice scenario. Such information is often not available within collected travel behaviour data. Therefore, an alternative method to capture choice set generation and, by extension, choice captivity is required to fill this gap in the literature.

Stopher (1980) showed that the failure to capture latent alternative consideration within a choice modelling structure often results in model misspecification. Further, Swait and Ben-Akiva (1986) highlighted the need to understand the effect of latent modal dependence on modal availabilities and choices in regions with rapidly changing social and economic structures. The dogit model was the first model structure that addressed the latent alternative reliance concept in the context of modal choice (Gaundry and Dagenais, 1979). The dogit model was originally formulated as an alternative to the logit structure to relax the IIA property inherent within the multinomial logit (MNL) formulation discussed earlier. Swait and Ben-Akiva (1987a, 1987b) further interpreted the dogit model as capable of predicting alternative reliance. The dogit model consists of two portions, an irrational[1] dependent portion and a rational portion. Swait and Ben-Akiva further expanded the constant-only captivity odds parameter present in the irrational portion of the original dogit model into a parameterized function. The analysis performed by Bordley (1990) suggested that the original form of the dogit model and the parameterized version are still valid for predicting choice behaviour under incomplete dependency conditions. This finding relaxes the concerns in prior applications of the model in which decision contexts were considered where complete alternative reliance was not guaranteed.

Although the first attempts to use the dogit and parameterized logit captivity (PLC) models proved to have significant improvements over corresponding MNL models estimated from the same data (Swait and Ben-Akiva, 1986, 1987a), limited applications of these models have been reported. Both the dogit and PLC models have closed forms of the likelihood functions, making the estimation process relatively straightforward and comparable to the standard MNL model. While there have been relatively few recent applications of the dogit and PLC models, a few notable attempts have been made to apply these models to different contexts. McCarthy (1997) applied the PLC model to intercity travel mode choice and found that the captivity model structure outperformed a model structure that did not consider dependency. Chu (2009) utilized the dogit generalized extreme value (DOGEV), originally proposed by Fry and Harris (2005) to model

[1] The use of “rational” and “irrational” terms in this context does not imply that individuals are classified to either category. Such models provide a probability that individuals’ choices are rational or irrational from the modeller’s perspective. It is assumed that a rational choice is made according to an explicit trade-off between the level-of-service attributes of the travel modes available to an individual. Therefore, the latent classification of irrational users only indicates that these users are not making the explicit trade-off between the available modes to them and as such not maximizing the specified model’s systematic utility. However, that does not imply that they are irrational individuals; they may be making a rational choice according to factors that unobserved to the modeller.

commuters' departure time choice. The DOGEV model structure replaces the rational conditional choice portion of the dogit model structure with a GEV model structure. In later applications, Chu (2010) examined the destination location choice using a form of the dogit model as a means to improve network-modelling capacity. Recently, the work of Habib and Weiss (2014) presented a version of the PLC model that was estimated with three repeated cross-sectional datasets for commuting mode choice. The scale parameter was parameterized within the rational portion of the choice model to account for heterogeneity across the different panel datasets as well as spatial variability. Consistent with other applications of the dogit and PLC models, the estimated model outperformed the more traditional MNL formulation.

In terms of model estimation, several estimation techniques are applied to estimate parameters' coefficients of the utility function, among which the maximum likelihood estimation is the most commonly used method. However, because of the complexity of other model structures, using advanced model estimation techniques has become an essential aspect of model development. As explained above, using more complex models (other than the MNL model), researchers can theoretically develop more behaviourally realistic models that account for taste variation and heterogeneity across the population. However, the practical implementation of many of these models is still an open and active area of research. The reason for this limitation is that some of these model formulations can no longer be solved in a closed-form routine. Therefore, simulation-assisted estimation techniques that provide a practical way to estimate complex logit models have been investigated (Train, 2008, 2009).

2.5 Current Practice in RP/SP Survey Design

2.5.1 The Need for RP/SP Surveys

Studying cross-regional travellers' behaviour requires exhaustive data on their trip patterns, including detailed information on each trip leg, such as access, transfer, and egress times. Typical commuting travel surveys do not provide sufficient data to conduct this type of analysis. This is the case for several reasons: cross-regional trips are often underrepresented in survey samples, the collected data do not provide the necessary level of detail on inter- and intra-modal trips, and the majority of typical travel surveys rely predominantly on revealed (observed) preference (RP) trip data. Previous research efforts have shown that RP data do not adequately capture the behavioural trade-offs involved in the travellers' decision-making process concerning new policy initiatives.

In other words, demand models developed based on RP data only are incapable of accurately forecasting individual choices in response to scenarios that do not exist (such as new transportation policies or the introduction of new modes that have never been used before) (Idris et al., 2012; Louviere et al., 2000).

To overcome the limitations of RP models, some researchers have replaced RP data with stated preference (SP) data. SP surveys are used to measure individuals' preferences towards hypothetical scenarios by asking the respondents questions about services or policies that do not exist (Hensher, 1994; Hensher et al., 1988; Louviere and Hensher, 1983). A summary of the advantages of using SP data over conventional RP data for travel behaviour analysis can be found in Kroes and Sheldon (1988). However, SP data have their own drawbacks. Previous studies have shown that individuals' stated preferences may not be consistent with their actual choices, which induces a systematic bias in the data (Wardman, 1988). Alternatively, using joint RP/SP data allows for scale adjustment of parameter estimates to correct the systematic bias of the SP data (Hensher et al., 2008). As it stands now, joint RP/SP surveys represent the state-of-the-art approach for travel behavioural data collection, in which, behavioural factors along with typical socioeconomic attributes are gathered to develop accurate econometric models that can explain the probabilistic response in accordance with changes in transportation level-of-service attributes as a result of the introduction of new policies.

2.5.2 SP Experimental Design

The purpose of performing an SP experiment is to quantify the independent effects of the design attributes on respondents' choices (Carson et al., 1994; Louviere and Hensher, 1983; Louviere and Woodworth, 1983). In general, the experiment design becomes more complex as the number of attributes and their levels increase. Therefore, it is desirable to keep the number of attribute levels as low as possible. Although the minimum number of attribute levels is two, if the attribute is expected to have a non-linear influence on the dependent variable, then at least three levels are required (Idris et al., 2012; Rose et al., 2008). In addition, maintaining balanced utilities and attribute-level ranges are desirable properties to increase the efficiency of the SP experiment design. Previous research has shown that, while it is statistically preferred to have a wide range of attribute levels, extremely wide ranges may result in choice situations with one dominant alternative (Carson et al., 1994; Caussade et al., 2005). As such, striking a balance among the alternatives' utilities helps reduce the chances of having a dominant alternative within any of the choice scenarios and therefore maximizes the information gathered from each choice task (Huber and Zwerina, 1996). In addition, maintaining a balance among the attribute levels by showing all levels equally the same time across the choice tasks provides sufficient data for parameter estimation (Bliemer and Rose, 2009; Caussade et al., 2005).

Several previous studies have relied on an orthogonal experimental design to develop SP surveys. However, recent studies have shown that efficient designs outperform orthogonal designs (Rose et al., 2008). In general, efficient designs aim at finding SP experiments that allow for parameter estimation with the lowest asymptotic standard error (Zwerina et al., 1996). Such designs require prior estimates of attribute parameters based on a specific model structure that can be obtained from similar studies or a pilot survey. While different measures of design efficiency have been used in the literature, the D-efficient design is the most common. The D-efficient design aims at finding designs that allow the model structure to estimate parameters with the smallest possible standard error. Details on efficient SP choice experiments and survey design can be found in Choice-Metrics (2012) and Idris et al. (2012).

2.5.3 Customized SP Experiments and Trip Planner Tools

The issue of generating an SP experiment (i.e. choice scenarios), including the choice alternatives, alternative attributes, and attribute factors, is crucial to the survey design. Therefore, the choice experiments have to be realistic, self-explanatory, and non-trivial to the respondent. The latter two are managed within the experimental design through the alternative attributes' definitions and balance, as explained in the previous section. However, to design a realistic choice situation for all respondents, the choice scenarios must be customized for each individual. In the context of cross-regional trips, as discussed previously, the universal choice set of the travel alternatives may include a huge number of options. Some of these options may be non-feasible to the survey respondents. Therefore, introducing an SP experiment where all alternatives are included for all survey respondents is not appropriate; it increases the respondents' burden and creates unrealistic choice situations. Using data collected based on unrealistic choice scenarios can lead to severe problems when used for modelling individuals' travel behaviour. Therefore, generating realistic choice scenarios is the key to ensuring the quality of the collected data.

In this context, to generate respondent-customized SP experiments, a trip planner tool is required. As such, for each individual, the tool generates only feasible choice alternatives with customized level-of-service attributes according to the individual's travel plans. Several studies have adopted the idea of generating respondent-customized SP experiments; however, most of those studies customized the scenarios by changing only the level-of-service attributes according to individuals' trip characteristics (Idris, 2013). In addition, the generation of level-of-service attributes is often done at the zone (i.e. aggregate) level without accurate consideration of individuals' trip origins/destinations. Moreover, the interaction between different modes of travel in the context of intermodal trips is not explicitly defined. One of the few studies that truly developed a respondent-customized SP choice scenario at the individual level was developed based on an online transit journey planner tool (Schmitt et al., 2014). However, as the name implies, the survey and the planner tool catered only to transit users. That is, a new method that integrates a multimodal trip planner tool within travel survey platforms to generate respondent-customized SP choice scenarios is essential. Further research is needed to fill this gap in in the area of travel survey methods to develop robust data collection tools for cross-regional trips, many of whom have viable intermodal options.

2.5.4 RP/SP Data and Mode Choice Models

Several studies have investigated individuals' travel behaviour using RP/SP data (Ben-Akiva and Morikawa, 1990; Ghosh, 2001; Hensher and Bradley, 1993; Imaz et al., 2014; Polydoropoulou and Ben-Akiva, 2001). In a recent study by Habib et al. (2014), data from an RP/SP survey on parking price levels at park-and-ride stations was used to develop a heteroscedastic mode choice model. The study showed that a relatively small RP/SP dataset can provide a good understanding of individuals' elasticity regarding policy changes. In another study, data from an RP/SP commuting survey were used to investigate the influence of transit service attributes on mode-switching behaviour (Osman et al., 2014). The results of SP-only and joint RP/SP mode shift models were compared, and the study concluded that the inclusion of joint RP/SP data has a positive effect on improving the goodness of fit and explanatory power of the SP-only model.

Whitehead et al. (2008) presented a review of several transportation-related studies on the joint estimation of RP/SP data. A classification of the type of data combination and econometric models, followed by a discussion about the advantages and disadvantages of each type, is presented. Based on their review, the study concluded that the combination of RP and SP data allows for utilizing the advantages of each dataset while avoiding the potential faults that may arise in relying solely on one of them. As discussed previously, RP data are limited to the current conditions that the individuals may encounter in existing travel scenarios. However, SP data extend this experience to conditions that do not currently exist, allowing decision-makers to use such data to forecast hypothetical scenarios. Therefore, jointly estimated RP/SP models have been found to outperform independently estimated RP and SP models. In addition, combining RP and SP data increases the efficiency and explanatory power of the estimated model.

In terms of using RP/SP data for discrete choice model development, several studies have used pooled RP/SP data (Haener et al., 2001). In other words, the estimation routine inherently assumes that the data come from two different populations or datasets. In addition, other studies have treated the repeated SP choice scenarios as independent records (Hensher et al., 2008; Whitehead et al., 2008). However, to estimate RP/SP models jointly, each data record should represent the RP information and all repeated SP choice scenarios for the same individual across the survey population. Brownstone et al. (2000) developed joint RP/SP MNL and simple mixed logit models for vehicle type choice. The study emphasizes the importance of combining RP/SP data to exploit the strengths and mitigate the weaknesses of each type. In addition, the model estimation results showed that advanced model structures are statistically better than the standard logit model. Similarly, Hensher et al. (2008) used pooled RP/SP data to develop nested and panel mixed logit mode choice models. Despite the wide range of applications of joint RP/SP models, most of the attention has been given to advances in model estimation without giving proper attention to the use of such models for forecasting. This limits the practical aspects of jointly estimated RP/SP models.

The development of joint RP/SP models may involve the estimation of coefficients that are uniquely determined by the RP or the SP data, coefficients that are equally identified within the two datasets (before applying the scale effect), and coefficients that are identified differently within the two datasets. Despite the recent advances in joint estimation of RP/SP models, there are no clear guidelines in the literature for classifying data-specific or pooled coefficients (Brownstone et al., 2000). In a few cases where all attributes are considered generic across both datasets, using a joint RP/SP model for prediction can be less problematic (Cherchi and Ortúzar, 2006). However, in most cases, data-specific coefficients must be estimated. For instance, SP attributes are often not available in the RP context and therefore cannot be estimated using the RP data. Similarly, different coefficients of the same attribute can be estimated across the two datasets, which can be tested empirically. Cherchi and Ortúzar (2006) presented a discussion of the development and application of joint RP/SP models when the RP and SP data show different systematic or random taste heterogeneity. Using a set of NL models, the empirical results showed that the estimation of SP interaction terms improved the models' forecasting performance. That is, a few studies have attempted to investigate the issue of using jointly estimated RP/SP models for policy analysis and forecasting; however, the literature in this area is clearly limited.

2.6 Chapter Summary

This chapter presents a review of the literature of intermodal travel behaviour in multimodal networks with an emphasis on cross-regional travel demand modelling. Similarly, the current practice of travel demand models with further emphasis on discrete choice models is presented. The literature on cross-regional travel demand modelling is evidently limited, and more research efforts are required to fill this gap. Because of the intermodal nature of various travel options available to them, cross-regional commuters may face complex choice situations in choosing their mode of travel. According to the cited literature, a comprehensive framework that encapsulates all possible decisions faced by individuals (such as mode, departure time, and route choices) in the context of cross-regional trips does not exist. Moreover, most of the existing regional demand models are developed using large-scale datasets that are heavily imbalanced towards intra-regional commuters. Therefore, the results of such models are highly influenced by the behaviour/travel patterns of intra-regional commuters. That is, there is a need to direct research efforts towards investigating the behaviour of this significant segment of the population (i.e. cross-regional commuters). Therefore, policy-sensitive mode choice models that can be used as a core component of a detailed travel demand framework for modelling cross-regional trips are required to fill this significant gap in the literature.

To develop behavioural models that are capable of capturing changes in individuals' decisions, exhaustive data on their travel plans are essential. In particular, choice models that are developed using stated preference (SP) data are capable of investigating newly introduced policies (i.e. policies that have never been applied before) that cannot be tested using models that are developed using revealed preference (RP) data. As such, an in-depth review of the current practice of data collection and survey design methods is presented. Several studies on the state-of-the-art data collection method (i.e. joint RP/SP surveys) and the corresponding types of choice models are reviewed. Finally, the practical aspects of using joint RP/SP models in forecasting are discussed.

CHAPTER 3

3 CONCEPTUAL FRAMEWORK

3.1 Chapter Overview

This chapter presents a conceptual framework of modelling cross-regional commuting trips. Section 3.2 provides an introduction and contextual background of the framework. In Section 3.3, the components of the framework are described and analyzed. Section 3.4 presents the elements of the conceptual framework to be investigated in this study. Finally, a summary of this chapter is provided in Section 3.5.

3.2 Introduction

As explained in Chapter 2, cross-regional trips are defined as trips that cross boundaries of municipal or regional jurisdictions that have different transit operators. Such trips are different than regular commuting trips because of the longer travel distance and/or the typical lack of service and fare coordination between transit service providers. Therefore, cross-regional commuting trips have distinct characteristics relative to other types of trips such as the longer trip distances and the potential usage of different travel modes. In addition, cross-regional commuters have distinct socioeconomic/demographic attributes such as high income and auto ownership levels (see Chapter 6). Such characteristics may influence individuals' travel decisions substantially.

Individuals' choices and their associated travel decisions are not straightforwardly determined. In fact, such decisions involve a series of trials by each individual until his/her travel plan is defined. In other words, assuming that travellers are rational, the decisions associated with their travel choices can be described as an iterative process in which individuals learn about the feasible travel options and accordingly choose the most favourable option based on their specific trip characteristics. Therefore, to investigate individuals' travel behaviour, it is desirable to develop a modelling framework that is capable of capturing inclusively all the important decisions faced by the travellers and the interactions among them. That is, to be capable of capturing individuals' decisions and how they choose their travel options, an ideal framework would consist of a series of disaggregate/agent-based model components that are interconnected to each other in a continuous feedback fashion. As such, at the end of each iteration, a pre-defined performance criterion (e.g. a percentage change in mode share or travel time) is assessed, and accordingly, the iterative procedure is either re-performed until convergence or terminated.

3.3 Conceptual Framework of Modelling Cross-Regional Commuting Trips

For typical cross-regional commuting trips with pre-defined home/work locations, commuters face three main choices in devising their regular travel activity schedules: departure time, travel mode, and route choices. Commuters define their departure time from home based primarily on their desired arrival time, the expected travel time of their chosen travel mode, and the corresponding travel route. Similarly, the choice of a travel mode depends on individuals' characteristics, the built environment, and the level-of-service attributes of each travel mode alternative at the chosen departure time. The route choice decision is determined based on the network attributes and/or transit schedules at the chosen departure time for a specific travel mode. Clearly, the three choices heavily influence each other, so individuals' travel plans are shaped jointly according to the three choices.

In light of the aforementioned discussion, a conceptual modelling framework that is capable of replicating similar choice situations faced by cross-regional commuters is presented herein. The modelling framework, shown in Figure 3-1, consists of a set of components starting with defining the travel demand and its distribution, followed by three main sub-models, including a departure time choice model, an enhanced mode choice model, and a multimodal trip assignment model. It should be noted that the order of applying these models is purely empirical and it does not imply that individuals are assumed to make their travel decisions in the same order. The framework takes individuals' work and home locations, their socioeconomic and demographic characteristics, and the network attributes and transport system constraints as inputs. To enhance the framework's realism and capability of replicating real-world scenarios, only individual-customized feasible travel alternatives, network attributes, and capacity constraints are considered within its different elements. Since the three decisions mutually affect each other, the framework is designed to allow the three corresponding components to complement each other seamlessly. Although this study places more emphasis on modelling the mode choice component, the interaction among the framework components is carefully considered. The decision of departure time choice is simulated based on a discrete choice model developed by Sasic and Habib (2013). An agent-based multimodal trip assignment that executes individuals' trips from home to work based on their simulated departure time and travel mode choices was developed by Weiss (2013). As such, traffic and transit assignments are performed simultaneously in a disaggregate setting, which overcomes the drawbacks of the conventional unimodal aggregate assignments in which the disaggregate inputs from preceding models were aggregated, sacrificing the fine details of their outcomes. Accordingly, the outcome of the trip assignment component includes individuals' travel plans and the corresponding attributes of their simulated choices. Based on the updated level-of-service attributes, new travel choices are simulated for each individual according to the developed models in an iterative procedure. Conceptually, after applying the framework multiple times until the target level of a performance criterion is reached, the final output is a detailed disaggregate probabilistic chain of decisions, including departure time, main mode, access mode (if any), access station location (if any), and driving and/or transit route choices. That is, the framework allows for the study of individuals' travel choices in a disaggregate fashion.

The travel demand and its distribution are assumed constant throughout the analysis period. Such an assumption is fair as long as the framework is used for modelling users' behaviour in the short to medium terms. In other words, the framework does not explicitly model long- or medium-term decisions such as changes in home or work locations and auto ownership levels. That is, this framework is appropriate for understanding individuals' behaviour and/or for studying the potential changes in their travel decisions as a result of policies targeted to influence short-term decisions such as departure time, travel mode, and route choice. Therefore, policy analysis tools developed using this framework are capable of providing decision-makers with relevant measures for delivering planning solutions in accordance with initiatives under investigation given the aforementioned constraints.

In summary, ideally the framework for modelling cross-regional commuting trips is designed as a series of continuously updated discrete choice/agent-based models. That is, for a specific target population, assuming fixed demand and home/work locations, the three choice components (i.e. departure-time, travel mode, and route choices) act as an integrated framework. As such, the framework can be used as a policy analysis tool for testing various transportation planning scenarios and their effects on travel demand.

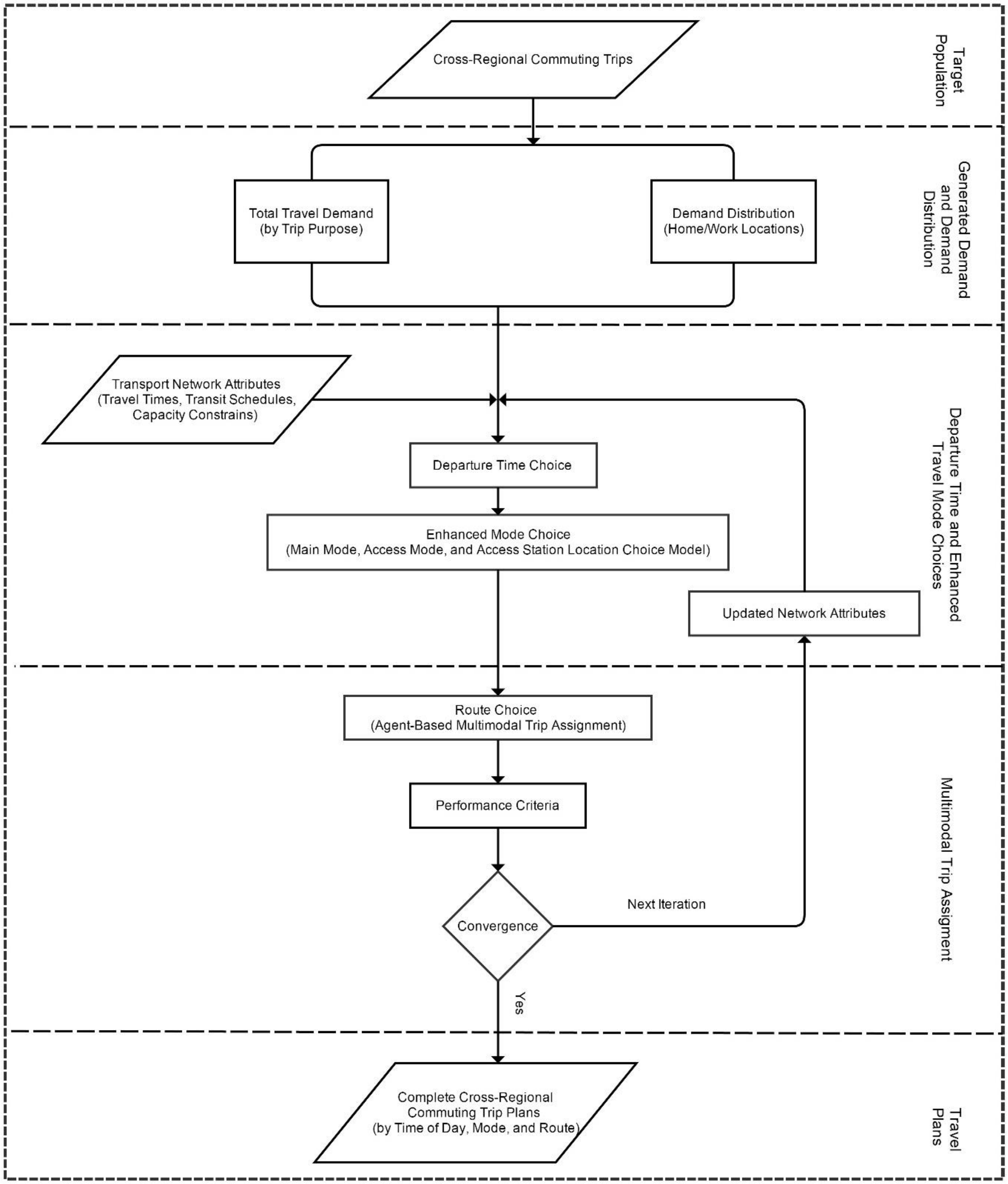


Figure 3-1 Conceptual Framework

3.4 Framework Components under Investigation

The primary reason for developing a conceptual framework for modelling cross-regional trips is to fill some of the aforementioned gaps in the literature on demand modelling in general and in the area of modelling cross-regional commuters' mode choice in particular. Therefore, the developed framework places a strong emphasis on the development of an enhanced mode choice component. In this section, the structure of the enhanced mode choice component is presented. As discussed in Chapter 2, previous research efforts have shown that models incorporating the integration of main mode and access mode choices or access mode and access station location choices in joint decision structures outperformed models developed assuming isolated decision structures. Many of the previous attempts to model similar decision structures considered the joint interaction between two of the three decisions faced by individuals. However, a joint decision structure that encapsulates the three choices (i.e. main mode, access mode, and access station) would provide a better understanding of cross-regional travellers' mode choice behaviour. As shown in Figure 3-2, instead of considering only two of the three choices (as shown on the left side), the conceptual model structure is composed of a joint tri-variate decision structure of main mode, access mode, and access station choices.

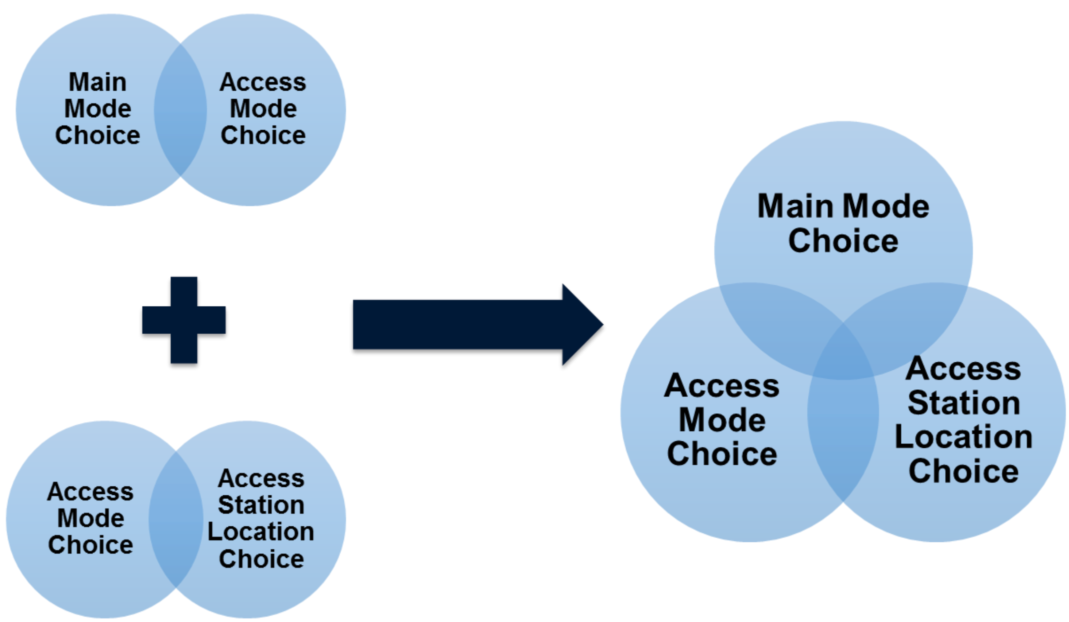


Figure 3-2 Tri-variate Joint Decision Structure

The three-level decision structure of the enhanced mode choice component can be presented as shown in Figure 3-3. Within the three-level joint structure, each choice affects the other two choices. For instance, assuming a trip in which commuter rail is chosen as the main mode, the choice of an automobile access mode or a non-motorized access mode affects the access station location choice. In addition, each level defines the universal choice set of alternatives for a certain choice decision, from which only a set of individual-specific feasible alternatives is considered based on attributes of both individuals and alternatives.

The first level defines the main mode choice. The main mode options are divided into three categories: automobile, transit, and non-motorized modes. The automobile modes include auto-driver and auto-passenger options, the transit modes include local transit and regional transit options, and the non-motorized modes include walking and biking options. The decision structure clearly distinguishes between regional and local transit services to account for the differences in the level-of-service characteristics of each type of transit service and, accordingly, their different targeted customers. In addition, according to the definition of cross-regional trips, non-motorized modes are considered non-feasible travel options and therefore excluded from individuals' feasible choice sets.

The second and third decision levels define the access mode and access station location choices, respectively. The second level defines the access mode choice conditional on choosing a transit mode as the main mode of travel. Both local and regional transit can be accessed by non-motorized modes, feeder local transit modes (operated by the same or a different service provider), and auto access, including park-and-ride (i.e. auto driver access) or kiss-and-ride (i.e. auto passenger access). Finally, the access station choice is defined at the third level based on the main mode and access mode choices. To define a feasible and realistic choice set of access station locations, a customized access station choice set is generated for each individual based on the chosen access mode. For instance, only regional transit park-and-ride stations are included within individuals' choice sets if the main mode is defined as regional transit and the access mode is defined as auto driving. Similarly, based on transit stations' proximity with respect to individuals' home/work locations, the closest station can be assumed to be the access station for transit options with non-motorized access.

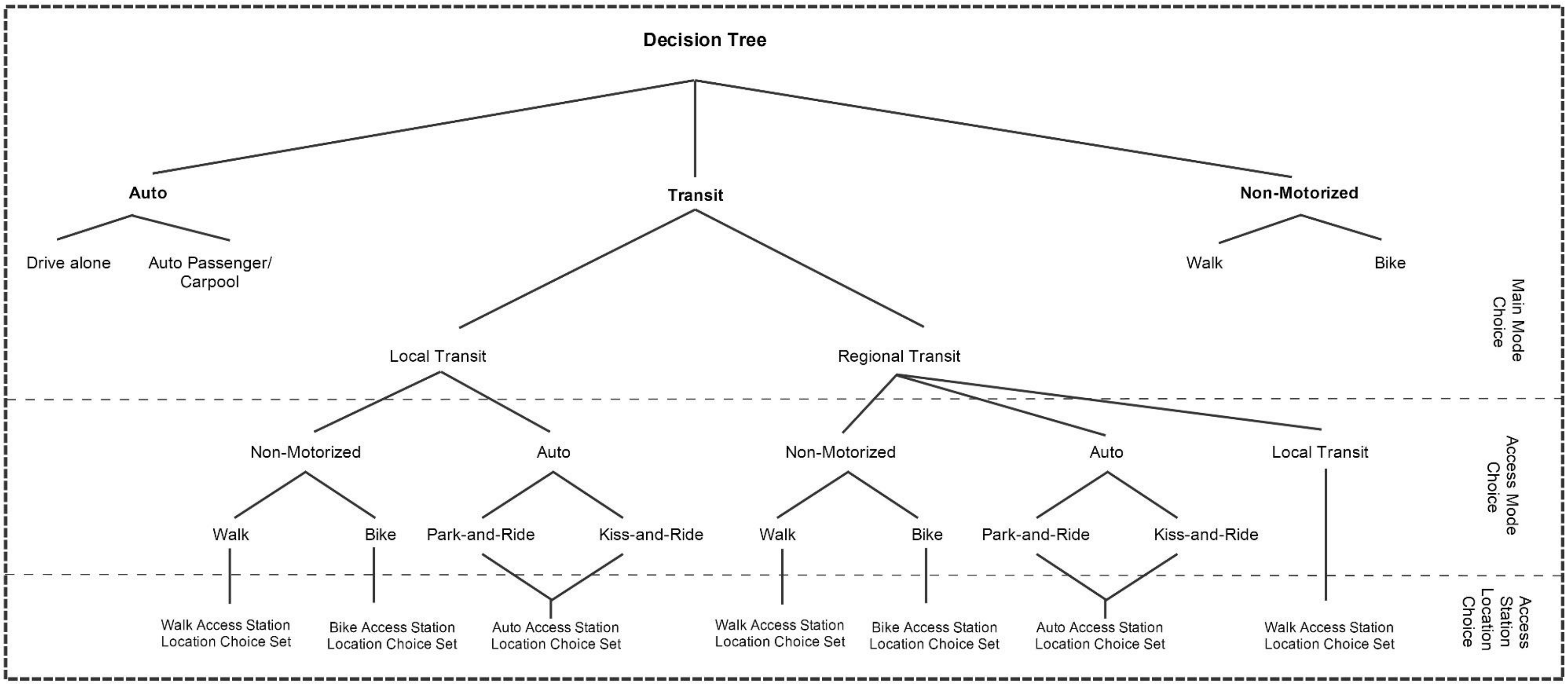


Figure 3-3 Decision Tree Structure

As explained in Chapter 2, choice models that are developed using stated preference (SP) data are capable of investigating policies that cannot be tested using models that are developed using revealed preference (RP) data. In other words, SP data along with detailed travel information are required to develop models that are capable of capturing changes in individuals' travel behaviour in response to hypothetical scenarios. These scenarios are presented to sample travellers to choose their most favourable travel option. The survey and its SP experiment are designed in view of policies that are required to be investigated, such as the introduction of new travel modes and technologies, the integration of services and fares among modes, and changes in the level-of-service attributes of current modes. The combination of all possible travel alternatives would increase the complexity and number of choice scenarios, add to the respondents' burden, and may violate the fundamentals of the choice theory. As such, only feasible travel options should be shown to the survey respondents.

However, developing such a three-level nesting structure would require extensive data on individuals' behaviour and may encounter various challenges in terms of the empirical estimation of the model. More importantly, this decision structure will expand the number of generated alternatives which might lead to cases where the choice situation is not consistent with behavioural aspects of the choice theory (i.e. by assuming that individuals consider as many alternatives when making a choice). Alternatively, for intermodal transit trips, access station choice is first predicted to generate feasible and customized travel options, then accordingly main mode and access mode choices are defined. That is, the access station choice is implicitly considered within the main mode and access mode choices.

In light of the above, the framework is developed over two phases, as shown in Figure 3-4. In phase I, data from an existing revealed preference travel survey are used along with detailed information about transit stations to develop an access station location choice model (Chapter 4). In phase II, utilizing the developed access location choice model and a multimodal trip planner tool, a revealed preference/stated preference (RP/SP) survey is designed and executed to collect data on commuters' current travel plans as well as their decisions in response to hypothetical scenarios (Chapters 5 and Chapter 6). Finally, the collected joint RP/SP data are used to develop the proposed enhanced mode choice model (Chapter 7).

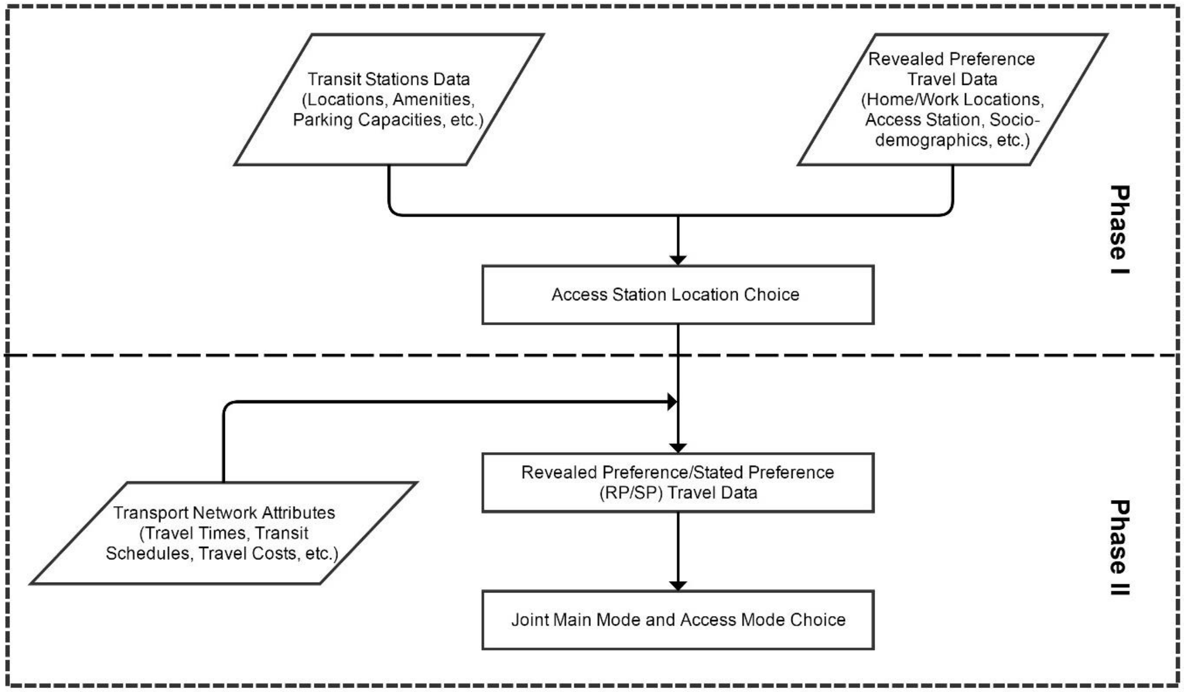


Figure 3-4 Framework Development

The framework offers a practice-ready modelling platform for cross-regional trips. The total demand and its distribution are loaded onto the travel demand model. Each individual is assigned a departure time for his/her trip from home to work according to the departure time choice model (Sasic and Habib, 2013). Afterwards, the access station location choice model provides a set of probabilities for station locations to be chosen by individuals who have transit modes with auto access available in their travel mode choice sets. Those probabilities are used to determine the expected level-of-service attributes for transit modes with auto access modes to be used in the joint main mode and access mode choice model. Then, the joint main mode and access mode choice model (which implicitly takes into account the access station location for transit modes with auto access) provides a set of probabilities for each combination of main mode and access mode (if any). The discrete choices of main mode, access mode, and access station location choices are randomly chosen based on the estimated probabilities before being passed to the multimodal trip assignment model (Weiss, 2013). Using individuals' simulated choices, the multimodal trip assignment model generates new level-of-service attributes of individuals' modes of travel. The new level-of-service attributes are fed back to simulate new choices in an iterative process until the model reaches convergence. Finally, the model output is a detailed disaggregate probabilistic chain of decisions, including departure time, main mode, access mode (if any), access station location (if any), and driving and/or transit route choices. That is, the framework allows for the study of individuals' travel choices in a disaggregate fashion.

3.5 Chapter Summary

This chapter presents a conceptual framework that is capable of capturing cross-regional commuters' travel behaviour. To achieve the objectives of this study, more behaviourally realistic models that account for the interaction among the different choice decisions faced by each individual, are required. However, according to the literature, traditional travel demand frameworks are not appropriate for modelling such complex behaviour, especially in the context of cross-regional travel because of the intermodal nature of various travel options available to them.

The framework consists of three main components, a departure time choice model, an enhanced mode choice model, and an agent-based multimodal trip assignment model. However, this study focuses on developing the enhanced mode choice model component while providing the necessary interactions with the pre-developed components of the framework. The framework takes individuals' work and home locations, their socioeconomic and demographic characteristics, and the network attributes and transport system constraints as inputs. The output of the framework provides a detailed disaggregate probabilistic chain of departure time, mode, and route choices.

The aim of developing such a framework is to fill some of the aforementioned gaps in the area of travel demand modelling in general and cross-regional commuters' mode choice modelling in particular. In terms of mode choice, a typical cross-regional trip may involve three main decisions: main mode, access mode, and access station location choices. Therefore, a joint tri-variate decision structure for the enhanced mode choice component is presented. This methodology embodies a more desirable and appealing approach to regional planning than previously developed models. The framework adopts a series of disaggregate models including policy-sensitive joint RP/SP discrete choice models that are developed using detailed disaggregate level-of-service attributes.

The framework is developed over two phases. In phase I, individuals' choice of access station location is analyzed, and outputs of this analysis are adopted to develop a multimodal planner tool to be utilized within an RP/SP data collection tool. This tool collects the required information to develop the enhanced mode choice model in phase II. This framework can be utilized to develop policy analysis tools that are capable of providing decision-makers with appropriate measures for delivering planning solutions in accordance with polices under investigation. Policies affecting pricing schemes (gas prices, transit fares, and parking costs), introducing new travel modes or technologies, and improving level-of-service attributes (travel times, information provision, and accessibility) can be evaluated using this framework.

CHAPTER 4

4 MODELLING ACCESS STATION CHOICE FOR CROSS-REGIONAL COMMUTER TRIPS

4.1 Chapter Overview

In this chapter, the access station choice of park-and-ride cross-regional commuters in the GTHA is examined. Three multinomial logit models are developed: a regional transit park-and-ride access station choice model, a local transit park-and-ride access station choice, and an access station location choice model for commuters with possible access to both regional and local transit park-and-ride stations. Section 4.2 provides a brief introduction of the motivation and purpose of this analysis. The following section, Section 4.3, provides details of the study area, the data description, and the conducted spatial analysis. In Section 4.4 and Section 4.5, the econometric modelling framework and empirical models are presented. Section 4.6 provides a discussion on the results of the developed models. Finally, a summary is provided in Section 4.7.

4.2 Introduction

The use of private cars for transit access (i.e. park-and-ride) is the most flexible form of transit modal integration; it provides access to transit for users who may not otherwise consider transit as a travel mode alternative. Park-and-ride (P&R) facilities have become an important component of urban transportation systems because they facilitate intermodal integration (Bos et al., 2005; Chakour and Eluru, 2014; Forsey et al., 2013; Holguín-Veras et al., 2012). Typically, park-and-ride facilities are located at major transit hubs that are characterized by wide spacing among facilities, making station access choice more complex. As such, access station choice is a key component of defining the choice of park-and-ride as a travel mode. More than 30% of the cross-regional commuters in the Greater Toronto Hamilton Area (GTHA) who access the rail transit by car choose a station other than the one closest to their home location (DMG, 2008). Therefore, further investigation is required to examine cross-regional park-and-ride commuters' behaviour in terms of their access station choice.

As explained in Chapter 3, the framework for modelling cross-regional trips involves a three-level decision structure in which individuals choose their main mode of travel, access mode, and access station location (for transit modes with automobile access). Therefore, as a first step toward developing this framework, a set of access station location models for park-and-ride/kiss-and-ride cross-regional commuters are developed and presented herein. The developed models are adopted in the data collection tool used to gather the required data for developing the enhanced mode choice model, as shown in Chapter 5.

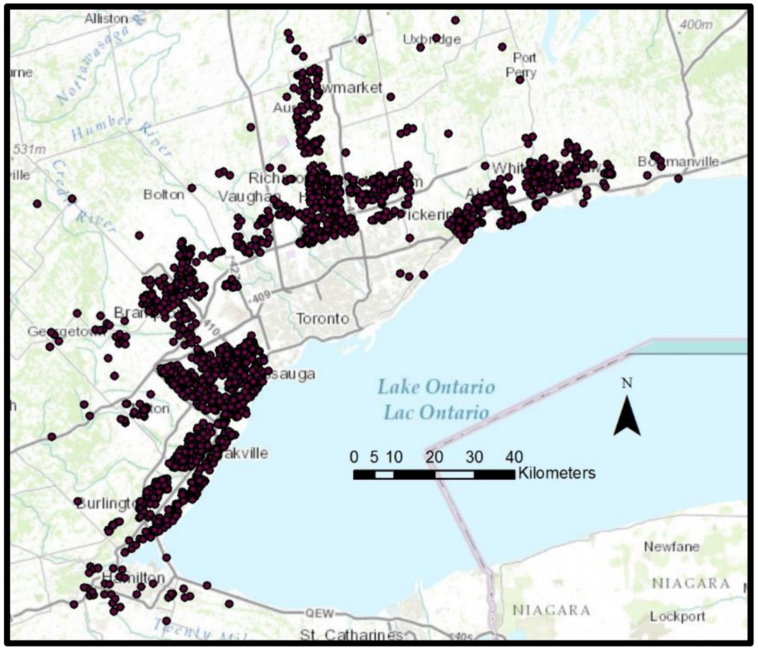
4.3 Study Area, Data Description, and Spatial Analysis

Situated directly to the northwest of Lake Ontario in the province of Ontario, the Greater Toronto Hamilton Area (GTHA) forms Canada's largest urban region. The GTHA consists of the City of Toronto, City of Hamilton, and four regional municipalities. In the GTHA, park-and-ride users choose between GO Transit park-and-ride stations, which serve trips across the region, and/or Toronto Transit Commission (TTC) Subway park-and-ride stations, which cater to trips toward the City of Toronto. The trip, demographic, and socioeconomic characteristics were extracted from the 2006 Transportation Tomorrow Survey (TTS) dataset. The TTS is a trip-based household survey conducted every five years in the GTHA among 5% of its population. Cross-regional commuter trips (i.e. trips that cross boundaries of municipal or regional jurisdictions that have different transit operators) represent around 35% of total commuting trips, which has grown by 7% since 1996 (DMG, 1998, 2008). The dataset provides detailed disaggregate individual trip records with the exact locations of households, employment, and park-and-ride access stations.

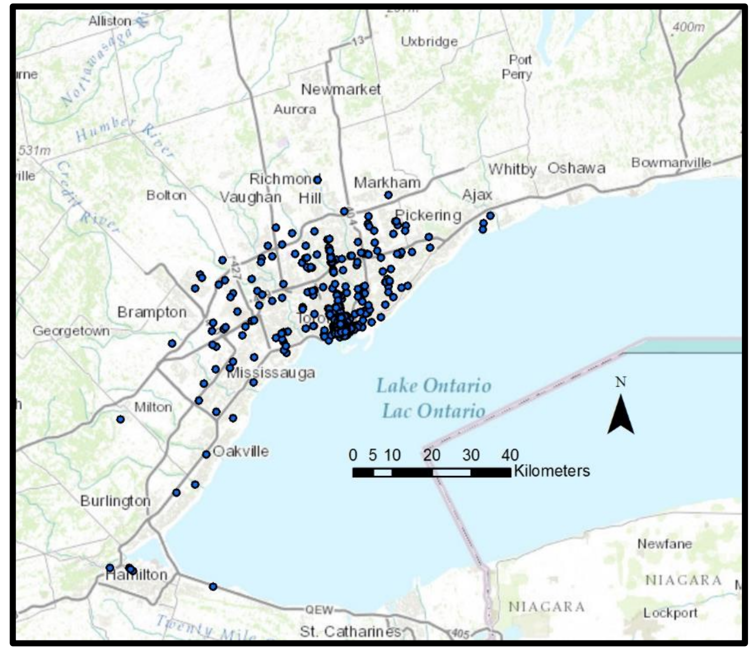
Data for cross-regional commuting trips in the morning peak period (6:00–9:00 a.m.) on regular weekdays were extracted from the TTS dataset. To study individuals' access station choice, trip records with transit as the chosen main mode and automobile as the access mode were used in this analysis. The total number of complete trip/passenger records used in this analysis is 1,884.

Figure 4-1 (a) shows the distribution of household locations (with at least one cross-regional park-and-ride commuting trip) across the GTHA, and Figure 4-1(b) shows the distribution of their corresponding employment locations. Data on park-and-ride station locations, parking lot capacities, parking costs, surrounding land use, and station amenities were obtained from the Toronto Transit Commission (TTC) and Metrolinx for the years 2010 and 2011, respectively. Since 2006 TTS data were used to develop access station location choice models, parking costs at park-and-ride stations were adjusted to match 2006 levels to account for the inflation rate. Other station amenities and packing capacities were assumed to have remained unchanged between 2006 and 2010/2011. Figure 4-2 presents park-and-ride locations for GO Train and TTC Subway across the region. Figure 4-2 (a) shows parking lot capacities and parking costs, while Figure 4-2 (b) shows the morning peak parking demand for cross-regional trips (the number of trip records in the dataset without applying expansion factors) at each location. It should be noted that all GO Train park-and-ride facilities provide free parking to their customers on a first-come-first-serve basis with options to purchase annual/monthly reserved parking spots. In contrast, park-and-ride facilities operated by TTC charge their users a daily parking fare ranging from \$3 to \$7. Figure 4-3 shows park-and-ride stations' catchment areas for cross-regional commuting trips. This defines an approximate area that covers households' access station choices based on the observed trip records. As shown in Figure 4-3, a substantial overlap between stations' catchment areas exists, which indicates that individuals whose home locations are situated in approximately the same area make different access station choices.

Since a significant percentage of park-and-ride/kiss-and-ride cross-regional commuters do not choose the closest access station to their home locations, a detailed investigation of factors that may influence their choices is conducted. Intuitively, the distance from the household location to the park-and-ride station location is an important factor. However, other spatial factors, such as the relative station direction between a straight line from the home location to the regular work location and a straight line from the home location to the park-and-ride station location may explain individuals' access station location if the station closest to their home locations was not chosen. As shown in Figure 4-4, the relative station direction (alpha) indicates whether the station in consideration is on the way to the workplace (S1) or not (S2). Stations with a higher alpha value (i.e. out of the way to the workplace) are expected to be less attractive to travellers. In addition, the dataset includes other variables that are expected to be useful in explaining individuals' choice of access stations, such as the station parking capacity, cost, amenities, etc.

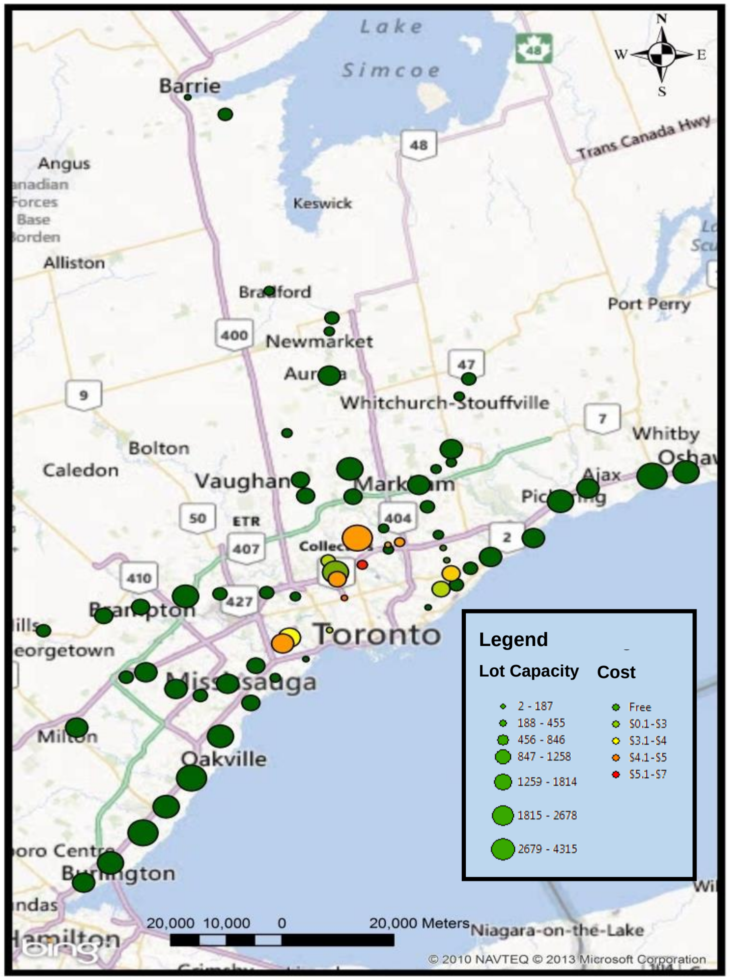


(a) Distribution of household locations

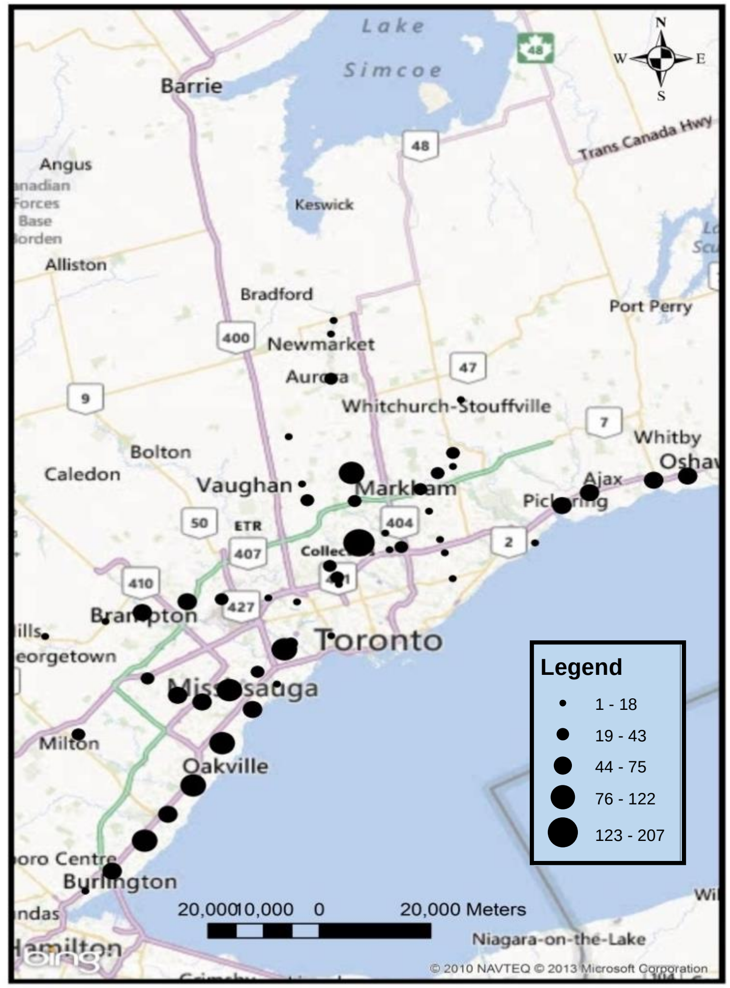


(b) Distribution of employment locations

Figure 4-1 Distribution of Cross-Regional Park-and-Ride Users' Home and Work Locations



(a) Park-and-ride lot capacities (size) and parking cost (color)



(b) Parking demand (number of trips) at park-and-ride locations

Figure 4-2 Park-and-Ride Parking Lot Capacities, Parking Cost, and Parking Demand

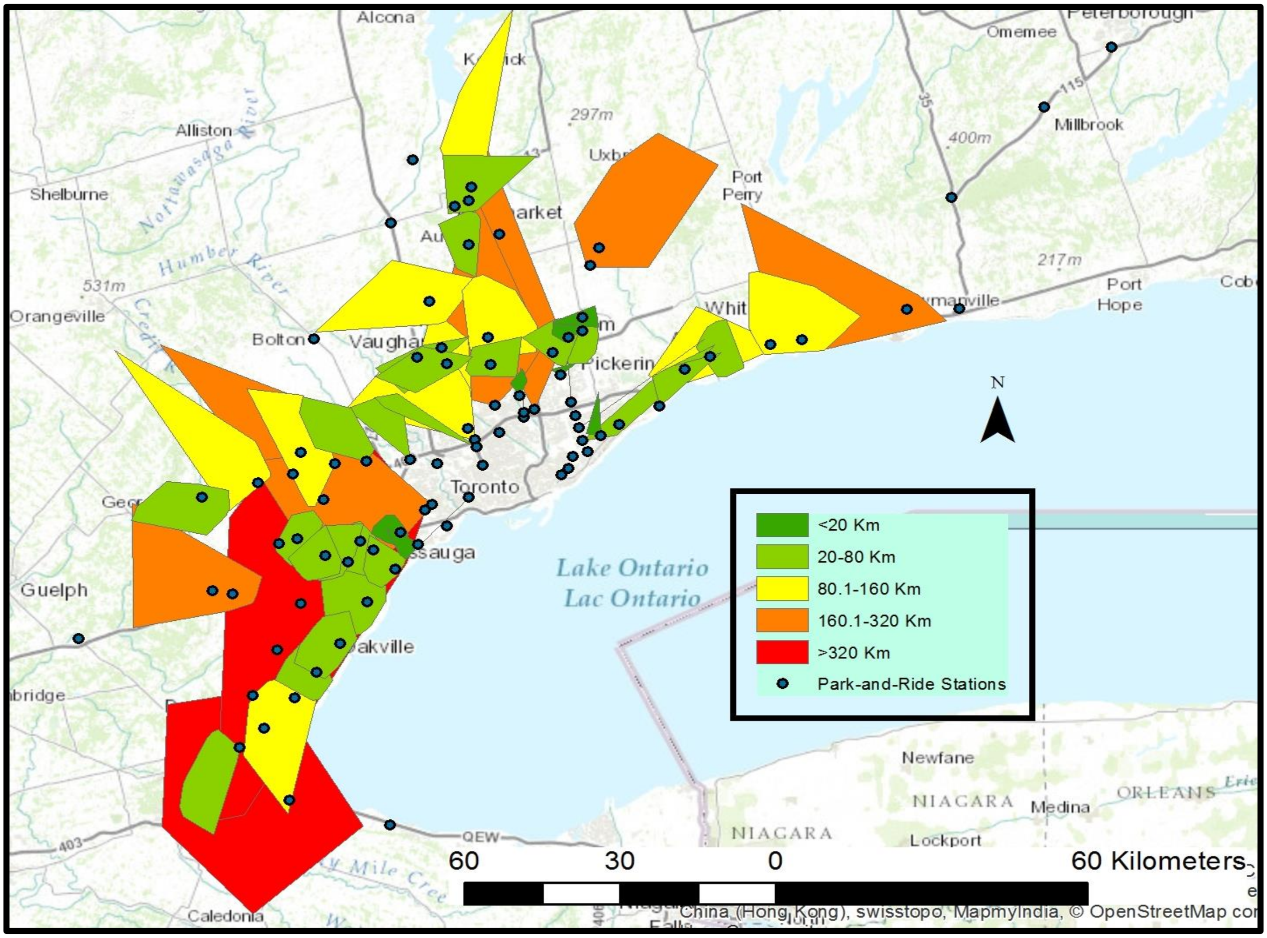


Figure 4-3 Park-and-Ride Catchment Areas

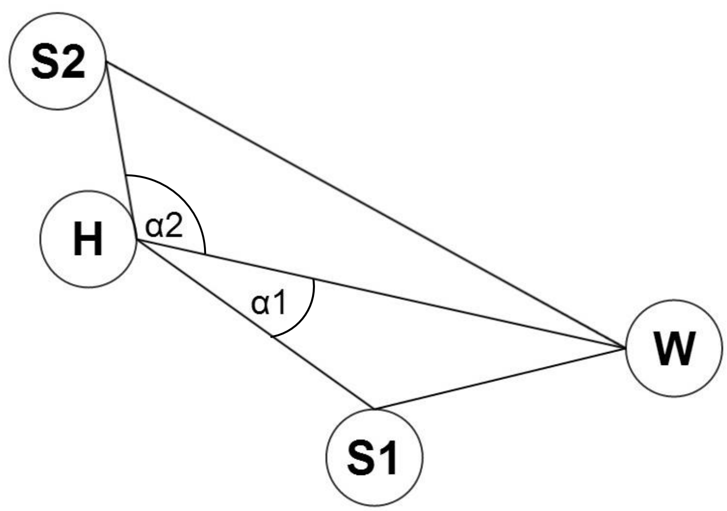


Figure 4-4 Relative Station Direction (Alpha αi) of Park-and-Ride Stations (Si) from Home Location (H) to Work Location (W)

4.4 Econometric Model

Three datasets were prepared for modelling and empirical investigations. The first dataset included all park-and-ride stations (including GO Train and TTC Subway stations) and their corresponding trip records. The second dataset included regional transit (GO Train) trip records, and the third dataset included local transit (TTC Subway) trip records. Household locations were used to define the choice set for each individual. The 53 park-and-ride GO Train stations are widely dispersed across the region; however, the 15 park-and-ride TTC Subway stations are concentrated in the City of Toronto, mainly near the terminal stations of the subway lines. Therefore, it was assumed that the five closest stations defined the access station choice set for GO Train users. In contrast, the three closest stations defined TTC Subway users' access station choice set. Respectively, 98% and 80% of cross-regional park-and-ride GO Train and TTC Subway users' observed access station choices fell within the pre-defined choice sets.

Individuals are considered to gain a certain level of utility by choosing one station from their pre-defined choice set. The utility function U for each station is composed of systematic V and random ε components. The systematic component explains the deterministic utility of choosing the corresponding alternative station as a linear-in-parameter function of the observed variables X and their corresponding coefficients β. The random utility component explains the unobserved random variations in choice. It is assumed that travellers are rational in selecting an access station among a set of feasible alternatives and to choose the alternative with the highest utility value. According to the random utility maximization (RUM) theory,

$$U_s = V_s + \varepsilon_s = (\beta x)_s + \varepsilon_s, \qquad [4-1]$$

where the subscript s indicates one of the stations in the choice set.

Model variables may affect the utility function at different levels. For instance, previous studies have shown that the access distance (i.e. from the trip origin to the station) is an important variable in access station choice models (Debrezion et al., 2007, 2009; Vijayakumar et al., 2011). In addition, station features such as the direction of the station relative to the home and work locations, parking lot capacity, parking cost, and other station amenities that may potentially affect individuals' station choice are introduced. Previous studies have often included service frequency as a variable in the utility functions. However, other studies showed that the effect of service frequency is relatively small compared to the effect of the distance between the home and station locations (Debrezion et al., 2007). In the GTHA, subway services run at high frequency levels (around two-minute headways) during the morning peak period, resulting in similar service frequencies among the stations under study. For regional transit services, we assume that individuals check the publicly available train schedules (which run at higher frequencies in the morning peak period compared to the rest of the day) and plan their trips accordingly based on the desired time to reach their destinations. Therefore, service frequency was not included as a variable in the station choice utility function.

To derive the probability function for access station choice, a distributional function of the random error component is assumed. We assume that the error term follows the independent and identical distribution (IID) of a Type I Extreme Value distribution. This assumption results in a multinomial logit (MNL) model (Ben-Akiva and Lerman, 1985) of the following form:

$$\Pr(s) = \frac{\exp(V_s)}{\sum_{s'=1}^{S} \exp(V_{s'})}, \qquad [4-2]$$

where the subscript *s* indicates the chosen station and *S* indicates the maximum number of stations under consideration within the individual choice set. For a sample of *Z* individuals, assuming that each individual's choice decision is independent, the joint likelihood function can be expressed as follows:

$$L(\beta) = \prod_{z}^{Z} (\Pr(s)_z{}^{y_{zs}}), \qquad [4-3]$$

where β is vector of variable coefficients that maximizes the likelihood function $L(\beta)$, and $y_{zs}=1$ if person *z* chooses station *s* from their choice set *S* and zero otherwise.

4.5 Empirical Models

Park-and-ride users are divided into three groups: individuals for whom only TTC Subway park-and-ride stations were within reasonable reach, individuals whom only GO Train stations were within reach, and individuals whom both TTC Subway and GO Train stations were within reach. Three multinomial logit models were estimated accordingly to better capture choice behavior of each of the three population segments. Table 4-1 presents definitions of the variables used in the models' development. As explained in Chapter 3, the local and regional transit park-and-ride access station location choice models are utilized in phase II of the framework development.

Table 4-1 Definitions of Variables

Variable Name	Description
Access Distance	Airline access distance in meters from the household location to the park-and-ride station location
Alpha (α)	Relative station direction (travel angle) in degrees between a straight line from the home location to the regular work location and a straight line from the home location to the park-and-ride station location, as shown in Figure 4-4
Lot Capacity	Park-and-ride lot capacity
Parking Cost	Parking cost at the morning peak period in Canadian dollars (CAD) at the park-and-ride location
Refresh Kiosk	=1 if station has a refreshment kiosk; =0 otherwise
Washrooms	=1 if station has a washroom facility; =0 otherwise
Reserved Parking	=1 if station has a reserved parking option; =0 otherwise
Regional Transit Station	=1 if station is a regional transit station; =0 otherwise
Connect to Local Transit	=1 if a GO Train station connects to local transit services; =0 otherwise
Connect to Regional Transit	=1 if a TTC Subway station connects to a regional service; =0 otherwise
TTC Pass	=1 if the individual possess a TTC Subway (Metro) pass; =0 otherwise
GO Pass	=1 if the individual possess a GO Transit pass; =0 otherwise

Various model structures and specifications were tested, and the final model specifications are reported in Tables 4-2, 4-3, and 4-4. Mixed logit models were examined for the regional commuter rail and local subway park-and-ride access station choice, with the assumption of access distance and relative station direction as random parameters (drawn from a lognormal distribution). However, the standard deviation for the random parameters was statistically insignificant. Therefore, a multinomial logit model structure was used. The empirical models were estimated using the MAXLIK component of the mlogit package in the R statistical software for maximum likelihood estimation (Croissant, 2012; Train and Croissant, 2012).

4.6 Results, Discussion, and Validation

Parameter estimation results for the GO Train and TTC Subway park-and-ride access station choice MNL model are reported in Table 4-2. Six parameters were estimated using a subset sample of 229 park-and-ride users who have at least one GO Train and one TTC Subway park-and-ride station within their choice set, which is composed of the five closest park-and-ride stations. All of the reported parameters in the final model specification are statistically significant, with t-statistics greater than 1.96 at the 95% confidence interval. The reported rho-squared value (Train, 2009), a measure of goodness of fit, is 0.55. This is a relatively high value considering the small sample size used in this analysis. In addition, the log likelihood ratio test shows a test statistics value of 352, which indicates that the reported model fit the data significantly better than the null model.

In general, all variable coefficients are estimated with the expected signs. Two variables are used to define stations' proximity to home locations considering the usual workplace: access distance and relative station direction. An increase in access distance and relative travel angle has a negative impact on access station choice. Since a significant percentage of park-and-ride locations in the GTHA run at capacity during morning peak periods, stations with larger parking capacities are more likely to be chosen because it increases individuals' chances of finding a parking spot. If a subway and a regional transit station happen to be in the same choice set of a cross-regional traveller, the regional transit station is less likely to be chosen. In the GTHA, a typical park-and-ride cross-regional trip destined for the City of Toronto could involve a drive to a nearby GO Train station to take the GO Train to Union Station and then to transfer to the TTC Subway to arrive at the final destination. Therefore, if a subway station is feasibly accessible from individuals' home locations, it will be more attractive than a similarly situated regional transit station. This will save the traveller an unnecessary mode transfer as well as time and costs. One of the other factors that may affect individuals' access station choice is whether or not the traveller possesses a transit pass. Two types of transit passes were considered: the first was the TTC Subway (metro) monthly pass, which entitled the holder unlimited rides on the TTC Subway. The second was the GO Transit pass (replaced in 2008 with the PRESTO smart card), which provides discounted pre-paid tickets for the GO Train. To account for transit pass type, a dummy variable was introduced into the model in association with access distance. The results showed that TTC Subway pass holders were more likely to drive distances to TTC Subway park-and-ride stations that were longer than those driven by GO Transit pass holders to GO Train park-and-ride stations.

Table 4-2 Regional and Local Transit Park-and-ride Access Station Choice MNL Model

Estimation Results		
Log-Likelihood (Full Model)	-141.76	
Log-Likelihood (Null Model)	-317.53	
Rho-squared Value	0.55	
Number of Observations	229	
Systematic Utility Function:	**Parameter**	**t-Statistics**
Access Distance	-0.00079	-6.51*
Alpha (α)	-0.01160	-3.21*
Lot Capacity	0.000536	4.28*
Regional	-2.35356	-4.41*
DIST * TTC Pass	0.000599	3.99*
DIST * GO Pass	-0.00115	-5.44*

Table 4-3 Regional Transit (GO Train) Park-and-ride Access Station Choice MNL Model

Estimation Results		
Log-Likelihood (Full Model)	-1,160	
Log-Likelihood (Null Model)	-1,541.4	
Rho-squared Value	0.25	
Number of Observations	1,655	
Systematic Utility Function:	**Parameter**	**t-Statistics**
Access Distance	-0.00072	-31.05*
Alpha (α)	-0.01252	-15.81*
Lot Capacity	0.000477	8.57*
Reserved Parking	0.327840	3.52*
Refresh Kiosk & Washrooms	0.401080	3.35*
Connect to Local Transit	1.717000	3.73*

Table 4-4 Local Transit (TTC Subway) Park-and-ride Access Station Choice MNL Model

Estimation Results		
Log-Likelihood (Full Model)	-224.01	
Log-Likelihood (Null Model)	-336.36	
Rho-squared Value	0.33	
Number of Observations	413	
Systematic Utility Function:	**Parameter**	**t-Statistics**
Access Distance	-0.00037	-4.06*
Alpha (α)	-0.03126	-3.40*
Parking Cost	-0.22126	-1.85▪
Refresh Kiosk & Washrooms	0.000295	1.89▪
Connect to Regional Transit	1.073700	5.35*

* Significant at the 95% confidence level; ▪ Significant at the 90% confidence level

Similarly, Table 4-3 and Table 4-4 show parameter estimation results for GO Train and TTC Subway park-and-ride access station choice MNL models, respectively. The reported parameters are estimated with the expected signs and found to be statistically significant (with t-statistics higher than 1.96) at the 95% confidence interval. The estimated parameter of parking cost at local transit park-and-ride stations was found to be statistically significant at the 90% confidence interval. As expected, park-and-ride locations with higher parking costs are less likely to be chosen over stations that provide parking at a lower cost or free. Similarly, stations that provide better local/regional service integration and better station amenities/facilities are preferred. Regional commuter rail and local subway park-and-ride access station choice models showed rho-squared values of 0.25 and 0.33, respectively. In addition, the log likelihood ratio tests show test statistics values of 763 and 225, respectively, which indicates that the reported models fit the data significantly better than the null models. A forecast of park-and-ride cross-regional travellers' access station choice was conducted and compared to actual observations. The percentage correctly predicted analysis, another goodness-of-fit measure (Train, 2009), showed that the combined model performs at a prediction accuracy of 76%, while the regional commuter rail and local subway park-and-ride access station choice models attain 75 % and 79%, respectively.

To investigate the sensitivity of park-and-ride users' to station access distance and relative station direction, the point elasticity of access distance and relative station direction were estimated for each individual in the three datasets. Average elasticity was calculated for the number of stations in each individual's choice set, and kernel densities were plotted as shown in Figure 4-5. Equation $4-4$ shows calculations for direct elasticity (Train, 2009):

$$E_{i_{Xni}} = \beta_X . X_{ni} . (1 - P_{ni}), \qquad [4-4]$$

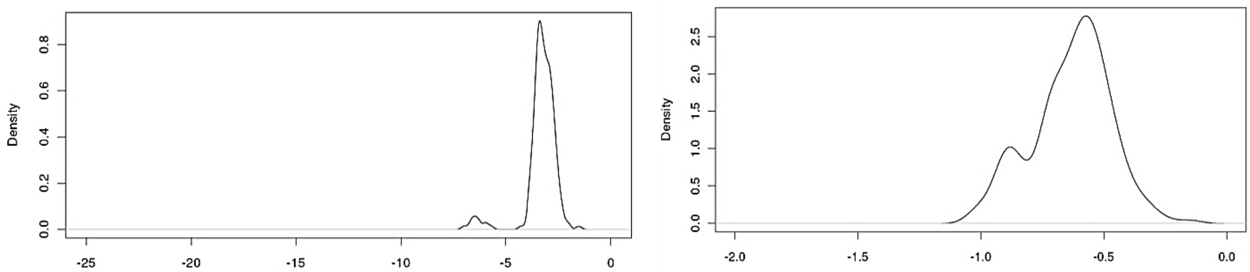
where E_{iXni} is the direct elasticity of a unit change in the observed factor X_{ni} with a parameter estimate β_X on the probability that individual n chooses alternative i, P_{ni}.

Figure 4-5 shows that the access distance has a bi-modal distribution; however, the relative station direction has a unimodal distribution. The bimodal distribution suggests two groups of users with different mean access distances. Such a distribution may result from different land use patterns near each station. Further investigation showed that different planning districts have different access distance distributions, which can be explained by the variation of station spacing in each region. Users from planning districts with higher station densities have more flexibility in driving to a more distant station within their acceptable access distance (with lower a marginal difference). However, as the station density decreases, users become more sensitive to drive longer distances to access a nearby station (e.g. City of Hamilton). This finding provides an explanation for the existence of two groups of users: those who are more flexible with respect to access distance and those who are more sensitive.

The sample average elasticity of the station access distance for the combined model is -3.3174 with a standard deviation of 0.7953, and the sample average elasticity of the relative station direction is -0.6325 with a standard deviation of 0.1605. This statistical result suggests that individuals are more sensitive to changes in station access distance than to changes in the station relative direction. Similarly, the regional and local transit park-and-ride access station choice models reported a sample average elasticity of station access distance of -5.1009 and -3.1116 with a standard deviation of 2.5183 and 1.8117, respectively. In contrast, the models reported a sample average elasticity of the relative station direction of -0.7191 and -0.3791 with a standard deviation of 0.2745 and 0.2547, respectively. These findings indicate that TTC Subway park-and-ride users are less sensitive to access distance than regional transit users. Hence, GO Train users' choices are more sensitive with respect to the station location than TTC Subway users. This can be explained by the wide spatial dispersion of park-and-ride GO Train stations across the GTHA compared to the fewer park-and-ride TTC Subway station options within the City of Toronto.

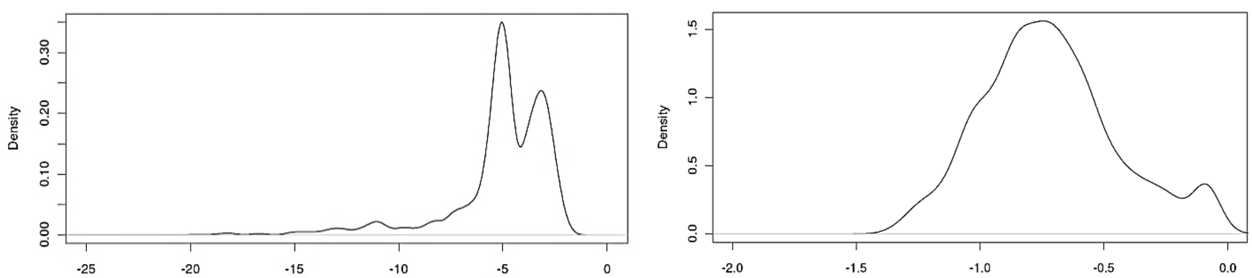
4.7 Chapter Summary

This chapter presents an access station choice model for park-and-ride cross-regional commuter trips in the GTHA. Two sources of data were used: data on cross-regional commuting trips during the morning peak period on regular weekdays were extracted from the 2006 TTS household travel survey, and data on park-and-ride station locations, parking lot capacities, parking costs, surrounding land use, and station amenities were obtained from transit service operators. The five and three closest stations defined the access station choice sets for GO Train and TTC Subway users, respectively. Three datasets were prepared for modelling and empirical investigation. Park-and-ride users were divided into three groups: individuals who have only TTC Subway stations within reasonable reach, individuals with only GO Train stations within reach, and individuals who have both TTC Subway and GO Train stations within reach. Various model structures and specifications were tested, and three multinomial logit models were estimated accordingly to capture the choice behaviour of each market segment.



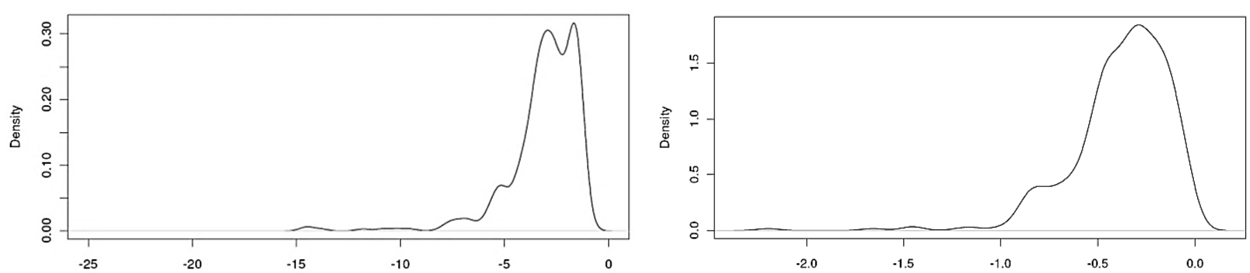
(a) Access distance average elasticity (b) Relative station direction (alpha) average elasticity

(i) Regional (GO Train) and local (TTC Subway) transit park-and-ride access station choice model



(c) Access distance average elasticity (d) Relative station direction (alpha) average elasticity

(ii) Regional transit (GO Train) park-and-ride access station choice model



(e) Access distance average elasticity (f) Relative station direction (alpha) average elasticity

(iii) Local transit (TTC Subway) park-and-ride access station choice model

Figure 4-5 Marginal Effects of Access Distance and Relative Station Direction

In general, the estimated parameters are statistically significant with the expected signs. Access distance and the relative station direction are the primary factors that affect individuals' choices. An increase in both access distance and relative travel angle has a negative impact on access station choice. In addition, station-level variables are included in the models. Stations with higher parking capacities are more likely to be chosen. Similarly, park-and-ride locations with higher parking costs are less likely to be chosen over stations that provide parking at a lower cost or free. In addition, stations that provide better local/regional service integration and better station amenities/facilities are preferred.

The developed regional commuter rail and local subway park-and-ride access station choice models are adopted in the data collection tool used to gather the required data for developing joint main mode and access mode choice models, as described in Chapter 5. That is, utilizing the developed models allows for the provision of realistic and customized travel alternatives for each individual, which in turn improves the quality of the collected data.

CHAPTER 5

5 SURVEY OF CROSS-REGIONAL INTERMODAL PASSENGER TRAVEL (*SCRIPT*)

5.1 Chapter Overview

This chapter presents the Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*), an on-line data collection tool. *SCRIPT* collects data on respondents' current commuting trips as well as their stated mode choice in response to hypothetical changes in the current mode attributes. An innovative multimodal trip planner tool is developed to generate feasible travel options for each choice experiment using information on households' auto ownership level, proximity to transit, work start time, and total travel time from home to work. The survey platform adopts the pre-developed access (see Chapter 4) location choice models to identify potential access stations for intermodal travel modes.

In Section 5.2, the need to develop a joint RP/SP interactive and customized survey is highlighted. Section 5.3 and Section 5.4 provide information on the survey study area and the travel mode options considered in this study. In light of the literature reviewed in Chapter 2, Section 5.5 explains the SP experimental design, adopting the D-efficient design technique. The following section, Section 5.6, provides details on the survey sampling design. In Section 5.7, details of the survey instrument design are presented, including the development of an innovative multimodal trip planner tool. Finally, Section 5.8 provides a summary of the chapter.

5.2 Introduction

To investigate individuals' mode choice preferences, transportation planners aim at developing probabilistic discrete mode choice models. Such models require detailed information about the travellers' detailed trip patterns and socio-demographic data. As mentioned in Chapter 2, there are two ways to quantify individuals' preferences: revealed preference (RP) and stated preference (SP) methods. Typically, travel surveys that rely on RP methods collect information about individuals' actual (observed) choices (e.g. details of their commuting trip on a typical working day). In contrast, SP surveys collect information on individuals' preferences toward a set of hypothetical scenarios (e.g. introducing a new travel mode or technology). That is, SP surveys provide more flexibility to examine new alternatives that have not been experienced before.

In particular, studying cross-regional travellers' behaviour requires extensive data on their trip patterns, including detailed information on each trip segment, such as access, transfer, and egress times. Typical commuting travel surveys do not provide sufficient data to conduct this type of analysis for several reasons: cross-regional trips are often underrepresented in survey samples, the collected data do not provide the necessary level of detail on inter- and intra-modal trips, and a majority of typical travel surveys rely predominantly on RP trip data. Previous research efforts have shown that RP data do not adequately capture the behavioural trade-offs involved in the travellers' decision-making process in response to changes in their travel modes' level-of-service attributes. Demand models developed based on only RP data are incapable of accurately forecasting individual choices in response to new transportation policies or the introduction of new modes that have never been used before (Idris et al., 2012; Louviere et al., 2000). Therefore, SP data are required to estimate the corresponding change in respondents' travel behaviour patterns in accordance with changes in the current level-of-service attributes or the introduction of new modes.

5.3 Survey Study Area

The survey is designed to be implemented in the Greater Toronto Hamilton Area (GTHA), Canada's largest urban region, although the framework is transferable to similar metropolitan areas. The GTHA's current population is over six million, with a projected growth to approximately 8.6 million by 2031 (Metrolinx, 2008). The GTHA, shown in Figure 5-1, consists of the City of Toronto, the City of Hamilton and four regional municipalities: York Region, Peel Region, Durham Region, and Halton Region. The GTHA has nine local transit and one regional transit (i.e. GO Transit) services operating under the administration of Metrolinx, "*an agency of the Government of Ontario, which was created to improve the coordination and integration of all modes of transportation*." As such, the GTHA provides a generic case study of a multimodal integrated transportation network.

Cross-regional trips are defined as trips that cross boundaries of municipal or regional jurisdictions that have different transit operators. A cross-regional trip may involve the use of multiple transit services or the interaction between two or more travel modes. For instance, a transit trip from the City of Mississauga to the City of Brampton involves the use of two transit services operated by two agencies (i.e. MiWay Transit and Brampton Transit). This trip is considered a cross-regional trip despite being conducted within the same region (i.e. Peel Region). In contrast, a trip within different local municipalities of the York Region is not considered a cross-regional trip in this study since it can be conducted using one transit system (i.e. York Region Transit).

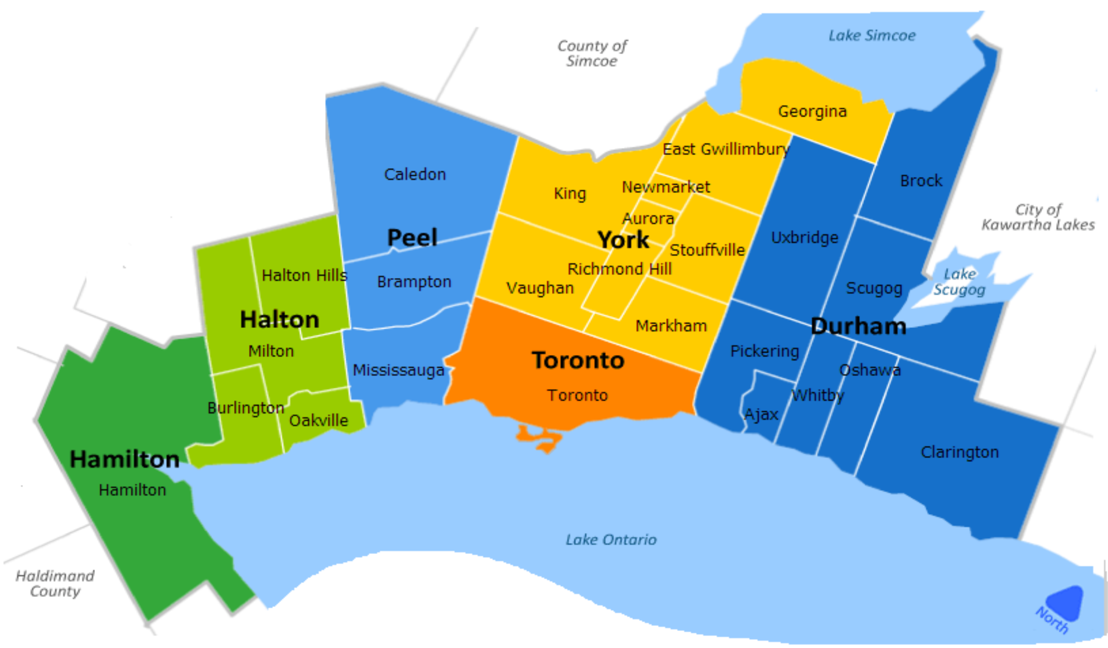


Figure 5-1 The Greater Toronto Hamilton Area[2]

Another factor that contributes to the complexity of studying cross-regional transit trips stems from the different transit fare systems across the region, such as flat fares, zone-based fares, and routes with special extra fares. Individuals may pay transit fares using cash, tokens, pre-purchased tickets, or smart cards/transit passes. In return, travellers receive a proof of payment that serves as a "transfer" that can be used to connect between multiple transit vehicles within or across transit agencies. Some agencies treat these transfers as "trip-based" proof of payments, and as such, they are only valid for an origin–destination trip, allowing for unlimited within-agency transfers toward

[2] http://findtheway.ca/en/

the trip destination without stopping over or doubling back. Other agencies accept transfers as "time-based" proof of payments, and as such, they are valid for a specific time (typically 90-120 minutes), allowing for unlimited within-agency transfers in any direction as long as the transfer is valid. Moreover, some agencies follow specific fare integration schemes that provide users with reduced transit fares (co-fares) for transferring between different agencies.

5.4 Travel Mode Options and Description

In this study, a universal choice set of nine travel modes is defined with a clear distinction between regional and local transit services as well as between the different access modes. The considered mode alternatives are auto driver, auto passenger/carpool, local transit with walk access, local transit with auto driver access (park-and-ride), local transit with auto passenger access (kiss-and-ride), regional transit with walk access, regional transit with auto driver access (park-and-ride), regional transit with auto passenger access (kiss-and-ride), and regional transit with local transit access. The terms "regional transit" and "GO Transit" are used interchangeably, and so are "local transit park-and-ride/kiss-and-ride" and "TTC Subway park-and-ride/kiss-and-ride." Table 5-1 shows a description of the nine travel options considered in this study.

5.5 Experimental Design

The purpose of performing an SP experiment is to quantify the independent effects of the design attributes (e.g. level-of-service attributes) on respondents' choices (Carson et al., 1994; Louviere and Hensher, 1983; Louviere and Woodworth, 1983). Typically, a number of choice tasks are presented to each respondent to select one alternative from a list of presented alternatives. The number of these alternatives can vary or be fixed during the experiment. Each alternative is defined by a set of attributes and attribute levels. Across the SP experiment, the alternatives' attribute levels are altered (according to the pre-designed SP experiment), so corresponding changes in individuals' responses can be quantified. As discussed in Chapter 2, the allocation of alternatives' attribute levels and their order of change across the different choice tasks play a major role in the quality of the experiment design. Therefore, the D-efficient design is adopted to develop the SP experiments for this study.

To design the SP experiment using the D-efficient design technique, prior parameter estimates of the design attributes as well as the model structure are required. Since such information is not typically available, a pilot survey was developed based on an orthogonal design and conducted among a random sample of the same target population sampling frame used later for the final survey. The number of choice scenarios generated by the orthogonal design to ensure attribute level balance is 108, which were divided into 18 blocks of six scenarios each. Data from the pilot survey were used to obtain the prior parameter estimates that were required to develop the D-efficient SP experimental design. After cleaning the dataset of incomplete or invalid records, 45 complete responses were used to estimate a pilot model. The model showed correctly signed coefficients for all variables. However, some parameters were statistically insignificant because of the small size of the dataset.

Table 5-1 Travel Modes and Description

Mode	Definition
Auto Driver	One passenger driving alone all the way
Auto Passenger/Carpool	Two or more passengers sharing the same car
Local Transit Walk Access	The use of one or more local transit service(s) with walk access.
Local Transit Auto Driver Access (TTC Park-and-ride)	The use of one or more local transit service(s) with auto driver access (park-and-ride); a person drives alone to a park-and-ride station, parks the car at the station, and takes public transit
Local Transit Auto Passenger Access (TTC Kiss-and-ride)	The use of one or more local transit service(s) with auto passenger access (kiss-and-ride); a person is dropped off at a park-and-ride station and takes transit; the other person leaves right after
Regional Transit Walk Access	The use of one or more regional transit (GO Transit) service(s) with walk access
Regional Transit Auto Driver Access (GO Transit Park-and-ride)	The use of one or more regional transit (GO Transit) service(s) with auto driver access (park-and-ride); a person drives alone to a park-and-ride station, parks the car at the station, and takes public transit
Regional Transit Auto Passenger Access (GO Transit Kiss-and-ride)	The use of one or more regional transit (GO Transit) service(s) with auto passenger access (kiss-and-ride); a person drives to a park-and-ride station, parks the car at the station, and takes transit
Regional Transit with Local Transit Access	The use of one or more regional transit (GO Transit) service(s) with local transit access.

5.6 Survey Sample Design

Data from the 2011/2012 Transportation Tomorrow Survey (TTS), a trip-based household survey conducted every five years in the GTHA among 5% of its population, were used to identify sampling probabilities based on spatial location, mode split, and gender (DMG, 2013). Other attributes, such as age, vehicle ownership level, and occupation type, were considered as well.

5.6.1 Target Population

The target population of this survey is identified as cross-regional commuters (i.e. travellers with home and work trip ends in different municipalities/regions that have different transit operators) 18 years and older within the study area. According to the 2011/2012 TTS data, cross-regional commuter trips represent around 35% of the total daily commuting trips in the GTHA. That is, the survey population is estimated to be 2,631,668 commuters.

5.6.2 Sampling Method, Sample Size, and Sample Allocation

Sampling methods can be classified into two main groups: probability and non-probability sampling methods. In general, probability sampling methods reduce selection bias (compared to non-probability sampling methods) by randomly selecting individuals based on a certain inclusion probability (Idris, 2013). That is, the survey population is divided into homogeneous strata (based on individuals' trip origin location, mode of travel, and gender), and the sampling probability of each stratum is defined according to the 2011/2012 TTS data. These probabilities are used as guidelines to ensure the collection of a representative sample of the target population. The sample selection was done based on simple random sampling.

The required sample size is defined based on the estimated survey population, an allowable margin of error (of the true value of the characteristics of the target population), and a design effect (DEFF) to account for the sampling method in use. The estimated survey population of 2,631,668, a marginal error equal to or less than 0.05, and a DEFF of 2.5 are used to determine the required sample size at 960 responses. Details on the sample size determination process can be found in Idris (2013). The estimated sample size is consistent with the required number of responses based on the developed D-efficient experiment design, which is estimated at between 800 and 1,000 records to estimate statistically significant variables at the 95% confidence interval. The N-proportional allocation method is used to determine the required sample size from each stratum.

5.7 Survey Instrument Design

The Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*) consists of three sections. Section A gathers RP information on the respondents' daily commuting trips and current travel attributes. In section B, individuals are asked to respond to hypothetical scenarios where travel modes' service attributes are different from the current state. Finally, section C collects data on respondents' socioeconomic and demographic characteristics. At the end of the survey, respondents are encouraged to give their feedback and suggestions for improvements in addition to ranking the survey's complexity level.

Numerous survey tests and trials were conducted to arrive at a survey design that provides the respondent with a user-friendly experience. Early trials helped identify potential improvements to the survey design, several of which were implemented in later trials. For example, to reduce high-density text and improve readability, survey instructions and sample answers were embedded in hyperlinks, appearing only by hovering the mouse arrow over the "Help?" buttons located alongside the survey questions. In addition, illustration figures and videos are presented to provide guidelines and walkthrough examples to the survey respondents. The average complexity of the survey dropped by more than 30% after these improvements were adopted (on a scale from 1 to 5, the average complexity of the survey dropped from 4.2 to 2.9). The engagement of professional software and website developers, comprehensive market research, survey respondents' feedback, and lessons learned from similar studies in the literature shaped the evolution of the survey interface from the early stages until the final design was set.

5.7.1 Survey Data Model

Figure 5-2 shows the survey's data model, which explains the logic behind building the survey structure with its specific question layout/order. The questions are tailored to accommodate all the possible travel mode options, including trips that may involve multiple trip segments. Within the survey questions, different trip segments are clearly defined using distinctive colour codes for each category of questions. Whether the trip is as complex as using three different travel modes or as simple as driving all the way from home to work, the survey is interactively adjusted to accommodate all varying trip components.

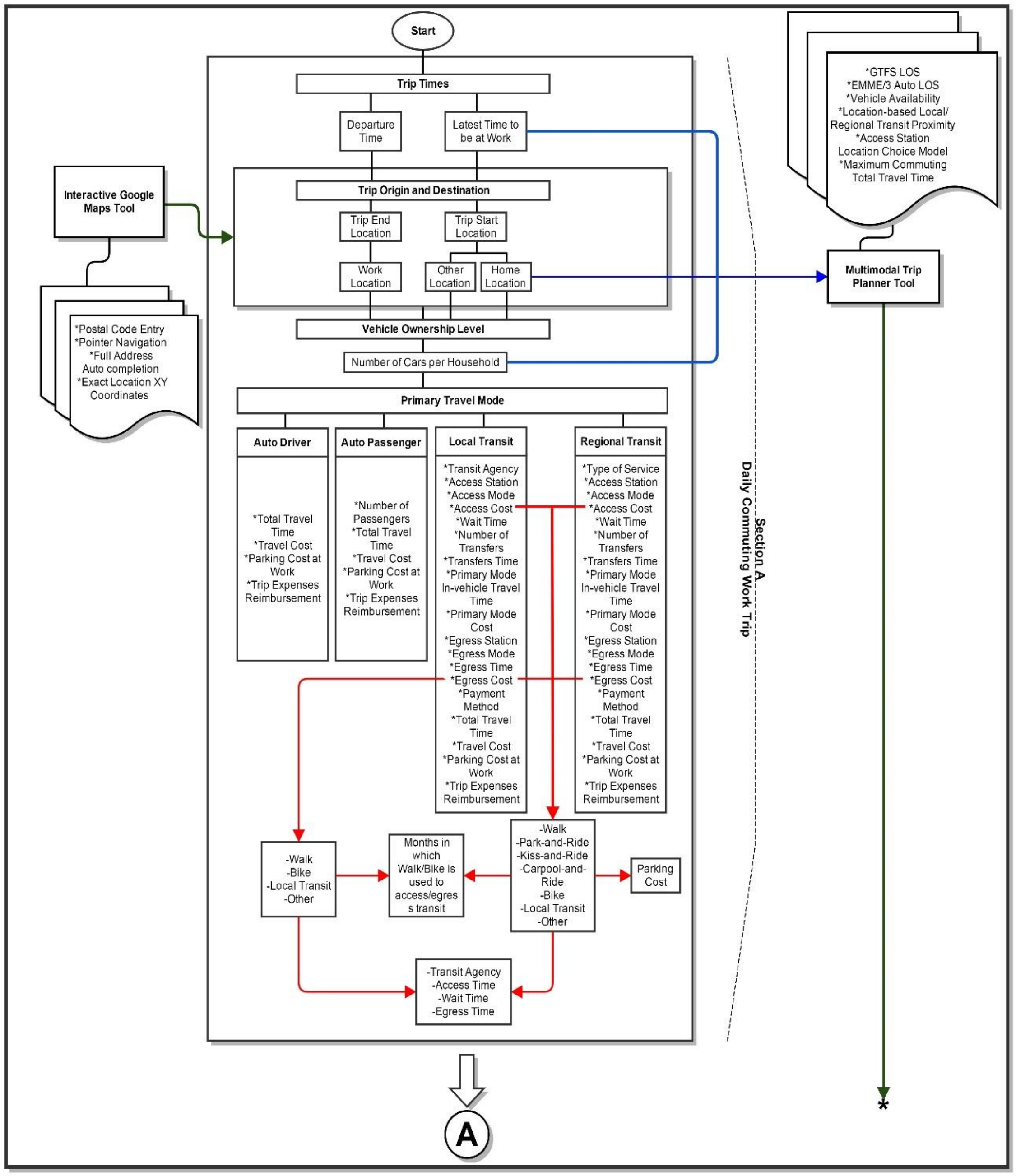


Figure 5-2 (a) SCRIPT Data Model

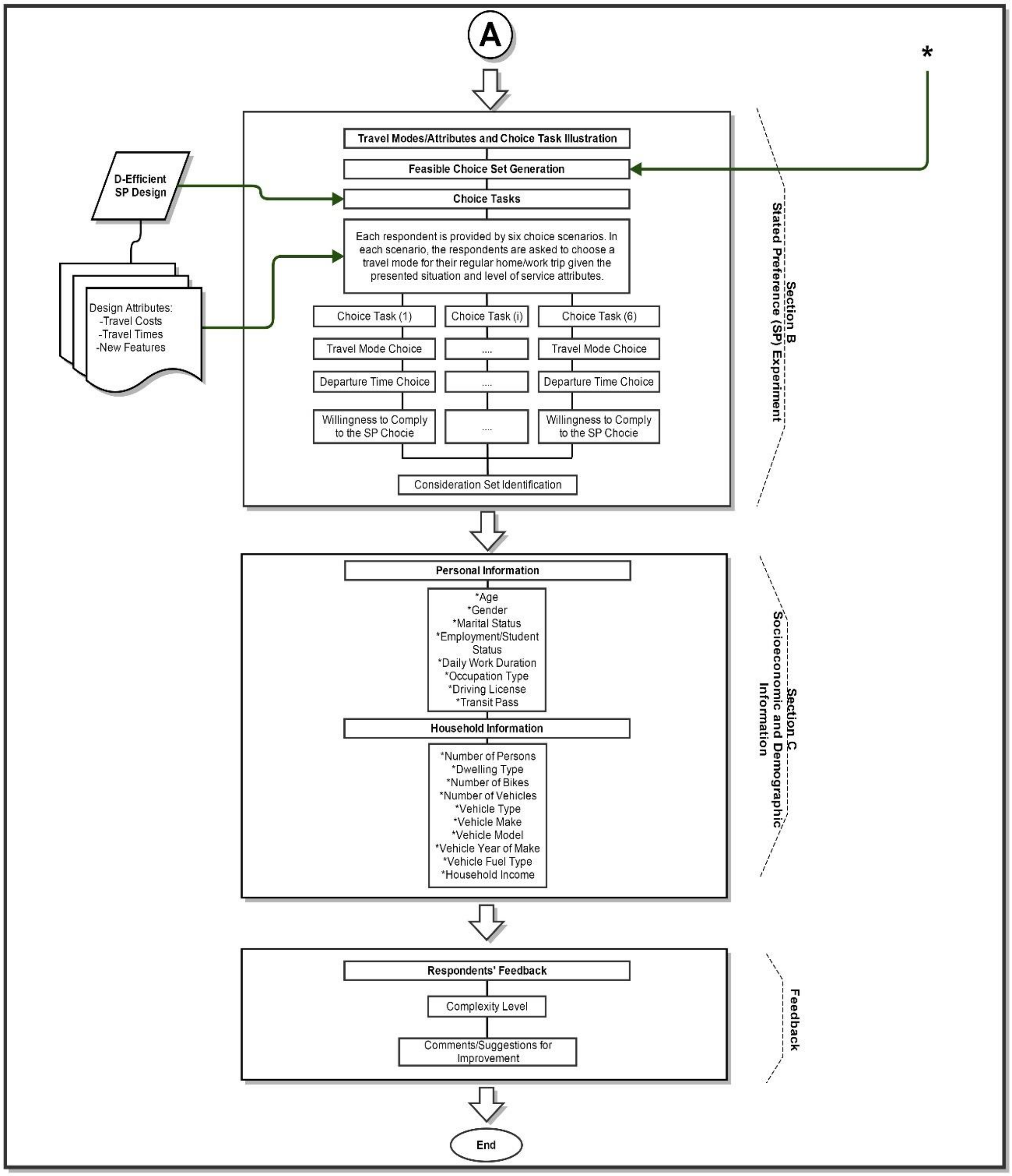


Figure 5-2 (b) SCRIPT Data Model

5.7.2 Multimodal Trip Planner Tool

The information gathered in section A feeds into an innovative multimodal trip planner tool. This tool was developed to generate all feasible travel options for each SP choice experiment for use in Section B of the survey. The tool is embedded within the survey platform to link sections A and B of the survey. Figure 5-3 summaries the steps of how the multimodal trip planner tool is utilized within *SCRIPT*'s platform. First, a set of feasible modes for each respondent based on vehicle ownership level, proximity to transit, and total travel time from home to work is defined. Second, level-of-service attributes, including different elements of travel cost and time components, are generated based on the specified arrival time at work for all feasible modes. Finally, those attributes are fed into the SP experiment design to adjust attribute levels before the final choice situations are presented to the respondents. For any given respondent, the tool finds his/her optimal wait and transfer times between different modes by adjusting the departure time from home while ensuring that he/she arrives at the final destination before the previously stated arrival time at work (elicited in the RP section). Table 5-2 shows the list of conditions under which travel modes are considered unfeasible options. It was assumed that the auto passenger mode is available for all individuals. The number of feasible alternatives typically ranges from two to nine options.

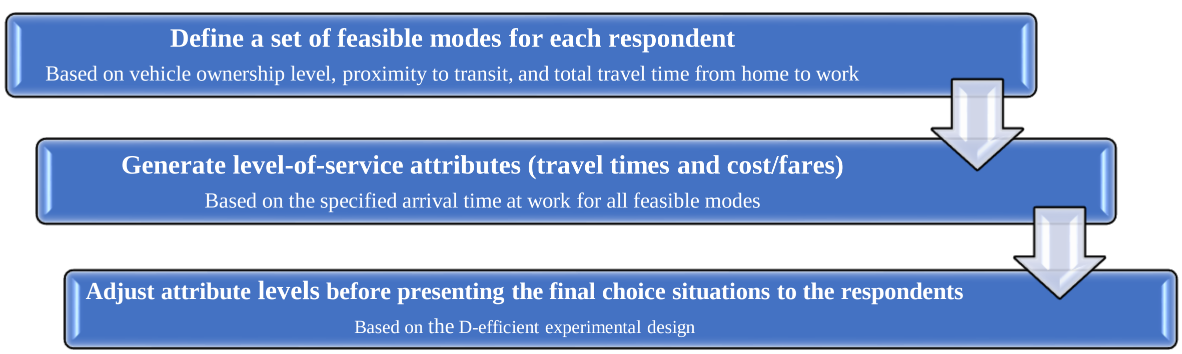


Figure 5-3 The Integration of the Multimodal Trip Planner Tool with *SCRIPT*

The travel time components for the auto driver and auto passenger modes are obtained from an offline origin–destination travel time matrix based on the 2012 EMME3 traffic assignment for the GTHA. However, the travel time components for all transit-based modes are generated based on General Transit Feed Specifications (GTFS) using the Google Maps application programming interface (API). Unlike existing similar tools, this trip planner generates intermodal travel alternatives such as park-and-ride or kiss-and-ride and transit trips that involve the use of two or more transit services (i.e. operated by different transit agencies), in addition to typical travel modes such as auto driving. Furthermore, the tool provides detailed travel cost components for all travel options. Driving cost is determined based on network-based travel distance and average gas cost per km according to the Canadian Automobile Association (CAA, 2014). Other cost components, such as parking costs at trip destinations or park-and-ride stations (if any), are included, along with transit fares, taking into consideration access and/or egress co-fares. For auto passenger/carpool travel modes, travel costs are defined as the auto driving mode divided by the number of travelers, if any; otherwise, a default value of 2 is used.

Table 5-2 Mode Feasibility

Condition	Unfeasible Modes
Zero vehicles in the household	Auto driver, local transit with auto driver access, and regional transit with auto driver access
Closest local transit stop is more than 2 Km (walking access time is more than 24 minutes)	Local transit with walk access and regional transit with local transit access
Closest regional transit stop is more than 2 Km (walking access time is more than 24 minutes)	Regional transit with walk access
Total travel time is more than 120 minutes	Case-specific

As explained in Chapter 4, the choice of transit access station for auto driver access (park-and-ride or kiss-and-ride) is not a straightforward decision of just choosing the nearest station. Among the cross-regional commuters in the GTHA, more than 30% of the park-and-ride users choose a station other than the closest one to their home locations (DMG, 2008). Variables such as access distance, direction of travel to the station relative to the home and work locations, parking cost, type of transit service, and surrounding land use are proven to affect users' choice of access stations. That is, access station location models for park-and-ride regional commuter rail (GO Transit) users and Toronto Transit Commission subway (TTC Subway) users are developed (see Chapter 4). These models are adopted in the multimodal trip planner tool and used to predict respondents' access station location choice for transit options with auto driver or passenger access. As such, for individuals who have park-and-ride and/or kiss-and-ride options available in their feasible choice sets, the trip planner tool selects commuter rail and/or subway access stations based on the pre-developed discrete choice models before generating the associated level-of-service attributes for presentation to the respondent. Therefore, the trip is divided into two main components: a driving access from the trip origin to the predicted station and a transit trip from that station to the final destination. That is, it provides realistic and customized travel alternatives for park-and-ride and/or kiss-and-ride users. To the author's knowledge, this is the first attempt to integrate a multimodal trip planner with a set of pre-developed econometric models to develop a comprehensive respondent-customized data collection tool.

5.7.3 Section A: Revealed Preference (RP) Data

In section A, the revealed preference section, the respondents are asked to provide detailed information about their work trip on the previous typical working day, including departure time, latest time they can be at work, trip start and end locations (location selection is done through an interactive map tool embedded in the survey Web page), vehicle ownership level, travel modes used, and other travel characteristics. The interactive map tool used in the survey is based on a Google Maps® interface that is enhanced with an auto-complete address/postal code option in addition to the typical map navigation capabilities. Based on the selected primary mode of travel, a list of customized questions is dynamically shown to each respondent to capture all necessary trip details for cross-regional commuters. This allows the gathering of detailed information on complex intermodal trips that involve the interactions of different transit modes/service providers and automobiles or non-motorized modes (i.e. walk and bike). Data on travel costs, including gas, parking and transit fares, and methods of payment, are also collected.

5.7.4 Section B: Stated Preference (SP) Data

As explained earlier, the main objective of conducting a stated preference experiment is to evaluate the effectiveness of various policy initiatives and to quantify the effect of changes in level-of-service attributes on individuals' choice of transit as a travel or access mode. These changes in level-of-service attributes are tied to policies under consideration. In this study, such policies aim at improving transit services with more emphasis placed on transit modal integration. Figure 5-4 shows the list of strategies considered in this study. In section B of the survey, the stated preference section, each respondent is provided with six choice scenarios. In each scenario, the respondents are asked to choose a mode of travel to conduct their commuting trip based on a set of attributes and mode characteristics. These attributes are altered according to the predefined polices in Figure 5-4. Table 5-3 shows the considered travel modes' attributes and their definitions. Similarly, Table 5-4 shows the travel modes' attribute levels (where applicable). The SP scenarios are presented in tables as shown in Figure 5-5. The respondents are provided with an embedded instructional video within the survey Web page that explains the SP experiment with a walkthrough example. The mode attributes are characterized into three categories: travel cost components, information provision and special features/amenities, and travel time components.

Figure 5-4 Polices under Investigation

In this study, each SP experiment presents up to nine travel mode options and up to 15 mode attributes, of which only nine attributes have different levels that vary across the experiments. The number of presented alternatives depends on the mode feasibility for each respondent, and accordingly, the corresponding mode attributes are shown or hidden. To maintain the requirements of D-efficient designs, as explained in Section 5.5, the Ngene® software package was used to develop the experiment design for this study (Choice-Metrics, 2012). Based on the mode feasibility conditions presented in Table 5-2, eight SP designs were generated to cover all possible cases. A customized SP design is presented to each respondent, and they are asked to choose a travel mode alternative for each of the six choice tasks.

Based on their chosen mode and its estimated travel time for each choice scenario, respondents are asked to provide their new departure time. This allows for studying the effect of mode choice on departure time choice. In addition, after each choice scenario, the respondents are asked to provide their level of confidence in making their future home–work trip based on the selected travel mode on a five-level scale: not confident, somewhat confident, neutral, confident, and strongly confident. After the sixth and last choice scenario, the respondents are asked to choose travel modes that were considered while making the choice. As described above, this study considers a universal choice set of nine travel alternatives. However, based on the conditions adopted within the trip planner tool, only feasible modes are presented to each respondent. However, respondents may not consider all of the feasible modes while making their choices. Such information helps in developing a customized choice set for each respondent and therefore reduces the error of unrealistically assuming a uniform choice set across all individuals.

5.7.5 Section C: Household and Personal Information

In section C, socioeconomic and demographic information is collected. On the individual level, the survey collects data on age, gender, marital status, employment/student status, daily work duration, occupation type, and the availability of a driving license or transit pass. Similarly, household information, such as the number of persons per household, dwelling type, number of bikes and vehicles per household, and household income is gathered.

Table 5-3 Travel Modes' Attributes

	Mode Attribute	Definition
Travel Cost Components	Travel cost/transit fare	Travel cost including fuel cost and/or transit fare(s) (Canadian dollars)
	Reserved parking at park-and-ride GO stations	The availability of a reserved parking option at a park-and-ride GO transit station; this attribute takes the value of "Yes" for available and "No" otherwise
	Daily/monthly parking cost at park-and-ride GO stations	Daily or monthly parking cost at park-and-ride GO transit stations (Canadian dollars); daily parking rates are provided if reserved parking option is not available and vice versa
	Parking cost at TTC Subway park-and-ride stations	Parking cost at park-and-ride TTC subway stations per day per person (Canadian dollars)
	Parking cost at trip destination	Daily parking cost at work location per person (Canadian dollars)
	Local transit to GO co-fare	Co-fare of local transit if local transit is used to access GO Transit (Canadian dollars)
	GO to local transit co-fare	Co-fare of local transit if local transit is used after GO Transit (Canadian dollars)
Information Provision and New Features	Next local transit vehicle information	The availability of information provision for the arrival of the next local transit vehicle; this attribute takes the value of "Yes" for available and "No" otherwise
	Wi-Fi on GO trains/buses	The availability of Wi-Fi services on regional (GO) trains/buses; this attribute takes the value of "Yes" for available and "No" otherwise
Travel Time Components	Transfer time(s)	Time taken to transfer between different transit lines, vehicles, or modes (minutes)
	Wait time	Time taken to wait for boarding a transit vehicle at the first (access) transit stop/station of the primary mode (minutes)
	Access time	Time taken to travel from the trip origin location to the first (access) transit stop/station of the primary mode (minutes)
	In-vehicle travel time	Time taken to travel from the first (access) transit stop/station to the last (egress) transit stop/station on a transit vehicle(s) of the primary mode (minutes)
	Egress time	Time taken to travel from the last (egress) transit stop/station of the primary mode to the final trip destination (minutes)
	Total trip time	Total trip time from the trip start location (origin) to the trip final destination (minutes)

Table 5-4 Travel Modes' Attribute Levels

Mode Attribute		Attribute Levels	
Travel Cost/Fare ($)	3	Low	Current
		Medium	+50% (car) +20% (transit)
		High	+75% (car) +30% (transit)
Reserved Parking (Regional Transit)	2	Yes	
		No	
Daily Parking Cost at Regional Transit Stations ($)	3	Low (current)	0
		Medium	4
		High	8
Monthly Parking Cost at Regional Transit Stations ($)	3	Low	40
		Medium (current)	80
		High	120
Parking Cost at Local Transit (TTC Subway) Park-and-ride Stations	NA	Current	
Parking Cost at Trip Destination	NA	Current	
Local Transit–Regional Transit Access Fare ($)	3	Low	0
		Medium	-50%
		High	Current
Regional Transit–Local Transit Egress Fare ($)	3	Low	0
		Medium	-50%
		High	Current
Next Local Transit Vehicle Information Provision	2	Yes	
		No	
Wi-Fi on Regional Transit Vehicles (Go Bus/Train)	2	Yes	
		No	
Transfer Time (at Transfer Stations between Local and Regional Transit)	3	Low	-50%
		Medium	Current
		High	+50%
Wait Time	3	Low	-50%
		Medium	Current
		High	+50%
Access Time	NA	Current	
In-vehicle Travel Time	NA	Current	
Egress Time	NA	Current	
Total Trip Time	NA		

Choice Scenario: 1

Mode Attribute	Auto Driver Help?	Auto Passenger/ Carpool Help?	Local Transit – Walk Access Help?	TTC Subway – Auto Driver Access Help?	TTC Subway – Auto Passenger Access Help?	GO Transit – Auto Driver Access Help?	GO Transit –Auto Passenger Access Help?
Travel Cost/Transit Fare -Help?	5.50	2.75	6.25	5.25	4.42	5.03	4.67
Reserved Parking at P&R GO Transit Stations -Help?	--	--	--	--	--	No	--
Daily/Monthly Parking Cost at P&R GO Transit Stations -Help?	--	--	--	--	--	4	--
Parking Cost at P&R TTC Subway Stations -Help?	--	--	--	5	--	--	--
Parking cost at trip destination -Help?	5	2.50	--	--	--	--	--
GO Transit to Local Transit co-fare (egress) - Help?	--	--	--	--	--	3	1.50
Next bus information - Help?	--	--	No	Yes	Yes	--	--
Wi-Fi on GO Transit -Help?	--	--	--	--	--	No	Yes
Transfer Time(s) - Help?	--	--	5	2	2	4	4
Waiting Time- Help?	--	--	4	2	2	1	1
Access Time - Help?	--	--	5	10	10	6	6
In-vehicle Travel Time - Help?	42	42	44	23	23	20	20
Egress Time - Help?	--	--	4	4	4	4	4
Total Trip Time - Help?	**42**	**42**	**62**	**41**	**41**	**35**	**35**
Choice	☑	☐	☐	☐	☐	☐	☐

Based on the chosen mode for your work trip, what is the expected departure time from home?

Please enter your departure time from home (Hour: Minute): 08 : 15

In the future, what would be your propensity to make your work trip using the option selected above?

☐ Not Confident ☐ Somewhat Confident ☐ Neutral ☑ Confident ☐ Strongly Confident

NEXT

Figure 5-5 Snapshot of a Sample SP Experiment

5.7.6 Respondents' Feedback

Finally, the survey ends with an acknowledgement of the respondents' efforts, followed by two optional questions. The respondents are asked to rank the complexity level of the survey on a scale from 1 (very easy) to 5 (very complex) and to leave comments/suggestions for improvements or to report any problems encountered. Adding the feedback section at the end of the survey was extremely useful in developing the survey structure/layout as well as the wording of questions. Various test runs were conducted, and respondents' feedback was considered carefully. As a result, the average complexity of the survey dropped from 4.2 (after the first pilot test) to 2.9 (after the final survey) based on respondent feedback.

5.8 Chapter Summary

This chapter presents the development of the Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*). *SCRIPT* collects information on the respondents' commuting trips (RP data) as well as their stated preferences toward different travel mode alternatives with improved and/or new level-of-service attributes (SP data). The survey is carefully designed to capture the changes in individuals' travel mode choices in response to policies that aim at improving transit services with more emphasis on transit modal integration. A pilot survey was conducted to test the survey platform, and a pilot sample was collected. This sample was used to develop a preliminary SP mode choice model. Using the preliminary model's estimated parameters as *a priori* estimates, the final SP experiments were developed based on the D-efficient design technique.

SCRIPT is respondent-customized; that is, the questions are tailored to accommodate all the possible travel mode options. The information gathered from the RP section feeds into an innovative multimodal trip planner tool. This tool generates only feasible travel options for each SP choice experiment based on households' auto ownership level, proximity to transit, work start time, and total travel time from home to work. For intermodal travel modes such as park-and-ride and/or kiss-and-ride, the tool selects the access stations based on the pre-developed access station location choice models before generating the associated level-of-service attributes for presentation to the respondent. Finally, socioeconomic and demographic information is collected. The following chapter, Chapter 6, provides details on the survey implementation, data collection process, and descriptive statistics of the collected data.

CHAPTER 6

6 SURVEY IMPLEMENTATION AND DATA COLLECTION

6.1 Chapter Overview

This chapter provides an overview of the survey implementation process. Section 6.2 discusses the data collection procedure. This is followed by detailed descriptive statistics and a preliminary analysis of the collected data, which are presented in Section 6.3. Finally, a chapter summary is provided in Section 6.4.

6.2 Data Collection

The data collection was done during the spring of 2014, and a supplementary subset was collected during the fall of the same year. The average time required to complete the survey was 20 minutes. Invitations to the online survey were randomly emailed to a panel of 35,000 respondents who previously agreed to be contacted for similar studies either by enrollment in air miles reward programs or by telephone requirement. A market research company conducted the survey on behalf of the research team. The sampling criteria and reward system were reviewed and approved by the Research Ethics and Protection committee at the University of Toronto.

Table 6-1 shows *SCRIPT's* implementation metrics. The total number of accepted invitations was 15,975, of which 2,986 respondents were qualified to participate in the study. The total number of complete responses was 1,003 with a completion rate (the ratio of the number of complete responses to the number of respondents who qualified to participate in the survey) of 33.6%. The overall response rate (the ratio of the number of complete responses to the number of respondents who attempted to participate in the survey) was 6.3%. This relatively low rate highlights the complex nature of the study due to the low qualification rate of the target population from the defined sample frame. After the data were cleaned (by removing invalid responses and/or records that do not belong to the study area's sample frame), 704 complete valid records were prepared to be used for empirical modelling.

Table 6-1 *SCRIPT* Implementation Metrics

Survey Metrics	
Number of sent invitations	35,000
Number of accepted invitations	15,975
Recruitment rate	45.6%
Number of qualified respondents	2,986
Qualification rate	18.7%
Number of complete responses	1,003
Completion rate	33.6%
Overall response rate	6.3%

6.3 Descriptive Data Statistics

6.3.1 Revealed Preference Data

Table 6-2 shows the aggregate mode shares of the collected RP data and their corresponding records from the 2011/2012 TTS.

Table 6-2 *SCRIPT* and TTS Sample Distribution by Mode Split*

	SCRIPT **Sample Distribution (%)**	**TTS Sample Distribution (%)**	**Difference (%)**
Auto	82	84	-2
Transit	18	16	+2

* Considering an aggregate modal distribution of auto and transit users.

Similarly, Table 6-3 shows the spatial distribution of trip origins from the collected RP data and their corresponding records from the 2011/2012 TTS. Figure 6-1 shows a density distribution of trip origins and destinations based on the collected sample.

Table 6-3 *SCRIPT* and TTS Sample Distribution by Trip Origin

	SCRIPT **Sample Distribution (%)**	**TTS Sample Distribution (%)**	**Difference (%)**
City of Toronto	25	25	0
Durham Region	13	10	+3
York Region	26	20	+6
Brampton	10	11	-1
Mississauga	15	14	+1
Halton Region	9	14	-5
Hamilton	2	6	-4

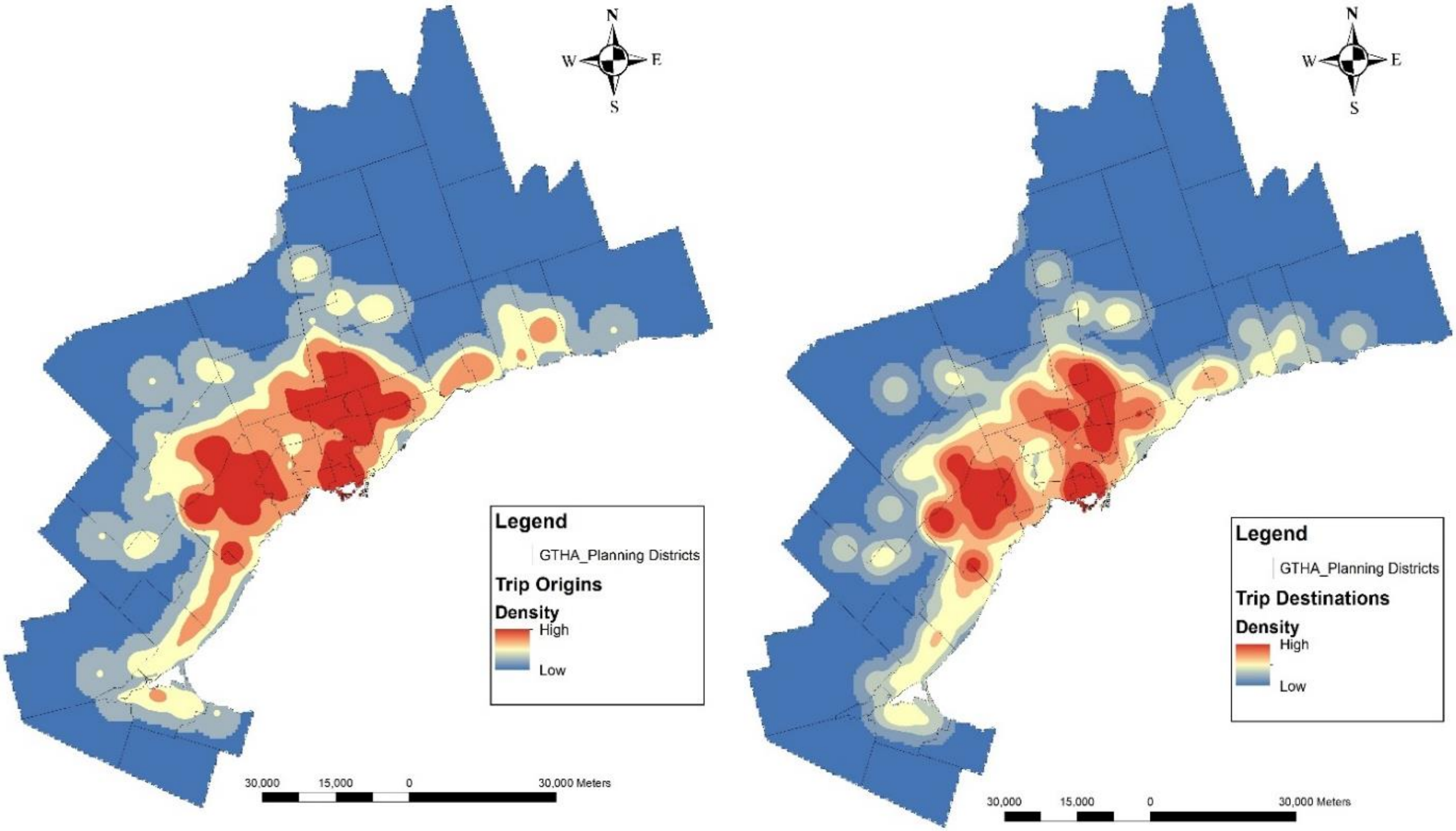
As mentioned earlier, cross-regional trips are long-distance trips by nature. More than 70% of automobile users and 95% of transit users in the dataset commute distances that are more than 30 minutes. In addition, the average travel time by local transit modes (with walk access) is 75 minutes. These statistics verify the validity of the aforementioned assumptions on the unfeasibility of non-motorized modes for cross-regional trips.

6.3.2 Stated Preference Data

Figure 6-2 shows detailed mode shares based on the TTS data for cross-regional commuters, *SCRIPT* RP data, and *SCRIPT* SP data. Figure 6-2 (a) and Figure 6-2 (b) show similar mode share distributions of the sample RP data and the corresponding TTS data. Figure 6-2 (c) shows a significant drop in the auto driver mode share and an increase in the auto passenger and transit mode shares because of changes in the alternatives' level-of-service attributes in the presented SP scenarios.

Figure 6-3 shows a distribution of the confidence level with which the survey respondents made their SP choices. As shown in Figure 6-3, more than 50% of the respondents made their SP choices with a high level of confidence. This indicates that the respondents would make similar choices in the future if similar choice situations arise.

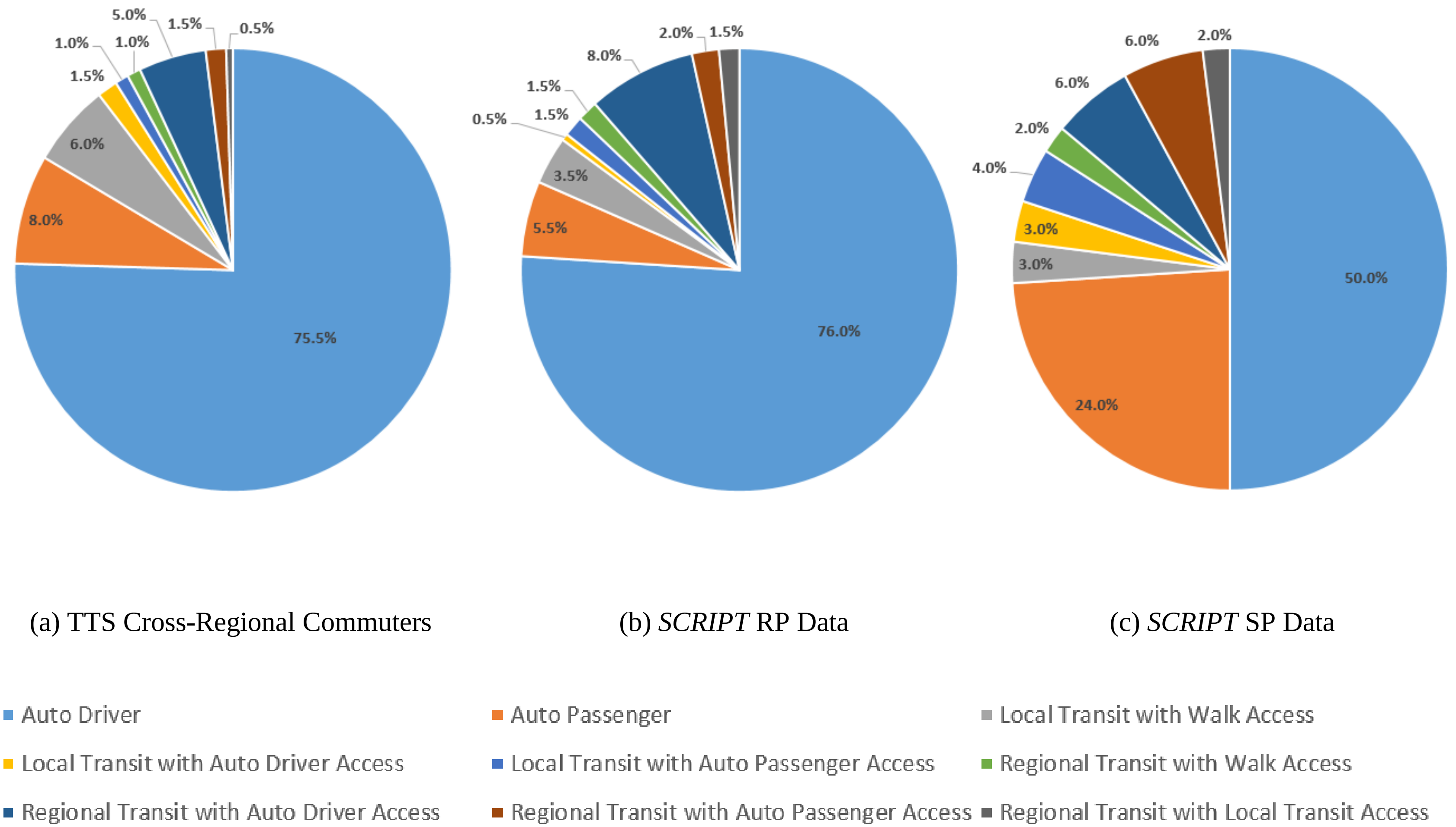
As explained in Chapter 5, a customized choice task was presented for each respondent based on the respondent's home/work location, household vehicle ownership level, household proximity to transit, and total travel time from home to work. The travel modes and their corresponding attributes were shown or hidden based on a set of predefined rules, as presented in Table 5-2. Figure 6-4 shows the distribution of the available modes across the collected sample.



(a) Trip Origins Density Distribution

(b) Trip Destinations Density Distribution

Figure 6-1 Spatial Distribution of Cross-Regional Trip End



(a) TTS Cross-Regional Commuters (b) *SCRIPT* RP Data (c) *SCRIPT* SP Data

Figure 6-2 Cross-Regional Trip Mode Shares

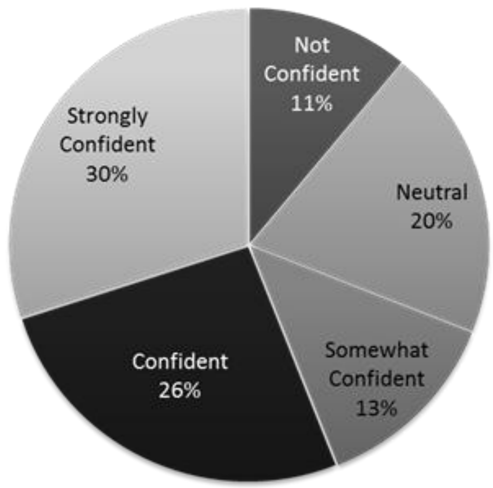


Figure 6-3 Level of Confidence in SP Choice

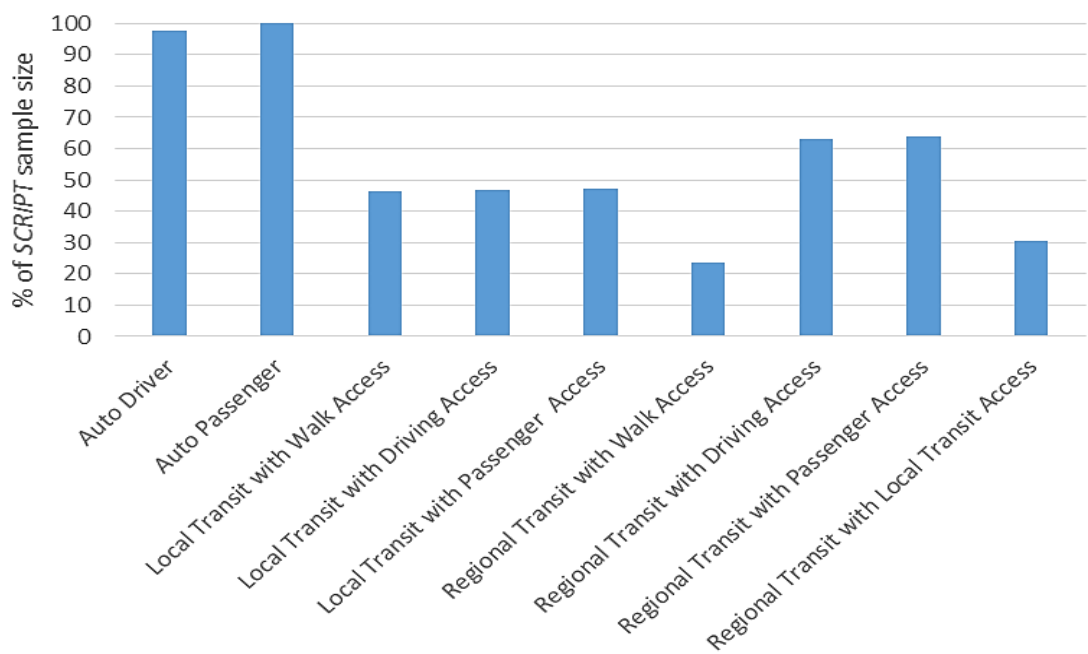


Figure 6-4 Mode Availability

6.3.3 Household and Personal Information

Table 6-4 shows the distribution of individuals by gender for the collected RP data and their corresponding records from the 2011/2012 TTS. Figure 6-5, Figure 6-6, Figure 6-7, Figure 6-8, and Figure 6-9 show the distributions of age, household size, vehicle ownership level, bike ownership level, and household yearly income, respectively. As shown, the age, household size, and vehicle ownership level distributions follow a Gaussian shape. In contrast, over 50% of the sampled households have less than one bike, which indicates that a majority of cross-regional commuters do not consider biking as a travel mode (or access mode). This is the case because of the relatively long commuting travel distance and the lack of bike infrastructure in areas other than downtown Toronto. In addition, the household yearly income level of more than 50% of the sampled households is $100,000 and over. This explains the high vehicle ownership level within the sample.

Table 6-4 *SCRIPT* and TTS Sample Distribution by Gender

	SCRIPT Sample Distribution (%)	TTS Sample Distribution (%)	Difference (%)
Male	59	62	-3
Female	41	38	+3

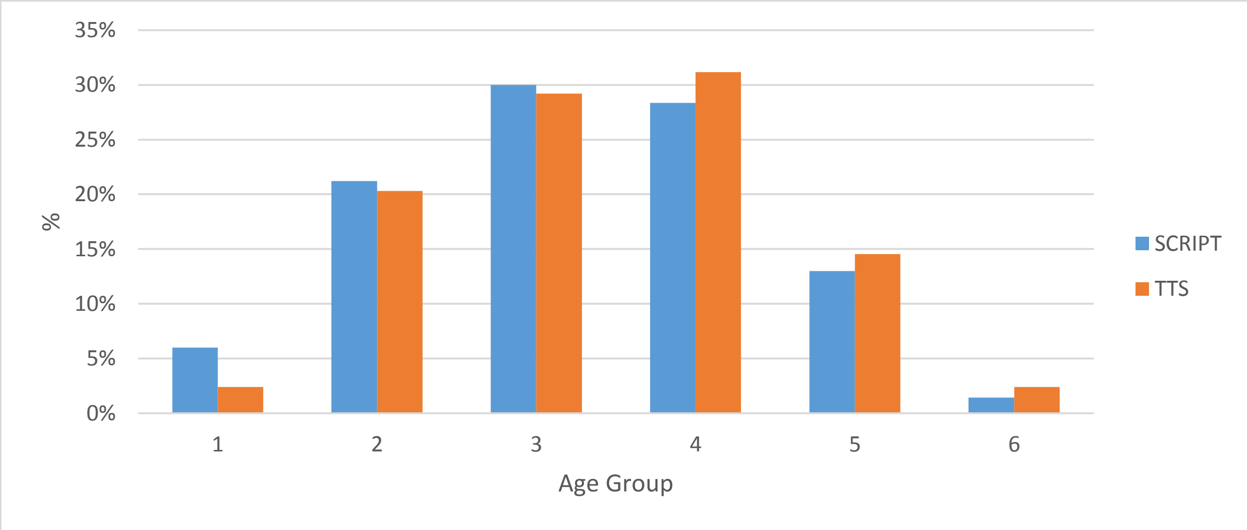


Figure 6-5 Age Distribution

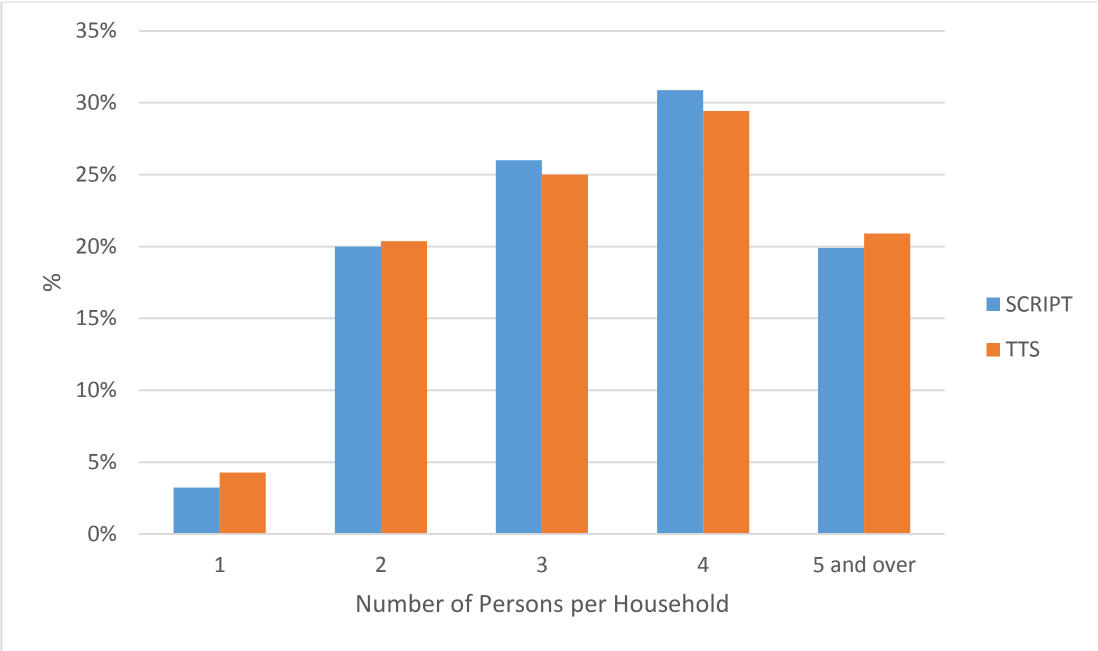


Figure 6-6 Household Size Distribution

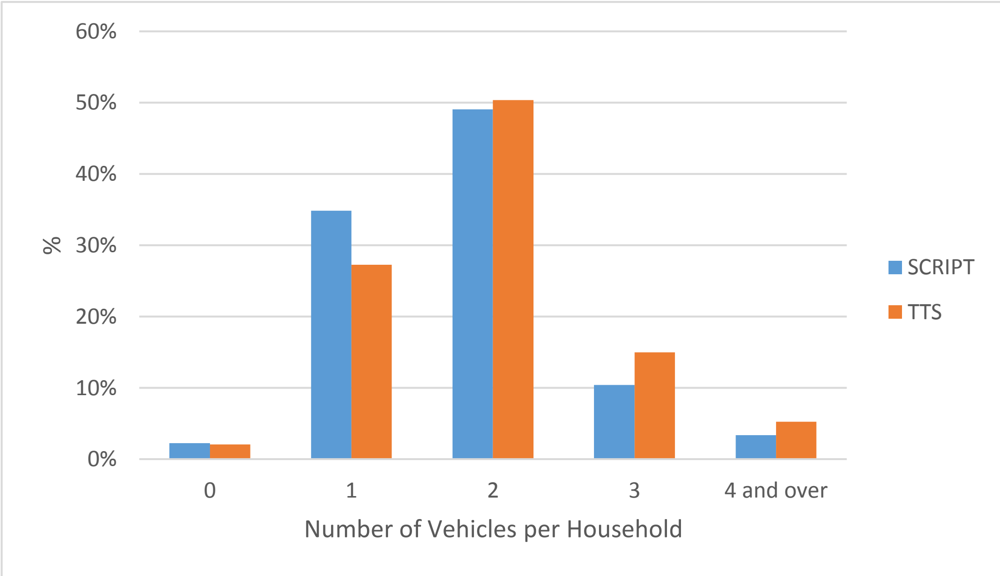


Figure 6-7 Vehicle Ownership Distribution

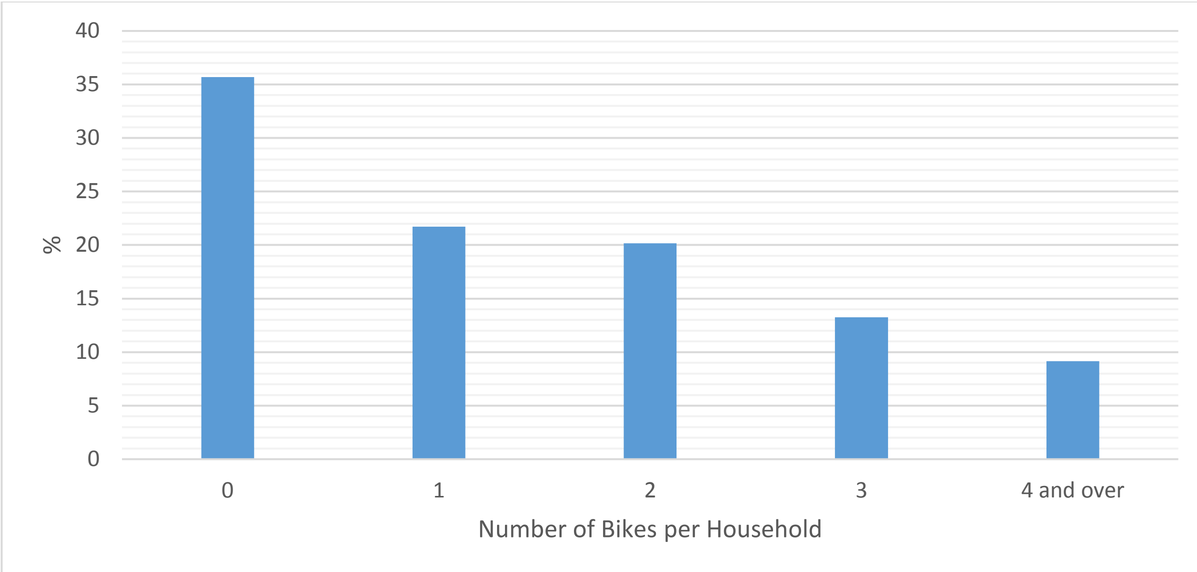


Figure 6-8 Bike Ownership Distribution

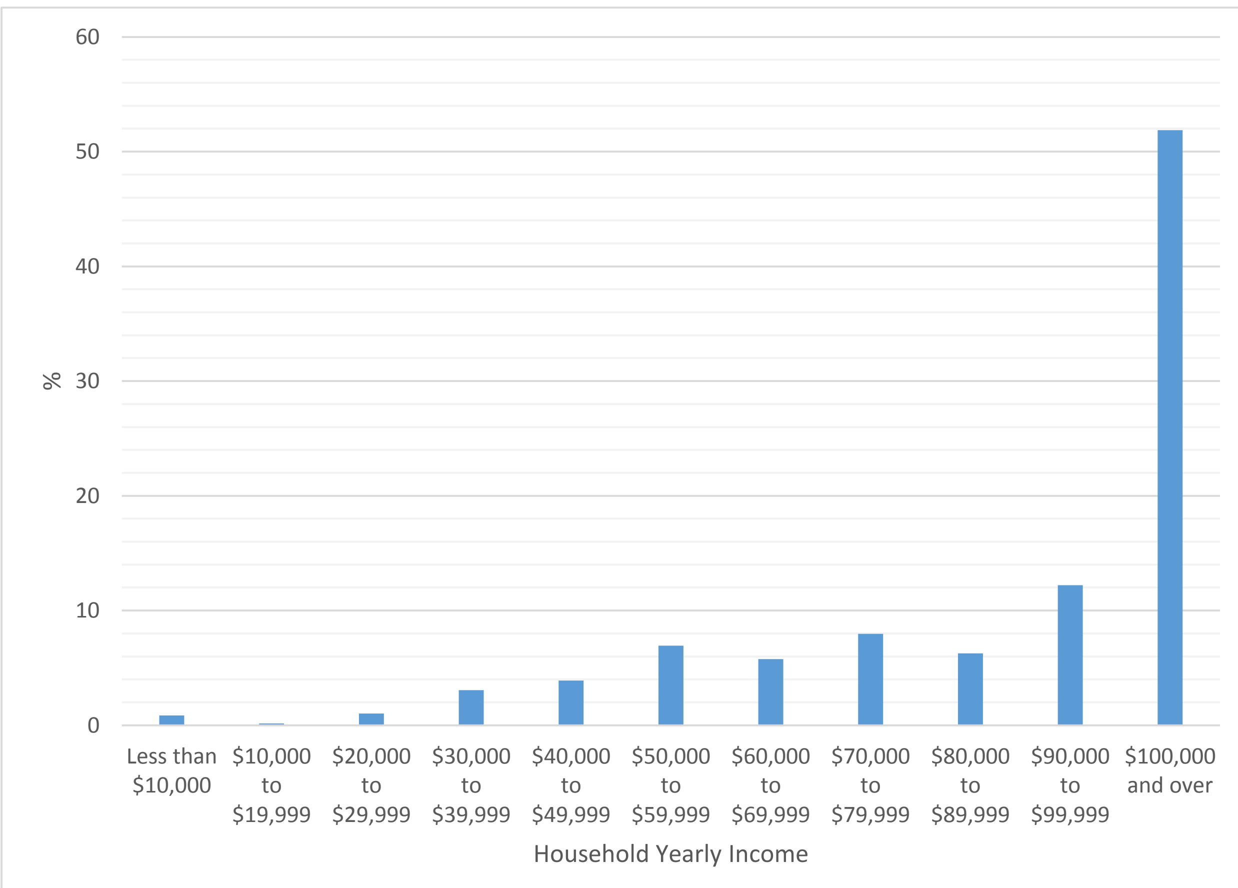


Figure 6-9 Household Yearly Income Distribution

Figure 6-10, Figure 6-11, and Figure 6-12 show distributions for household dwelling type, individuals' marital status, and occupation types. A majority of the sampled households are families (married respondents) living in houses composed of two or more persons. In addition, over 25% of the sampled individuals are employed within the professional/management/technical occupation group. The sample includes 1% cross-regional students.

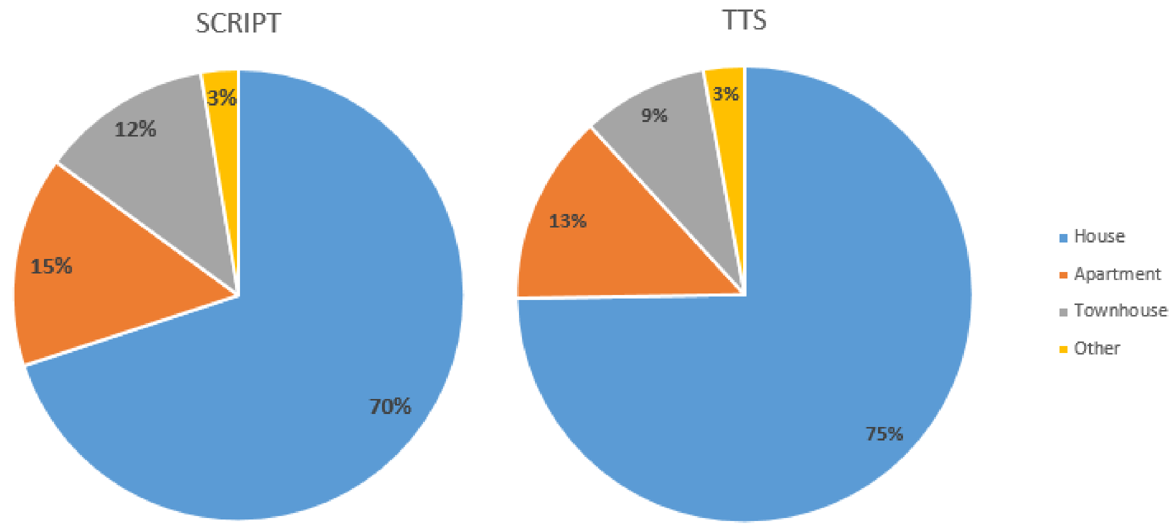


Figure 6-10 Dwelling Type Distribution

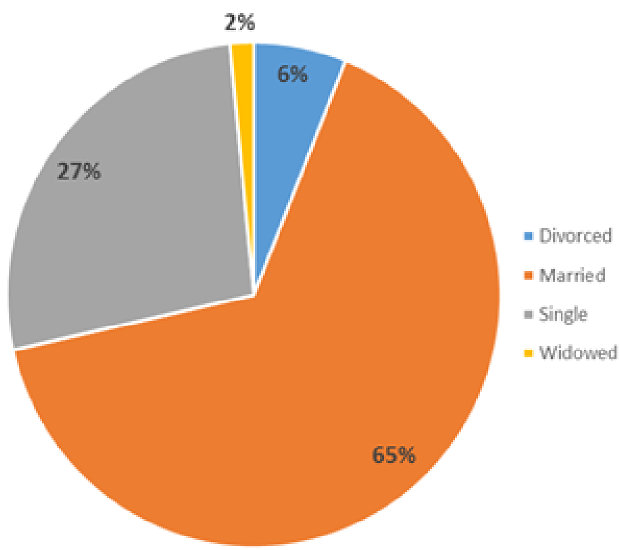


Figure 6-11 Marital Status Distribution

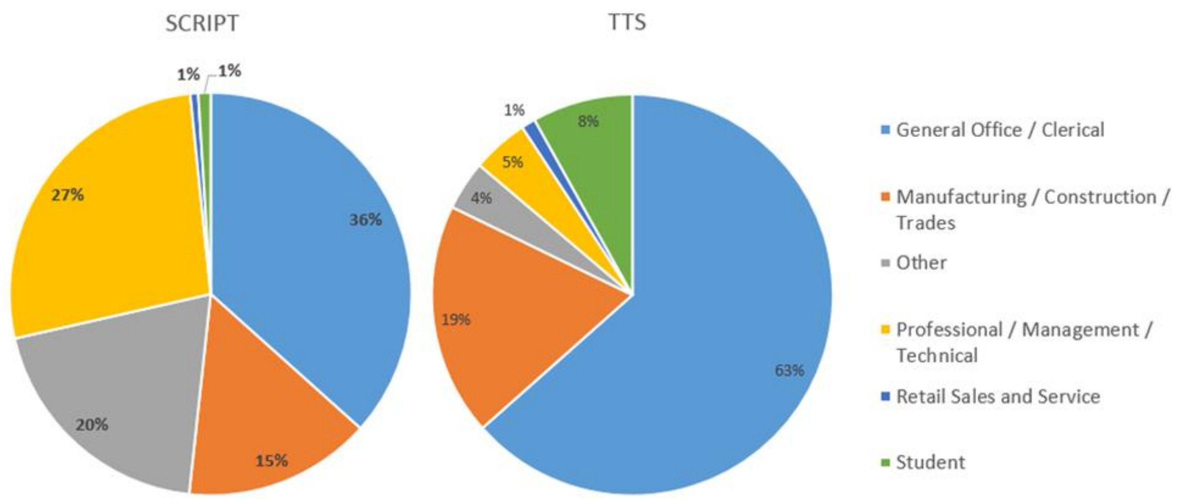


Figure 6-12 Occupation Type Distribution

In accordance with the high vehicle ownership level, 97% of the respondents possess a driving license, as shown in Figure 6-13. Similarly, around 74% of the sampled individuals do not hold transit passes. Figure 6-14 shows the distribution of transit pass possession within the sample dataset. This explains the auto-oriented mode split of the RP data, with around 82% of individuals relying on automobiles as their commuting mode of travel. Interestingly, 25% of individuals who have access to at least one car per household possess a transit pass. These individuals may contribute to the observed 12% share of transit modes with auto driver or passenger access (i.e. park-and-ride and/or kiss-and-ride).

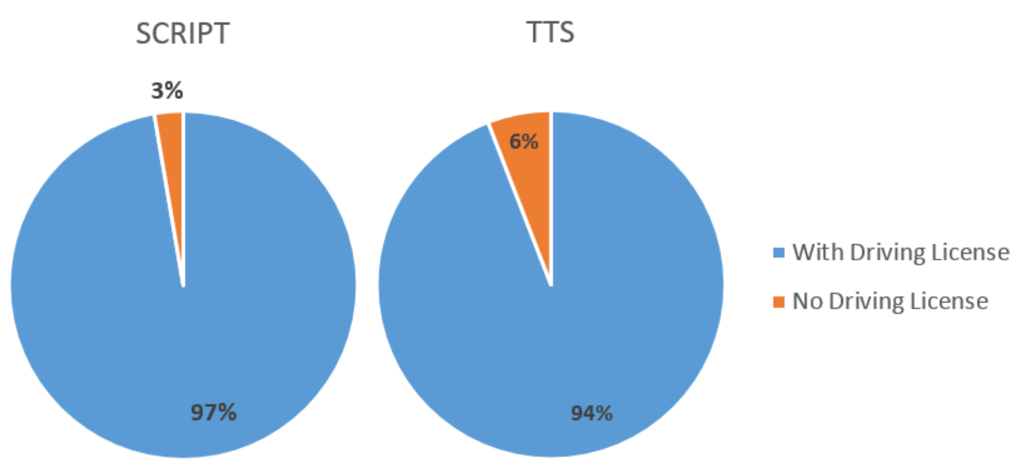


Figure 6-13 Driving Licence Possession Distribution

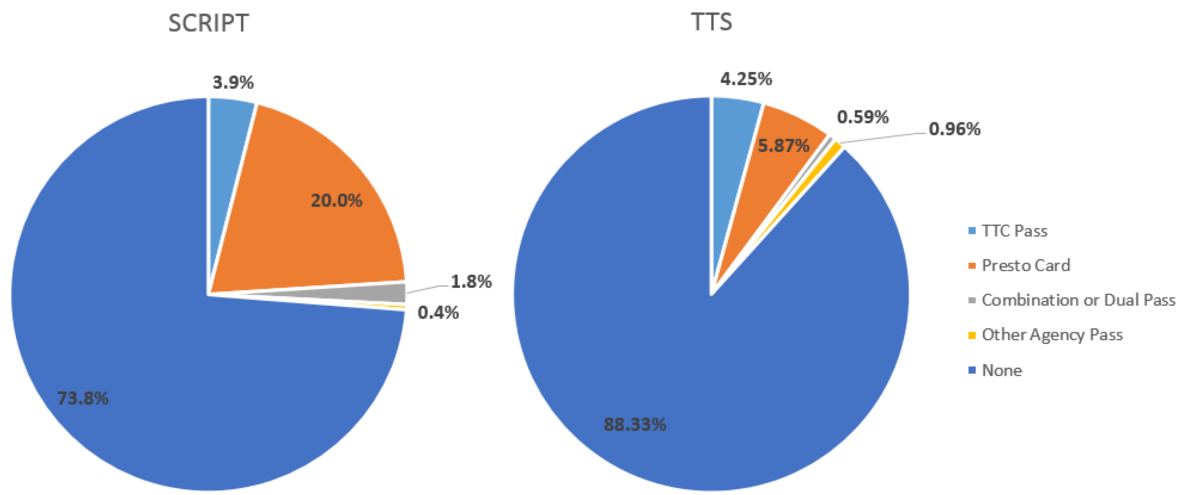


Figure 6-14 Transit Pass Possession Distribution

6.4 Chapter Summary

This chapter provides details on *SCRIPT's* implementation and data collection processes. Descriptive statistics and preliminary analysis of the collected data are presented. The results showed that the collected data provide a representative sample with only marginal differences compared to the 2012/2013 TTS records. The reason for these minor changes is perhaps caused by the temporal changes and difference in sampling methods between the two surveys. This verifies the validity of the survey design/implementation process and the quality of the collected data. The following chapter, Chapter 7, presents the econometric modelling frameworks developed empirically using *SCRIPT*'s data.

CHAPTER 7

7 MODE CHOICE MODELLING OF CROSS-REGIONAL COMMUTING TRIPS

7.1 Chapter Overview

This chapter presents the development of a set of cross-regional commuters' joint main mode and access mode choice models using revealed preference (RP) and stated preference (SP) data from the Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*) and corresponding level-of-service attributes data. Section 7.2 provides details of *SCRIPT* data preparation for empirical modelling, including RP/SP data fusion and the generation of level-of-service attributes. Further, Section 7.3 presents the econometric framework of the estimated empirical models, which are presented in Section 7.4 and Section 7.5. Section 7.6 provides a comparison between the models developed. Section 7.7, presents the development of an interactive policy analysis tool and showcases its practical implications. Finally, Section 7.8 provides a summary of this chapter.

7.2 *SCRIPT* Data

SCRIPT data are used to develop a set of econometric mode choice models to help in understanding cross-regional commuters' travel behaviour. As explained in Chapter 5, *SCRIPT* provides detailed information about respondents' current travel options and their stated preferences concerning different travel alternatives in response to changes in the current level-of-service attributes. This section presents how level-of-service attributes were generated, how RP and SP data were prepared, and how the complete joint dataset was used for the empirical models' estimation and validation.

7.2.1 Generating Level-of-service Attributes

Level-of-service attributes of the considered travel modes are generated using the multimodal trip planner tool (as described in section 5.7.2). The travel time components for the auto driver and auto passenger modes are obtained from an offline origin–destination travel time matrix based on the 2012 EMME3 traffic assignment for the GTHA. However, the travel time components for all transit-based modes are generated based on the General Transit Feed Specifications (GTFS) using the Google Maps application programming interface (API). In addition, driving travel costs are determined based on network-based travel distance and average gas cost per km according to the Canadian Automobile Association (CAA). In contrast, transit travel costs are determined based on a pre-developed transit fare matrix that takes into consideration access and/or egress co-fares between different transit services. Parking costs at work locations are defined based on the stated values by the survey respondents, while parking costs at park-and-ride stations are defined based on the parking cost charged at each station by the transit service operators. The RP data are joined with the current level-of-service attributes. However, the SP data are joined with the modified level-of-service attributes based on changes in attributes' levels according to the experimental design. That is, respondents' SP mode choices are associated with the level-of-service attributes shown to them during the SP experiment.

7.2.2 Data Fusion of RP/SP Information

The collected revealed preference, stated preference, and household and personal information is prepared to construct a complete database for empirical modelling development. Three datasets are prepared: RP-only, SP-only, and joint RP/SP data. Each dataset includes corresponding level-of-service attributes and household and personal information.

The joint RP/SP dataset encompasses seven data points for each individual: one RP and six SP records. In other words, each row of the dataset represents an individual's observed mode choice and the six stated mode choices as well as their associated level-of-service attributes. In addition, since the choice set varies across the individuals within the dataset (based on the feasible modes for each respondent), data on the availability of all travel modes are included.

7.2.3 Estimation and Validation Datasets

The total number of complete and valid records that are prepared for empirical modelling is 704. A randomly selected subset of the full sample dataset is prepared for model estimation using data for 560 individuals (80% of the full sample dataset size); the remaining records are retained for model validation. That is, an independent holdout sample of 144 trip records is reserved to validate the developed models. Retaining a subset of the full dataset for model validation is a practical practice in which the predictive performance of the developed model is rigorously tested.

However, more advanced/complex model structures that can provide in-depth insights into individuals' behaviour in making travel choices may require additional data. Therefore, such models are often estimated and validated using the full dataset.

7.3 Econometric Modelling Frameworks

In this section, the econometric modelling frameworks of the estimated empirical models are presented. The frameworks can be classified into two groups: traditional mode choice models, which, as explained in Chapter 2, suffer from the independence and irrelevant alternatives (IIA) property, and advanced modelling frameworks, which relax this assumption. These frameworks take into account the full RP/SP information within the model estimation routines. Traditional mode choice models include the multinomial logit (MNL) model, which is the most common discrete choice model formulation. To relax the IIA property, two advanced choice model structures are adopted: the nested logit (NL) and the parameterized logit captivity (PLC) models.

7.3.1 Multinomial Logit Model

Each respondent has a customized choice set that contains from two to nine alternatives. Individuals are assumed to achieve a certain level of utility by choosing one alternative from their choice set. According to the random utility maximization (RUM) theory,

$$U_n = V_n + \varepsilon_n = (\beta \cdot x)_n + \varepsilon_n, \qquad [7-1]$$

where U is the utility function, the subscript *n* indicates one of the mode alternatives in the choice set, V is the systematic utility component, which is a linear-in-parameter function of the observed variables χ and their corresponding coefficients β, and ε is the random error component, which is assumed to follow the independent and identical distribution (IID) of a Type I Extreme Value distribution. It was assumed that travellers are rational in making their decisions by choosing the alternative with the highest utility value among a set of feasible alternatives. This assumption results in a multinomial logit (MNL) model (Ben-Akiva and Lerman, 1985) of the following form:

$$\Pr(n) = \frac{\exp(\mu \cdot V_n)}{\sum_{n'=1}^{N} \exp(\mu \cdot V_{n'})}, \qquad [7-2]$$

where Pr(n) is the probability of choosing alternative n, N indicates the total number of alternatives under consideration by each respondent, and μ is the scale parameter. As explained above, individuals have variable choice sets depending on the availability of travel modes. As such, the probability calculations takes into consideration the varying choice set across the individuals.

To jointly estimate an RP/SP model, an artificial tree structure, as shown in Figure 7-1, is assumed to identify the differences between the two datasets: RP and SP data. This is captured through a scale parameter that is estimated for the SP data relative to a normalized (fixed to 1) scale parameter for the RP data. In addition, the scale parameter allows for capturing the heteroscedasticity in individuals' responses. As such, the scale parameter was parameterized as an exponential function of the respondents' attributes (Habib et al., 2014):

$$\mu = \exp(\alpha \cdot \tau), \qquad [7-3]$$

where τ refers to attributes that can explain scale variation (e.g. socio-demographic or land use attributes), and α refers to their corresponding coefficients.

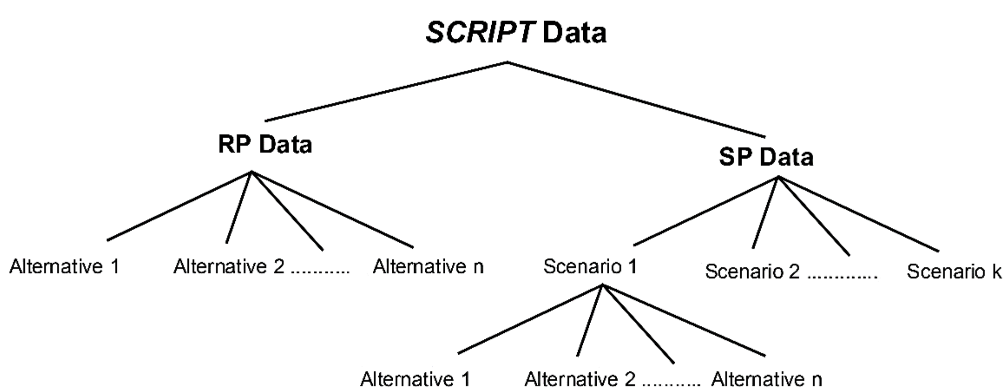


Figure 7-1 The Artificial Nested Structure of Joint RP/SP MNL Model

The repeated SP observations are treated as panel data to allow for the correction of the group behaviour (i.e. each observation is represented by six hypothetical choice scenarios) (Cherchi and Ortúzar, 2006). For a sample of Z individuals with multiple responses (an RP choice and six SP choices), the joint likelihood function can be expressed as follows:

$$L(\beta) = \prod_{z}^{Z} \prod_{d}^{D} \prod_{n}^{C} (\Pr(n)_{dz})^{y_{ndz}}, \qquad [7-4]$$

where β is a vector of variable coefficients that maximizes the likelihood function $L(\beta)$, d represents either the RP data or one of the six choice situations of the SP data, and y_{ndz} =1 if person z chooses alternative n from a variable choice set C in dataset d and zero otherwise. SP choice scenarios where low confidence levels were reported by the respondents are not considered in the estimation process.

7.3.2 IIA-relaxed Models

This section presents the econometric modelling framework of two selected models to relax the IIA property of the MNL model: the NL and PLC models.

7.3.2.1 Nested Logit Model

Unlike the MNL model structure, where equal competition between all pairs of alternatives is assumed, the NL model structure, as shown in Figure 7-2, categorizes common alternatives in "nests." The underlying assumption of the NL model formulation considers a partially common error term component for within-nest alternatives. That is, the error term component of nested alternatives can be divided into two portions: within-nest and alternative-specific error terms. Therefore, the utility function of the MNL model, as shown in equation 7 – 1, can be reformulated to adopt this nesting structure. For n alternatives and j nests, m represents the number of nested alternatives in nest j where $m<n$, the utility equation of alternative m is as follows:

$$U_m \mid j = V_m + \varepsilon_m + \varepsilon_j = (\beta \cdot x)_m + \varepsilon_m + \varepsilon_j, \qquad [7-5]$$

The choice probabilities can be interpreted based on the assumed multi-level decision structure. That is, marginal and conditional choice probabilities of the upper and lower nests, respectively, can be identified. It is important to note that the choice probabilities do not imply that individuals make their decisions based on the assumed nesting order. The assumption of such an artificial decision tree is purely analytical and can only be verified empirically, as explained in Section 7.5.

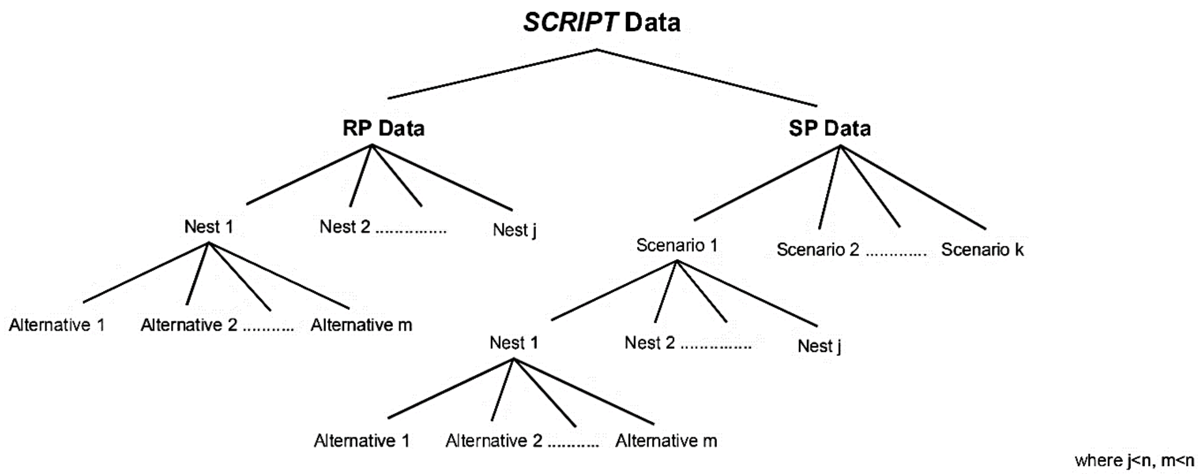


Figure 7-2 The Artificial Nested Structure of the Joint RP/SP NL Model

Assuming a two-level decision tree as shown in Figure 7-2, the conditional choice probabilities of the lower-level nests are calculated as follows:

$$\Pr(m \mid j) = \frac{\exp(\mu_j \cdot V_m)}{\sum_{m'=1}^{M} \exp(\mu_j \cdot V_{m'})}, \qquad [7-6]$$

where $\Pr(m|\ j)$ is the probability of choosing alternative m within nest j, M indicates the total number of alternatives within nest j, and μ_j is the scale parameter of nest j.

The marginal choice probabilities of the upper nests are calculated as follows:

$$\Pr(j) = \frac{\exp(\mu \cdot V_j)}{\sum_{j'=1}^{J} \exp(\mu \cdot V_{j'} + [\mu / \mu_j] \cdot I_j)}, \qquad [7-7]$$

where $\Pr(j)$ is the probability of choosing alternative from nest j, V_j is a function of common attributes within nest j (if any), J indicates the total number of nests j, μ is the scale parameter, and I is the logsum variable of nest j.

The logsum represents the expected utility of the within-nest alternatives. This value is calculated as the log of the sum of exponents of the nested utilities, as shown in equation 7 − 8 (Koppelman and Bhat, 2006):

$$I_j = \log\left\{\sum_{m'=1}^{M} \exp(\mu_j \cdot V_{m'})\right\} \qquad [7-8]$$

Finally, the unconditional probabilities of any mode alternative within individuals' choice sets can be obtained as follows:

$$\Pr(m) = \Pr(m \mid j) \cdot \Pr(j) \qquad [7-9]$$

To account for the differences across the RP and SP datasets, an artificial tree structure, as shown in Figure 7-2, is assumed. This difference is captured through a scale parameter that is estimated for the SP data relative to a normalized (fixed to 1) scale parameter for the RP data. The scale parameter is parameterized as an exponential function of the respondents' attributes:

$$\mu = \exp(\alpha \cdot \tau), \qquad [7-10]$$

where τ refers to attributes that can explain scale variation (e.g. socio-demographic or land use attributes), and α refers to their corresponding coefficients. In addition, the nest-specific (upper level) scale parameter μ_j is identified similarly using a constant-only τ to capture the difference in within-nest variation across the nests.

The value of the nest-specific scale parameter μ_j is constrained between the dataset-specific scale parameter μ and positive infinity. This indicates a non-zero correlation between within-nest alternatives. A higher value of μ_j relative to μ indicates lower variation across nest alternatives (i.e. higher substitution between within-nest alternatives). Such constraints are adopted in the models' estimation routines to test different nesting structures empirically, and thereby, they are either accepted or rejected.

Similarly, for a sample of Z individuals with multiple responses (an RP choice and six SP choices), the joint likelihood function can be expressed as follows:

$$L(\beta) = \prod_{z}^{Z} \prod_{d}^{D} \prod_{m}^{C} (\Pr(m)_{dz})^{y_{mdz}}, \qquad [7-11]$$

where β is a vector of variable coefficients that maximizes the likelihood function $L(\beta)$, d represents either the RP data or one of the six choice situations of the SP data, and y_{mdz} =1 if person z chooses alternative m from a variable choice set C in dataset d and zero otherwise. SP choice scenarios where low confidence levels were reported by the respondents are not considered in the estimation process.

7.3.2.2 Parameterized Logit Captivity Model

The parameterized logit captivity (PLC) model structure presented below is a modified version of the dogit model structure originally proposed by Gaundry and Dagenais (1979). As a latent class model, according to an inclusion probability, individuals are either classified as rational or captive users. Rational users are assumed to make a choice according to the level-of-service attributes, while in contrast, captive users are assumed to rely on a specific alternative under all circumstances, irrespective of the level-of-service attributes of other options. The artificial nested structure of the joint RP/SP PLC model structure is presented in Figure 7-3.

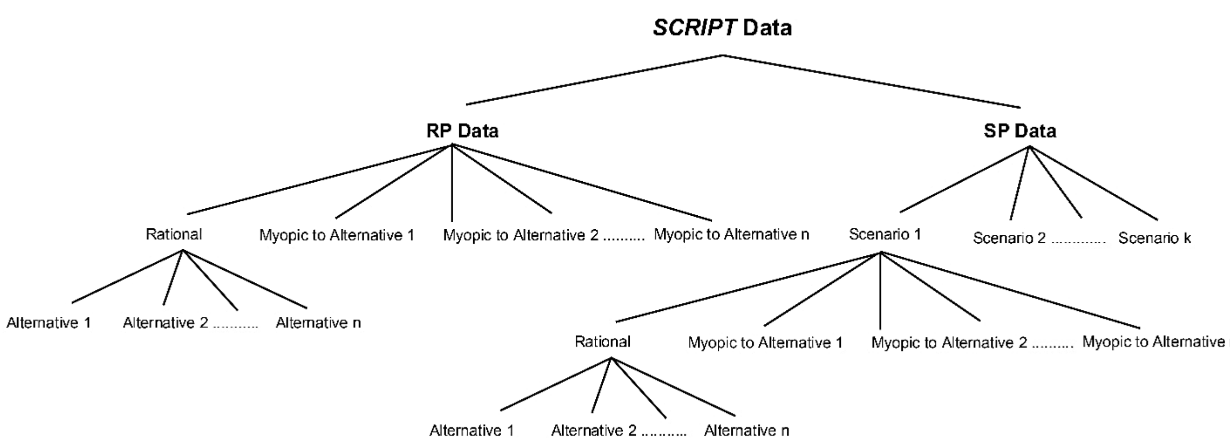


Figure 7-3 The Artificial Nested Structure of Joint RP/SP PLC Models

Based on this structure, Pr(n), the probability of selecting mode n can be defined as follows:

$$\Pr(n) = P_{captive-n} + P_{n|rational} \cdot P_{rational}, \qquad [7-12]$$

$$where\ P_{rational} = (1 - \sum_{n'=1}^{N} P_{captive-n'}), \qquad [7-13]$$

Here, $P_{captive-n}$ refers to the probability of being captive to mode n, and $P_{n|rational}$ refers to the probability of selecting mode n given that the decision-maker is choosing from his/her full choice set. Therefore, the dogit model can be formulated as follows:

$$\Pr(n) = \frac{c_n}{1 + \sum_{n'=1}^{N} c_{n'}} + \frac{1}{1 + \sum_{n'=1}^{N} c_{n'}} \cdot P_{n|rational}, \qquad [7-14]$$

where C_n is the captivity odds parameter of alternative n. Further, by parameterizing the captivity odds function, the PLC model formulation can be expressed as follows:

$$\Pr(n) = \frac{\exp(\mu \cdot D_n)}{1 + \sum_{n'=1}^{N} \exp(\mu \cdot D_{n'})} + \frac{1}{1 + \sum_{n'=1}^{N} \exp(\mu \cdot D_{n'})} \cdot P_{n|rational,} \qquad [7-15]$$

where D_n is a linear-in-parameter function of alternative specific or decision-maker socioeconomic variables s and their corresponding parameter coefficients λ:

$$D_n = (\lambda \cdot s)_n, \qquad [7-16]$$

Assuming that individuals are rational in making their decisions by choosing the alternative with the highest utility value among a set of feasible alternatives, the $P_{n|rational}$ term follows the same multinomial logit model structure, which can be expressed as follows:

$$P_{n|rational} = \frac{\exp(\mu \cdot V_n)}{\sum_{n'=1}^{N} \exp(\mu \cdot V_{n'})}, \qquad [7-17]$$

where μ is the scale parameter, and V is the systematic utility component, which is a linear-in-parameter function of the observed variables χ and their corresponding coefficients β:

$$V_n = (\beta \cdot x)_n, \qquad [7-18]$$

It should be noted that the use of socio-demographic or land use variables to parameterize the captivity odds function (which is the dependent portion of the model) stems from the probability of dependence on a given mode is based more on external or internal values than on comparisons of level-of-service values. However, in the rational portion of the model, the decision-maker is assumed to be making a deliberate trade-off between the different alternatives based on their characteristics (i.e. level-of-service attributes).

To account for the differences across the RP and SP datasets, an artificial tree structure, as shown in Figure 7-3, is assumed. This difference is captured through a scale parameter that is estimated for the SP data relative to a normalized (fixed to 1) scale parameter for the RP data. Typical applications of the PLC or dogit models have an assumed constant scale parameter. However, to account for heteroscedasticity across the sample population, the scale parameter can be parameterized as a function of the socioeconomic variables of each decision-maker. Therefore, the scale parameter can be expressed as follows:

$$\mu = \exp(\alpha \cdot \tau), \qquad [7-19]$$

where τ refers to attributes that can explain scale variation, and α refers to their corresponding coefficients. As such, the PLC model formulation can be expressed as follows:

$$\Pr(n) = \frac{\exp(\mu \cdot D_n)}{1 + \sum_{n'=1}^{N} \exp(\mu \cdot D_{n'})} + \frac{1}{1 + \sum_{n'=1}^{N} \exp(\mu \cdot D_{n'})} \cdot \frac{\exp(\mu \cdot V_n)}{\sum_{n'=1}^{N} \exp(\mu \cdot V_{n'})} \qquad [7-20]$$

Similarly, for a sample of *Z* individuals with multiple responses (an RP choice and six SP choices), the joint likelihood function can be expressed as follows:

$$L(\beta) = \prod_{z}^{Z} \prod_{d}^{D} \prod_{n}^{C} (\Pr(n)_{dz})^{y_{ndz}}, \qquad [7-21]$$

where β is a vector of variable coefficients that maximizes the likelihood function $L(\beta)$, *d* represents either the RP data or one of the six choice situations of the SP data, and y_{ndz}=1 if person *z* chooses alternative *n* from a variable choice set *C* in dataset *d* and zero otherwise. SP choice scenarios where low confidence levels were reported by the respondents are not considered in the estimation process.

7.4 Empirical MNL Models

The full sample size used for model estimation and validation is 704 individuals. As the names imply, the RP-only and SP-only models are restricted models that include only RP and SP data, respectively, while the joint RP/SP model considers both RP and SP information. The variables used in model specification are presented in Table 7-1 along with the corresponding polices tied to each variable (if any).

Table 7-1 Definitions of Variables

Variable Name	Description	Corresponding Policy
Travel Cost/Fare	Travel cost including gas and parking cost at the work location or transit fares	Increase driving cost/reduce transit fares
P&R Cost at TTC Stations	Parking cost at TTC park-and-ride locations per day	Increase parking cost at TTC park-and-ride stations
P&R Cost at GO Stations	Parking cost at GO Transit park-and-ride locations per day	Introduce pay parking at GO Transit park-and-ride stations
In-vehicle Travel Time (no Wi-Fi)	In-vehicle travel time if Wi-Fi is not available on GO Transit vehicles	Reduce transit travel time
In-vehicle Travel Time (Wi-Fi)	In-vehicle travel time if Wi-Fi is available on GO Transit vehicles	Provide new transit features (Wi-Fi on board)
Access and Egress Times	Sum of access and egress travel times	Improve transit accessibility
Wait and Transfer Times	Sum of waiting and transfer travel times	Reduce waiting and transfer times
Next Local Transit Vehicle Information Provision	1 if the next transit vehicles' arrival time information is available; 0 otherwise	Provide accurate transit information
Need for a 2nd Transfer	1 if the individual needs to make more than one transfer between GO Transit and other travel modes; 0 otherwise	Reduce number of intermodal transfers
Number of Vehicles per Household	Number of vehicles per household	N/A
Transit Pass Possession	1 if the individual owns a transit pass; 0 otherwise	N/A
Trip O/D: City of Toronto	1 if the trip origin/destination is from/to the City of Toronto; 0 otherwise	N/A
Age	Individuals age	N/A
Gender (Male)	1 if male; 0 otherwise	N/A

These variables are carefully selected in accordance with the policies under investigation (see Figure 5-4). For instance, the parking cost at park-and-ride stations variable is defined as a separate cost component (i.e. not aggregated with the total travel cost). Therefore, the developed models are sensitive to corresponding policies, such as the introduction of pay parking at park-and-ride stations. Similarly, the out-of-vehicle travel time components are divided into access and egress times and wait and transfer times. As such, the effectiveness of policies that aim at improving different level-of-service attributes can be appropriately quantified. The modelling framework considers nine alternatives; it explicitly distinguishes between auto driver and auto passenger modes, regional and local transit modes, and different access modes. Each travel mode has its own level-of-service characteristics and therefore different target customers. That is, various policy initiatives that target specific travel modes such as high-occupancy vehicle (HOV) policies for auto passenger modes or the introduction of transit mode-specific features can be analyzed.

The final empirical model specifications are presented in Table 7-2, Table 7-3, and Table 7-4. In addition to the models presented below, several model specifications were tested to find the best specifications while providing the highest explanatory power and statistical significance. In the following sub-sections, results of the three models are presented. The results for the joint RP/SP model are discussed in greater detail, and analogous conclusions can be drawn for the two other models.

The empirical models were estimated using codes written in GAUSS® using the MAXLIK component for maximum likelihood estimation (Aptech-Systems, 2012).

7.4.1 RP-only MNL Model

Table 7-2 presents the parameter estimates of the RP-only MNL mode choice model using the RP data on individuals' actual mode choices. Seventeen parameters are estimated using the full sample size. The rho-squared (with respect to a constant-only model), a measure of the model's goodness of fit, indicates how much of an improvement the estimated model offers over a naïve model that assumes that all parameters are zero while allowing for constants. The reported rho-squared value is 0.31, which indicates an acceptable goodness of fit. According to the literature on discrete choice models, a rho-squared value in the range of 0.15 to 0.4 is considered a good fit. In addition, the log likelihood ratio test shows a test statistics value of 336, which indicates that the reported models fit the data significantly better than the constant-only model.

All the reported parameters are estimated with the expected signs and found to be statistically significant (with t-statistics higher than 1.96) at the 95% confidence interval except for the parking cost at local transit park-and-ride locations, caused by the low number of park-and-ride users within the RP sample. However, it was retained in the model because it was estimated with the expected sign and to ensure specification consistency with the SP models. The relative values of the estimated parameters indicate that out-of-vehicle travel times are perceived to be 1.2 to 2 times higher than in-vehicle travel time.

These findings show consistency with corresponding mode choice models that verify the validity of the survey design, sampling procedure, and data quality. As such, the RP-only model sets the ground as the first step toward developing policy-sensitive behavioural models. However, as explained in Chapter 2, RP-only models are incapable of accurately forecasting individual choices in response to scenarios that do not currently exist (such as new transportation policies targeted at integrating services and fares among contiguous local transit agencies, local and regional transit systems, or between automobiles and local or regional transit systems). In other words, predicting/forecasting users' behaviour caused by changes in level-of-service attributes beyond the ranges observed in the RP data, the addition of new service features (e.g. on-board Wi-Fi), or the introduction of new travel modes are beyond the scope of traditional RP-only models. Therefore, policy-sensitive SP models are developed to capture the associated changes in travel demand with respect to changes in the level-of-service attributes, as shown in the following sections.

Table 7-2 RP-only MNL Model

MNL Logit Model		RP-only Model	
Log Likelihood of Full Model		-380.225	
Log Likelihood of Constant-only Model		-548.111	
Rho-squared Value		0.306	
Number of Observations		704	
Variable	**Mode**	**Parameter**	**t-Statistics**
Systematic Utility Function:			
Alternative Specific Constant	Auto driver	2.6051	6.839
Alternative Specific Constant	Auto passenger	0 (fixed)	0 (fixed)
Alternative Specific Constant	Local transit with walk access	-0.4003	-0.597
Alternative Specific Constant	Local transit with auto driver access (TTC park-and-ride)	-4.3841	-1.297
Alternative Specific Constant	Local transit with auto passenger access (TTC kiss-and-ride)	-2.6383	-1.456
Alternative Specific Constant	Regional transit with walk access	-0.9951	-1.272
Alternative Specific Constant	Regional transit with auto driver access (GO park-and-ride)	-0.9548	-1.299
Alternative Specific Constant	Regional transit with auto passenger access (GO kiss-and-ride)	-1.4855	-2.166
Alternative Specific Constant	Regional transit with local transit access	-0.6875	-0.868
Travel Cost/Fare	All modes	-0.0838	-4.519
P&R Cost at TTC Stations	Local transit with auto driver access (TTC park-and-ride)	-0.0458	-0.066
In-vehicle Travel Time	All modes	-0.022	-2.977
Access and Egress Times	All transit alternatives	-0.044	-2.513
Wait and Transfer Times	All transit alternatives	-0.0259	-1.974
Need for a 2nd Transfer	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	-0.8691	-2.475
Number of Vehicles per Household	Auto driver, local transit with auto driver access (TTC park-and-ride), and regional transit with auto driver access (GO park-and-ride)	0.4598	2.359
Transit Pass Possession	All transit alternatives	3.1088	9.651
Trip O/D: City of Toronto	All transit alternatives	1.3774	2.738

7.4.2 SP-only MNL Model

Table 7-3 presents the results of the SP-only MNL mode choice model. The estimation routine takes into consideration the repeated observations by each individual across the six SP choice situations. The reported model is consistent with the RP-only model specification except for the new features that were added to the SP experiment, including the introduction of parking cost at park-and-ride GO Transit stations, the availability of Wi-Fi on GO Transit vehicles, and the provision of real-time information on local transit vehicles' arrival times. All parameters are estimated with the expected signs and relative values and are statistically significant at the 95% confidence interval except for the provision of information of local transit vehicles' arrival times, which is statistically significant at the 90% confidence interval. The reported rho-squared value is 0.112, and the log likelihood ratio test shows a test statistics value of 915, which indicates that the reported models fit the data significantly better than the constant-only model.

Table 7-3 SP-only MNL Model

MNL Logit Model		SP-only Model	
Log Likelihood of Full Model		-3,622.91	
Log Likelihood of Constant-only Model		-4,080.26	
Rho-squared Value		0.112	
Number of Observations		560	
Variable	**Mode**	**Parameter**	**t-Statistics**
Systematic Utility Function:			
Alternative Specific Constant	Auto driver	0.5999	5.808
Alternative Specific Constant	Auto passenger	0 (fixed)	0 (fixed)
Alternative Specific Constant	Local transit with walk access	-0.6315	-2.472
Alternative Specific Constant	Local transit with auto driver access (TTC park-and-ride)	-0.0202	-0.02
Alternative Specific Constant	Local transit with auto passenger access (TTC kiss-and-ride)	-1.008	-3.466
Alternative Specific Constant	Regional transit with walk access	-0.6263	-2.142
Alternative Specific Constant	Regional transit with auto driver access (GO park-and-ride)	-0.751	-2.692
Alternative Specific Constant	Regional transit with auto passenger access (GO kiss-and-ride)	-0.7095	-3.034
Alternative Specific Constant	Regional transit with local transit access	-0.8398	-2.64
Travel Cost/Fare	All modes	-0.0493	-7.081
P&R Cost at TTC Stations	Local transit with auto driver access (TTC park-and-ride)	-0.4301	-2.305
P&R Cost at GO Stations	Regional transit with auto driver access (GO park-and-ride)	-0.0699	-2.244
In-vehicle Travel Time (no Wi-Fi)	All modes	-0.0284	-9.775
In-vehicle Travel Time (Wi-Fi)	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	-0.0234	-6.773
Access and Egress Times	All transit alternatives	-0.0679	-9.066
Wait and Transfer Times	All transit alternatives	-0.0253	-3.866
Next Local Transit Vehicle Information Provision	Local transit with walk access, local transit with driving access (TTC park-and-ride), local transit with passenger access (TTC kiss-and-ride), and regional transit with local transit access	0.1963	1.698
Need for a 2nd Transfer	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	-0.4538	-3.197

Number of Vehicles per Household	Auto driver, local transit with auto driver access (TTC park-and-ride), and regional transit with auto driver access (GO park-and-ride)	0.2707	5.3
Transit Pass Possession	All transit alternatives	1.6337	14.416
Trip O/D: City of Toronto	All transit alternatives	0.4124	2.644

7.4.3 Joint RP/SP MNL Model

The jointly estimated RP/SP MNL mode choice model results are shown in Table 7-4. The model is estimated using both RP and SP mode choice data. Various model specifications are tested and compared to one another until the reported final model specification with the highest explanatory power is reached. The estimation routine takes into consideration the repeated observations by each individual across the RP data and the six SP choice situations as shown in Equation 7-4. As mentioned before, the survey respondents were asked to provide information about their level of confidence in making their SP choices. Choice scenarios where low confidence levels were reported by the respondents were not considered in the estimation process.

Thirty-one parameters were estimated with the expected signs and relative values and are statistically significant at the 95% confidence interval except for the provision of information on local transit vehicles' arrival times, which is statistically significant at the 90% confidence interval. The reported rho-squared value is 0.155, which is better than the reported value for the SP-only model. Clearly, the use of the combined RP/SP data enhanced the goodness of fit and explanatory power of the joint model. Generally, RP models fit the data better because they represent actual behaviour and the SP data encompasses an induced variation because of the nature of the hypothetical choice experiment in which the survey respondents were making their decisions while some attributes of the current transportation systems' elements were altered. That explains why SP and RP/SP models were estimated with lower rho-squared value compared to the RP-only model. However, as explained earlier, SP data allow for testing new scenarios that have never been experienced before. Therefore, using RP/SP data exploit the strengths and mitigate the weaknesses of each type by complementing one another. The log likelihood ratio test shows a test statistics value of 1,279, which indicates that the reported models fit the data significantly better than the constant-only model.

The estimated parameters of the joint RP/SP model are classified into three groups: parameters exclusively estimated by either the RP or the SP datasets, parameters estimated with different coefficients in each dataset, and parameters estimated with the same coefficient (before taking the scale parameter effect into consideration) in the RP and the SP datasets. Typically, alternative specific constants (ASC) dataset-specific coefficients are estimated, while variables that belong to one dataset and scale parameter factors are uniquely estimated by one dataset (usually the SP data)

(Brownstone et al., 2000). Other level-of-service attributes that appear in both datasets as well as socioeconomic attributes are estimated with the same coefficients. As such, different RP/SP coefficients are estimated for ASC; exclusive SP coefficients are estimated for the introduction of parking cost at park-and-ride GO Transit stations, the availability of Wi-Fi on GO Transit vehicles and scale parameter factors; and pooled RP/SP coefficients (expect for the scale parameter effect) are estimated for all other variables, as shown in Table 7-4. Since the specification of the utility functions of the choice alternatives may vary across the RP and the SP datasets, the associated data-specific ASC may have different values or signs. In fact, the ASC captures the average impact of the unobserved factors in the model specification on the utility function of each alternative. Therefore, the relative values and/or signs of one data-specific ASC compared to the corresponding estimates of the other dataset's ASC do not have a particular implication on the interpretation of the estimated model's results.

Assuming a unit scale parameter for the RP data, the SP scale parameter is relatively estimated as a parameterized exponential function of a constant, gender, and a logarithm of individuals' age. The estimated parameters of the SP scale factor are statistically significant, verifying the assumed tree structure of the two datasets. The SP scale parameter is estimated to be less than 1 (for all individuals in the dataset). In other words, the SP scale parameter is lower than the RP scale parameter, which indicates that the variance within the SP data is higher than in the RP data. This typical finding explains that the SP data encompasses an induced variation according to the introduced changes in some attributes of current transportation systems' elements. As explained earlier, the SP scale is parameterized using socio-demographic attributes to capture the heteroscedasticity across the sample. For instance, the parameter estimates of the SP scale factor suggest that young males have a lower scale factor than old females which indicates that the latter group have lower variance across their SP choices, and vice versa.

It is worth mentioning that former tested specifications (prior to the final reported model) considered access, egress, and transfer times as one variable and waiting time as a separate variable. Such specifications showed that waiting time is statistically insignificant, while other travel time components are statistically significant. This may result from the fact that cross-regional commuters who use regional transit or local transit services (other than TTC), which are relatively low-frequency services, plan their trips according to service schedules to avoid long waiting times. In addition, the trip planner tool that was used to develop the SP choice scenarios optimized respondents' departure times to minimize waiting times at transit stations. Hence, another model structure was tested to capture the effect of all travel time components on mode choice. This model structure considers two major travel time components: in-vehicle and out-of-vehicle travel times. The latter is divided into two sub-components: access and egress times and transfer and wait times. In the context of cross-regional trips, it seems that the presented travel time component structure is more suitable to represent the data.

The effects of the introduction of paid parking at GO Transit park-and-ride stations, the provision of information on local transit vehicles' arrival time, and the introduction of Wi-Fi service on GO Transit vehicles on the probability of commuting mode choice are captured through the SP environment of the model. Different model specifications have not shown a significant difference between the monthly and the daily parking schemes at GO Transit park-and-ride stations. Therefore, equivalent daily parking rates were used for monthly parking schemes. The model results showed that parking cost at GO Transit park-and-ride stations is perceived to be approximately 1.4 times higher than the total travel cost.

Initially, the provision of information on local transit vehicles' arrival time variable was estimated with an exclusive SP coefficient since it represents a new feature that is assumed not to be currently available and therefore cannot be captured through the RP data. However, this feature is currently provided by a few transit service agencies within the study area. Thus, to account for the current un/availability of information on local transit vehicles' arrival time, this parameter is estimated as a pooled coefficient from the RP and the SP datasets. The model results showed that providing individuals with real-time information on transit vehicles' arrival time would increase the likelihood that they will choose local transit as their travel mode.

Table 7-4 Joint RP/SP MNL Model

MNL Logit Model – Joint Estimation		RP/SP Model			
Log Likelihood of Full Model		-3,477.48			
Log Likelihood of Constant-only Model		-4,116.38			
Rho-squared Value		0.155			
Number of Observations		560			
Variable	**Mode**	**Parameter**	**t-Statistics**	**Parameter**	**t-Statistics**
Systematic Utility Function:		**RP Coefficients**		**SP Coefficients**	
Alternative Specific Constant	Auto driver	2.8022	10.548	0.8137	5.269
Alternative Specific Constant	Auto passenger	0 (fixed)	0 (fixed)	0 (fixed)	0 (fixed)
Alternative Specific Constant	Local transit with walk access	1.7581	3.888	-0.7417	-2.156
Alternative Specific Constant	Local transit with auto driver access (TTC park-and-ride)	0.0999	0.082	-0.6783	-0.686
Alternative Specific Constant	Local transit with auto passenger access (TTC kiss-and-ride)	-0.648	-1.033	-1.3036	-3.329
Alternative Specific Constant	Regional transit with walk access	1.3622	2.445	-0.7658	-1.961
Alternative Specific Constant	Regional transit with auto driver access (GO park-and-ride)	0.8796	2.025	-0.8114	-2.282
Alternative Specific Constant	Regional transit with auto passenger access (GO kiss-and-ride)	0.385	0.839	-0.8418	-2.714
Alternative Specific Constant	Regional transit with local transit access	1.0062	1.575	-0.7131	-1.688
Travel Cost/Fare	All modes	-0.0716	-6.853	-0.0716	-6.853
P&R Cost at TTC Stations	Local transit with auto driver access (TTC park-and-ride)	-0.4409	-2.225	-0.4409	-2.225
P&R Cost at GO Stations	Regional transit with auto driver access (GO park-and-ride)	0 (fixed)	0 (fixed)	-0.0983	-2.31
In-vehicle Travel Time (no Wi-Fi)	All modes	-0.0404	-8.363	-0.0404	-8.363
In-vehicle Travel Time (Wi-Fi)	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	0 (fixed)	0 (fixed)	-0.0338	-6.518

Access and Egress Times	All transit alternatives	-0.0952	-8.187	-0.0952	-8.187
Wait and Transfer Times	All transit alternatives	-0.0484	-5.587	-0.0484	-5.587
Next Local Transit Vehicle Information Provision	Local transit with walk access, local transit with driving access (TTC park-and-ride), local transit with passenger access (TTC kiss-and-ride), and regional transit with local transit access	0.2126	1.798	0.2126	1.798
Need for a 2nd Transfer	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	-0.6191	-3.322	-0.6191	-3.322
Number of Vehicles per Household	Auto driver, local transit with auto driver access (TTC park-and-ride), and regional transit with auto driver access (GO park-and-ride)	0.3796	5.174	0.3796	5.174
Transit Pass Possession	All transit alternatives	2.3783	9.305	2.3783	9.305
Trip O/D: City of Toronto	All transit alternatives	0.4679	2.272	0.4679	2.272
Exponential Function of Scale Parameter:					
Constant	SP scale factor	0 (fixed)	0 (fixed)	-1.2816	-3.584
Ln (Age)	SP scale factor	0 (fixed)	0 (fixed)	0.3064	3.346
Gender	SP scale factor	0 (fixed)	0 (fixed)	-0.1786	-3.677

Similarly, introducing Wi-Fi service for regional transit users is expected to increase the modal share of GO Transit. Several model specifications were tested to quantify the effect of the availability of Wi-Fi on commuters' mode choice. The results of preliminary models showed that the introduction of Wi-Fi on GO Transit modes is only statistically significant for individuals who spend 40 minutes or more on GO Transit vehicles (i.e. individuals whose in-vehicle travel time is greater than or equal to the average in-vehicle travel time for GO Transit users within the sample). This finding triggered further investigation of in-vehicle travel time interaction with the availability of Wi-Fi on GO Transit vehicles. The final model specification shows two in-vehicle travel time parameters, one if Wi-Fi is available and the other if Wi-Fi is not. The two coefficients are estimated with the expected negative sign and are statistically significant[3]. The estimated coefficient of in-vehicle travel time if Wi-Fi is available has a smaller negative effect on the probability of choosing GO Transit modes than the estimated coefficient of in-vehicle travel time if Wi-Fi is not available. This indicates that individuals are more likely to choose GO Transit if Wi-Fi service is available on GO Transit vehicles.

In terms of personal and household attributes, the model results showed that the number of vehicles per household has a positive impact on the probability of choosing car-dependent modes such as auto driver and park-and-ride. Similarly, transit pass possession increases the probability of the selection of transit as a travel mode. Further, individuals who commute from/to the City of Toronto are more likely to use transit. This is likely driven by the city's unique multimodal transit system, high-density land use, and supportive transit policies.

7.4.3.1 Value of Travel Time Savings

For linear-in-parameter specifications of the utility function (similar to the formulation used in this study), the marginal rate of substitution (i.e. the trade-off) of attributes is estimated as the ratio of the estimated coefficient of each variable to the estimated coefficient of travel cost. As such, the willingness to pay (i.e. the amount that an individual would pay for a particular good) to receive real-time information on local transit vehicles' arrival time is calculated and found to be up to $3

[3] The coefficients are statistically significant from each other. The test statistics value is found to be -2.19 which passes t-statistics of -1.96 at the 95% confidence interval.

per trip. Similarly, the value of travel time savings (VOT), which is the extra cost that a person would be willing to pay to save one hour of travel time (Train, 2009), is calculated for the different travel time components as shown in Table 7-5. The estimated VOT (based on the joint RP/SP model) is \$33.85/hr., which is a reasonable value given the average wage rates within the sample data (more than 50% of the sample's household yearly income is \$100,000 and above). Similarly, the VOT for GO Transit modes if Wi-Fi is available is lower than the VOT if Wi-Fi is not available. The results indicate that individuals are willing to pay up to nine cents per minute to use Wi-Fi on GO Transit vehicles.

Table 7-5 Value of Travel Time Savings

Value of Travel Time Savings (VOT) (\$/hr.)	RP-only Model	SP-only Model	Joint RP/SP Model
In-vehicle Travel Time (No Wi-Fi)	15.75	34.56	33.85
In-vehicle Travel Time (Wi-Fi)	N/A	28.48	28.32

7.4.3.2 Marginal Effects of Travel Cost and Time

To investigate the sensitivity to travel cost and in-vehicle travel time of each mode alternative, the direct point elasticities are estimated based on the joint RP/SP model results, and kernel densities are plotted as shown in Figure 7-4. The kernel density estimator provides an approximate probability density function (PDF) using the data observations. Therefore, a kernel density distribution has the same properties as the probability density function. In other words, the area under the curve is equal to 1, and the value of the density function is proportional to the probability that a data point is approximately equal to its corresponding value. The disaggregate point direct elasticity (of each individual) is estimated for linear-in-parameter utility functions based on equation 7 – 23. For the SP data, the average of six elasticity values corresponding to the six SP scenarios is estimated for each individual. The direct elasticity estimates the effect of a percentage increase/decrease in an observed factor of a choice alternative on the probability that this alternative will be chosen.

$$E_{i_{Xni}} = \beta_X . X_{ni} . (1 - P_{ni}), \qquad [7-22]$$

where E_{iXni} is the direct elasticity of a unit change in the observed factor X_{ni} with a parameter estimate β_X on the probability that individual n will choose alternative i, P_{ni} (Train, 2009).

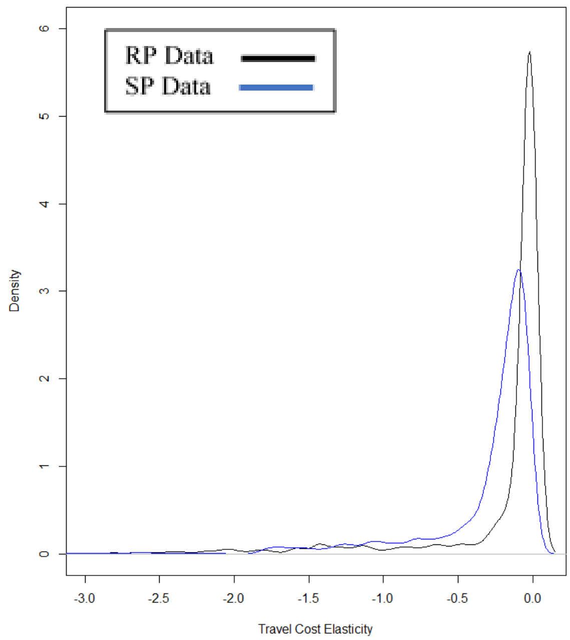
As expected, SP elasticity density charts show more variance than RP elasticity density charts. Nevertheless, both RP and SP elasticity density charts are consistent in terms of their distributions across the two samples. For instance, the elasticity density charts of travel cost and in-vehicle travel time of the auto driver mode, presented in Figure 7-4 (a) and Figure 7-4 (b), show that the majority of the RP sample is inelastic (less sensitive) to changes in travel cost and in-vehicle travel time, which is an expected result given the high modal share of the auto driver mode within the sample as well as the high vehicle ownership level per sampled household. However, the corresponding density charts of the SP sample show considerably flatter distributions, which indicates that individuals are likely to become more elastic to such changes when other mode alternatives are presented with altered level-of-service attributes.

Table 7-6 presents the average direct elasticities of travel cost and in-vehicle travel time for the nine mode alternatives across both RP and SP datasets. For instance, the estimated average direct elasticity of travel cost of the auto driver mode (based on the SP dataset) is -0.32. That is, for a 1% increase in the auto driver travel cost, there would be a decrease of 0.32% in the auto driver mode share. Similarly, the elasticity density charts of in-vehicle travel time for GO Transit modes show the average SP elasticity of in-vehicle travel time when Wi-Fi is available or unavailable. These charts indicate that individuals are less sensitive to changes in-vehicle travel times in the presence of a Wi-Fi service.

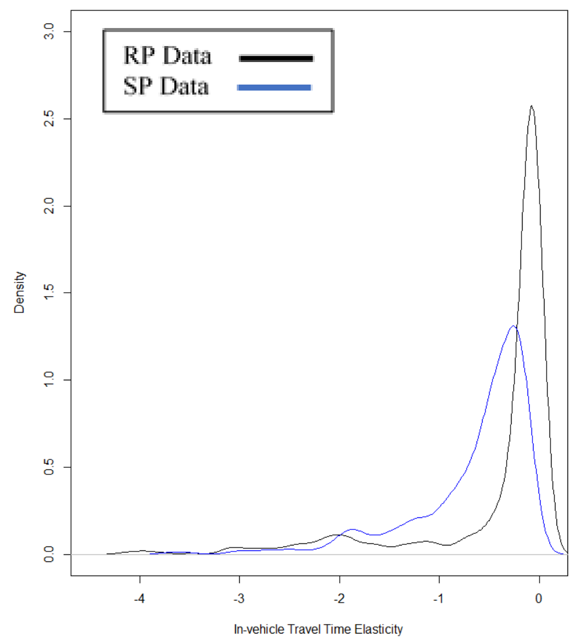
The elasticity density chart of travel cost for the local transit with walk access mode, as presented in Figure 7-4 (e), shows a clear bi-modal distribution, which suggests two groups of users with different mean travel costs. A further investigation showed that such a distribution results from the co-fare structure across the different transit agencies. Group 1 is identified as individuals whose commuting trips originate in or end in the City of Toronto, and thereby, pay two "full" transit fares. Group 2 is identified as individuals who commute from/to other cities, and thereby, pay one "full" transit fare and a "co-fare." As shown in Figure 7-4 (s), the two groups are perfectly identified, indicating that individuals of group 1 are more sensitive to changes in travel cost than those of group 2. The elasticity density chart for parking cost at regional transit park-and-ride stations is plotted as shown in Figure 7-4 (t). Similarly, the density distribution indicates two groups of users/parking schemes: free parking and paid parking. The average elasticity of parking cost at GO Transit park-and-ride stations is -0.28.

Table 7-6 Average Direct Elasticities for Travel Cost and In-vehicle Travel Time

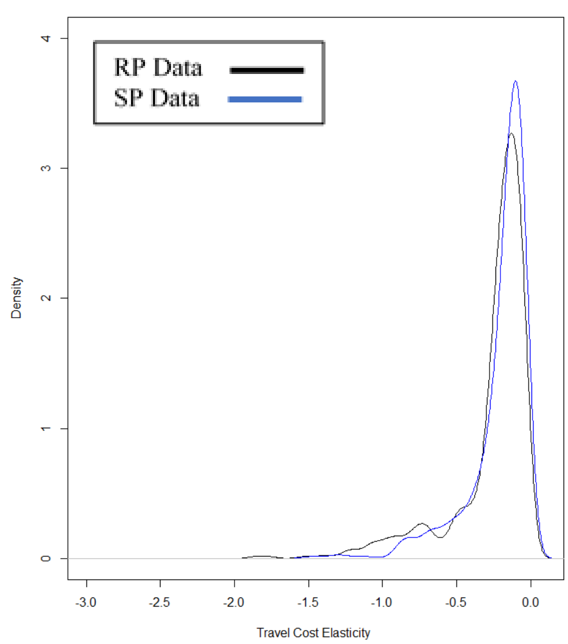
Mode Alternative	RP Data		SP Data	
	Travel Cost	In-vehicle Travel Time	Travel Cost	In-vehicle Travel Time
Auto Driver	-0.22	-0.43	-0.32	-0.66
Auto Passenger	-0.27	-1.54	-0.21	-0.97
Local Transit with Walk Access	-0.30	-1.74	-0.30	-1.44
Local Transit with Auto Driver Access	-0.46	-1.15	-0.46	-0.88
Local Transit with Auto Passenger Access	-0.46	-1.14	-0.32	-0.86
Regional Transit with Walk Access	-0.42	-1.42	-0.34	-1.04
Regional Transit with Auto Driver Access	-0.37	-1.33	-0.34	-0.97
Regional Transit with Auto Passenger Access	-0.39	-1.47	-0.32	-0.99
Regional Transit with Local Transit Access	-0.49	-1.64	-0.36	-1.20



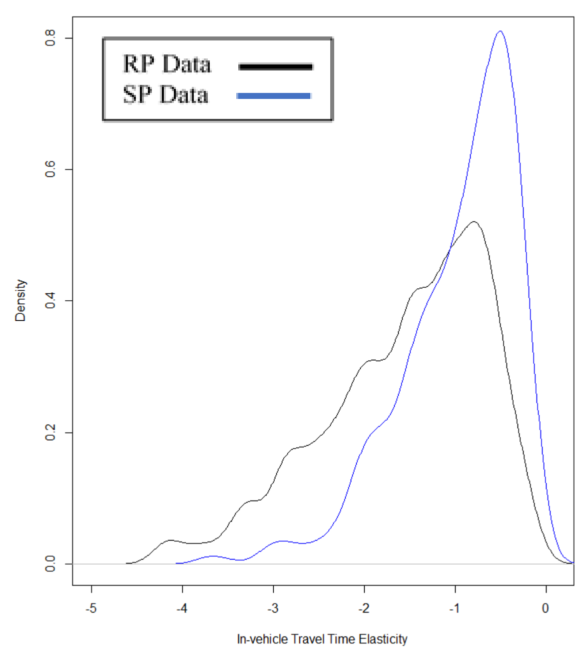
(a) Marginal Effects of Travel Cost for Auto Driver Mode



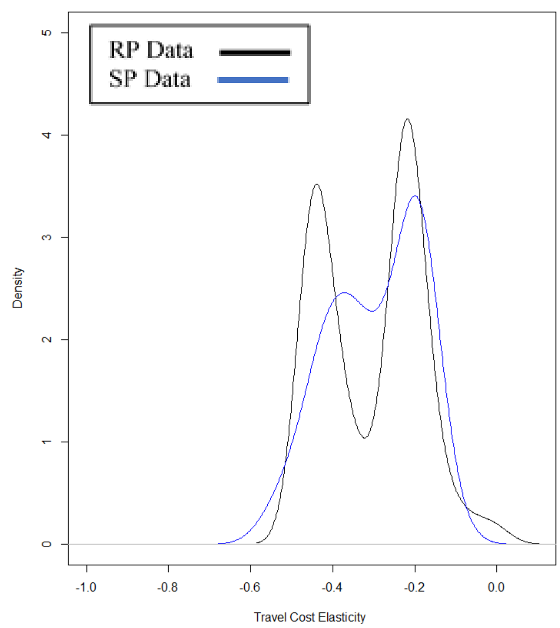
(b) Marginal Effects of In-vehicle Travel Time for Auto Driver Mode



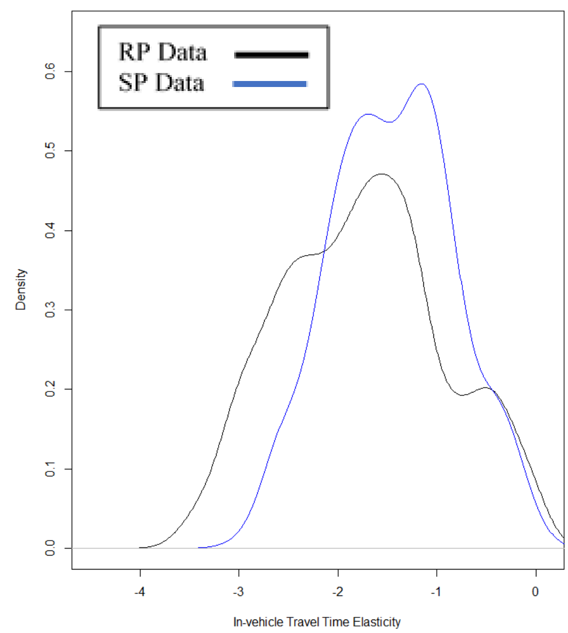
(c) Marginal Effects of Travel Cost for Auto Passenger Mode



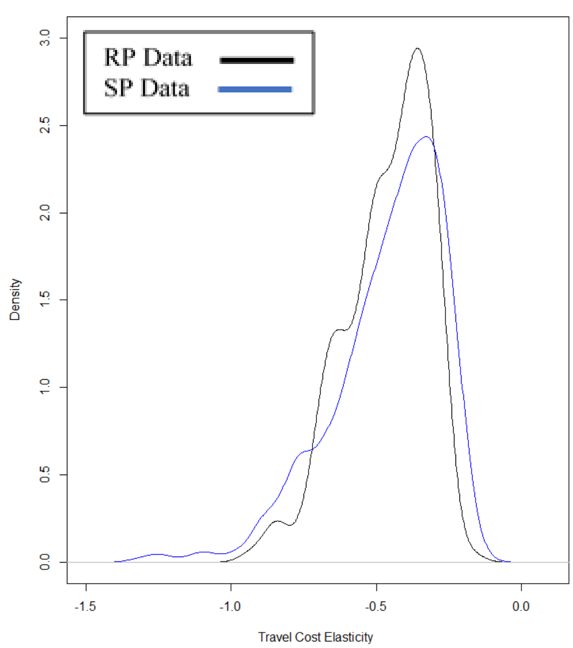
(d) Marginal Effects of In-vehicle Travel Time for Auto Passenger Mode



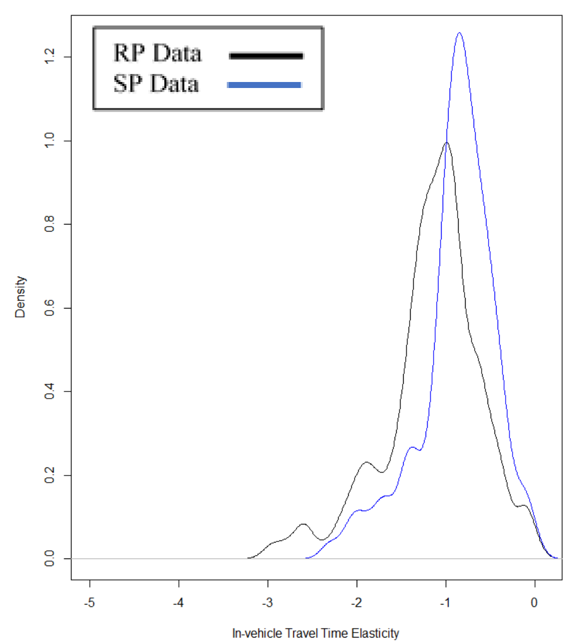
(e) Marginal Effects of Travel Cost for Local Transit with Walk Access Mode



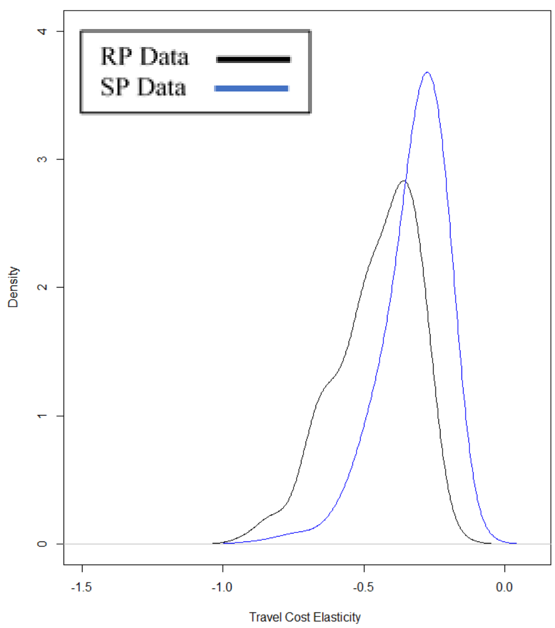
(f) Marginal Effects of In-vehicle Travel Time for Local Transit with Walk Access Mode



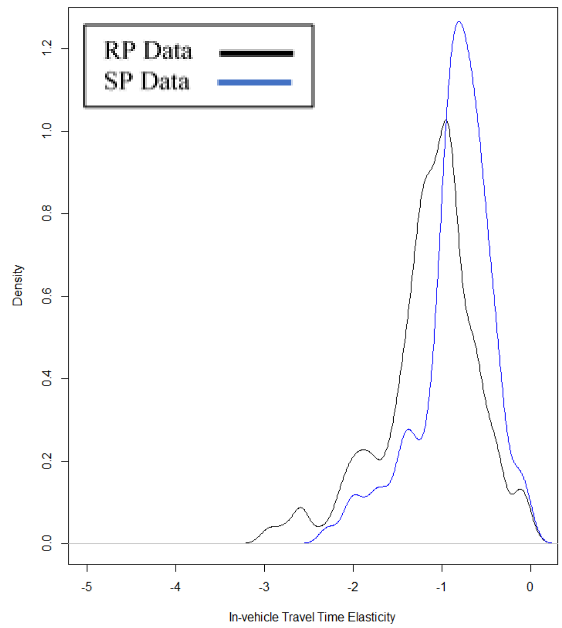
(g) Marginal Effects of Travel Cost for Local Transit with Auto Driver Access Mode



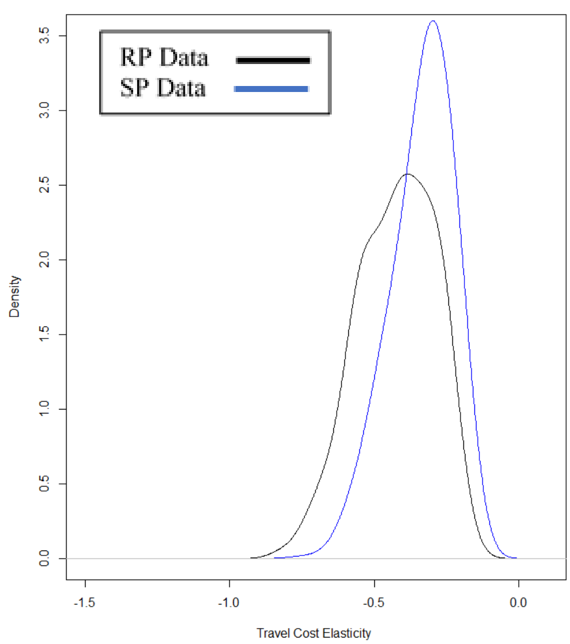
(h) Marginal Effects of In-vehicle Travel Time for Local Transit with Auto Driver Access Mode



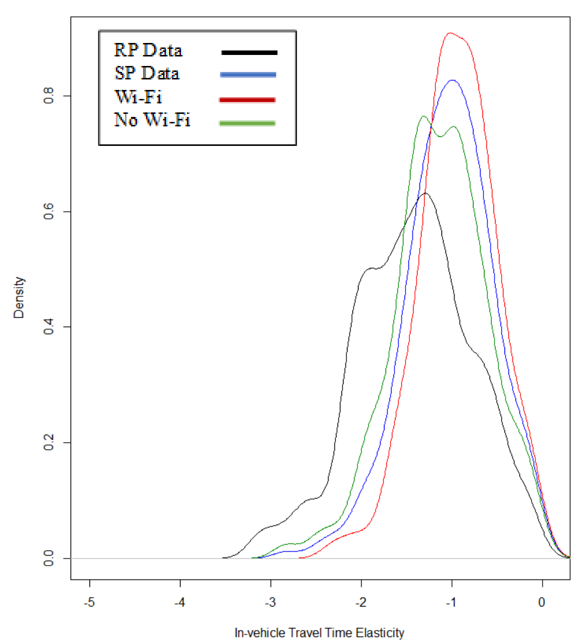
(i) **Marginal Effects of Travel Cost for Local Transit with Auto Passenger Access Mode**



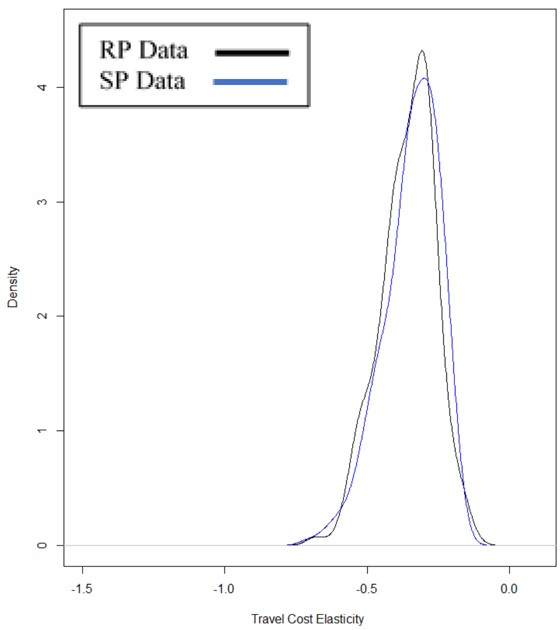
(j) **Marginal Effects of In-vehicle Travel Time for Local Transit with Auto Passenger Access Mode**



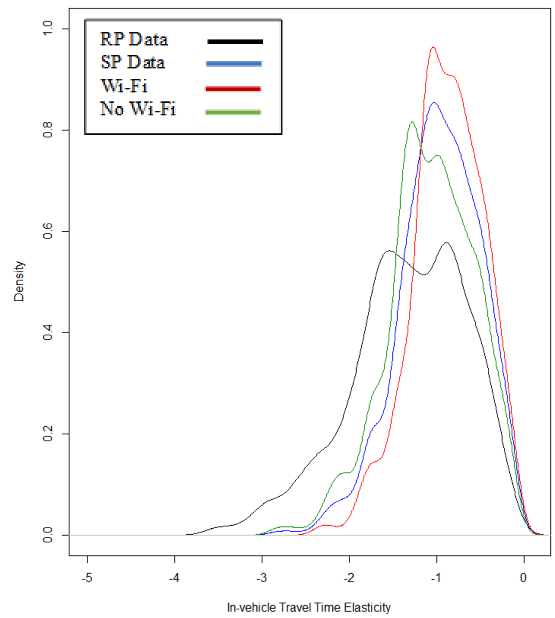
(k) **Marginal Effects of Travel Cost for Regional Transit with Walk Access Mode**



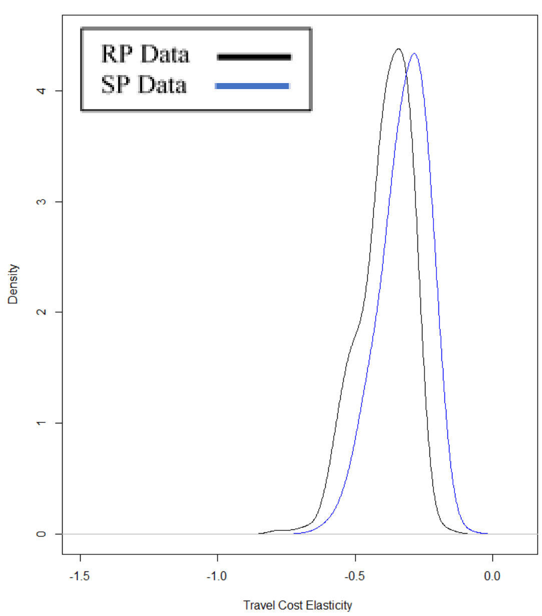
(l) **Marginal Effects of In-vehicle Travel Time for Regional Transit with Walk Access Mode**



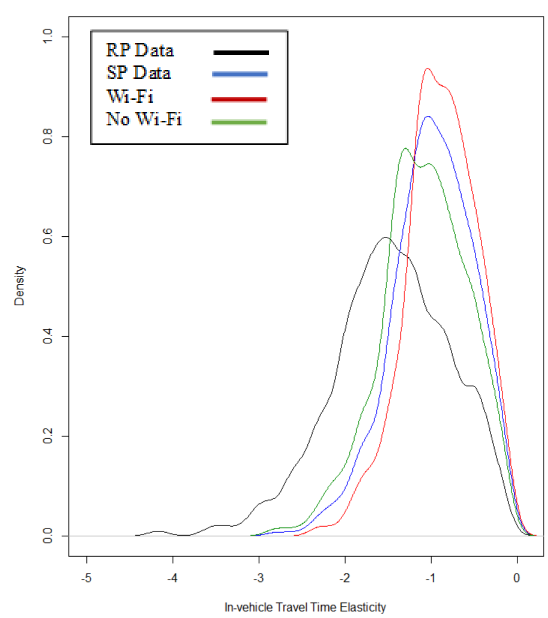
(m) Marginal Effects of Travel Cost for Regional Transit with Auto Driver Access Mode



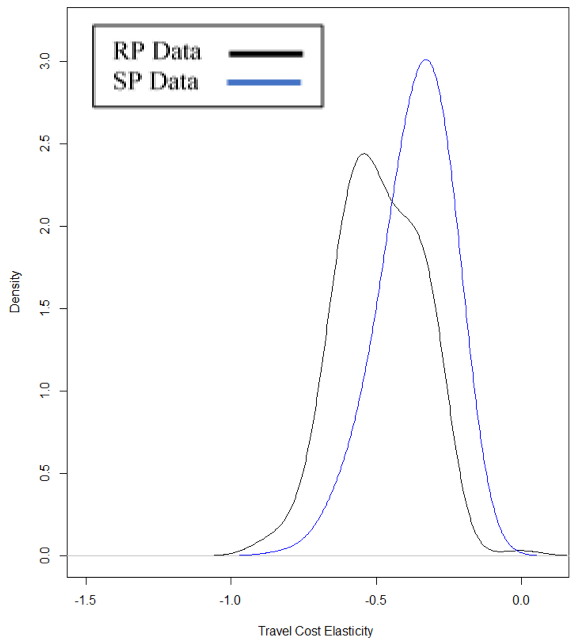
(n) Marginal Effects of In-vehicle Travel Time for Regional Transit with Auto Driver Access Mode



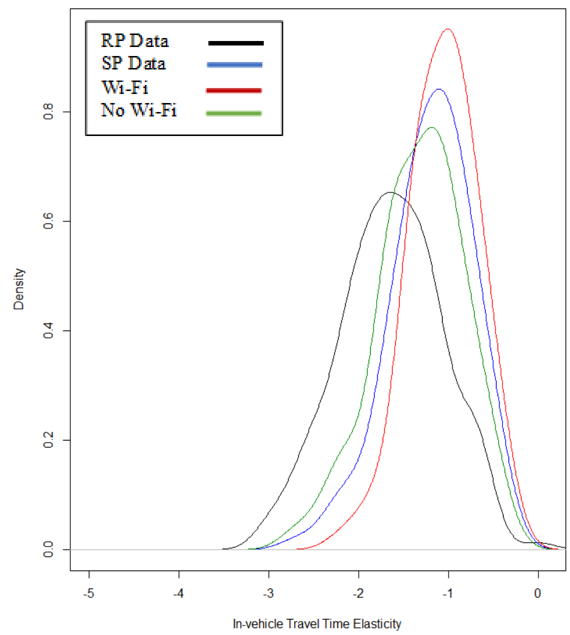
(o) Marginal Effects of Travel Cost for Regional Transit with Auto Passenger Access Mode



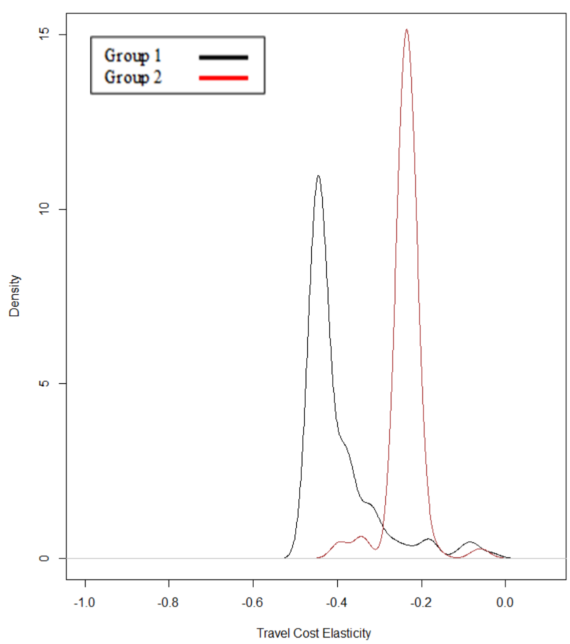
(p) Marginal Effects of In-vehicle Travel Time for Regional Transit with Auto Passenger Access Mode



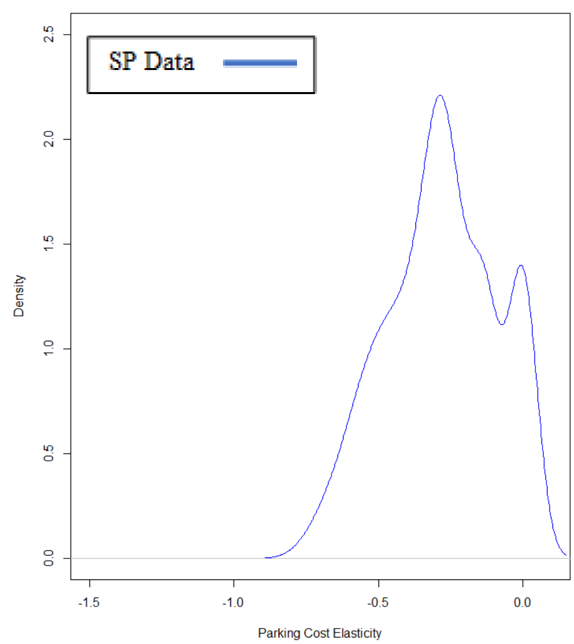
(q) Marginal Effects of Travel Cost for Regional Transit with Local Transit Access Mode



(r) Marginal Effects of In-vehicle Travel Time for Regional Transit with Local Transit Access Mode



(s) Marginal Effects of RP Travel Cost for Local Transit with Walk Access Mode (Identification of Sample Segments)



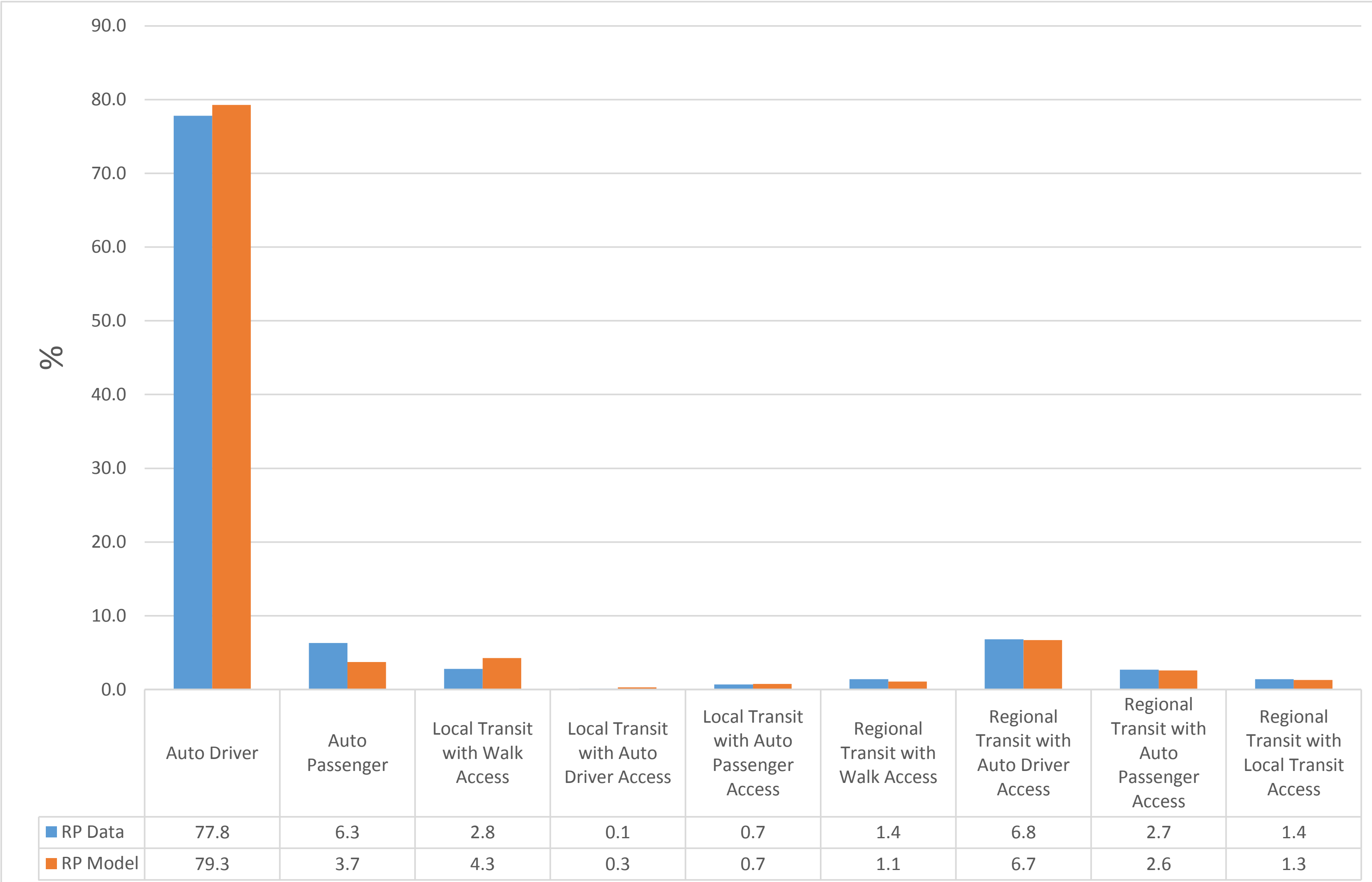
(t) Marginal Effects of Parking Cost at Park-and-Ride Stations for Regional Transit with Auto Driver Access Mode

Figure 7-4 Marginal Effects

7.4.4 Model Validation

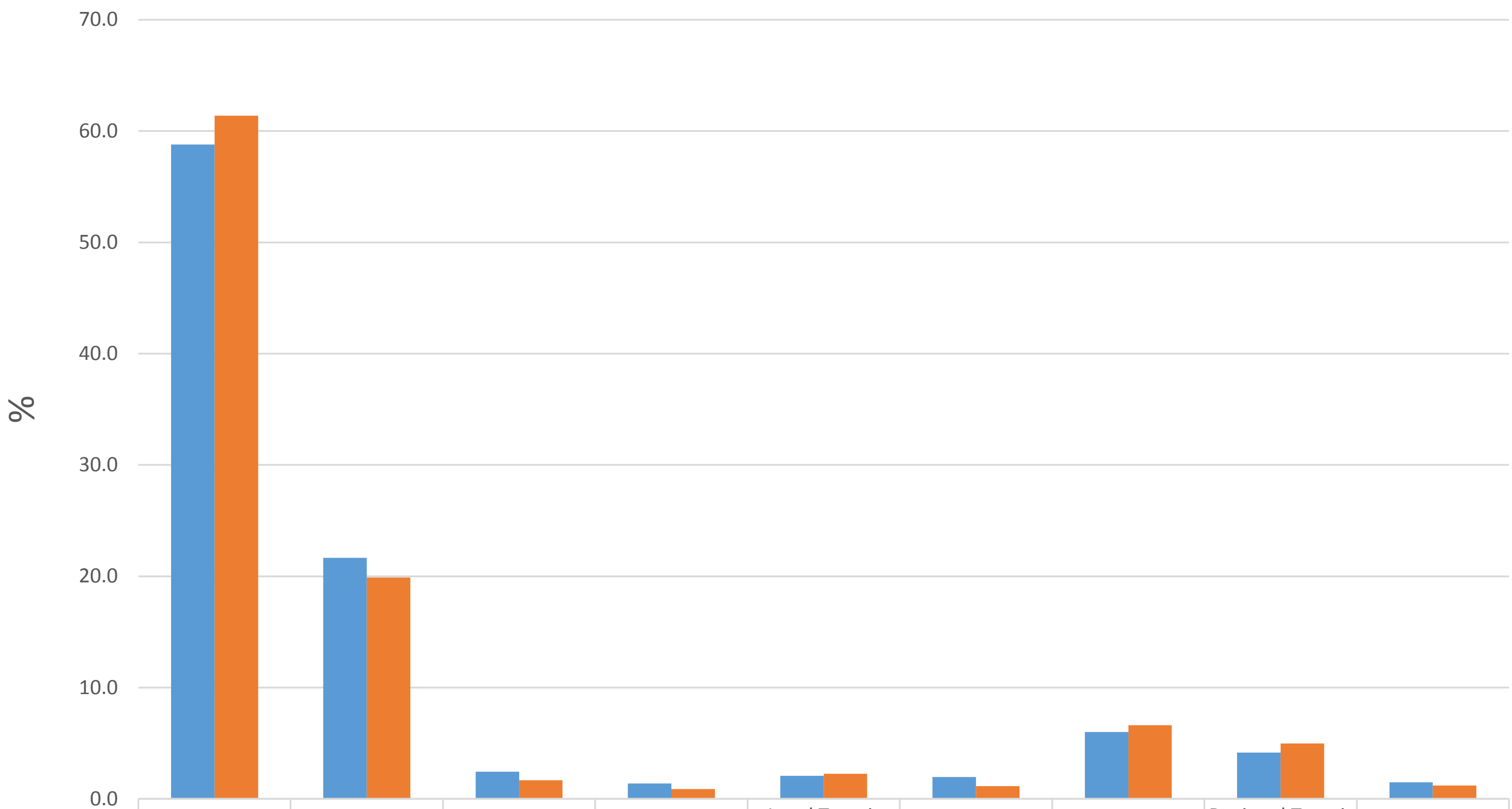
An independent subset of the collected RP/SP data was randomly selected and retained for model validation. This subset was not used in the joint RP/SP model estimation process to accurately investigate the predictive performance of the developed model against a holdout sample. The subset consists of 144 individuals (around 20% of the full sample dataset). In general, the estimation and validation datasets are consistent in terms of their modal shares across the RP and SP datasets.

To investigate the predictive performance of the joint RP/SP model, the developed model was used to predict the modal share using the holdout sample of the RP and SP datasets. The resulting aggregate modal shares of the joint RP/SP model and the actual modal shares of the dataset are shown in Figure 7-5. The developed model appears to predict the observed modal shares accurately, with minor variations in the auto driver and auto passenger modes. In addition, the holdout sample was used to build a confusion matrix to evaluate the prediction accuracy of the developed model. The overall accuracy (the correct predictions divided by all predictions) is found to be 80%. The precision and recall were estimated for all mode alternatives. The precision and recall of classifying the auto driving mode are 91% and 89%, respectively. However, the precision and recall of classifying other modes are relatively lower, which indicates that the numbers are slightly skewed by having an imbalance in the number of values across the alternatives. This is because the model was estimated with a higher percentage of auto driving mode observations, which represents the market share of this mode alternative across the target population.



	Auto Driver	Auto Passenger	Local Transit with Walk Access	Local Transit with Auto Driver Access	Local Transit with Auto Passenger Access	Regional Transit with Walk Access	Regional Transit with Auto Driver Access	Regional Transit with Auto Passenger Access	Regional Transit with Local Transit Access
RP Data	77.8	6.3	2.8	0.1	0.7	1.4	6.8	2.7	1.4
RP Model	79.3	3.7	4.3	0.3	0.7	1.1	6.7	2.6	1.3

Figure 7-5 (a) Joint RP/SP Model Validation – RP Data



	Auto Driver	Auto Passenger	Local Transit with Walk Access	Local Transit with Auto Driver Access	Local Transit with Auto Passenger Access	Regional Transit with Walk Access	Regional Transit with Auto Driver Access	Regional Transit with Auto Passenger Access	Regional Transit with Local Transit Access
SP Data	58.8	21.6	2.4	1.4	2.1	2.0	6.0	4.2	1.5
SP Model	61.38	19.88	1.67	0.88	2.25	1.14	6.62	4.97	1.21

Figure 7-5 (b) Joint RP/SP Model Validation – SP Data

7.5 Empirical IIA-Relaxed Models

The models presented above reveal meaningful insights into cross-regional commuters' behaviour. The joint RP/SP MNL model is validated against a holdout sample data. The results endorse the expected predictive capabilities of the developed model and verify its capacity for practical implications. However, further in-depth behavioural insights can be captured through the application of advanced model structures. The MNL model structure inherently assumes that all feasible alternatives are independent and irrelevant of each other. This assumption can be relaxed by considering the correlation between subgroups of alternatives or by identifying different groups of individuals. That is, an empirical investigation of two model structures that relax the underlying assumptions of the MNL is conducted. This section presents the application of two empirical models: nested logit (NL) and parameterized captivity logit (PLC) models.

Such model structures require exhaustive data to accurately capture the effects of the attributes under investigation as a result of the highly non-linear likelihood function of the PLC model structure and the restriction of having at least one available alternative within each alternative nest for all individuals within the dataset to maintain a valid formulation of the NL model structure. Therefore, the two models are estimated using the full sample joint RP/SP dataset.

The empirical models were estimated using codes written in GAUSS® using the MAXLIK component for maximum likelihood estimation (Aptech-Systems, 2012).

7.5.1 Joint RP/SP NL Model

In this section, a joint RP/SP NL model is developed based on the econometric model presented in Section 7.3.2.1. The model is estimated using the full sample, including both RP and SP data. The estimation routine takes into consideration the repeated observations of each individual across the RP data and the six SP choice situations. SP choice scenarios where low confidence levels were reported by the respondents are not considered within the estimation process.

Various nesting structures are empirically tested. In principle, alternatives are grouped into nests according to presumed similarities across nested alternatives. That is, alternatives are nested based on their types (e.g. auto-based and transit-based), transit service provider (e.g. local transit and regional transit), or transit access modes (e.g. auto access or walk/transit access). Based on the estimated logsum parameter values and level of significance of each nesting structure, the developed NL models are either accepted or rejected. As explained earlier, assuming a specific nesting structure does not imply that individuals make their decisions accordingly. Rather, it only implies the inherent assumption of the alternatives' commonality within each nest.

The nesting structure shown in Figure 7-6 is adopted, and Table 7-7 shows the estimated parameters of the empirical NL model. The estimated logsum parameter value and its level of significance suggest that the nesting structure should be accepted because it shows consistency with the NL model's theoretical derivation. In particular, this nesting structure classifies the nine modes under consideration into three groups: drive alone, shared drive, and transit. Each of the drive alone and the shared drive groups encompasses one travel mode: auto driver and auto passenger/carpool, respectively. In contrast, the transit group includes seven transit mode alternatives. That is, it is assumed that all transit alternatives share a common component of their error term that partially relaxes the IIA property of the MNL model by taking into account the similarities between the nested alternatives.

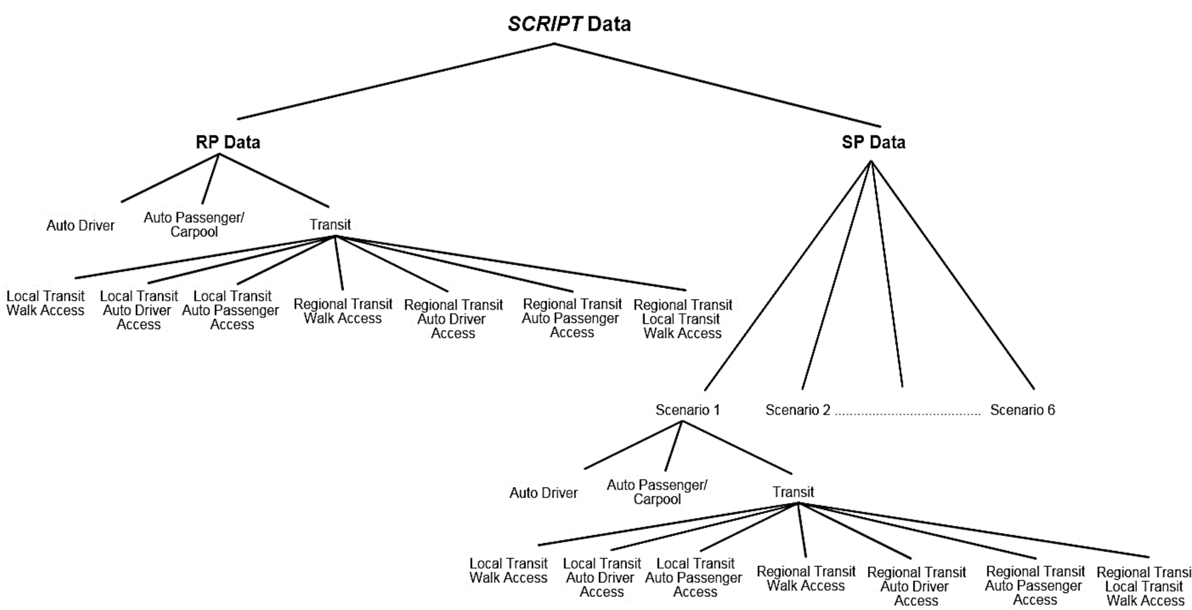


Figure 7-6 Joint RP/SP NL Empirical Model Structure

Thirty parameters are estimated with the expected signs and relative values, and are statistically significant at the 95% confidence interval except for the provision of information on local transit vehicles' arrival times, which is statistically significant at the 90% confidence interval. The full dataset consists of 704 trip records, but only 630 trip records are used for estimation to ensure the feasibility of at least one mode within each nest for all individuals in the dataset. The reported rho-squared value is 0.161. The log likelihood ratio test shows a test statistics value of 1,595, which indicates that the reported models fit the data significantly better than the constant-only model.

The estimated parameters are classified into the same three groups used in the model specification of the joint RP/SP MNL model: RP-specific, SP-specific, and pooled RP/SP coefficients. The developed model shows consistency in general with previous research findings and in particular with the previously presented models. The relative values of the estimated parameters indicate the perceived relative effect of different travel cost and time components by travellers. The unit parking cost at park-and-ride locations is perceived as having greater influence than the unit travel cost. Similarly, out-of-vehicle travel times (including access, egress, wait, and transfer times) are perceived to be 1.6 times higher than in-vehicle travel time. In addition, the estimated coefficient of in-vehicle travel time if Wi-Fi is available has a smaller negative effect on the probability of choosing GO Transit modes than the estimated coefficient of in-vehicle travel time if Wi-Fi is not available. Similarly, the need for a second transfer between GO Transit and other travel modes has a significant negative effect on the probability of choosing GO Transit modes, while the provision of transit vehicles' arrival time has a positive impact on the probability of choosing local transit modes. As expected, an increase in the number of vehicles per household results in an increase in the probability of using the auto driver mode. In addition, possessing a transit pass and having one end of the trip in the City of Toronto increase the probability of transit dependence.

Consistent with the findings of the developed joint RP/SP MNL model, the SP scale parameter is estimated to be less than 1 (i.e. lower than the RP scale parameter), which indicates that the variance within the SP data is higher than that of the RP data. In addition, the scale parameter of the transit nest is estimated to be greater than the corresponding upper-level scale parameters (i.e. RP and SP data scale parameters). In other words, the variance across the nested alternatives is lower than the variance across all mode alternatives, which indicates that the correlation between the nested alternatives is higher than the correlation between all alternatives. This is consistent with the underlying assumption of the NL model formulation; therefore, the empirical model is accepted.

Table 7-7 Joint RP/SP NL Model

NL Logit Model – Joint Estimation		RP/SP Model			
Log Likelihood of Full Model		-4,156.844			
Log Likelihood of Constant-only Model		-4,954.521			
Rho-squared Value		0.161			
Number of Observations		704 (630)			
Variable	**Mode**	**Parameter**	**t-Statistics**	**Parameter**	**t-Statistics**
Systematic Utility Function:		**RP Coefficients**		**SP Coefficients**	
Alternative Specific Constant	Auto driver	4.8019	6.544	1.6428	2.286
Alternative Specific Constant	Auto passenger	3.2428	4.483	-0.4277	-0.66
Alternative Specific Constant	Local transit with walk access	1.873	2.471	0.5527	0.799
Alternative Specific Constant	Local transit with auto driver access (TTC park-and-ride)	0 (fixed)	0 (fixed)	0 (fixed)	0 (fixed)
Alternative Specific Constant	Local transit with auto passenger access (TTC kiss-and-ride)	0.6285	0.858	-1.2924	-1.993
Alternative Specific Constant	Regional transit with walk access	2.7149	3.529	-0.178	-0.272
Alternative Specific Constant	Regional transit with auto driver access (GO park-and-ride)	3.0243	4.305	0.1223	0.187
Alternative Specific Constant	Regional transit with auto passenger access (GO kiss-and-ride)	1.8666	2.63	-0.4047	-0.625
Alternative Specific Constant	Regional transit with local transit access	2.7889	3.611	-0.1934	-0.293
Travel Cost/Fare	All modes	-0.0802	-8.264	-0.0802	-8.264
P&R Cost at TTC Stations	Local transit with auto driver access (TTC park-and-ride)	-0.4494	-3.088	-0.4494	-3.088
P&R Cost at GO Stations	Regional transit with auto driver access (GO park-and-ride)	0 (fixed)	0 (fixed)	-0.0715	-2.23
In-vehicle Travel Time (No Wi-Fi)	All modes	-0.0363	-8.972	-0.0363	-8.972
In-vehicle Travel Time (Wi-Fi)	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	0 (fixed)	0 (fixed)	-0.0314	-7.434

Out-of-vehicle Travel Time	All transit alternatives	-0.059	-8.679	-0.059	-8.679
Next Local Transit Vehicle Information Provision	Local transit with walk access, local transit with driving access (TTC park-and-ride), local transit with passenger access (TTC kiss-and-ride), and regional transit with local transit access	0.2072	1.703	0.2072	1.703
Need for a 2nd Transfer	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	-0.478	-3.504	-0.478	-3.504
Number of Vehicles per Household	Auto driver	0.2348	3.732	0.2348	3.732
Transit Pass Possession	All transit alternatives	2.2591	10.137	2.2591	10.137
Trip O/D: City of Toronto	All transit alternatives	0.7167	3.824	0.7167	3.824
Exponential Function of SP Scale Parameter:					
Constant	SP scale factor	0 (fixed)	0 (fixed)	-0.1198	-1.263
Gender	SP scale factor	0 (fixed)	0 (fixed)	-0.1878	-3.986
Exponential Function of Nesting Scale Parameter:*					
Constant	Transit nest scale factor	-2.0926	-3.125	-2.0926	-3.125

*The scale parameter of the lower nest is estimated as an exponential function in addition to the scale parameter of the upper nest to ensure that the nesting structure is consistent with the theoretical derivation of the nested logit model (i.e. the lower nest's scale parameter is higher than the upper nest's scale parameter; both scale parameters are positive).

7.5.2 Joint RP/SP PLC Model

The PLC model is estimated using the full sample of 704 trip records, including both RP and SP data. Similar to the previously presented joint RP/SP models, the estimation routine takes into consideration the repeated observations of each individual across the RP data and the six SP choice situations. SP choice scenarios where low confidence levels were reported by the respondents are not considered within the estimation process. To the author's knowledge, this application of the PLC model using joint RP/SP data is unique.

Because the likelihood function of the PLC model structure is highly non-linear, it is important to have realistic initial parameter assumptions. This is essential for optimal parameter estimation, as a large number of local maxima or apparent flat surfaces within the likelihood function may result in parameter estimates that do not maximize the likelihood function. Therefore, a similar MNL model specification model was estimated first to provide rough estimates for initial parameter assumptions for the PLC model estimation.

Table 7-8 shows the results of the empirical model following the PLC model structure shown in Figure 7-3. Forty-eight parameters were estimated with the expected signs and relative values and are statistically significant at the 95% confidence interval except for the provision of information on local transit vehicles' arrival times, which is statistically significant at the 90% confidence interval. The reported rho-squared value is 0.16, which is almost equal to the reported rho-squared value of the NL model. The log likelihood ratio test shows a test statistics value of 1,711, which indicates that the reported models fit the data significantly better than the constant-only model.

Typical level-of-service attributes (travel times and travel costs/fares) are included in the systematic utility functions of the model, which represents the rational choice component. Such attributes contribute to the trade-offs between alternatives that are made by rational individuals prior to making a decision. In contrast, socio-demographic and mode-specific attributes are added to the parameterized captivity functions. These factors are more likely to influence individuals' dependency on one specific alternative. The estimated parameters are classified into the same three groups used in the model specifications of the joint RP/SP MNL and NL models: RP-specific, SP-specific, and pooled RP/SP parameters. The systematic utility and parameterized captivity functions are estimated with different sets of alternative specific constants for both RP and SP datasets. In addition, the scale factor of the systematic utility function is parameterized to capture the heteroscedasticity in individuals' responses. Similarly, the estimated SP scale factor is lower than the RP scale parameter (fixed to 1), which verifies that the variance within the SP data is higher than that of the RP data.

Despite the dissimilarities between the PLC and the previously presented model formulations, the empirical results are consistent with the aforementioned findings of the MNL and NL models. The systematic utility function parameters of the PLC model, including travel costs (driving cost and/or transit fares, as well as parking costs) and different travel time components (in-vehicle and out-of-vehicle travel times) are estimated with the expected negative signs. The relative values of the estimated parameters indicate the perceived relative effect of different travel cost and time components by travellers. The unit parking cost at park-and-ride locations is perceived as having a greater influence compared to the unit travel cost. Similarly, out-of-vehicle travel times (including access, egress, wait, and transfer times) are perceived to be 1.9 times higher than in-vehicle travel time. In addition, the need for a second transfer between GO Transit and other travel modes has a significant negative effect on the probability of choosing GO Transit modes, while the provision of transit vehicles' arrival time has a positive impact on the probability of choosing local transit modes.

The estimated parameters of the parameterized captivity function provide insights into individuals' modal reliance/captivity. Based on the estimated parameters of the parameterized captivity function, the model suggests that 75% of individuals in the sample dataset make their choices based on the rational trade-off between all feasible mode alternatives that are available to them (i.e. by comparing alternatives' level-of-service attributes). However, the remaining 25% of individuals in the sample are classified as captive users who may depend on one specific mode alternative and inattentive to the rational trade-off between other alternatives. Out of which, 17% are found to be car-dependent users which explains the high RP mode share in the dataset. The parameter estimates of the parameterized captivity function show that an increase in the number of vehicles per household results in an increase in the probability of auto driver reliance. Similarly, individuals with a higher number of vehicles compared to the number of persons per household are less likely to consider auto passenger as their sole travel mode. In contrast, possessing a transit pass and having one end of the trip in the City of Toronto increases the probability of transit dependency.

Table 7-8 Joint RP/SP PLC Model

PLC Logit Model – Joint Estimation		RP/SP Model			
Log Likelihood of Full Model		-4,491.302			
Log Likelihood of Constant-only Model		-5,346.789			
Rho-squared Value		0.160			
Number of Observations		704			
Variable	**Mode**	**Parameter**	**t-Statistics**	**Parameter**	**t-Statistics**
Systematic Utility Function:		**RP Coefficients**		**SP Coefficients**	
Alternative Specific Constant	Auto driver	2.9507	4.206	2.0285	4.99
Alternative Specific Constant	Auto passenger	0 (fixed)	0 (fixed)	0 (fixed)	0 (fixed)
Alternative Specific Constant	Local transit with walk access	6.4198	3.688	3.0635	3.46
Alternative Specific Constant	Local transit with auto driver access (TTC park-and-ride)	5.1629	2.673	2.8293	1.857
Alternative Specific Constant	Local transit with auto passenger access (TTC kiss-and-ride)	2.973	2.46	1.4805	1.968
Alternative Specific Constant	Regional transit with walk access	2.2227	0.616	3.126	3.026
Alternative Specific Constant	Regional transit with auto driver access (GO park-and-ride)	2.7899	2.178	3.0047	3.435
Alternative Specific Constant	Regional transit with auto passenger access (GO kiss-and-ride)	-3.1359	-0.247	1.8054	2.493
Alternative Specific Constant	Regional transit with local transit access	-0.1359	-0.023	3.4659	3.063
Travel Cost/Fare	All modes	-0.2203	-5.536	-0.2203	-5.536
P&R Cost at TTC Stations	Local transit with auto driver access (TTC park-and-ride)	-0.6239	-2.141	-0.6239	-2.141
P&R Cost at GO Stations	Regional transit with auto driver access (GO park-and-ride)	0 (fixed)	0 (fixed)	-0.19	-2.111
In-vehicle Travel Time (No Wi-Fi)	All modes	-0.1093	-4.759	-0.1093	-4.759
In-vehicle Travel Time (Wi-Fi)	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	0 (fixed)	0 (fixed)	-0.0972	-4.374

Out-of-vehicle Travel Time	All transit alternatives	-0.2032	-5.14	-0.2032	-5.14
Next Local Transit Vehicle Information Provision	Local transit with walk access, local transit with driving access (TTC park-and-ride), local transit with passenger access (TTC kiss-and-ride), and regional transit with local transit access	0.3795	1.697	0.3795	1.697
Need for a 2nd Transfer	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	-0.7511	-2.019	-0.7511	-2.019
Exponential Function of Scale Parameter:					
Constant	SP scale factor	0 (fixed)	0 (fixed)	-0.3186	-2.597
Gender	SP scale factor	0 (fixed)	0 (fixed)	-0.207	-4.034
Parameterized Captivity Function:					
Alternative Specific Constant	Auto driver	0.1348	0.181	-2.2695	-4.623
Alternative Specific Constant	Auto passenger	-1.4374	-1.702	-1.7724	-3.246
Alternative Specific Constant	Local transit with walk access	-5.6489	-7.617	-8.2565	-9.593
Alternative Specific Constant	Local transit with auto driver access (TTC park-and-ride)	-15.6694	-0.589	-10.0521	-9.158
Alternative Specific Constant	Local transit with auto passenger access (TTC kiss-and-ride)	-8.5856	-7.624	-9.1894	-9.474
Alternative Specific Constant	Regional transit with walk access	-6.0311	-7.596	-8.4754	-9.068
Alternative Specific Constant	Regional transit with auto driver access (GO park-and-ride)	-5.1156	-6.866	-7.7924	-9.474
Alternative Specific Constant	Regional transit with auto passenger access (GO kiss-and-ride)	-6.333	-8.522	-8.0629	-9.532
Alternative Specific Constant	Regional transit with local transit access	-6.1803	-7.997	-8.9118	-9.34
Number of Vehicles per Household	Auto driver	0.3749	2.47	0.3749	2.47
Ratio of Number of Vehicles to Number of Persons	Auto passenger	-2.9066	-2.808	-2.9066	-2.808
Transit Pass Possession	All transit alternatives	4.7976	10.301	4.7976	10.301
Trip O/D: City of Toronto	All transit alternatives	1.8845	5.015	1.8845	5.015

7.6 Comparison of Joint RP/SP MNL, NL, and PLC Models

A joint RP/SP MNL model is estimated with model specifications that are consistent with the IIA-relaxed model specifications, as shown in Table 7-9. The joint RP/SP MNL model is estimated using the full sample to compare the results with those of the joint RP/SP IIA-relaxed models (i.e. the NL and PLC models).

In general, the findings of the three models are consistent. As such, the relative values of the estimated parameters are analogous. For instance, out-of-vehicle travel times (including access, egress, wait, and transfer times) are perceived to be 1.7, 1.6, and 1.9 times higher than in-vehicle travel time according to the MNL, NL, and PLC model results, respectively. Similarly, the value of travel time savings (VOT) is calculated for the different travel time components according to the estimated travel time and cost parameters of the three developed models, as shown in Table 7-10. In terms of the estimated models' goodness of fit, both the NL and PLC models showed higher rho-squared values than the MNL model. In particular, the NL model showed a slightly higher rho-squared value than the PLC model. Therefore, the NL is recommended to be used for policy analysis as discussed in the following section.

Table 7-9 Joint RP/SP MNL Model for NL/PLC Comparison

MNL Logit Model – Joint Estimation		RP/SP Model			
Log Likelihood of Full Model		-4,497.462			
Log Likelihood of Constant-only Model		-5,341.4035			
Rho-squared Value		0.158			
Number of Observations		704			
Variable	**Mode**	**Parameter**	**t-Statistics**	**Parameter**	**t-Statistics**
Systematic Utility Function:		**RP Coefficients**		**SP Coefficients**	
Alternative Specific Constant	Auto driver	2.6427	11.011	0.6422	4.442
Alternative Specific Constant	Auto passenger	0 (fixed)	0 (fixed)	0 (fixed)	0 (fixed)
Alternative Specific Constant	Local transit with walk access	1.0894	2.711	-1.3721	-4.141
Alternative Specific Constant	Local transit with auto driver access (TTC park-and-ride)	-0.5885	-0.572	-1.0607	-1.389
Alternative Specific Constant	Local transit with auto passenger access (TTC kiss-and-ride)	-1.7627	-3.246	-2.3267	-6.399
Alternative Specific Constant	Regional transit with walk access	0.723	1.466	-1.2086	-3.53
Alternative Specific Constant	Regional transit with auto driver access (GO park-and-ride)	0.8518	2.351	-0.9091	-3.114
Alternative Specific Constant	Regional transit with auto passenger access (GO kiss-and-ride)	-0.4608	-1.118	-1.4898	-5.209
Alternative Specific Constant	Regional transit with local transit access	0.8244	1.661	-1.2861	-3.469
Travel Cost/Fare	All modes	-0.0762	-8.291	-0.0762	-8.291
P&R Cost at TTC Stations	Local transit with auto driver access (TTC park-and-ride)	-0.4633	-2.97	-0.4633	-2.97
P&R Cost at GO Stations	Regional transit with auto driver access (GO park-and-ride)	0 (fixed)	0 (fixed)	-0.0756	-2.158
In-vehicle Travel Time (No Wi-Fi)	All modes	-0.0355	-9.206	-0.0355	-9.206
In-vehicle Travel Time (Wi-Fi)	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	0 (fixed)	0 (fixed)	-0.0306	-7.252

Out-of-vehicle Travel Time	All transit alternatives	-0.0616	-9.785	-0.0616	-9.785
Next Local Transit Vehicle Information Provision	Local transit with walk access, local transit with driving access (TTC park-and-ride), local transit with passenger access (TTC kiss-and-ride), and regional transit with local transit access	0.2395	1.793	0.2395	1.793
Need for a 2nd Transfer	Regional transit with walk access, regional transit with auto driver access (GO park-and-ride), regional transit with auto passenger access (GO kiss-and-ride), and regional transit with local transit access	-0.4614	-3.26	-0.4614	-3.26
Number of Vehicles per Household	Auto driver	0.1934	3.133	0.1934	3.133
Ratio of Number of Vehicles to Number of Persons	Auto passenger	-0.7046	-4.11	-0.7046	-4.11
Transit Pass Possession	All transit alternatives	2.4008	10.823	2.4008	10.823
Trip O/D: City of Toronto	All transit alternatives	0.6106	3.692	0.6106	3.692
Exponential Function of Scale Parameter:					
Constant	SP scale factor	0 (fixed)	0 (fixed)	-0.1042	-1.145
Gender	SP scale factor	0 (fixed)	0 (fixed)	-0.1781	-4.101

Table 7-10 Value of Travel Time Savings

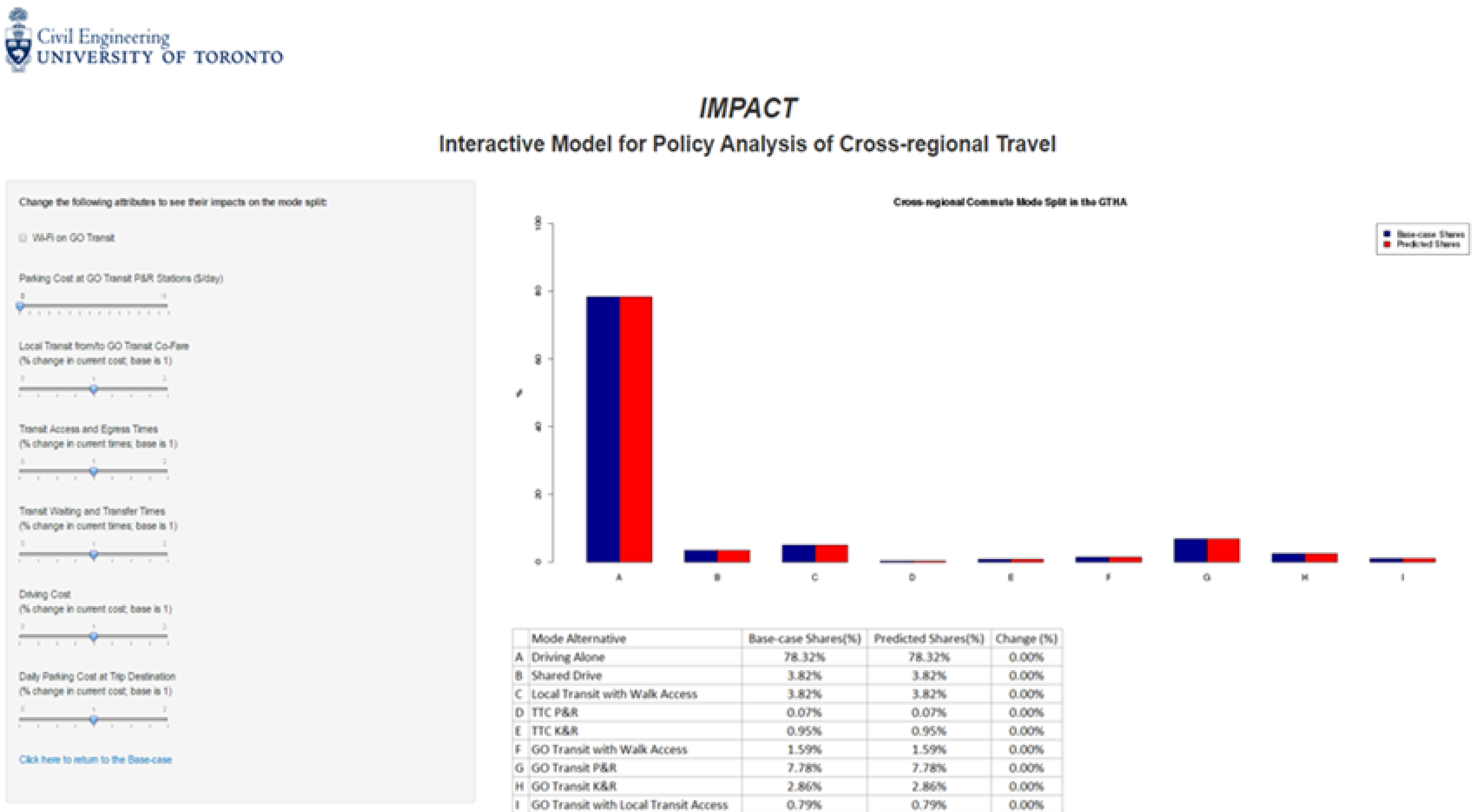
Value of Travel Time (VOT) Savings ($/hr.)	MNL Model	NL Model	PLC Model
In-vehicle Travel Time (No Wi-Fi)	27.95	27.16	29.77
In-vehicle Travel Time (Wi-Fi)	24.09	23.49	26.47

7.7 Policy Analysis Tool and Practical Implications

According to the comparison of the three developed joint RP/SP models, the IIA- relaxed models showed higher rho-squared values than the MNL model. In particular, the NL model has the highest rho-squared value among the three developed model. In addition, the NL model structure is a more theoretically sound model than the MNL model. Although the PLC model provides valuable insights in terms of the latent segmentation of individuals within the sample and therefore it identifies a probabilistic weight to individuals' choices to account for their level on dependency on specific modes, the application of this model for policy analysis is an open area of research. The limitation of applying the PLC model for forecasting stems from the complexity associated with the model calibration.

In light of the above discussion, the developed joint RP/SP NL model is used to predict corresponding changes in aggregate modal shares in response to new transportation policies. To effectively use the developed model for predicting market shares, the RP environment is used (because it represents the actual behaviour) and the SP specific variables (i.e. variables that are identified using the SP data which reflect the newly introduced policies) are moved to the RP utilities (without adjusting for the scale effect) to develop a "hybrid" model that can be used for forecasting. The common practice is to adjust the ASC to reproduce market shares that are sufficiently close to those of the sample data (Cherchi and Ortúzar, 2006; Train, 2009). After calibrating the ASC, the model can be used in predicting demand changes in response to changes in the explanatory variables.

The NL model was calibrated using the full sample RP data and used to test the effectiveness of various policy initiatives. The Interactive Model for Policy Analysis of Cross-Regional Travel (*IMPACT*) was developed to facilitate testing the policies under investigation. Figure 7-7 shows a snapshot of the Web-based policy analysis tool with the base-case aggregate modal shares of the full sample RP data. To showcase the model prediction functionality, four independent (one policy at a time) policies are investigated. In each policy analysis, corresponding explanatory variables are adjusted and the predicted model shares are compared to the base-case aggregate modal shares. It should be noted that the MNL and PLC models were calibrated and used to conduct a similar policy analysis. In general, the results of the policy analysis using the three models are consistent with slightly different predicted market shares.



	Mode Alternative	Base-case Shares(%)	Predicted Shares(%)	Change (%)
A	Driving Alone	78.32%	78.32%	0.00%
B	Shared Drive	3.82%	3.82%	0.00%
C	Local Transit with Walk Access	3.82%	3.82%	0.00%
D	TTC P&R	0.07%	0.07%	0.00%
E	TTC K&R	0.95%	0.95%	0.00%
F	GO Transit with Walk Access	1.59%	1.59%	0.00%
G	GO Transit P&R	7.78%	7.78%	0.00%
H	GO Transit K&R	2.86%	2.86%	0.00%
I	GO Transit with Local Transit Access	0.79%	0.79%	0.00%

Figure 7-7 Snapshot of *IMPACT*

Policy 1: Introducing Wi-Fi on GO Transit

In this policy analysis, it was assumed that all individuals in the dataset will have access to Wi-Fi service on all GO Transit modes. Based on the model's predicted modal shares, the modal share of GO Transit modes has increased from 13.0% to 14.2%. This increase in the GO Transit modal share is associated with a 0.9% decrease in the driving modal shares.

Policy 2: Introducing Pay Parking at GO Transit Park-and-Ride Stations

The effect of introducing pay parking at GO Transit park-and-ride stations (which is currently free) is investigated. In general, the predicted modal shares show a decrease in the GO Transit modal share and, in particular, a decrease in the GO Transit park-and-ride with auto driver access modal share. Table 7-11 shows the predicted modal shares of GO Transit modes in response to the introduction of different parking costs at GO Transit park-and-ride stations.

Table 7-11 Predicted GO Transit Mode Shares

Parking Cost at GO Transit Park-and-Ride Stations	GO Transit Modes*	GO Transit Park-and-Ride with Auto Driver Access Mode**
$1	12.8%	7.4%
$3	12.3%	6.7%
$5	11.9%	6.0 %

*Base-case mode share is 13.0%
** Base-case mode share is 7.8%

Policy 3: Reducing Transit Co-Fares to/from GO Transit

In this policy analysis, GO Transit users are exempted from paying the co-fare when local transit is used for access/egress. Accordingly, the model predictions show an increase in GO Transit modal shares of approximately 0.76%.

Policy 4: Increasing Driving Cost

Similarly, the current driving and parking costs at work locations for the driving mode alternatives (i.e. auto driver and auto passenger) are increased by 50%. Based on the new estimated modal shares, the modal shares of the driving mode alternatives decrease by 2.8%, and, interestingly, a corresponding increase in GO Transit with driving access (park-and-ride and kiss-and-ride) modal shares of 1.7% is captured.

The above policy analysis is provided for demonstration purposes to illustrate how the model can be used to investigate individual policies. The model is capable of examining the likely effect of different changes than those used above for each policy. Additionally, the model can predict the changes in mode choices as a result of introducing more than one policy at a time.

7.8 Chapter Summary

This chapter presents the development of a set of cross-regional commuter mode choice models using *SCRIPT* RP and SP data. As a first step toward developing behavioural models, an RP-only multinomial logit (MNL) model is developed using *SCRIPT* RP data. The estimation results show consistency with corresponding operational mode choice models, demonstrating the validity of the survey design, sampling procedure, and data quality. Nevertheless, RP-only models are incapable of accurately forecasting individuals' choices in response to the policies under investigation. Therefore, policy-sensitive SP models are developed to capture the associated changes in travel demand with respect to changes in level-of-service attributes. A joint RP/SP MNL model is developed using *SCRIPT's* RP and SP data. The estimated model outperformed the corresponding SP-only model, which indicates the effect of incorporating the full information (i.e. the combined RP/SP data) on capturing changes in individuals' preferences according to policy implications. The estimated parameters are reported with the expected signs and are statistically significant (with t-statistics higher than 1.96) at the 95% confidence interval except for a few variables. An independent subset of the collected RP/SP data was randomly selected and retained for model validation. The developed model appears to predict the observed modal share accurately with minor variations.

To relax the independent and irrelevant alternative (IIA) property that is inherent within the traditional MNL model formulation, nested logit (NL) and parameterized logit captivity (PLC) models are estimated using the joint RP/SP data. The NL model allows for the capture of different substitution patterns across nests (e.g. automobile and transit modes). In contrast, the PLC model formulation allows for the capture of the latent segmentation of captive (myopic) and rational individuals. That is, by parameterizing the latent captivity function, valuable insights into the effect of different variables on individuals' reliance on a specific mode of travel are provided. In general, based on the empirical results, adopting model structures in which the IIA property is relaxed improves the explanatory power of the developed models while providing consistent findings. Furthermore, the developed joint RP/SP NL model is calibrated and used to develop the Interactive Model for Policy Analysis of Cross-Regional Travel (*IMPACT*) to predict corresponding changes in aggregate modal shares in response to a sample of four transportation policies. In each policy analysis, corresponding explanatory variables are adjusted within the dataset, and the predicted model shares are compared to the base-case aggregate modal shares.

CHAPTER 8

8 CONCLUSIONS, RESEARCH CONTRIBUTIONS, AND FUTURE RESEARCH

8.1 Chapter Overview

This chapter provides a summary of the research presented in Section 8.2. In Section 8.3, the main research contributions are highlighted. Finally, Section 8.4 discusses possible directions for future research.

8.2 Research Summary

As explained in Chapter 2, cross-regional trips are defined as trips that cross boundaries of municipal or regional jurisdictions that have different transit operators. Such trips have unique characteristics because of the longer trip distances and/or the typical lack of service coordination between transit service providers. Furthermore, cross-regional travellers have distinctive socioeconomic/demographic attributes, such as high income and auto ownership levels. The continuous growth in the number of cross-regional commuting trips, which are mostly dominated by private automobile drivers, has produced persistent pressure that contributes to the peak-period traffic congestion in metropolitan areas. Therefore, different travel demand management and land use policies have been under investigation by regional authorities to find sustainable transportation solutions to alleviate this pressure. Improving transit modal integration at either the intramodal or intermodal level is one of the promising strategies toward such sustainable transportation solutions. However, a proper understanding of travellers' behaviour is required prior to investigating the effects of such strategies. Accordingly, a conceptual demand modelling framework that is sensitive to the characteristics of intermodal commuting trips is developed.

The presented modelling framework starts with input data on travel demand and its distribution, followed by three main components, a departure time choice model, an enhanced mode choice model, and a multimodal trip assignment model. This framework is appropriate for understanding individuals' behaviour and/or studying the potential changes in their travel decisions as a result of policies targeted at short- to medium-term decisions such as departure time, travel mode, and route choice decisions. At the end of each iteration, a pre-defined performance criterion (e.g. a percentage change in mode share or travel time) is assessed, and accordingly, the iterative procedure is either re-performed until convergence or terminated. After the last iteration, the framework produces a detailed set of discrete choices of cross-regional travel decisions. These decisions include individuals' expected departure time choice, joint decisions on main travel mode choice and access mode choice (if any), access station location choice (if any), and driving and/or transit route choices.

This framework places more emphasis on the mode choice model component. According to the literature on intermodal travel (see Chapter 2), the decision structure of such trips is defined by three choices: main mode, access mode, and access station location choice. To develop models that are capable of capturing intermodal travellers' current and possible future behaviour changes within the context of multimodal networks, the three choices and their interactions are carefully considered in the different phases of the model development. The literature highlights the importance of considering the joint effect of the three choices on individuals' travel decisions. However, developing such a three-level nesting structure would require extensive data on individuals' behaviour and encounter various challenges in terms of the empirical estimation of the model. More importantly, the decision structure and the number of generated alternatives might lead to cases where the choice situation is not consistent with behavioural aspects of the choice theory. Therefore, the framework is developed over two phases. In phase I, data from an RP travel survey are used along with detailed information on transit stations to develop an access station location choice model. In phase II, an innovative multimodal trip planner tool that adopts the developed access location choice model is integrated with an RP/SP survey to collect data on commuters' current travel plans as well as their decisions in response to hypothetical scenarios. The collected RP/SP data are used to develop a joint main mode and access mode choice model that implicitly takes into account the access station location for transit modes with auto access.

The access location station choice is examined for individuals using a transit mode with auto access. In contrast, users who walk to transit stations are assumed to access the transit system *via* the closest station. Therefore, the developed access location choice models investigate the key factors that affect park-and-ride commuters' behaviour of access station location choice. Using these models, a tool is developed to predict the access station location choice of individuals who have park-and-ride as an available mode in their choice sets. Two sources of data were used for model development: data on morning peak-period cross-regional commuting trips, extracted from the 2006 TTS household travel survey, and data on park-and-ride station-level attributes such as locations, parking lot capacities, parking costs, surrounding land use, and station amenities, obtained from transit service operators. The five and three closest stations defined the access station choice set for regional and local transit park-and-ride users, respectively. Three datasets were prepared for modelling and empirical investigations to capture the difference in the choice behaviour of each market segment. Three MNL logit models are developed: a model for all park-and-ride users, a model for regional transit (GO Train) park-and-ride users, and a model for local transit (TTC Subway) park-and-ride users.

In general, the estimated parameters are statistically significant with the expected correct signs for the three models. Access distance (between individuals' home and access station locations) and relative station direction (between individuals' home and work locations) are the primary factors that affect individuals' choices. An increase in both access distance and relative travel angle has a negative impact on access station choice. Other factors, such as parking capacity and parking costs, provide behavioural insights in terms of explaining individuals' choices. The regional and local transit park-and-ride access station location choice models are adopted within the design of the RP/SP survey in phase II of the development of the presented framework.

To develop behavioural models that are capable of capturing changes in individuals' decisions, exhaustive data on their travel plans are essential. In particular, choice models that are developed using stated preference (SP) data are capable of investigating newly introduced policies (i.e. policies that have never been applied before) that cannot be tested using models that are developed using revealed preference (RP) data. This allows for the evaluation of the effectiveness of the proposed policy initiatives since SP experiments provide more flexibility in examining such policies by collecting information on individuals' preferences toward a set of hypothetical scenarios that do not currently exist. However, individuals' stated preferences may not be consistent with their actual choices, which may induce a systematic bias in the data. Alternatively, using joint RP/SP data allows for the scale adjustment of parameter estimates to correct the systematic bias of the SP data. That is, The Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*) is designed to collect RP/SP data that allow for the evaluation of the effectiveness of policy initiatives targeted at enhancing transit services and quantifying the effect of changes in level-of-service attributes on individuals' travel choices.

SCRIPT collects information on the respondents' commuting trips (RP data) as well as their stated preferences toward mode choices with improved and/or new levels of service attributes (SP data). *SCRIPT* is an online respondent-customized survey. As such, the survey questions are tailored to accommodate all possible travel mode options. The survey is carefully designed to capture the changes in individuals' travel mode choices in response to policies aimed at improving transit services with more emphasis on transit modal integration. The survey consists of three sections, collecting revealed preference data, stated preference data, and household and personal information, respectively. The information gathered from the RP section feeds into an innovative multimodal trip planner tool. This tool is developed to generate only feasible travel options for each SP choice experiment based on households' auto ownership level, proximity to transit, work start time, and total travel time from home to work. For intermodal travel modes such as park-and-ride and/or kiss-and-ride, the tool selects the access stations based on the pre-developed access station location choice models before generating the associated level-of-service attributes for presentation to the respondent. Finally, socioeconomic and demographic information is collected.

A pilot survey is conducted to test the survey platform, and a pilot sample is collected from the same sample frame that was used to collect data for the final survey. This sample is used to develop a preliminary SP mode choice model. Using the preliminary model's estimated parameters as *a priori* estimates, the final SP experiments are developed based on the D-efficient design technique. To design the survey sample, the survey population is divided into homogeneous strata and the 2011/2012 TTS data are used to identify the survey sampling probabilities of each stratum based on spatial location, mode split, and gender. These probabilities are used as guidelines to ensure the collection of a representative sample of the target population. The sample selection is done based on simple random sampling. Accordingly, the sample size is determined and found to be consistent with the required number of complete responses based on the developed D-efficient experiment design, which is estimated to be between 800 and 1,000 records to estimate statistically significant variables at the 95% confidence interval. The N-proportional allocation method is used to determine the required sample size from each stratum.

The data collection was done during the spring and fall seasons of 2014. The average time required to complete the survey was 20 minutes. The total number of complete responses was 1,003 with a completion rate (the ratio of the number of complete responses to the number of respondents who qualified to participate in the survey) of 33.6%. After the data were cleaned (by removing invalid responses and/or records that do not belong to the study area's sample frame), 704 complete valid records were prepared to be used for empirical modelling.

Using *SCRIPT'*s RP/SP data, a set of econometric joint main mode and access mode choice models are developed to help meet the research objectives. Three types of models are developed: RP-only, SP-only, and joint RP/SP. The RP-only and SP-only models are restricted models based on the RP and SP data, respectively, while the joint RP/SP model considers both RP and SP data jointly. The developed models reveal meaningful insights into understanding cross-regional commuters' mode choice behaviour by explaining the probabilistic responses to changes in level-of-service attributes as a result of the introduction of new policies.

A comparison between conventional multinomial logit (MNL) RP-only and SP-only models against a joint MNL RP/SP model is presented. The estimation routine of the joint RP/SP model takes into consideration the repeated observations by each individual across the RP data and the different SP choice situations. In addition, the scale parameter, which is estimated for the SP data relative to a normalized (fixed to 1) scale parameter for the RP data, was parameterized as an exponential function of the respondents' attributes to capture the heteroscedasticity in individuals' responses across the sample. The joint RP/SP model shows consistency in general with previous research findings and in particular with the developed RP-only and SP-only models. However, the values of the estimated parameters of the RP-only and SP-only models are different from their corresponding parameters in the RP/SP model. This indicates the effect of data enrichment (i.e. the combined RP/SP data) on capturing the scaled (corrected) relative influence of each variable on the probability of individuals' mode choices. The use of the combined RP/SP data enhanced the goodness of fit and explanatory power of the joint RP/SP model compared to the corresponding SP-only model. Using a holdout sample, the developed models appear to predict the observed modal share accurately with only minor variations.

Furthermore, to relax the independent and irrelevant alternative (IIA) property that is inherent within the traditional MNL model formulation, nested logit (NL) and parameterized logit captivity (PLC) models are estimated using the joint RP/SP data. The NL model allows for the capture of different substitution patterns across nests (e.g. automobile and transit modes). Alternatives are grouped into three nests according to presumed similarities across nested alternatives: drive alone, shared drive, and transit. In contrast, the PLC model formulation allows for capturing the latent segmentation of two types of individuals; those who consider more than one alternative when they make the choice of travel mode and those who depend on one specific mode of travel. That is, by parameterizing the latent captivity function, valuable insights into the effects of different variables on individuals' reliance on a specific mode of travel are provided. In general, based on the empirical results, adopting model structures in which the IIA property is relaxed improves the explanatory power of the developed models.

Finally, a comparison between the three models is presented and accordingly recommendations of using the developed nested logit (NL) model for policy analysis is made. The NL model is calibrated to be used for the prediction of corresponding changes in aggregate modal shares in response to the introduction of new transportation policies. The calibrated model is used to develop a web-based policy analysis tool, *IMPACT*.

8.3 Research Contributions

The development of the RP/SP survey, *SCRIPT*, is considered a significant contribution toward developing comprehensive respondent-customized surveys. The survey adopts an innovative multimodal trip planner tool developed to generate the SP experiments of the survey. The tool generates individual-specific feasible travel options and their level-of-service attributes pivoted on individuals' RP data. It defines a set of feasible modes for each respondent based on vehicle ownership level, proximity to transit, time required to be at work, and total travel time from home to work. Then, the tool generates detailed level-of-service attributes (including all travel cost and time components) based on the specified arrival time at work for all feasible modes. Finally, it adjusts attribute levels before presenting the final choice situations to the respondents based on the D-efficient experimental design.

Unlike existing similar tools, the tool generates customized/realistic intermodal travel options such as park-and-ride or transit trips that involve the use of two or more transit services, in addition to the typical travel modes. One of the innovative features of the multimodal trip planner tool is that it adopts access station location choice models to predict the access station location for individuals who have transit modes with auto driver or passenger access available in their choice sets. The access station location choice models are developed for park-and-ride users with a clear distinction between regional and local transit users. In addition, the models consider both travellers' and station-level attributes such as access distance (between individuals' home and access station locations), relative station direction (between individuals' home and work locations), parking capacity, parking cost, and other station amenities. That is, the developed models provide disaggregate-level models that are sensitive to individuals' home/work locations and park-and-ride station characteristics. Therefore, the generated SP experiments are customized, and the generated level-of-service attributes for each mode are realistic. To the author's knowledge, this is the first attempt to integrate a multimodal trip planner with a set of pre-developed econometric models to develop a comprehensive respondent-customized data collection tool.

The development of joint RP/SP models has become common in the literature of discrete choice models. However, the use of such models in practice is limited, with evidently limited guidelines for how joint RP/SP models are validated and calibrated to be used for policy analysis. As shown in Chapter 7, using the collected RP/SP data, three model structures are adopted to develop joint main mode and access mode choice models. The three joint RP/SP models are: multinomial logit (MNL), nested logit (NL), and parameterized logit captivity (PLC). The NL (which assumes a nesting structure that classifies the travel modes into three groups, drive alone, shared drive, and transit) and the PLC (which provides an inclusion probability that latently classifies individuals as rational or captive users) models relax the IIA property of the MNL model.

The developed PLC model provides behavioural insights into individuals' modal reliance/captivity by parametrizing the captivity odds function (the captive portion of the model) as a function of socio-demographic and/or mode-specific variables, while parametrizing the systematic utility (the rational portion of the model) as a function of level-of-service attributes. This classification stems from the fact that the probability of dependence on a given mode is based more on factors other than the level-of-service attributes of travel modes. However, rational individuals are assumed to be making deliberate trade-offs between the different alternatives based on their characteristics. To the author's knowledge, this application of the PLC model using joint RP/SP data is unique.

Although the PLC model provides such valuable insights, the application of this model for policy analysis is an open area of research. The limitation of applying the PLC model for forecasting stems from the complexity associated with the model calibration. Therefore, the developed joint RP/SP NL model is recommended to be used for policy analysis. The developed joint main mode and access mode NL model is calibrated to develop a web-based policy analysis tool. The Interactive Model for Policy Analysis of Cross-Regional Travel (*IMPACT*) is developed to facilitate testing the policies under investigation. The tool is used to predict corresponding changes in aggregate modal shares in response changes in the explanatory variables in response to new transportation policies. The development, validation, and calibration of the presented models are considered another significant contribution of the conducted research.

In general, the presented travel demand model provides a generic modelling framework for intermodal trips within the context of cross-regional travel in multimodal networks. The framework attempts to capture individuals' observed travel behaviour by adopting a continuous updating process. It consists of a series of disaggregate/agent-based model components that are interconnected with each other in a continuous feedback fashion to be capable of capturing individuals' decision structures and how they choose their travel options. The framework provides a detailed disaggregate probabilistic chain of decisions, including a departure time choice, a joint main mode and access mode (if any) choice, an access station location choice (if any), and a route choice. That is, for each commuting trip (at the disaggregate level), the framework provides a detailed set of discrete choices that defines commuters' travel decisions. The framework offers a practice-ready modelling platform for cross-regional trips by connecting the three model components. The total demand and its distribution are loaded onto the travel demand model. Each individual is assigned a departure time for his/her trip from home to work according to the departure time choice model (Sasic and Habib, 2013). Afterwards, the access station location choice model provides a set of probabilities for station locations to be chosen by individuals who have transit modes with auto access available in their travel mode choice sets. Those probabilities are used to determine the expected level-of-service attributes for transit modes with auto access modes to be used in the joint main mode and access mode choice model. Then, the joint main mode and access mode choice model (which implicitly takes into account the access station location for transit modes with auto access) provides a set of probabilities for each combination of main mode and access mode (if any). The discrete choices of main mode, access mode, and access station location choices are randomly chosen based on the estimated probabilities before being passed to the multimodal trip assignment model (Weiss, 2013). Using individuals' simulated choices, the multimodal trip assignment model generates new level-of-service attributes of individuals' modes of travel. The new level-of-service attributes are fed back to simulate new choices in an iterative process until the model reaches convergence. Finally, the model output is a detailed disaggregate probabilistic chain of decisions, including departure time, main mode, access mode (if any), access station location (if any), and driving and/or transit route choices. That is, the framework allows for the study of individuals' travel choices in a disaggregate fashion.

8.4 Directions for Future Research

The presented framework consists of three separately developed model components. Applying the framework as an operational travel demand model is the next step of this research. The model components may be calibrated using 2011/2012 TTS data as a further validation step to move the model components towards being practice ready. Subsequently, the calibrated models can be applied to back-cast 2006 TTS data to assess the model forecasting capabilities. Although the presented framework provides a comprehensive method to model cross-regional commuting trips by capturing changes in individuals' decisions regarding their departure times, travel modes, and travel route choices, the framework has some limitations. Specifically, because the travel demand and its distribution are assumed to be fixed, the framework is incapable of capturing changes in individuals' decisions caused by changes in their home or work locations. Such long-term decisions are not considered within the scope of this study. Considering individuals' long-term decisions as well as investigating the effects of household interactions and day-long activity scheduling on cross-regional travellers' behaviour is a promising direction for extending this research.

Further applications using the collected data on individuals' travel behaviour (*SCRIPT*'s RP/SP data) may be considered. For instance, data on individuals' revealed mode choice and self-reported corresponding level-of-service attributes were collected. The multimodal trip planner tool was used to generate level-of-service attributes for all travel modes, including the chosen mode. Comparing the self-reported and generated level-of-service attributes can provide insights into individuals' perceptions of the modes they use. In addition, the survey collected information on individuals' expected departure time in response to the selected travel mode after each SP scenario. Investigating the effect of mode choice on adjusting the departure time choice can provide behavioural insights into the interaction between travel mode and departure choices. Similarly, the interaction between travel mode and route choices of cross-regional commuters, especially in the context of multimodal networks, is another open area of research. Finally, although the SP experiments were customized by showing only feasible travel modes for each respondent, individuals might not consider all options while making the choice. Data on the individuals' consideration set were collected at the end of each SP experiment. Examining individuals' characteristics that affect the formation of their consideration set can provide insights into understanding the key factors that contribute to the elimination of feasible options from their choice sets.

9 REFERENCES

Aptech-Systems. GAUSS Engine and MAXLIK, 2012. http://www.aptech.com.

Beimborn, E. A., M. J. Greenwald, and X. Jin. Accessibility, Connectivity, and Captivity: Impacts on Transit Choice, *Transportation Research Record*, 2003, pp. 1–9.

Ben-Akiva, M., and T. Morikawa. Estimation of Switching Models from Revealed Preferences and Stated Intentions. *Transportation Research Part A: General*, Vol. 24, No. 6, 1990, pp. 485–495.

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

Bergman, A., J. Gliebe, and J. Strathman, J. Modelling Access Mode Choice for Inter-suburban Commuter Rail. *Journal of Public Transportation*, Vol. 14, No. 4, 2011, pp. 23–42.

Bhat, C. R. Covariance Heterogeneity in Nested Logit Models: Econometric Structure and Application to Intercity Travel. *Transportation Research Part B: Methodological*, Vol. 31, No. 1, 1997, pp. 11–21.

Bhat, C. R. Accommodating Flexible Substitution Patterns in Multi-dimensional Choice Modelling: Formulation and Application to Travel Mode and Departure Time Choice. *Transportation Research Part B: Methodological*, Vol. 32, No. 7, 1998, pp. 455–466.

Bliemer, M. C., and J. M. Rose. Efficiency and Sample Size Requirements for Stated Choice Experiments. *Transportation Research Board 88th Annual Meeting*, 2009.

Bordley, R. F. The Dogit Model Is Applicable Even without Perfectly Captive Buyers. *Transportation Research Part B: Methodological*, Vol. 24, No. 4, 1990, pp. 315–323.

Bos, I., R. Van Der Heijden, E. Molin, and H. J. P. Timmermans. Traveler Preference for Park-and-ride Facilities: Empirical Evidence of Generalizability. *Transportation Research Record*, 2005, pp. 126–134.

Bos, I. D. M., R. E. C. M. Van der Heijden, E. J. E. Molin, and H. J. P. Timmermans. The Choice of Park and Ride Facilities: An Analysis Using a Context-dependent Hierarchical Choice Experiment. *Environment and Planning A*, Vol. 36, No. 9, 2004, pp. 1673–1686.

Brownstone, D., D. S. Bunch, and K. Train. Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-fuel Vehicles. *Transportation Research Part B: Methodological*, Vol. 34, No. 5, 2000, pp. 315–338.

CAA. Canadian Automobile Association, 2014. http://www.caa.ca/.

Cairns, M. R. The Development of Park and Ride in Scotland. *Journal of Transport Geography*, Vol. 6, No. 4, 1998, 295–307.

Carson, R. T., J. J. Louviere, D. A. Anderson, P. Arabie, D. S. Bunch, D. A. Hensher, R. M. Johnson, W. F. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans, and J. B. Wiley. Experimental Analysis of Choice. *Marketing letters,* Vol. 5, No. 4, 1994, pp. 351–367.

Caussade, S., J. de Dios Ortúzar, L. I. Rizzi, and D. A. Hensher. Assessing the Influence of Design Dimensions on Stated Choice Experiment Estimates. *Transportation Research Part B: Methodological.* Vol. 39, No. 7, 2005, pp. 621–640.

Cervero, R., and K. Kockelman. Travel Demand and the 3Ds: Density, Diversity, and Design. *Transportation Research Part D: Transport and Environment,* Vol. 2, No. 3, 1997, pp. 199–219.

Chakour, V., and N. Eluru. Analyzing Commuter Train User Behavior: A Decision Framework for Access Mode and Station Choice. *Transportation,* Vol. 41, No. 1, 2014, pp. 211–228.

Cherchi, E., and J. D. D. Ortúzar. Use of Mixed Revealed-preference and Stated-preference Models with Nonlinear Effects in Forecasting. *Transportation Research Record*, 2006, pp. 27–34.

Choice-Metrics. Ngene 1.1. 1 User Manual and Reference Guide, 2012. http://www.choice-metrics.com/documentation.html.

Chorus, C. G., and B. G. Dellaert. Travel choice inertia: the joint role of risk aversion and learning. *Journal of Transport Economics and Policy (JTEP), Volume 46, Number 1, January 2012, pp. 139-155(17).*

Chu, Y.-L. Work Departure Time Analysis Using Dogit Ordered Generalized Extreme Value Model. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2132, No. 1, 2009, pp. 42–49.

Chu, Y.-L. A Combined Destination and Route Choice Model for Capturing both Compulsory and Discretionary Trips. *Transportation Research Board 89th Annual Meeting*, 2010.

Croissant, Y. Estimation of Multinomial Logit Models in R: The Mlogit Packages. *R Package Version 0.2-2,* 2012. http://cran.r-project.org/web/packages/mlogit/vignettes/mlogit.pdf.

Debrezion, G., E. Pels, and P. Rietveld. Choice of Departure Station by Railway Users. 2007.

Debrezion, G., E. Pels, and P. Rietveld. Modelling the Joint Access Mode and Railway Station Choice. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 45, No. 1, 2009, pp. 270–283.

Data Management Group. 1996 Transportation Tomorrow Survey. Joint Program in Transportation, University of Toronto, 1998 http://www.dmg.utoronto.ca/transportationtomorrowsurvey/origindestination.html.

Data Management Group. 2006 Transportation Tomorrow Survey. Joint Program in Transportation, University of Toronto, 2008. http://www.dmg.utoronto.ca/transportationtomorrowsurvey/origindestination.html.

Data Management Group. 2011 Transportation Tomorrow Survey. Joint Program in Transportation, University of Toronto, 2012.

Fan, K.-S., E. J. Miller, and D. Badoe. Modelling Rail Access Mode and Station Choice. *Transportation Research Record*, Vol. 1413, 1993.

Foote, P. J. Chicago Transit Authority Weekday Park-and-ride Users: Choice Market with Ridership Growth Potential. *Transportation Research Record*, 2000, pp. 158–168.

Forsey, D., K. M. N. Habib, E. Miller, and A. Shalaby. An Evaluation of the Impacts of Introducing a New Transit System on Commuting Mode Choice and Transit Ridership: A Case Study of the VIVA Bus Transit System in Toronto. *Transportation Research Part A*, 2013.

Fry, T. R., and M. N. Harris. The Dogit Ordered Generalized Extreme Value Model. *Australian & New Zealand Journal of Statistics*, Vol. 47, No. 4, 2005, pp. 531–542.

García, R., and A. Marín. Parking Capacity and Pricing in Park'n Ride Trips: A Continuous Equilibrium Network Design Problem. *Annals of Operations Research*, Vol. 116, No. 1–4, 2002, pp. 153–178.

Gaundry, M.J., and M. G. Dagenais. The Dogit Model. *Transportation Research Part B: Methodological*, Vol. 13, No. 2, 1979, pp. 105–111.

Ghosh, A. Valuing Time and Reliability: Commuters' Mode Choice from a Real Time Congestion Pricing Experiment. University of California Transportation Center, 2001.

Givoni, M., and P. Rietveld. The Access Journey to the Railway Station and Its Role in Passengers' Satisfaction with Rail Travel. *Transport Policy*, Vol. 14, No. 5, 2007, pp. 357–365.

Goetz, A. R., and T. M. Vowles. Progress in Intermodal Passenger Transportation: Private Sector Initiatives. *Transp. LJ*, Vol. 27, 2000, p. 475.

Graham, D. Modelling Intermodal Transportation Systems: Establishing a Common Language. *Transp. LJ* 27, 2000, p. 353.

Habib, K. M. N. A Joint Discrete-continuous Model Considering Budget Constraint for the Continuous Part: Application in Joint Mode and Departure Time Choice Modelling. *Transportmetrica A: Transport Science*, Vol. 9, No. 2, 2013, pp. 149–177.

Habib, K. M. N., N. Day, and E. J. Miller. An Investigation of Commuting Trip Timing and Mode Choice in the Greater Toronto Area: Application of a Joint Discrete-continuous Model. *Transportation Research Part A: Policy and Practice*, Vol. 43, No. 7, 2009, pp. 639–653.

Habib, K. M. N., M. S. Mahmoud, and J. Coleman. Effect of Parking Charges at Transit Stations on Park-and-ride Mode Choice: Lessons Learned from Stated Preference Survey in Greater Vancouver, Canada. *Transportation Research Record*, 2014, pp. 163–170.

Habib, K. M. N., and A. Weiss. Evolution of Latent Modal Captivity and Mode Choice Patterns for Commuting Trips: A Longitudinal Analysis Using Repeated Cross-sectional Datasets. *Transportation Research Part A: Policy and Practice*, Vol. 66, No. 1, 2014, pp. 39–51.

Haener, M. K., P. C. Boxall, and W. L. Adamowicz. Modelling Recreation Site Choice: Do Hypothetical Choices Reflect Actual Behavior? *American Journal of Agricultural Economics*, Vol. 83, No. 3, 2001, pp. 629–642.

Hensher, D. A. Stated Preference Analysis of Travel Choices: The State of Practice. *Transportation*, Vol. 21, No. 2, 1994, pp. 107–133.

Hensher, D. A., P. O. Barnard, and T. P. Truong. The Role of Stated Preference Methods in Studies of Travel Choice. *Journal of Transport Economics and Policy*, 1988, pp. 45–58.

Hensher, D. A., and M. Bradley. Using Stated Response Choice Data to Enrich Revealed Preference Discrete Choice Models. *Marketing Letters*, Vol. 4, No. 2, 1993, pp. 139–151.

Hensher, D. A., J. M. Rose, and W. H. Greene. Combining RP and SP Data: Biases in Using the Nested Logit 'Trick' – Contrasts with Flexible Mixed Logit Incorporating Panel and Scale Effects. *Journal of Transport Geography*, Vol. 16, No. 2, 2008, pp. 126–133.

Holguín-Veras, J., J. Reilly, F. Aros-Vera, W. Yushimito, and J. Isa. Park-and-ride Facilities in New York City. *Transportation Research Record*, 2012, pp. 123–130.

Huber, J., and K. Zwerina. The Importance of Utility Balance in Efficient Choice Designs. *Journal of Marketing Research*, 1996, pp. 307–317.

Idris, A. O. Modal Shift Forecasting Models for Transit Service Planning. Toronto, University of Toronto, 2013.

Idris, A. O., K. M. N. Habib, and A. S. Shalaby. Modal Shift Forecasting Model for Transit Service Planning: Survey Instrument Design. *The 12th International Conference on Advanced Systems for Public Transport (CASPT)*, 2012.

Imaz, A., K. M. N. Habib, A. Shalaby, and A. O. Idris. Investigating the Factors Affecting Transit User Loyalty. *Public Transport*, Vol. 7, No. 1, 2014, pp. 39–60.

Jacques, C., K. Manaugh, and A. M. El-Geneidy. Rescuing the Captive [Mode] User: An Alternative Approach to Transport Market Segmentation. *Transportation*, Vol. 40, No. 3, 2013, pp. 625–645.

Kastrenakes, C. R. *Development of a Rail Station Choice Model for NJ Transit*, 1988.

Khan, O. A. Modelling Passenger Mode Choice Behaviour Using Computer Aided Stated Preference Data, 2007.

Koppelman, F. S., and C. Bhat. A Self instructing Course in Mode Choice Modelling: Multinomial and Nested Logit Models. *US Department of Transportation, Federal Transit Administration*, Vol. 31, 2006.

Koppelman, F. S. and C. H. Wen. The Paired Combinatorial Logit Model: Properties, Estimation and Application. *Transportation Research Part B: Methodological*, Vol. 34, No. 2, 2000, pp. 75–89.

Korf, J., and M. Demetsky. Analysis of Rapid Transit Access Mode Choice. *Transportation Research Record*, Vol. 817, 1981.

Kroes, E. P., and R. J. Sheldon. Stated Preference Methods: An Introduction. *Journal of Transport Economics and Policy*, 1988, pp. 11–25.

Kumar, A., and Y. Gur. *Consideration of Alternative Access, Egress, and Line-Haul Travel Choices within UTPS Framework*. 1982.

Lee, R. W. Exploration of Long-distance Interregional Commuting Issues: Analysis of Northern California Interregional Commuters Using Census Data and Focus Group Interviews. *Transportation Research Record*, 1996, pp. 29–36.

Li, Z.-C., W. H. Lam, S. Wong, D.-L. Zhu, and H.-J. Huang, H. Modelling Park-and-ride Services in a Multimodal Transport Network with Elastic Demand. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1994, No. 1, 2007, pp. 101–109.

Liao, F., T. Arentze, and H. Timmermans. Supernetwork Approach for Multimodal and Multiactivity Travel Planning, *Transportation Research Record*, 2010, pp. 38–46.

Litman, T. Introduction to Multi-modal Transportation Planning. *Victoria Transport Policy Institute*, Vol. 15, 2011.

Louviere, J. J., and D. A. Hensher. Using Discrete Choice Models with Experimental Design Data to Forecast Consumer Demand for a Unique Cultural Event. *Journal of Consumer Research*, 1983, pp. 348–361.

Louviere, J. J., D. A. Hensher, and J. D. Swait. *Stated Choice Methods: Analysis and Applications*. Cambridge, Cambridge University Press, 2000.

Louviere, J. J., and G. Woodworth. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. *Journal of Marketing Research*, 1983, pp. 350–367.

Lu, A., and A. N. Marsh. Zen and the Art of Commuter Rail Operations: 2 Taiwan Railways Administration's Design, Operations, and Philosophy 3. *Notes*, Vol. 43, 2011, p. 44.

Lu, L. The Vital Role of Metropolitan Access in Commuter, Regional, Intercity and Overnight Rail Passenger Transportation – And Its Relationship to Technology. Cambridge, Massachusetts Institute of Technology, 2003.

Manski, C. F. The Structure of Random Utility Models. *Theory and Decision*, Vol. 8, No. 3, 1977, pp. 229–254.

Manski, C. F., and D. McFadden. *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MIT Press, 1981.

McCarthy, P. S. The Role of Captivity in Aggregate Share Models of Intercity Passenger Travel. *Journal of Transport Economics and Policy*, 1997, pp. 293–308.

McFadden, D. Conditional Logit Analysis of Qualitative Choice Behavior. 1973.

McFadden, D. Econometric Analysis of Qualitative Response Models. *Handbook of econometrics*, Vol. 2, 1984, pp. 1395–1457.

McFadden, D. The Choice Theory Approach to Market Research. *Marketing Science*, Vol. 5, No. 4, 1986, pp. 275–297.

Meek, S., S. Ison, and M. Enoch. Stakeholder Perspectives on the Current and Future Roles of UK Bus-based Park and Ride. *Journal of Transport Geography*, Vol. 17, No. 6, 2009, pp. 468–475.

Metrolinx. The Regional Transportation Plan for the Greater Toronto and Hamilton Area (GTHA): "The Big Move," 2008. http://www.metrolinx.com/thebigmove/en/default.aspx.

Metrolinx. GO Transit Rail Passenger Survey Report, 2011.

Meyer, M. D., and E. J. Miller. *Urban Transportation Planning: A Decision-oriented Approach*. 2001.

Miller, E. J. The Greater Toronto Area Travel Demand Modelling System, Version 3.0, Vol. I: Model Overview. Joint Program in Transportation, University of Toronto, Toronto, 2007.

Mukundan, S. *An Access Mode and Station Choice Model for the Washington DC Metrorail System*. 1991.

Öhman, M., and U. Lindgren. Who Is the Long-distance Commuter? Patterns and Driving Forces in Sweden. *Cybergeo: European Journal of Geography*, 2003.

Ortúzar, J., and L. G. Willumsen. *Modelling Transport*. John Wiley & Sons, 2011.

Osman, A., A. Shalaby, and K. M. N. Habib. Dissecting the Role of Transit Service Attributes in Attracting Commuters: Lessons from a Comprehensive RP-SP Study on Commuting Mode Switching Behaviour in Toronto. *Transportation Research Board 93nd Annual Meeting*, Washington, D.C., 2014.

Outwater, M., K. Tierney, M. Bradley, E. Sall, A Kuppam, and V. Modugula. California Statewide Model for High-speed Rail. *Journal of Choice Modelling*, Vol. 3, No. 1, 2010, pp. 58–83.

Park, S., J. Kang, and K. Choi. Finding Determinants of Transit Users' Walking and Biking Access Trips to the Station: A Pilot Case Study. *KSCE Journal of Civil Engineering*, Vol. 18, No. 2, 2014, pp. 651–658.

Polydoropoulou, A., and M. Ben-Akiva. Combined Revealed and Stated Preference Nested Logit Access and Mode Choice Model for Multiple Mass Transit Technologies. *Transportation*

Research Record: Journal of the Transportation Research Board, Vol. 1771, No. 1, 2001, pp. 38–45.

Revelt, D., and K. Train. Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics*, Vol. 80, No.4, 1998, pp. 647–657.

Roorda, M. J., D. Passmore, and E. J. Miller. Including Minor Modes of Transport in a Tour-based Mode Choice Model with Household Interactions. *Journal of Transportation Engineering*, Vol. 135, No. 12, 2009, pp. 935–945.

Rose, J. M., M. C. Bliemer, D. A. Hensher, and A. T. Collins. Designing Efficient Stated Choice Experiments in the Presence of Reference Alternatives. *Transportation Research Part B: Methodological*, Vol. 42, No, 4, 2008, pp. 395–406.

Sandow, E., and K. Westin. The Persevering Commuter – Duration of Long-distance Commuting. *Transportation Research Part A: Policy and Practice*, Vol. 44, No. 6, 2010, pp. 433–445.

Sasic, A., and K. N. Habib. Modelling Departure Time Choices by a Heteroskedastic Generalized Logit (Het-GenL) Model: An Investigation on Home-based Commuting Trips in the Greater Toronto and Hamilton Area (GTHA). *Transportation Research Part A: Policy and Practice*, Vol. 50, 2013, pp. 15–32.

Schmitt, L., S. Harris, and G. Currie. Adapting an Online Transit Journey Planner into a Low-cost Travel Survey Tool. *Transportation Research Record*, 2014, pp. 8–15.

Sobieniak, J., R. Westin, T. Rosapep, and T. Shin. Choice of Access Mode to Intercity Terminals. *Transportation Research Record*, Vol. 728, 1979.

Stopher, P. R. Captivity and Choice in Travel-behavior Models. *Transportation Engineering Journal of the American Society of Civil Engineers*, Vol. 106, No. 4, 1980, pp. 427–435.

Swait, J., and M. Ben-Akiva. Analysis of the Effects of Captivity on Travel Time and Cost Elasticities. *Behavioural Research for Transport Policy*, 1986, pp. 119–134.

Swait, J., and M. Ben-Akiva. Empirical Test of a Constrained Choice Discrete Model: Mode Choice in Sao Paulo, Brazil. *Transportation Research Part B: Methodological*, Vol. 21, No. 2, 1987, pp. 103–115.

Swait, J., and M. Ben-Akiva. Incorporating Random Constraints in Discrete Models of Choice Set Generation. *Transportation Research Part B: Methodological*, Vol. 21, No. 2, 1987, pp. 91–102.

Tardiff, T. J. *Perception of the Availability of Transportation Alternatives for Various Trip Purposes*. 1976.

Titheridge, H., and P. Hall. Changing Travel to Work Patterns in South East England. *Journal of Transport Geography*, Vol. 14, No. 1, 2006, pp. 60–75.

Train, K., and Y. Croissant. Kenneth Train's Exercises Using the Mlogit Package for R. R Package Version 0.2-3, 2012. http://cran.r-project.org/web/packages/mlogit/index.html.

Train, K. E. EM Algorithms for Nonparametric Estimation of Mixing Distributions. *Journal of Choice Modelling,* Vol. 1, No, 1, 2008, pp. 40–69.

Train, K. E. *Discrete Choice Methods with Simulation*. Cambridge, Cambridge University Press, 2009.

Tsamboulas, D., J. Golias, and M. Vlahoyannis. Model Development for Metro Station Access Mode Choice. *Transportation,* Vol. 19, No. 3, 1992, pp. 231–244.

Tsang, F. W. K., A. S. Shalaby, and E. J. Miller. Improved Modelling of Park-and-ride Transfer Time: Capturing the Within-day Dynamics. *Journal of Advanced Transportation,* Vol. 39, No. 2, 2005, pp. 117–137.

Van Ham, M., and P. Hooimeijer. Regional Differences in Spatial Flexibility: Long Commutes and Job Related Migration Intentions in the Netherlands. *Applied Spatial Analysis and Policy,* Vol. 2, No. 2, 2009, pp. 129–146.

Vijayakumar, N., A. M. El-Geneidy, and Z. Patterson. Driving to suburban rail stations: Understanding variables that affect driving distance and station demand. *Transportation Research Record*, 2011, pp. 97–103.

Wardman, M. A Comparison of Revealed Preference and Stated Preference Models of Travel Behaviour. *Journal of Transport Economics and Policy*, 1988, pp. 71–91.

Wardman, M., and G. Whelan. Using Geographical Information Systems to Improve Rail Demand Models. *Final Report to Engineering and Physical Science Research Council*. 1999.

Washbrook, K., W. Haider, and M. Jaccard. Estimating Commuter Mode Choice: A Discrete Choice Analysis of the Impact of Road Pricing and Parking Charges. *Transportation,* Vol. 33, No. 6, 2006, pp. 621–639.

Weiss, A. Framework for the Integration of a Parameterized Logit Captivity Model for Morning Commuting in the Greater Toronto and Hamilton Area with an Agent Based Dynamic Traffic Micro Simulation. 2013.

Wells, S. S. Analysis of the Differential Impacts of Transport Modes on Travel Behaviour in the Greater Toronto Area. *Transportation Research. Part A, Policy and Practice,* Vol. 31, No. 1, 1997, pp. 84–84.

Wen, C. H., W. C., Wang, and C. Fu, C. Latent Class Nested Logit Model for Analyzing High-speed Rail Access Mode Choice. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 48, No. 2, 2012, pp. 545–554.

Whitehead, J. C., S. K. Pattanayak, G. L. Van Houtven, and B. R. Gelso. Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Ecological Services: An Assessment of the State of the Science. *Journal of Economic Surveys,* Vol. 22, No. 5, 2006, pp. 872–908.

Zhang, M., and B. Chen. Understanding Emerging Commuting Trends in a Weekly Travel Decision Frame: Implications for Mega-region Transportation Planning. Center for Transportation Research, 2011.

Zwerina, K., J. Huber, and W. Kuhfeld. *A General Method for Constructing Efficient Choice Designs*. Durham, NC: Fuqua School of Business, Duke University, 1996.

APPENDIX A

Survey of Cross-Regional Intermodal Passenger Travel (*SCRIPT*)

HOME PAGE

(Survey invitation and consent to participate)

Study of Commuter Travel Mode Choices in the Greater Toronto and Hamilton Area

This Survey is Intended to Measure Cross-Regional Commuters' Mode Choice Preferences

Dear Survey Respondent,

You have been randomly selected to participate in a research study conducted by the Department of Civil Engineering at the University of Toronto and sponsored by Metrolinx, an agency of the Government of Ontario responsible for improving the coordination and integration of all modes of transportation in the Greater Toronto and Hamilton Area. This study aims at achieving a better understanding of cross-regional commuters' mode choice preferences (e.g., personal car, local transit, regional transit, park-and-ride, etc.) in response to changes in transit service attributes.

We are contacting a random sample of commuters in the Greater Toronto and Hamilton Area (GTHA) to gather information on their personal attitudes and habits associated with daily commuting work trips, with special focus placed on inter-regional work trips (those crossing regional boundaries). The outcome of the study will help your Provincial and Municipal governments make informed and evidence-based improvements to roads and public transit in your area. All collected information will be stored securely at the University of Toronto and will be processed with the utmost confidentiality and for research purposes only.

The survey is divided into three sections: **Section** A will gather information on your daily travel habits and current travel options; **Section B** will present a set of hypothetical scenarios where travel modes' service attributes are different than the current state, and you will be asked to indicate your preferences; and, **Section C** will collect your socioeconomic and demographic characteristics. We kindly ask you to participate in this web survey so that your opinion is represented in our study. This survey is designed to be as short as possible and will take approximately 20 minutes to complete. Please answer every question in each section in order to proceed to subsequent sections.

You may choose not to complete the survey at any time without any penalty. You can withdraw from the survey by simply closing the webpage anytime. Keep in mind, however, that the responses submitted in previous sections are not retrievable, and therefore will still be anonymously included in the final survey results. Please note that there is no related risk involved with your participation in this study. All the collected information will be stored securely at the University and will be processed with the utmost confidentiality and for academic purposes only. The results will be reported in aggregate form, with no references to specific participants of the survey. Your cooperation is highly appreciated.

Should you have any questions about the study, please feel free to e-mail us at mohamed.mahmoud@utoronto.ca. For any questions regarding your rights as a respondent in this survey, you are free to contact the office of Research Ethics, University of Toronto, McMurrich Building, 2nd floor, 12 Queen's Park Crescent West Toronto, ON M5S 1S8, Tel: (416) 946-3273, Fax: (416) 946-5763, Email: ethics.review@utoronto.ca

Consent of Participant

By pressing the "Login" button, you will indicate to us that you agree, of your own freewill, to voluntarily participate in this study after carefully reading and fully understanding the information presented in the introductory section of the survey.

Login

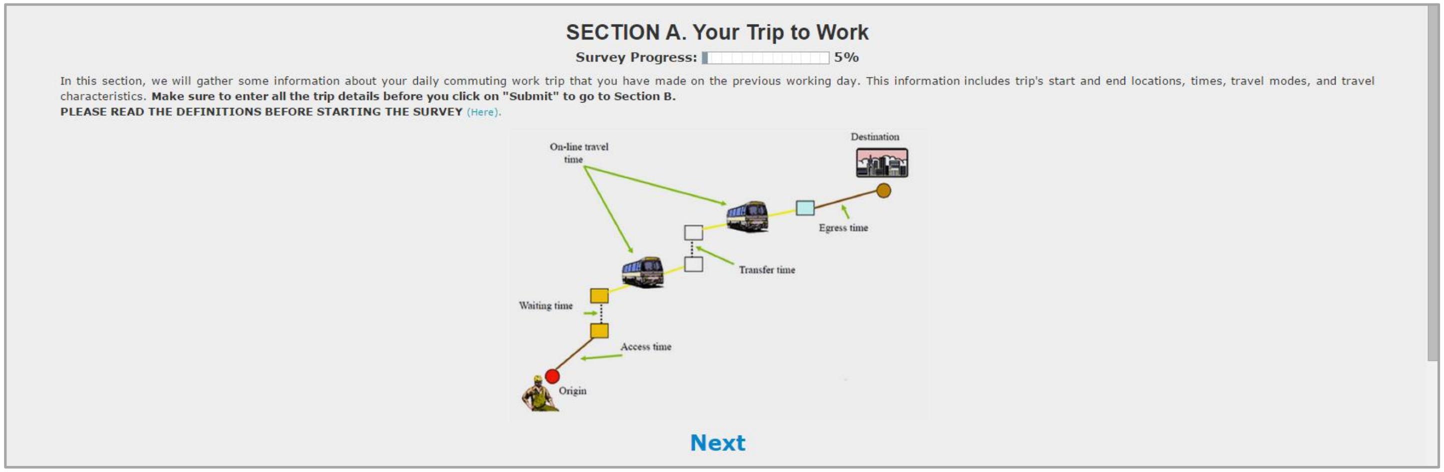
Posted by: Mohamed Mahmoud

Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

SECTION A

Introduction to Section A (RP Survey)



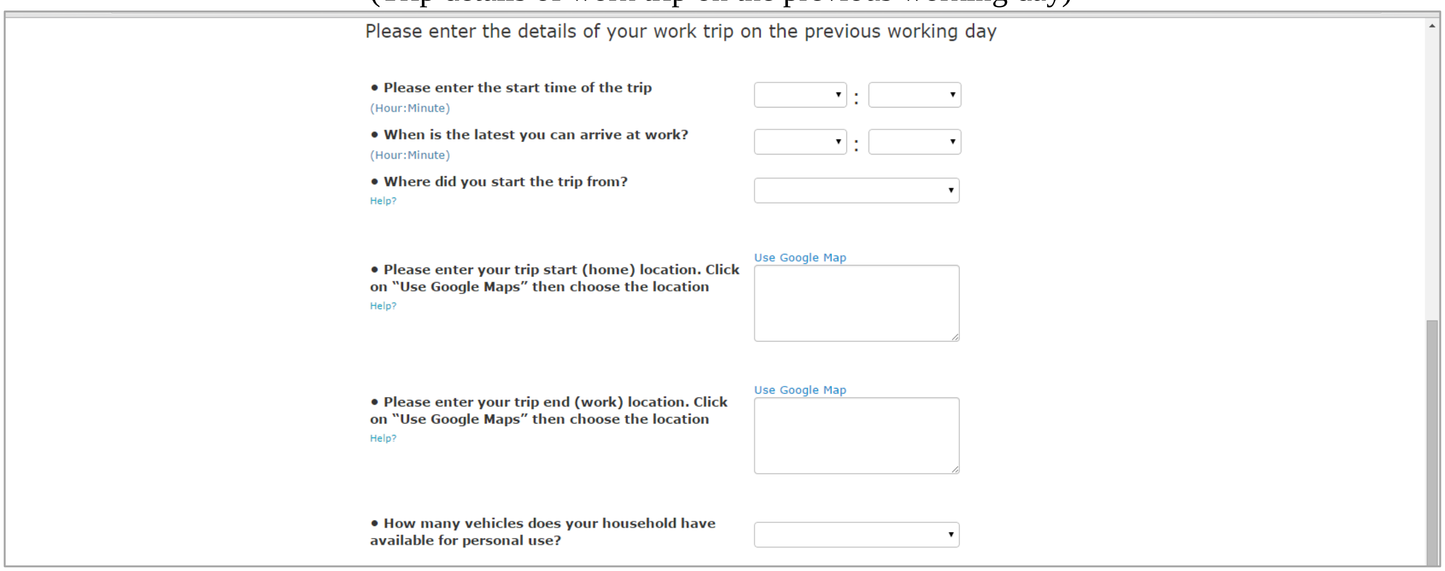
SECTION A. Your Trip to Work

Survey Progress: 5%

In this section, we will gather some information about your daily commuting work trip that you have made on the previous working day. This information includes trip's start and end locations, times, travel modes, and travel characteristics. **Make sure to enter all the trip details before you click on "Submit" to go to Section B.**
PLEASE READ THE DEFINITIONS BEFORE STARTING THE SURVEY (Here).

Next

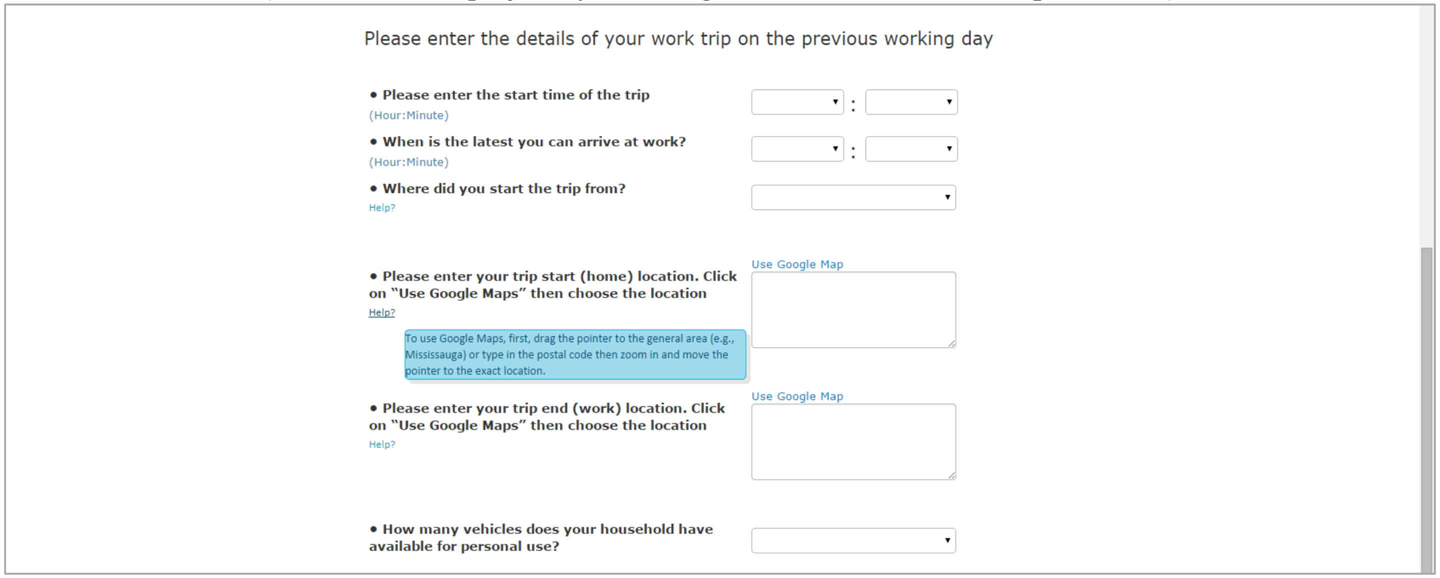
Section A (RP Survey) – Part 1
(Trip details of work trip on the previous working day)



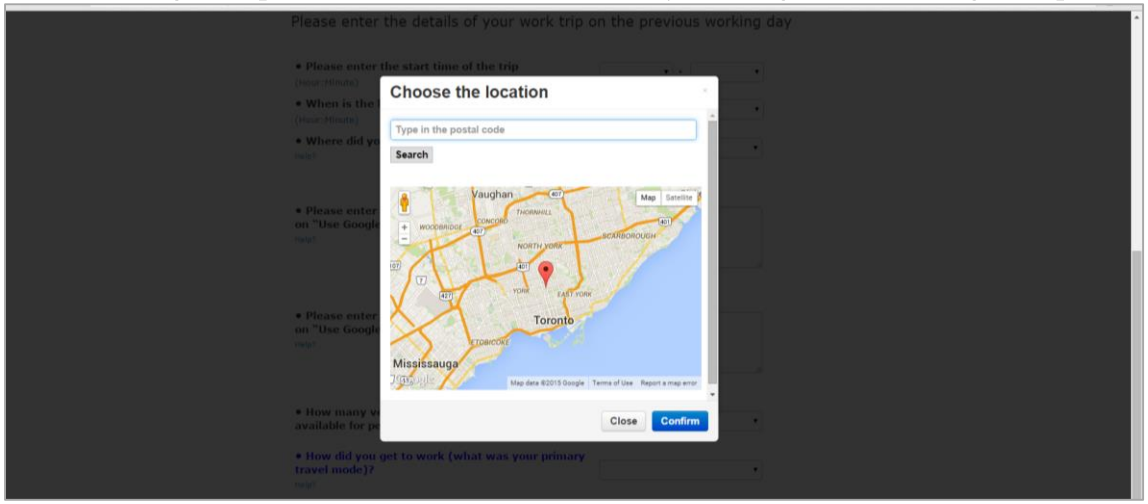
Please enter the details of your work trip on the previous working day

- **Please enter the start time of the trip**
(Hour:Minute)
- **When is the latest you can arrive at work?**
(Hour:Minute)
- **Where did you start the trip from?**
Help?
- **Please enter your trip start (home) location. Click on "Use Google Maps" then choose the location**
Help?
Use Google Map
- **Please enter your trip end (work) location. Click on "Use Google Maps" then choose the location**
Help?
Use Google Map
- **How many vehicles does your household have available for personal use?**

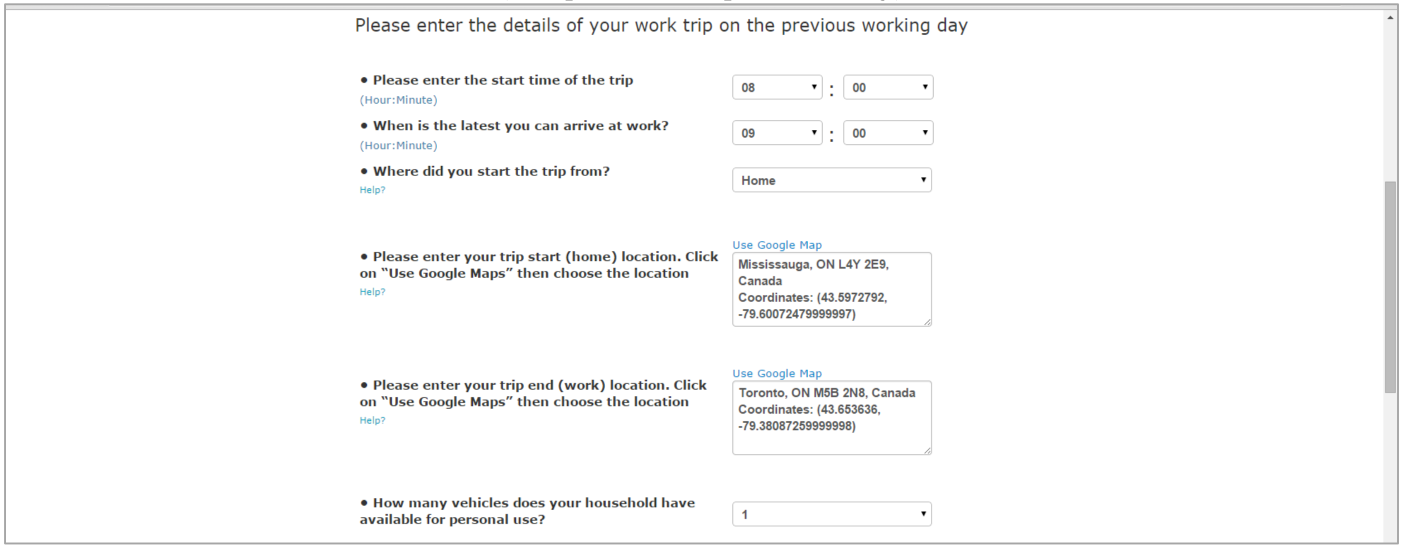
Section A (RP Survey) – Part 1
(Instructions displayed by hovering the mouse over the "Help?" button)



Section A (RP Survey) – Part 1
(Interactive Google Map shows for location selection by clicking on "Use Google Map" button)



Section A (RP Survey) – Part 1
(Sample of a completed survey)



Please enter the details of your work trip on the previous working day

• Please enter the start time of the trip
(Hour:Minute)
08 : 00

• When is the latest you can arrive at work?
(Hour:Minute)
09 : 00

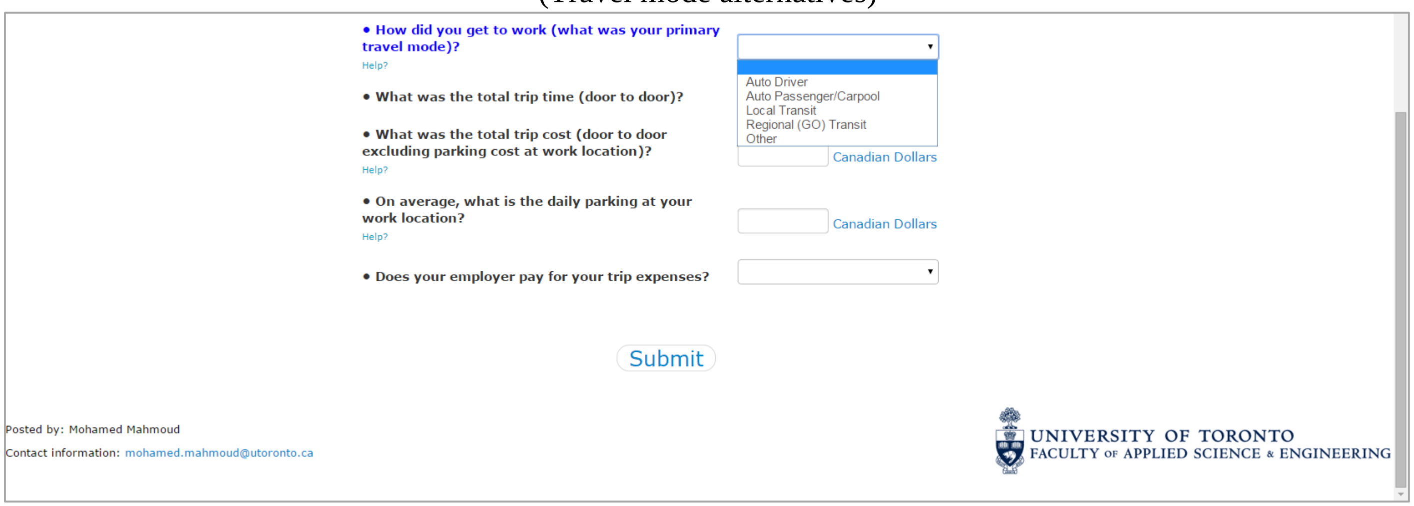
• Where did you start the trip from?
Help?
Home

• Please enter your trip start (home) location. Click on "Use Google Maps" then choose the location
Help?
Use Google Map
Mississauga, ON L4Y 2E9, Canada
Coordinates: (43.5972792, -79.60072479999997)

• Please enter your trip end (work) location. Click on "Use Google Maps" then choose the location
Help?
Use Google Map
Toronto, ON M5B 2N8, Canada
Coordinates: (43.653636, -79.38087259999998)

• How many vehicles does your household have available for personal use?
1

Section A (RP Survey) – Part 2
(Travel mode alternatives)



• How did you get to work (what was your primary travel mode)?
Help?
Auto Driver
Auto Passenger/Carpool
Local Transit
Regional (GO) Transit
Other

• What was the total trip time (door to door)?

• What was the total trip cost (door to door excluding parking cost at work location)?
Help?
Canadian Dollars

• On average, what is the daily parking at your work location?
Help?
Canadian Dollars

• Does your employer pay for your trip expenses?

Submit

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Section A (RP Survey) – Part 2
(Travel mode selection – auto driver mode)

• How did you get to work (what was your primary travel mode)? — Auto Driver
Help?

• What was the total trip time (door to door)? — Minutes

• What was the total trip cost (door to door excluding parking cost at work location)? — Canadian Dollars
Help?

• On average, what is the daily parking at your work location? — Canadian Dollars
Help?

• Does your employer pay for your trip expenses?

Submit

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Section A (RP Survey) – Part 2
(Travel mode selection – auto passenger/carpool mode)

• How did you get to work (what was your primary travel mode)? — Auto Passenger/Carpool
Help?

• How many passengers, including yourself, were using the car?

• What was the total trip time (door to door)? — Minutes

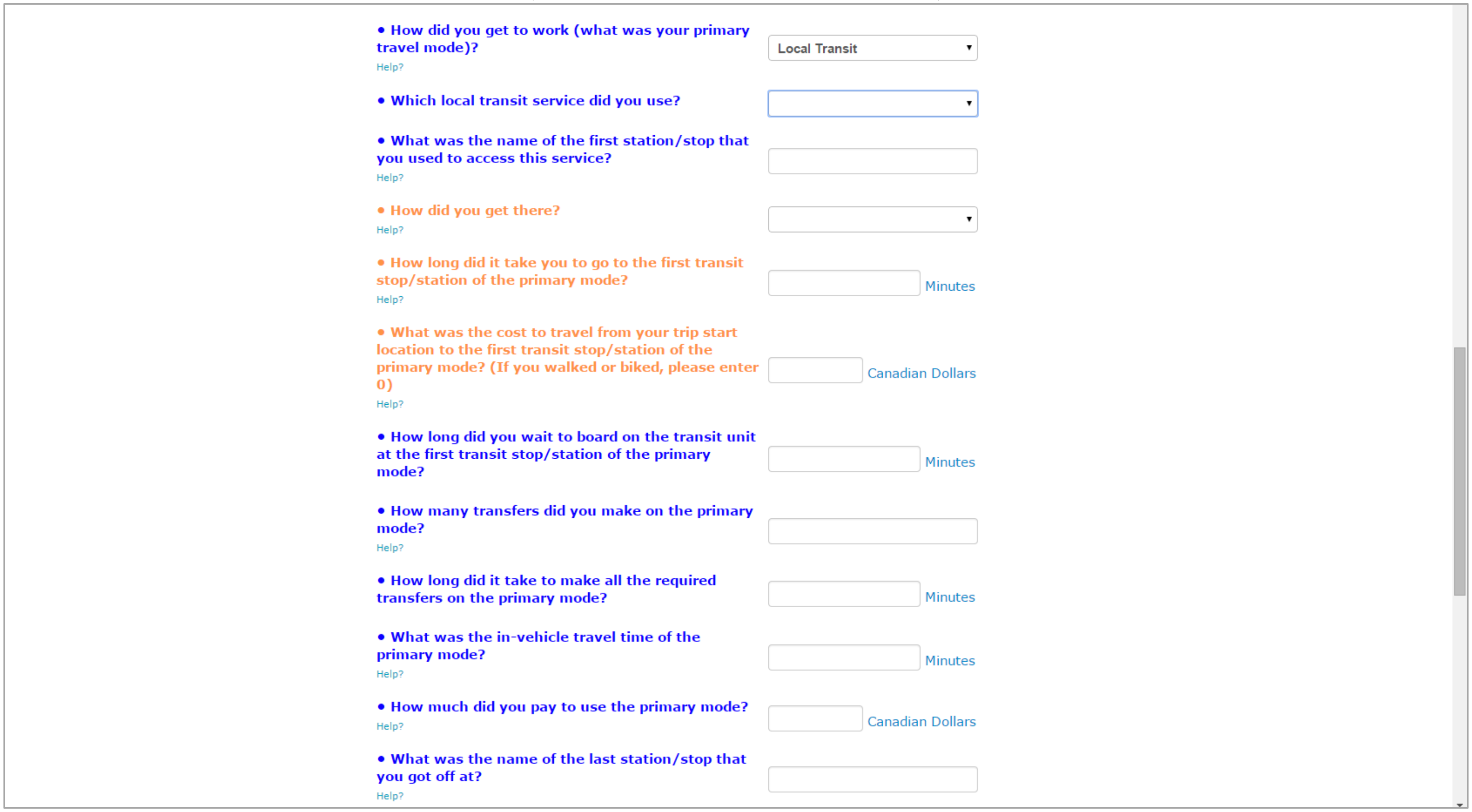
• What was the total trip cost (door to door excluding parking cost at work location)? — Canadian Dollars
Help?

• On average, what is the daily parking at your work location? — Canadian Dollars
Help?

• Does your employer pay for your trip expenses?

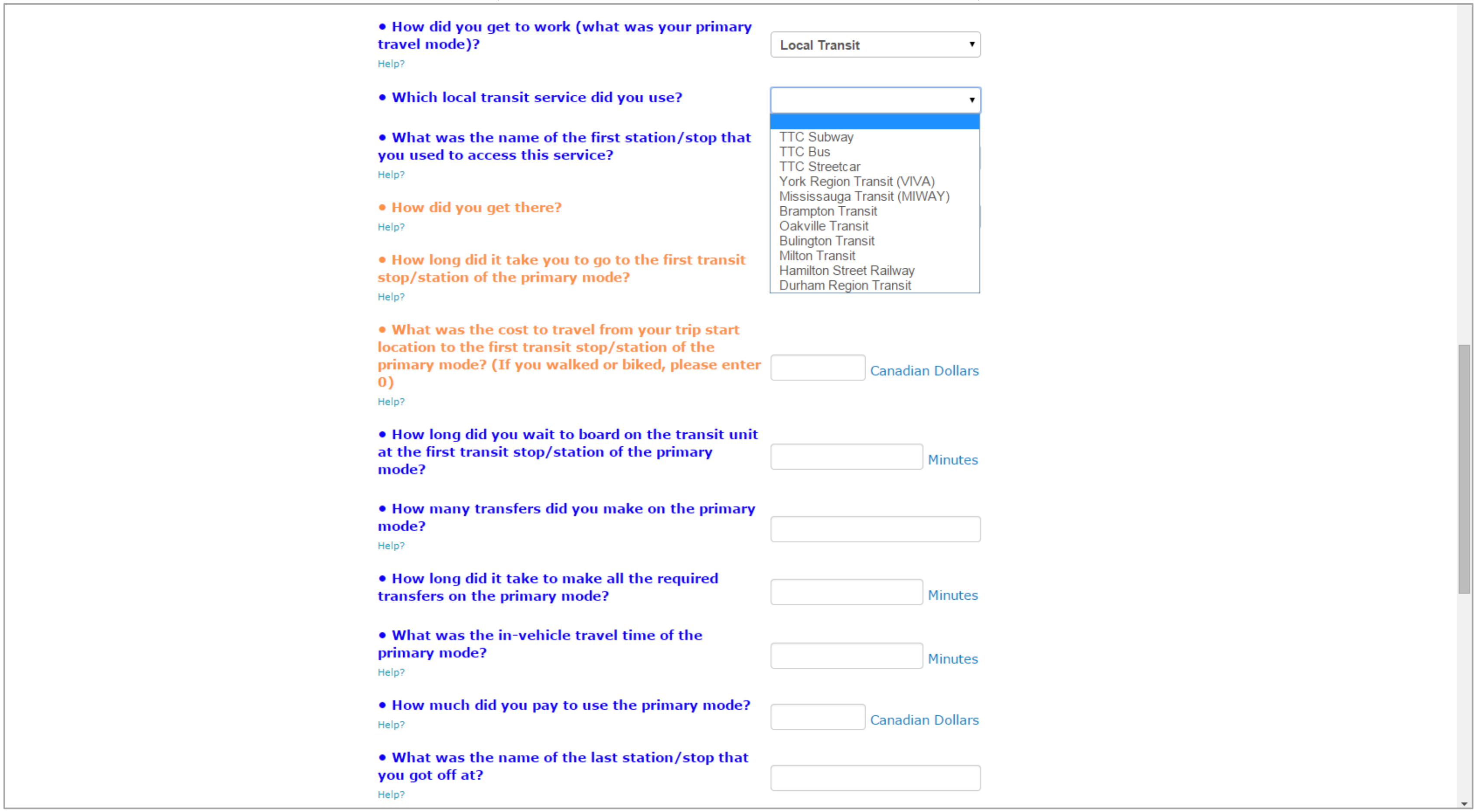
Submit

Section A (RP Survey) – Part 2
(Travel mode selection – local transit)



- How did you get to work (what was your primary travel mode)? Help? — Local Transit
- Which local transit service did you use?
- What was the name of the first station/stop that you used to access this service? Help?
- How did you get there? Help?
- How long did it take you to go to the first transit stop/station of the primary mode? Help? — Minutes
- What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0) Help? — Canadian Dollars
- How long did you wait to board on the transit unit at the first transit stop/station of the primary mode? — Minutes
- How many transfers did you make on the primary mode? Help?
- How long did it take to make all the required transfers on the primary mode? — Minutes
- What was the in-vehicle travel time of the primary mode? Help? — Minutes
- How much did you pay to use the primary mode? Help? — Canadian Dollars
- What was the name of the last station/stop that you got off at? Help?

Section A (RP Survey) – Part 2
(Travel mode selection – local transit cont'd)



• How did you get to work (what was your primary travel mode)? Local Transit
Help?

• Which local transit service did you use?

TTC Subway
TTC Bus
TTC Streetcar
York Region Transit (VIVA)
Mississauga Transit (MIWAY)
Brampton Transit
Oakville Transit
Bulington Transit
Milton Transit
Hamilton Street Railway
Durham Region Transit

• What was the name of the first station/stop that you used to access this service?
Help?

• How did you get there?
Help?

• How long did it take you to go to the first transit stop/station of the primary mode?
Help?

• What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0) Canadian Dollars
Help?

• How long did you wait to board on the transit unit at the first transit stop/station of the primary mode? Minutes

• How many transfers did you make on the primary mode?
Help?

• How long did it take to make all the required transfers on the primary mode? Minutes

• What was the in-vehicle travel time of the primary mode? Minutes
Help?

• How much did you pay to use the primary mode? Canadian Dollars
Help?

• What was the name of the last station/stop that you got off at?
Help?

Section A (RP Survey) – Part 2
(Travel mode selection – local transit cont'd)

• **How did you get to work (what was your primary travel mode)?** Local Transit

Help?

• **Which local transit service did you use?** TTC Subway

• **What was the name of the first station/stop that you used to access this service?**

Help?

• **How did you get there?**

Help?

- Walk
- Park-and-Ride
- Kiss-and-Ride (as a Driver or Passenger)
- Carpool-and-Ride
- Bike
- Local Transit
- Other

• **How long did it take you to go to the first transit stop/station of the primary mode?**

Help?

• **What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0)** Canadian Dollars

Help?

• **How long did you wait to board on the transit unit at the first transit stop/station of the primary mode?** Minutes

• **How many transfers did you make on the primary mode?**

Help?

• **How long did it take to make all the required transfers on the primary mode?** Minutes

• **What was the in-vehicle travel time of the primary mode?** Minutes

Help?

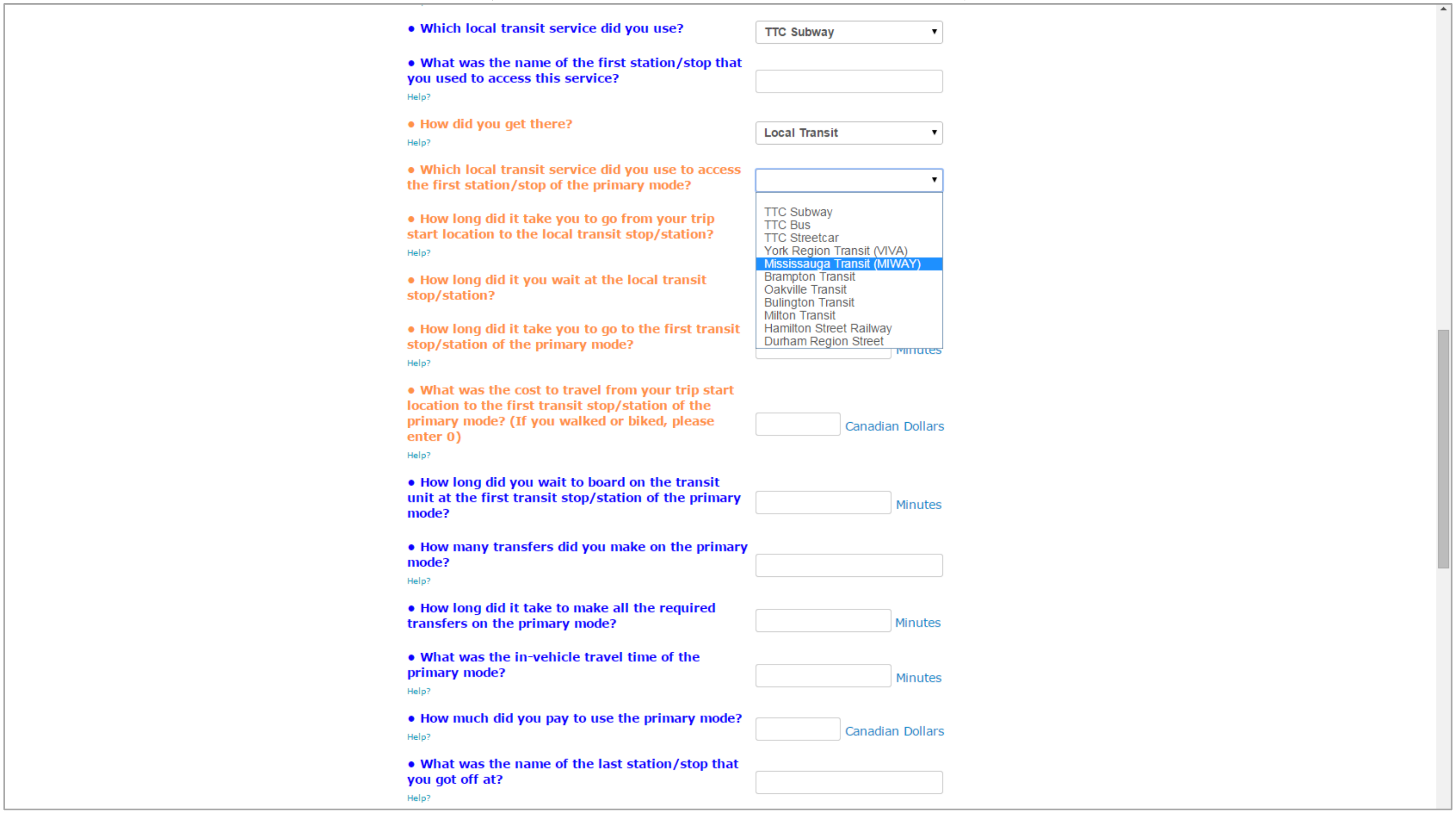
• **How much did you pay to use the primary mode?** Canadian Dollars

Help?

• **What was the name of the last station/stop that you got off at?**

Help?

Section A (RP Survey) – Part 2
(Travel mode selection – local transit cont'd)



• Which local transit service did you use? TTC Subway

• What was the name of the first station/stop that you used to access this service?
Help?

• How did you get there? Local Transit
Help?

• Which local transit service did you use to access the first station/stop of the primary mode?

TTC Subway
TTC Bus
TTC Streetcar
York Region Transit (VIVA)
Mississauga Transit (MIWAY)
Brampton Transit
Oakville Transit
Bulington Transit
Milton Transit
Hamilton Street Railway
Durham Region Street

• How long did it take you to go from your trip start location to the local transit stop/station?
Help?

• How long did it you wait at the local transit stop/station?

• How long did it take you to go to the first transit stop/station of the primary mode? Minutes
Help?

• What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0) Canadian Dollars
Help?

• How long did you wait to board on the transit unit at the first transit stop/station of the primary mode? Minutes

• How many transfers did you make on the primary mode?
Help?

• How long did it take to make all the required transfers on the primary mode? Minutes

• What was the in-vehicle travel time of the primary mode? Minutes
Help?

• How much did you pay to use the primary mode? Canadian Dollars
Help?

• What was the name of the last station/stop that you got off at?
Help?

Section A (RP Survey) – Part 2
(Travel mode selection – local transit cont'd)

- **How did you go to your work location from there?**
  Help?
- **How long did it take you to go to your work location from there?** Minutes
- **What was the cost to travel from the last transit stop/station of the primary mode to your work location? (If you walked or biked, please enter 0)** Canadian Dollars
- **What was the total trip time (door to door)?** Minutes
- **What was the total trip cost (door to door excluding parking cost at work location)?** Canadian Dollars
  Help?
- **How do you typically pay your transit fare?**
  Help?
- **On average, what is the daily parking at your work location?** Canadian Dollars
  Help?
- **Does your employer pay for your trip expenses?**

Submit

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY of APPLIED SCIENCE & ENGINEERING

Section A (RP Survey) – Part 2
(Travel mode selection – local transit cont'd)

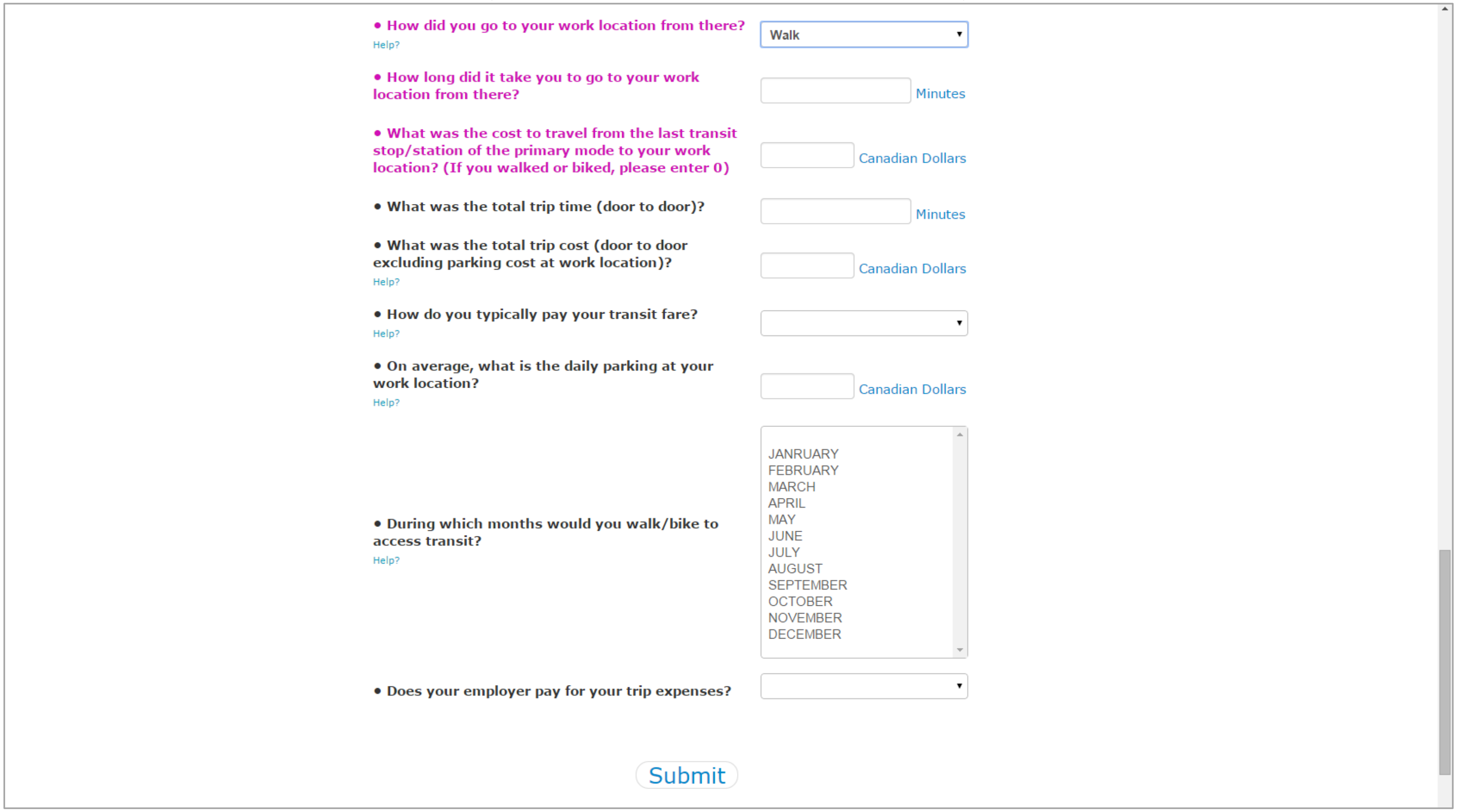
- **How did you go to your work location from there?**
  Help?
  Walk
  Bike
  Local Transit
  Other
- **How long did it take you to go to your work location from there?**
- **What was the cost to travel from the last transit stop/station of the primary mode to your work location? (If you walked or biked, please enter 0)** Canadian Dollars
- **What was the total trip time (door to door)?** Minutes
- **What was the total trip cost (door to door excluding parking cost at work location)?** Canadian Dollars
  Help?
- **How do you typically pay your transit fare?**
  Help?
- **On average, what is the daily parking at your work location?** Canadian Dollars
  Help?
- **Does your employer pay for your trip expenses?**

Submit

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Section A (RP Survey) – Part 2
(Travel mode selection – local transit cont'd)



• How did you go to your work location from there? Walk
Help?

• How long did it take you to go to your work location from there? Minutes

• What was the cost to travel from the last transit stop/station of the primary mode to your work location? (If you walked or biked, please enter 0) Canadian Dollars

• What was the total trip time (door to door)? Minutes

• What was the total trip cost (door to door excluding parking cost at work location)? Canadian Dollars
Help?

• How do you typically pay your transit fare?
Help?

• On average, what is the daily parking at your work location? Canadian Dollars
Help?

• During which months would you walk/bike to access transit?
Help?

JANRUARY
FEBRUARY
MARCH
APRIL
MAY
JUNE
JULY
AUGUST
SEPTEMBER
OCTOBER
NOVEMBER
DECEMBER

• Does your employer pay for your trip expenses?

Submit

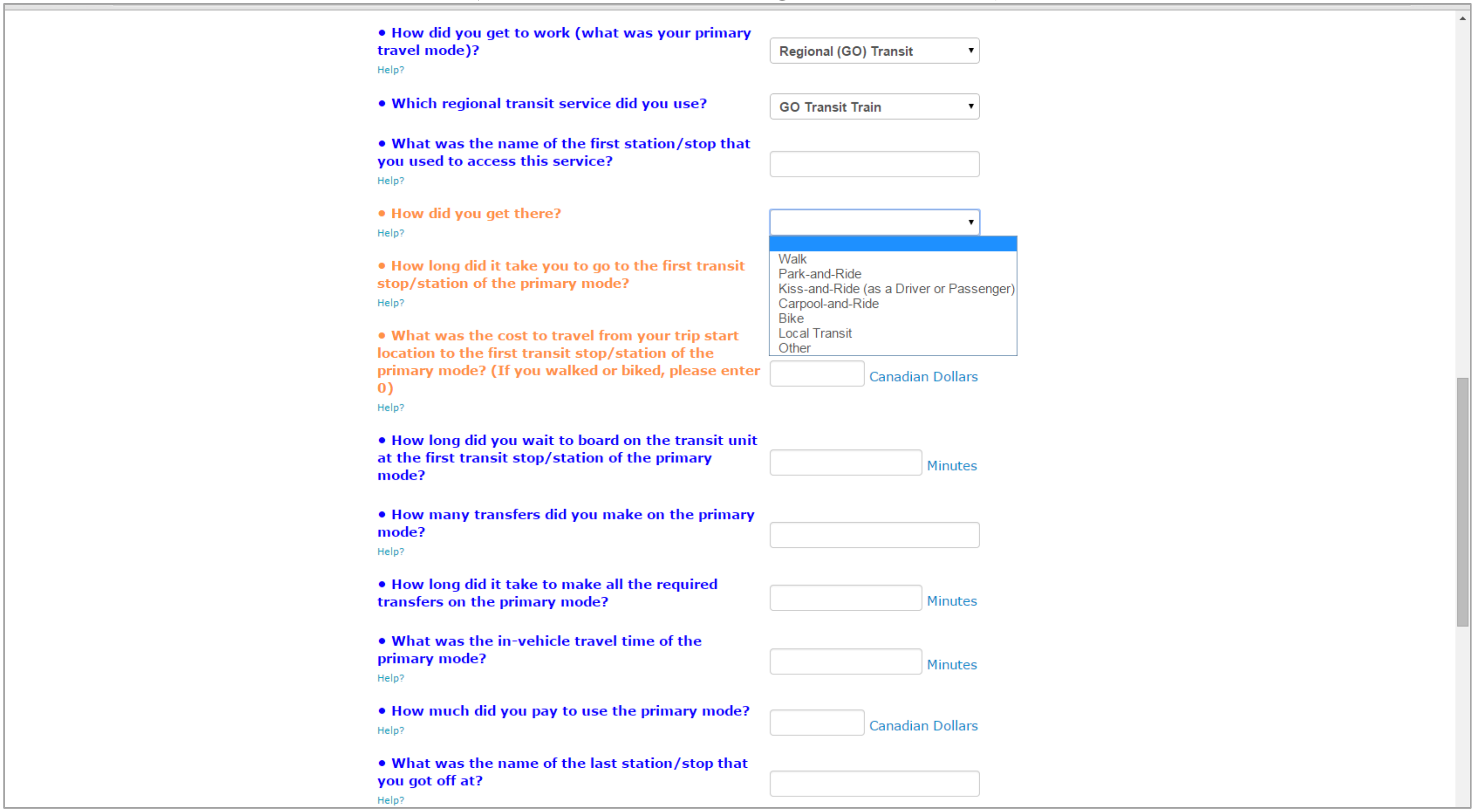
Section A (RP Survey) – Part 2
(Travel mode selection – regional transit)

- How did you get to work (what was your primary travel mode)? — Regional (GO) Transit
  Help?
- Which regional transit service did you use?
- What was the name of the first station/stop that you used to access this service?
  Help?
- How did you get there?
  Help?
- How long did it take you to go to the first transit stop/station of the primary mode? — Minutes
  Help?
- What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0) — Canadian Dollars
  Help?
- How long did you wait to board on the transit unit at the first transit stop/station of the primary mode? — Minutes
- How many transfers did you make on the primary mode?
  Help?
- How long did it take to make all the required transfers on the primary mode? — Minutes
- What was the in-vehicle travel time of the primary mode? — Minutes
  Help?
- How much did you pay to use the primary mode? — Canadian Dollars
  Help?
- What was the name of the last station/stop that you got off at?
  Help?

Section A (RP Survey) – Part 2
(Travel mode selection – regional transit cont'd)

- How did you get to work (what was your primary travel mode)? — Regional (GO) Transit
  Help?
- Which regional transit service did you use? — GO Transit Bus / GO Transit Train
- What was the name of the first station/stop that you used to access this service?
  Help?
- How did you get there?
  Help?
- How long did it take you to go to the first transit stop/station of the primary mode? — Minutes
  Help?
- What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0) — Canadian Dollars
  Help?
- How long did you wait to board on the transit unit at the first transit stop/station of the primary mode? — Minutes
- How many transfers did you make on the primary mode?
  Help?
- How long did it take to make all the required transfers on the primary mode? — Minutes
- What was the in-vehicle travel time of the primary mode? — Minutes
  Help?
- How much did you pay to use the primary mode? — Canadian Dollars
  Help?
- What was the name of the last station/stop that you got off at?
  Help?

Section A (RP Survey) – Part 2
(Travel mode selection – regional transit cont'd)



• How did you get to work (what was your primary travel mode)? — Regional (GO) Transit
Help?

• Which regional transit service did you use? — GO Transit Train

• What was the name of the first station/stop that you used to access this service?
Help?

• How did you get there?
Help?

- Walk
- Park-and-Ride
- Kiss-and-Ride (as a Driver or Passenger)
- Carpool-and-Ride
- Bike
- Local Transit
- Other

• How long did it take you to go to the first transit stop/station of the primary mode?
Help?

• What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0) — Canadian Dollars
Help?

• How long did you wait to board on the transit unit at the first transit stop/station of the primary mode? — Minutes

• How many transfers did you make on the primary mode?
Help?

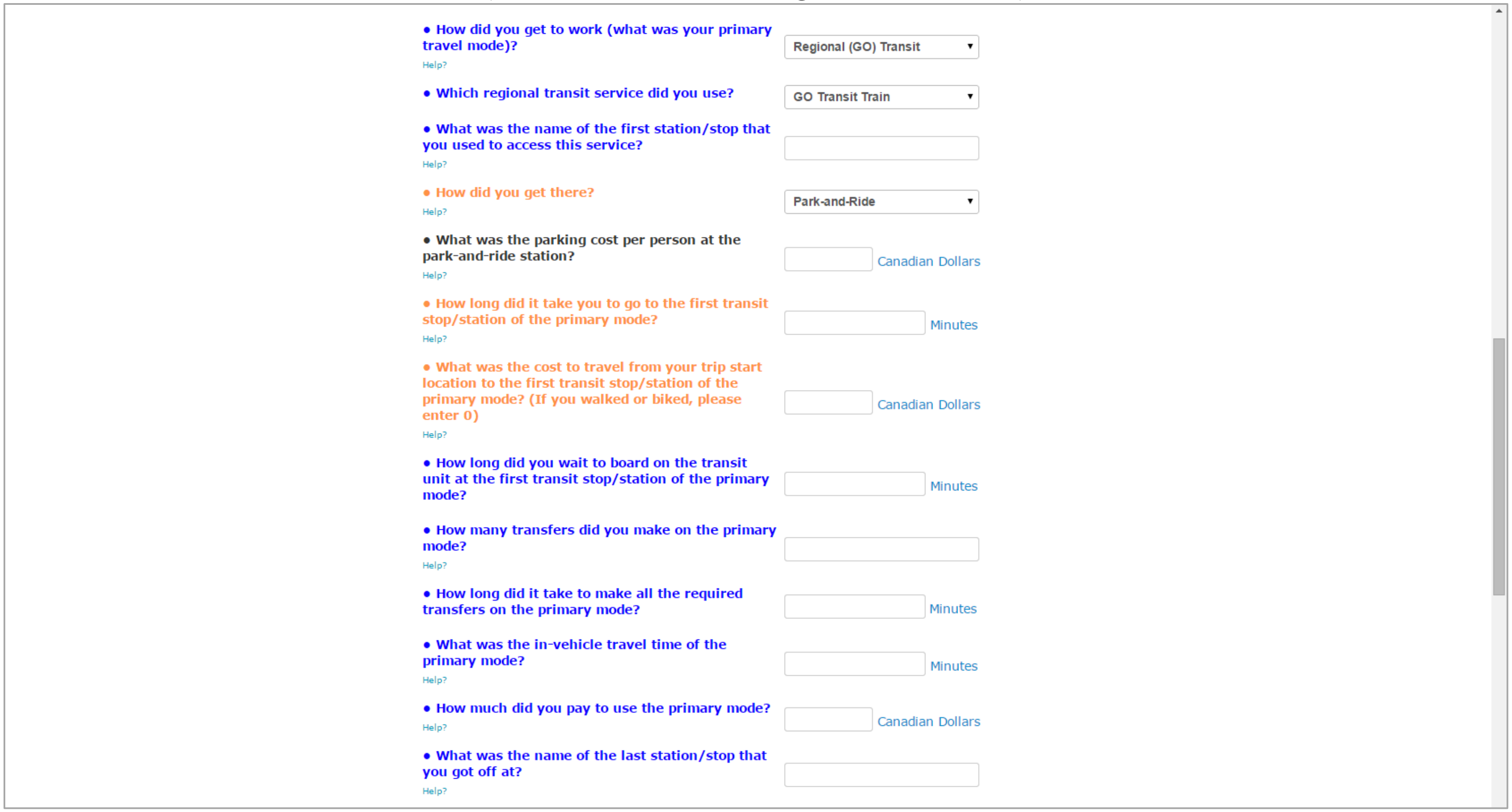
• How long did it take to make all the required transfers on the primary mode? — Minutes

• What was the in-vehicle travel time of the primary mode? — Minutes
Help?

• How much did you pay to use the primary mode? — Canadian Dollars
Help?

• What was the name of the last station/stop that you got off at?
Help?

Section A (RP Survey) – Part 2
(Travel mode selection – regional transit cont'd)



• How did you get to work (what was your primary travel mode)? — Regional (GO) Transit
Help?

• Which regional transit service did you use? — GO Transit Train

• What was the name of the first station/stop that you used to access this service?
Help?

• How did you get there? — Park-and-Ride
Help?

• What was the parking cost per person at the park-and-ride station? — Canadian Dollars
Help?

• How long did it take you to go to the first transit stop/station of the primary mode? — Minutes
Help?

• What was the cost to travel from your trip start location to the first transit stop/station of the primary mode? (If you walked or biked, please enter 0) — Canadian Dollars
Help?

• How long did you wait to board on the transit unit at the first transit stop/station of the primary mode? — Minutes

• How many transfers did you make on the primary mode?
Help?

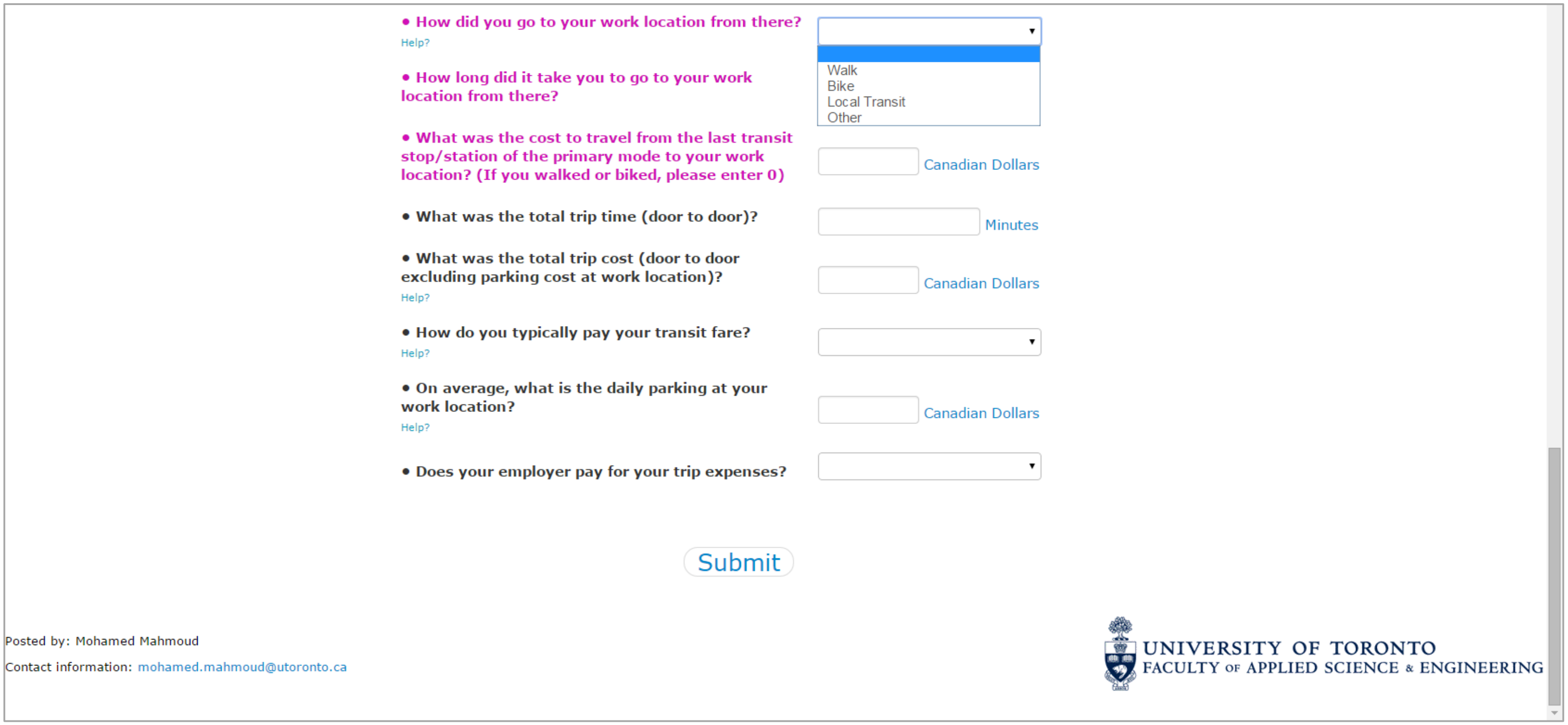
• How long did it take to make all the required transfers on the primary mode? — Minutes

• What was the in-vehicle travel time of the primary mode? — Minutes
Help?

• How much did you pay to use the primary mode? — Canadian Dollars
Help?

• What was the name of the last station/stop that you got off at?
Help?

Section A (RP Survey) – Part 2
(Travel mode selection – regional transit cont'd)



• How did you go to your work location from there?
Help?

Walk
Bike
Local Transit
Other

• How long did it take you to go to your work location from there?

• What was the cost to travel from the last transit stop/station of the primary mode to your work location? (If you walked or biked, please enter 0) Canadian Dollars

• What was the total trip time (door to door)? Minutes

• What was the total trip cost (door to door excluding parking cost at work location)? Canadian Dollars
Help?

• How do you typically pay your transit fare?
Help?

• On average, what is the daily parking at your work location? Canadian Dollars
Help?

• Does your employer pay for your trip expenses?

Submit

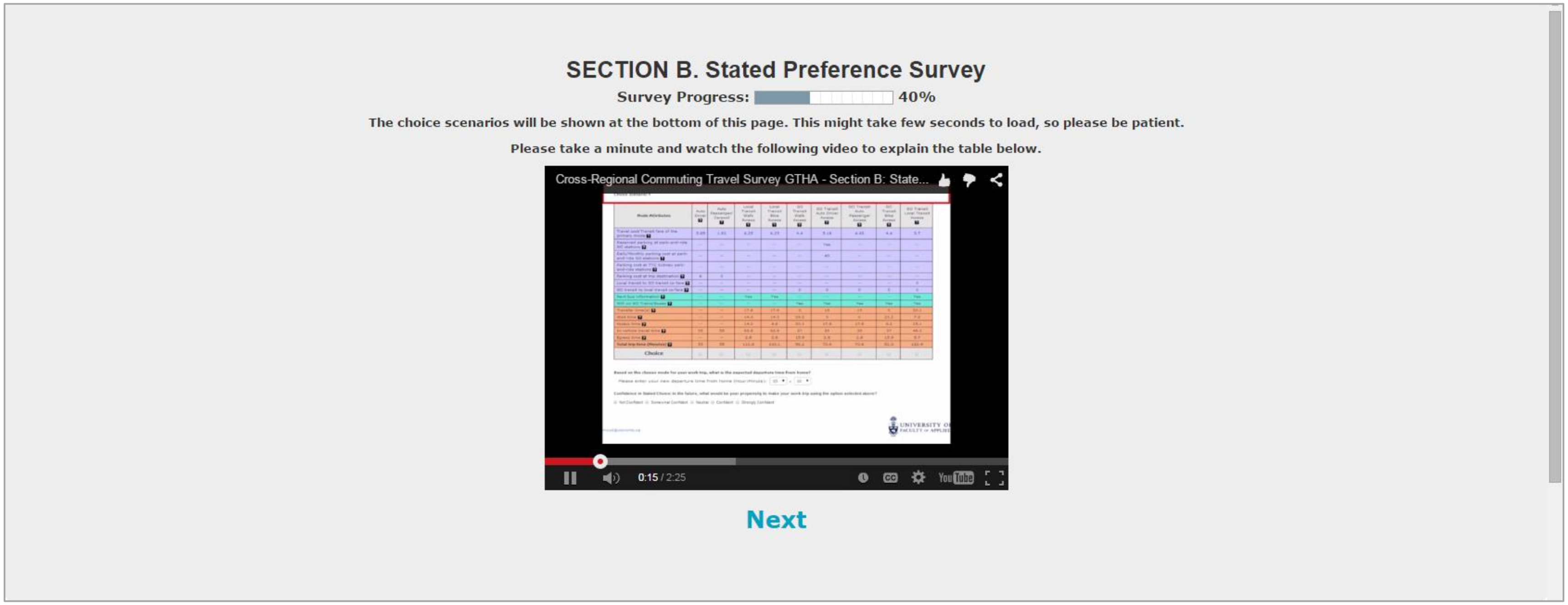
Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

SECTION B

Introduction to Section B (SP Survey)

(Instructional video with a walkthrough example of the SP experiment)



Section B (SP Survey)
(Example of a typical choice scenario)

Choice Scenario: 1

Mode Attributes	Auto Driver Help?	Auto Passenger/ Carpool Help?	Local Transit - Walk Access Help?	TTC Subway - Auto Driver Access Help?	TTC Subway - Auto Passenger Access Help?	GO Transit - Auto Driver Access Help?	GO Transit - Auto Passenger Access Help?	GO Transit - Local Transit Access Help?
Travel cost/Transit fare of the primary mode Help?	5.28	2.64	6.25	5.29	4.45	5.88	5.58	5.28
Reserved parking at park-and-ride GO stations Help?	--	--	--	--	--	No	--	--
Daily/Monthly parking cost at park-and-ride GO stations Help?	--	--	--	--	--	0.0	--	--
Parking cost at TTC Subway park-and-ride stations Help?	--	--	--	5	--	--	--	--
Parking cost at trip destination Help?	15	7.5	--	--	--	--	--	--
Local transit to GO transit co-fare (access) Help?	--	--	--	--	--	--	--	0.7
GO transit to local transit co-fare (egress) Help?	--	--	--	--	--	0.0	1.50	0.00
Next bus information Help?	--	--	No	Yes	Yes	--	--	No
Wi-Fi on GO Trains/Buses Help?	--	--	--	--	--	No	Yes	No
Transfer time(s) Help?	--	--	9	4	4	4	4	7
Wait time Help?	--	--	10	2	2	3	3	4
Access time Help?	--	--	10	11	11	5	5	22
In-vehicle travel time Help?	42	42	52	34	34	28	28	50
Egress time Help?	--	--	3	3	3	3	3	14
Total trip time (Minutes) Help?	**42**	**42**	**84**	**54**	**54**	**43**	**43**	**97**
Choice	☐	☐	☐	☐	☐	☐	☐	☐

Based on the chosen mode for your work trip, what is the expected departure time from home?

Please enter your new departure time from home (Hour:Minute): 00 : 00

In the future, what would be your propensity to make your work trip using the option selected above?

☐ **Not Confident** ☐ **Somewhat Confident** ☐ **Neutral** ☐ **Confident** ☐ **Strongly Confident**

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Section B (SP Survey)
(Instructions displayed by hovering the mouse over the “Help?” button)

Choice Scenario: 1

Mode Attributes	Auto Driver Help?	Auto Passenger/ Carpool Help?	Local Transit - Walk Access Help?	TTC Subway - Auto Driver Access Help?	TTC Subway - Auto Passenger Access Help?	GO Transit - Auto Driver Access Help?	GO Transit - Auto Passenger Access Help?	GO Transit - Local Transit Access Help?
Travel cost/Transit fare of the primary mode Help?	5.28	2.64	6.25	5.29	4.45	5.88	5.58	5.28
Reserved parking at park-and-ride GO stations Help?	--	--	--	--	--	No	--	--
Daily/Monthly parking cost at park-and-ride GO stations Help?	--	--	--	--	--	0.0	--	--
Parking cost at TTC Subway park-and-ride stations Help?	--	--	--	5	--	--	--	--
Parking cost at trip destination Help?	15	7.5	--	--	--	--	--	--
Local transit to GO transit co-fare (access) Help?	--	--	--	--	--	--	--	0.7
GO transit to local transit co-fare (egress) Help?	--	--	--	--	--	0.0	1.50	0.00
Next bus information Help?	--	--	No	Yes	Yes	--	--	No
Wi-Fi on GO Trains/Buses Help?					--	No	Yes	No
Transfer time(s) Help?					4	4	4	7
Wait time Help?	--	--	10	2	2	3	3	4
Access time Help?	--	--	10	11	11	5	5	22
In-vehicle travel time Help?	42	42	52	34	34	28	28	50
Egress time Help?	--	--	3	3	3	3	3	14
Total trip time (Minutes) Help?	**42**	**42**	**84**	**54**	**54**	**43**	**43**	**97**
Choice	☐	☐	☐	☐	☐	☐	☐	☐

The availability of Wi-Fi services on regional (GO) Train/Buses. Takes the values of (Yes) for available and (NO) for not available.

Based on the chosen mode for your work trip, what is the expected departure time from home?

Please enter your new departure time from home (Hour:Minute): 00 : 00

In the future, what would be your propensity to make your work trip using the option selected above?

☐ **Not Confident** ☐ **Somewhat Confident** ☐ **Neutral** ☐ **Confident** ☐ **Strongly Confident**

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Section B (SP Survey)
(Example of a complete choice scenario – scenario 1)

Choice Scenario: 1

Mode Attributes	Auto Driver Help?	Auto Passenger/ Carpool Help?	Local Transit - Walk Access Help?	TTC Subway - Auto Driver Access Help?	TTC Subway - Auto Passenger Access Help?	GO Transit - Auto Driver Access Help?	GO Transit - Auto Passenger Access Help?	GO Transit - Local Transit Access Help?
Travel cost/Transit fare of the primary mode Help?	5.28	2.64	6.25	5.29	4.45	5.88	5.58	5.28
Reserved parking at park-and-ride GO stations Help?	--	--	--	--	--	No	--	--
Daily/Monthly parking cost at park-and-ride GO stations Help?	--	--	--	--	--	0.0	--	--
Parking cost at TTC Subway park-and-ride stations Help?	--	--	--	5	--	--	--	--
Parking cost at trip destination Help?	15	7.5	--	--	--	--	--	--
Local transit to GO transit co-fare (access) Help?	--	--	--	--	--	--	--	0.7
GO transit to local transit co-fare (egress) Help?	--	--	--	--	--	0.0	1.50	0.00
Next bus information Help?	--	--	No	Yes	Yes	--	--	No
Wi-Fi on GO Trains/Buses Help?	--	--	--	--	--	No	Yes	No
Transfer time(s) Help?	--	--	9	4	4	4	4	7
Wait time Help?	--	--	10	2	2	3	3	4
Access time Help?	--	--	10	11	11	5	5	22
In-vehicle travel time Help?	42	42	52	34	34	28	28	50
Egress time Help?	--	--	3	3	3	3	3	14
Total trip time (Minutes) Help?	**42**	**42**	**84**	**54**	**54**	**43**	**43**	**97**
Choice	☐	☐	☐	☐	☐	☑	☐	☐

Based on the chosen mode for your work trip, what is the expected departure time from home?

Please enter your new departure time from home (Hour:Minute): 08 : 15

In the future, what would be your propensity to make your work trip using the option selected above?

☐ **Not Confident** ☐ **Somewhat Confident** ☐ **Neutral** ☑ **Confident** ☐ **Strongly Confident**

Next

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Section B (SP Survey)
(Example of a complete choice scenario – scenario 6)

Choice Scenario: 6

Mode Attributes	Auto Driver Help?	Auto Passenger/ Carpool Help?	Local Transit - Walk Access Help?	TTC Subway - Auto Driver Access Help?	TTC Subway - Auto Passenger Access Help?	GO Transit - Auto Driver Access Help?	GO Transit - Auto Passenger Access Help?	GO Transit - Local Transit Access Help?
Travel cost/Transit fare of the primary mode Help?	6.16	3.08	8.13	4.13	3.57	6.8	6.6	5.28
Reserved parking at park-and-ride GO stations Help?	--	--	--	--	--	No	--	--
Daily/Monthly parking cost at park-and-ride GO stations Help?	--	--	--	--	--	8	--	--
Parking cost at TTC Subway park-and-ride stations Help?	--	--	--	5	--	--	--	--
Parking cost at trip destination Help?	15	7.5	--	--	--	--	--	--
Local transit to GO transit co-fare (access) Help?	--	--	--	--	--	--	--	0.50
GO transit to local transit co-fare (egress) Help?	--	--	--	--	--	3	3	0.0
Next bus information Help?	--	--	Yes	Yes	No	--	--	No
Wi-Fi on GO Trains/Buses Help?	--	--	--	--	--	Yes	Yes	Yes
Transfer time(s) Help?	--	--	9	4	4	4	4	21
Wait time Help?	--	--	10	2	2	3	3	4
Access time Help?	--	--	10	11	11	5	5	22
In-vehicle travel time Help?	42	42	52	34	34	28	28	50
Egress time Help?	--	--	3	3	3	3	3	14
Total trip time (Minutes) Help?	**42**	**42**	**84**	**54**	**54**	**43**	**43**	**111**
Choice	☐	☐	☐	☑	☐	☐	☐	☐

Based on the chosen mode for your work trip, what is the expected departure time from home?

Please enter your new departure time from home (Hour:Minute): 08 : 05

In the future, what would be your propensity to make your work trip using the option selected above?

☐ Not Confident ☐ Somewhat Confident ☐ Neutral ☐ Confident ☑ Strongly Confident

Which of these modes did you consider when you made your choices?

- ☑ Auto Driver
- ☐ Auto Passenger/ Carpool
- ☐ Local Transit with Walk Access
- ☑ Local Transit with Auto Driver Access (Park-and-Ride)
- ☐ Local Transit with Auto Passenger Access (Kiss-and-Ride)
- ☑ GO Transit with Auto Driver Access (Park-and-Ride)
- ☐ GO Transit with Auto Passenger Access (Kiss-and-Ride)
- ☐ GO Transit with Local Transit Access

Submit

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

SECTION C

Section C (personal and household information)

SECTION C. Personal Information

Survey Progress: 80%

- **What is your gender?**
- **How old are you?**
- **What is your marital status?**
- **What is your employment/student status?**
- **What is your average work duration per day (hours)?**
- **What is your current occupation?**
- **How many people, including yourself, live in your household?**
- **Do you have a driver's licence?**
- **Do you have a transit pass?**
- **What is your housing (dwelling unit) type?**
- **How many bikes does your household have available for personal use?**
- **How many vehicles does your household have available for personal use?**
- **What is the total household income level in Canadian Dollars before tax per year (gross income)?**

Done

Posted by: Mohamed Mahmoud

Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Section C (personal and household information)
(Example of a complete survey)

SECTION C. Personal Information

Survey Progress: 80%

• What is your gender?	Male
• How old are you?	29
• What is your marital status?	Married
• What is your employment/student status?	Full time
• What is your average work duration per day (hours)?	8
• What is your current occupation?	Professional / Managem
• How many people, including yourself, live in your household?	2
• Do you have a driver's licence?	Yes
• Do you have a transit pass?	Presto Card
• What is your housing (dwelling unit) type?	Apartment
• How many bikes does your household have available for personal use?	1
• How many vehicles does your household have available for personal use?	1
• Please enter the details of the car that you use	
• What is the car type?	Sedan
• What is the car make? (e.g., Ford, Toyota, Chevy, Honda, Volvo)	BMW
• What is the car model? (e.g., Escort, Prius, Camaro, Civic, 240DL)	i3
• What is the year of your car make?	2014
• What is the car fuel type?	Conventional
• What is the total household income level in Canadian Dollars before tax per year (gross income)?	$100,000 and over

Done

Posted by: Mohamed Mahmoud
Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING

Respondents' Feedback

Please rate the level of complexity of the survey:

1 (Very Simple) 2 3 4 5 (Very Complex)

If you have any comments or suggestions for improvements, please let us know:

Submit

Thank you for completing the survey!

For more information, please contact Mohamed Mahmoud at mohamed.mahmoud@utoronto.ca

Posted by: Mohamed Mahmoud

Contact information: mohamed.mahmoud@utoronto.ca

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING