Titre: Title:	Enhancing the Modelling of Travel Demand Using an Activity-Based Approach
Auteur: Author:	Farhana Yasmin
Date:	2016
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Yasmin, F. (2016). Enhancing the Modelling of Travel Demand Using an Activity-Based Approach [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/2117/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/2117/
Directeurs de recherche: Advisors:	Catherine Morency, & Matthew Roorda
Programme: Program:	Génie civil

UNIVERSITÉ DE MONTRÉAL

ENHANCING THE MODELLING OF TRAVEL DEMAND USING AN ACTIVITY-BASED APPROACH

FARHANA YASMIN
DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIAE DOCTOR
(GÉNIE CIVIL)
AVRIL 2016

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

ENHANCING THE MODELLING OF TRAVEL DEMAND USING AN ACTIVITY-BASED APPROACH

présentée par: YASMIN Farhana

en vue de l'obtention du diplôme de : Philosophiae Doctor

a été dûment acceptée par le jury d'examen constitué de :

M. FAROOQ Bilal, Ph. D., président

Mme MORENCY Catherine, Ph. D., membre et directrice de recherche

M. ROORDA Matthew, Ph. D., membre et codirecteur de recherche

M. TRÉPANIER Martin, Ph. D., membre

M. MOHAMMADIAN Abolfazl, Ph. D., membre

DEDICATION

To my dear parents and beloved family

ACKNOWLEDGEMENTS

I have been fortunate to have so many helpful people in my life during this Ph.D. study. Though it will not be enough to express my gratitude in words to all these people for their support, encouragement, and love, I would still like to offer my heartfelt appreciation to these people.

First and foremost, I would like to offer my deep and sincere gratitude to my supervisor, Prof. Catherine Morency for her valuable guidance, timely insights, and endless encouragement throughout this incredible journey. Her invaluable suggestions and confidence in my abilities have greatly inspired and encouraged me to persevere and succeed. The energy and enthusiasm she has for her research was also truly motivational for me.

I would also like to express my sincere appreciation to my co-supervisor, Prof. Matthew Roorda for his valuable advices, encouragement, and constructive feedbacks. I am indebted to him for my significantly improved research skills, especially analytical writing and presenting results. I would also like to acknowledge him for all his efforts to improve the quality of the articles and eventually this dissertation.

I would like to thank my committee members, Prof. Bilal Farooq, Prof. Martin Trépanier, and Prof. Abolfazl Mohammadian for offering their valuable time and for providing thoughtful comments and suggestions to improve this dissertation. I would also like to convey my heartfelt thanks to all of my teachers for providing me with all the knowledge, moral support and precious life lessons.

I also wish to thank Dr. Clarke Wilson for providing me the ClustalG software and quick responses for all of my questions on this software. Special thanks to Professor Bilal Farooq for teaching the short course on the discrete choice modelling in Winter 2014, which inspired me to estimate the models in BIOGEME.

I am grateful to Hubert Verreault for helping me in several occasions, especially with data preparation and micro-level validation analysis, Pierre-Leo Bourbonnais for travel time calculation of transit trips, and Kinan Bahbouh for his constant encouragement and help. I am also grateful to my former and current colleagues at Polytechnique - Amira, Pegah, Mohamed, Hamzeh, Marc-André, Jean-Simon and many others for the motivational discussions, and for all the funs we have had in the last couple of years. Special thanks go to my dear friends, in particular Sadia, Sabiha, Sajeda, Samira, Rifat, Nusrat and Sabreena for their motivation and encouragement throughout this journey.

I would like to express my appreciation to all the groups and organizations for funding which made my Ph.D. work possible. I gratefully acknowledge the Ministère des transports du Québec (Service de la modélisation des systèmes de transport), the Ville de Montréal, the Agence métropolitaine de transport (AMT), and the Société de transport de Montréal (STM) for their support and contributions. My sincere thanks also go to the Mobilité research chair for its support. I also wish to thank the technical committee on O-D travel surveys in the Montreal region for providing access to data for research purposes.

I would also like to thank my dear parents and siblings (Professor Ahsanul Habib, Zinnatul, and Farzana) for their unconditional love, support and encouragement throughout my life.

At last, with all my heart, I would like to thank my beloved husband, Dr. Kabir who has been always there for me in every situation. There will never be a right word to express my appreciation for his unconditional love and support. I also want to express my deep love to my daughter, Suhaila for bringing such joy and happiness in our lives and sharing our dreams.

RÉSUMÉ

Cette thèse vise à enrichir le processus actuel de la modélisation de la demande de transport dans la Grande Région de Montréal (GRM) en utilisant une approche basée sur le modèle d'activités TASHA (*Travel Activity Scheduler for Household Agents*). TASHA a été développé en se basant sur des données provenant des déplacements de l'enquête *Transportation Tomorrow Survey* (TTS) de 1996 pour la Grande Région de Toronto (GRT). Cette recherche vise à appliquer le modèle TASHA dans le contexte montréalais en utilisant l'enquête de déplacements Origine-Destination (O-D) de 2003 et les données du recensement canadien de 2001. TASHA simule, pour un jour typique de la semaine, les horaires quotidiens d'activités (individuelle et combiné) de l'ensemble des personnes dans la région. Cette étude vise à évaluer la transférabilité du modèle TASHA à une autre région métropolitaine en comparant les caractéristiques des activités observées et simulées par TASHA (fréquence d'activité, heure de début, durée et distance) pour cinq activités différentes (travail, étude, magasinage, retour à domicile et autres). Ces comparaisons sont effectuées à trois niveaux d'agrégation : niveau macro pour l'ensemble de la population, niveau méso par segments de population et niveau micro à l'échelle individuelle.

Les résultats obtenus lors de la validation au niveau macro et méso semblent très prometteurs. TASHA permet de reproduire avec succès, les comportements d'activités dans un contexte différent de celui de Toronto, du moins pour les activités régulières (travail, étude). Cependant, TASHA génère de grandes différences pour certaines activités qui sont plus flexibles (magasinage et autres). Bien que la validation au niveau micro fait partie intégrante de la culture de modélisation dans la région de Montréal, la nécessité d'une telle validation pour ce type de modèle n'est pas nécessaire et ne donne pas de bons résultats.

Cette étude cherche également à expliquer les raisons derrière les grandes différences constatées pour certains cas à différents niveaux de résolution. En examinant les différences au niveau spatial, temporel ou autres entre l'île de Montréal et la région de Toronto, le potentiel de transférabilité à différents niveaux (surtout au niveau méso et macro) semblent très prometteurs. Toutefois, nous recommandons la ré-estimation des paramètres du modèle et l'utilisation, si disponible, des distributions locales des attributs des activités (fréquence, heure de début et durée) lors de la transférabilité du modèle TASHA d'un contexte à l'autre.

Les résultats obtenus de TASHA et l'analyse des grandes différences, constatées dans certains cas, nous a permis d'approfondir certains éléments et de proposer certaines amélioration du cadre de la modélisation (par exemple, l'évolution de la production des activités dans le temps et le développement d'une procédure plus systématique pour la segmentation de la population).

Dans TASHA, les distributions observées des caractéristiques des activités (fréquence des différents types d'activité, heure de début et la durée) à partir de 1996 (TTS de la GRT) sont utilisés comme données entrantes dans la composante de génération d'activités. De plus, ces distributions sont supposées constantes dans le temps. Dans cette étude, les mêmes hypothèses sont considérées et donc, les distributions observées des d'activité sont utilisées tout en transférant le modèle TASHA au contexte de l'île de Montréal. Cette recherche examine également l'hypothèse de la stabilité temporelle des attributs de génération d'activité au fil du temps et elle examine empiriquement les changements dans les distributions de génération d'activité (fréquence, heure de début et durée) sur une période de 10 ans en utilisant les enquêtes sur les déplacements O-D 1998, 2003 et 2008 de la GMR. L'analyse des tendances révèle que les distributions d'activité pour le travail, l'étude, le magasinage et autres activités ont considérablement changé dans le temps. Nous suggérons donc d'incorporer ces changements dans les distributions afin de les appliquer dans TASHA. Une telle modification permettrait de mieux refléter les changements temporels des comportements de déplacements dans la GRM.

Dans TASHA, la segmentation de la population utilisée pour développer des distributions d'activité est intuitive. Ce développement n'a pas été basé sur un processus de segmentation systématique. Nous supposons alors qu'une approche de segmentation systématique de la population permettrait d'améliorer le développement des distributions d'attributs d'activité. Cette recherche développe une méthodologie combinant deux méthodes (méthode d'alignement multiple de séquence (SAM) et le modèle logit multinomial (MNL)) pour segmenter la population de l'île de Montréal en utilisant l'enquête O-D de 2003. Premièrement, la méthodologie SAM est appliquée pour segmenter la population en fonction de la similitude des patrons d'activités quotidiens individuels (représentés par trois caractéristiques : la fréquence, l'heure de début de l'activité et la durée). Par la suite, des modèles logit multinomiaux sont estimés pour les segments distincts de la population en fonction de leurs caractéristiques sociodémographiques. Ce modèle simple de segmentation, développé dans cette recherche, peut être intégré dans le processus de modélisation de TASHA ou dans un autre cadre de modélisation basée sur les activités. Le modèle peut également être appliqué pour identifier un groupe spécifique pour simuler leurs calendriers d'activités et/ou de procédé à une enquête détaillée de leurs comportements de déplacements.

L'application d'un modèle basé sur les activités, comme TASHA, dans l'ile de Montréal offre des résultats prometteurs avec quelques limitations. Nous estimons qu'avec quelques améliorations, par exemple l'utilisation des paramètres locaux et des distributions locales d'attributs ainsi que l'intégration d'un nouveau modèle de segmentation sociodémographique, le modèle TASHA pourrait mieux performer dans la simulation des horaires d'activité de la population Montréal.

ABSTRACT

This thesis aims to enhance the current modelling approach of travel demand of the Greater Montreal Area (GMA) using an activity-based approach, TASHA (Travel Activity Scheduler for Household Agents). It is also an effort to contribute to the validation and enhancement in the activity-based modelling framework by demonstrating a validation procedure of activity-based models and proposing some improvements. TASHA has been developed based on trip diary data from the 1996 Transportation Tomorrow Survey (TTS) for the Greater Toronto Area (GTA), Canada. This research applies TASHA in the context of the Island of Montreal, Canada, using the 2003 Origin-Destination (O-D) travel survey and the 2001 Canadian Census. TASHA simulates daily schedules of activities (individual and joint) for all individuals in this region. This research assesses the spatial transferability, as a validation test, of the TASHA model by comparing model simulated to observed activity attributes (activity frequency, start time, duration, and distance) for different activities. The validation has been performed at three different levels, macro-level (aggregation of the entire population), meso-level (aggregation by population segments by age group and gender, and by home location), and micro-level (individuals).

Validation results at different levels (specially at macro- and meso-level) seem quite promising. TASHA can successfully reproduce activity behaviours of another context, at least for fixed activities (work, school) with few exceptions. However, TASHA provides large differences for some activity attributes of flexible activities (shopping, other).

This research also discusses the potential reasons behind the large differences found in some cases at different levels. Considering spatial, temporal and other differences between the Montreal Island and the Toronto Area, the transferability results at different levels (especially at macro- and meso-level) seem quite promising. However in general, we recommend to re-estimate model parameters and to use local activity attributes distributions (frequency, start time and duration) if available, when transferring the TASHA model from one context to another.

The TASHA application results and the analysis of the potential reasons behind the large differences found in some cases at different levels direct us towards further investigations and/or improvements of some elements of the activity-based modelling framework (for instance the evolution of activity generation attributes over time and development of more systematic procedure for population segmentation).

In TASHA, observed distributions of activity attributes (activity frequency, start time and duration) from the 1996 TTS of the GTA are used as inputs in the activity generation component and these distributions are assumed to remain constant over time. We keep the same assumption and thus use these observed distributions of activity attributes while transferring TASHA to the context of the Montreal Island. However, this research further examines the hypothesis of temporal stability of activity generation attributes over time. It empirically investigates the changes in the distributions of activity generation attributes (frequency, start time and duration) over a 10-year period using the 1998, 2003 and 2008 O-D surveys of the GMA. The trend analysis reveals that distributions of activity attributes for work, school, shopping, and other activities are significantly changing over time. We suggest preparing activity attributes for the application of an activity-based model, TASHA, such that they reflect temporal changes in travel behaviour of the GMA.

In TASHA, population segmentation to develop activity attributes distributions was based on intuition and testing. We assume that if we systematically segment the population to develop activity attributes distributions, it may improve the TASHA model performance. This research develops a methodology combining two methods (multiple sequence alignment method (SAM) and multinomial logit model (MNL)) to segment the population of the Montreal Island using the 2003 O-D survey. First, it applies the multiple SAM to segment the population based on the similarities in individuals' daily activity patterns (represented by three activity attributes, frequency, start time and duration). Second, it estimates the MNLs for the distinct segments of individuals based on their socio-demographic characteristics. The simple population segmentation model developed in this research can be integrated in the TASHA modelling or other activity-based modelling frameworks to segment the population to develop activity generation attributes. The model can also be applied to identify a relevant group of interest to simulate their activity schedules and/or conduct detail investigation of their travel behaviours.

The activity-based model (i.e. TASHA) application in Montreal provides some promising results along with some limitations. It is felt that with some improvements (for instance using local parameters and activity attributes distributions, and integrating newly developed population segmentation model), TASHA would perform better in simulating individuals' activity schedules in Montreal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMT	Agence Métropolitaine de Transport
CBD	Central Business District
CHAID	Chi-squared Automatic Interaction Detector
CPM	Computational Process Model
CSA	Connection Scan Algorithm
ES-3	Entrants - Sortants à 3 niveaux
FMS	Flexible Model Structure
GMA	Greater Montreal Area
GTA	Greater Toronto Area
HOV	High Occupancy Vehicle
ICT	Information and Communication Technologies
IIA	Independence from Irrelevant Alternatives
ILUTE	Integrated Land Use, Transportation, Environment
ITS	Intelligent Transportation Systems
K-S test	Kolmogorov-Smirnov Test
LCM	Latent Class Model
MADEDE	Modèle d'Analyse Désagrégée de la Demande
MADITUC	Modèle d'Analyse Désagrégée des Itinéraires de Transport Urbain Collectif

MNL	Multinomial Logit Model
MTQ	Ministère des Transports du Québec
NL	Nested Logit Model
O-D survey	Origin-Destination Survey
SAM	Sequence Alignment Method
SOLA	Second-Order Linear Approximation
TASHA	Travel Activity Scheduler for Household Agents
TAZ	Traffic Analysis Zone
TDM	Travel Demand Management
TTS	Transportation Tomorrow Survey

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

1.1 Background and motivation

Transportation planning involves forecasts of the travel behaviours of the people in a region to provide adequate facilities and services to meet the future demand. Travel demand models have been developed in the 1950s for future travel demand forecasting and employed as decision support tools for transportation planning over the last several decades. For efficient planning, these models must have the ability to respond to the changes in the attributes of the transportation system and the people, who use the system, with changes in policies and strategies in an area.

First generation travel demand models are known as the traditional four-stage models or classical sequential procedures. These trip-based models have been developed to assess capital intensive infrastructure investment projects. However, changes in urban, environmental and energy policy in the 1970s brought a big change in the focus of travel demand forecasting (McNally, 2007a); as energy-constrained environment with increased energy price, energy shortage, conservation policies and so on, and environmental concerns in this era made a shift in supply-oriented focus of transportation planning (Grist & Demetsky, 1981). Besides, due to the increase of different challenges such as urban sprawl, auto dependency, women participation in workforce and congestion, travel behaviour of individuals and households became complex (Shiftan & Ben-Akiva, 2011). The traditional approach has been greatly criticized due to its inability to represent underlying travel behaviours of individuals because of its static and aggregate nature (Beinborn, 1995; McNally, 2007a). Moreover, different policy instruments in the field of transportation planning, namely Travel Demand Management (TDM), Intelligent Transportation Systems (ITS) technology and, High Occupancy Vehicle (HOV) lanes require more efficient decision support tools than traditional four-stage models (Pinjari & Bhat, 2011; Roorda, 2005; Shiftan et al., 2003). Therefore, researchers and practitioners have sought a new approach; they have tried to improve traditional trip-based, aggregate models with the disaggregate approach based on the discrete choice theory since the 1970s. However, these trip-based approaches cannot address the interrelationship between trips and activities, temporal constraints and dependencies of activity scheduling, or the underlying activity behaviours (Roorda, 2005). In addition, the notion that travel demand is a derived demand from participating in out-of-home activities changed the focus of travel demand modelling from trip-based to activity-based. The activity-based approach has been introduced in transportation planning to overcome limitations observed in the traditional model since the 1980s, though the concept emerged in the literature in the 1970s. The evolution of different modelling approaches is shown in Figure 1.1.

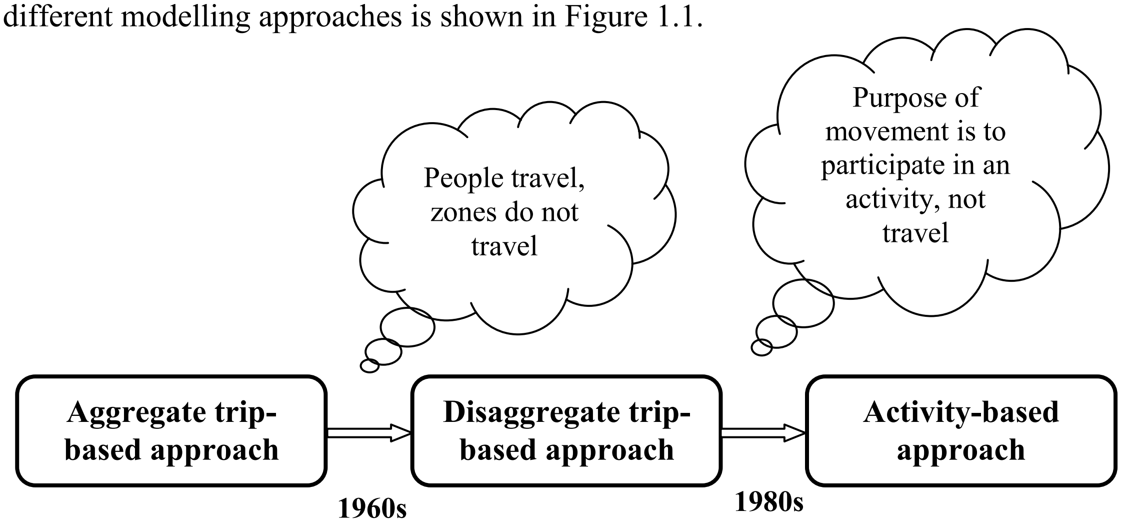


Figure 1.1: Evolution of travel demand modelling approaches

(adapted from (Bhat, 1998a))

The importance of the activity-based approach has already been well recognized in the travel demand analysis over traditional four-stage approach (Bhat & Lawton, 2000; Kitamura, 1996; Kitamura, Pas, Lula, Lawton, & Benson, 1996; Roorda, 2005; Shiftan et al., 2003). However, traditional models are still most commonly used in practice. There must be extensive validation of activity-based models before practical applications in real world contexts. To date, activity-based models are validated using different methods including testing the spatial transferability of the models (Yasmin, Morency, & Roorda, 2015a). However, validation of activity-based models using such a test is not that frequent (Arentze, Hofman, Van Mourik, & Timmermans, 2002; Bowman, Bradley, Castiglione, & Yoder, 2014). Also from a practical point of view, a model with good transferability would save time, cost and expertise required to develop a model for a new context; if the model can be implemented in the context.

In Montreal, Canada, despite its rich datasets from extensive Origin-Destination (O-D) travel surveys (Agence Métropolitaine de Transport [AMT], 2015), the modelling tools (which are a combination of aggregate and disaggregate approaches) used by the Ministère des transports du Québec (MTQ) in the Greater Montreal Area (GMA) are still trip-based (Ministère des transports du Québec [MTQ], 2014; Tremblay, 2007). There has been no attempt to develop or apply an activity-based modelling framework in this region. Therefore, this PhD research is an effort to enhance the current approach of modelling of travel demand of the GMA using an activity-based approach, TASHA (Travel Activity Scheduler for Household Agents) by applying the model in this region. TASHA, a disaggregate model, simulates activity schedules and travel patterns for a 24-hour typical weekday for all individuals in a household. This model has been developed by Miller and Roorda (2003) for the Greater Toronto Area (GTA), Canada using the trip diary data from the 1996 Transportation Tomorrow Survey (TTS) of this region. This dataset is very similar to that of the Montreal O-D survey allowing an opportunity to apply TASHA in the context of Montreal and to test its spatial transferability.

Activity-based models are coming into practice in various parts of North America in recent years (Bradley & Bowman, 2006; Davidson, Vovsha, Freedman, & Donnelly, 2010; Miller, Vaughan, King, & Austin, 2015; Shiftan & Ben-Akiva, 2011). Researchers have argued that TASHA, like other activity-based models, will also be the next generation model with some validations and modifications. There are two major components, activity generation and activity scheduling in the TASHA modelling framework (Roorda, Miller, & Habib, 2008). Activity generation in TASHA is based on the observed distributions of activity attributes (frequency of different types of activity, start time and duration) from the 1996 TTS of the GTA (Miller & Roorda, 2003). A total of 262 frequency distributions of the activity attributes are developed from the 1996 TTS of the GTA, cross-classified by activity type, individual, household and schedule attributes (Roorda, 2005). TASHA assumes that these distributions of activity attributes remain constant over time. In the Toronto area, the validation results indicate that TASHA successfully reproduces activity/travel patterns in the GTA, at least for short-term forecasts, as activity attributes distributions have remained quite stable (thus, temporally transferable) (Roorda et al., 2008). However, this assumption of temporal stability of the activity generation attributes over time may not be the case for the Montreal region. Additionally in TASHA, the population segmentation, using different variables, for each activity type to develop activity attributes distributions has been done based on intuition and testing; the selection of variables to generate activities is not through a systematic process and these segments are fixed over time. Thus, this PhD research also focuses on the evolution of the activity generation behaviours in the Montreal region. It also proposes an approach of segmenting the population to develop activity attributes distributions by combining a clustering approach and discrete choice method. The ultimate goal is to contribute to the validation and enhancement in the activity-based modelling framework (such as TASHA) to increase the practical application of such model in the real world contexts.

1.2 Objectives

This research has two goals. The first is to enhance the current modelling approach of travel demand of the Greater Montreal Area (GMA) using an activity-based approach. The second is to validate and enhance the activity-based modelling framework. To fulfill the main goals of this thesis, specific objectives are as follows:

- To apply an activity-based travel demand model called TASHA (Travel Activity Scheduler for Household Agents) in the context of Montreal.
- To compare TASHA simulated to observed activity attributes (activity frequency, start time, duration, and distance) at three different levels of aggregation (macro-, meso-, and micro-level).
- To empirically examine daily activity generation behaviours and their evolutions over time in Montreal.
- To propose a systematic way of segmenting the population, using a clustering approach and discrete choice method, to develop activity generation attributes distributions for an activity-based travel demand model.

1.3 Original contributions

The major contributions of this thesis are as follows:

- Empirical assessment of the spatial transferability, as a validation test, of an activity-based travel demand model, TASHA (Travel Activity Scheduler for Household Agents).
- Demonstration of a validation procedure of the activity-based travel demand models.
- Examination of the changes in the distributions of activity generation attributes (activity frequency, start time and duration) over time in the Greater Montreal Area (GMA) and investigation of the possible reasons for these changes over time.
- Development of a methodology to segment the population to develop activity generation attributes distributions which are the input of an activity-based travel demand model.

- Propose how to enhance both the current process of modelling of travel demand of Montreal and the activity-based modelling framework.

1.4 Thesis structure

This thesis consists of nine chapters.

Chapter 1 presents the purpose and motivation behind this research, the objectives, and the original research contributions.

Chapter 2 provides a review of relevant literature to this research work and identifies existing research gaps and limitations.

Chapter 3 presents the methodology followed to achieve the objectives of this research. Most of the results obtained from this research are presented in the form of scientific papers (published or, submitted) in the following chapters (Chapter 4, 5, and 6). The research on population segmentation process, presented in Chapter 7 will be submitted in publication in the near future.

Chapter 4 presents the article entitled "Assessment of spatial transferability of an activity-based model, TASHA". This article has been published in the journal "Transportation Research Part A: Policy and Practice".

Chapter 5 includes the article titled "Macro-, meso-, and micro-level validation of an activity-based travel demand model", which has been submitted for publication in the journal "Transportmetrica A: Transport Science".

Chapter 6 presents the article entitled "Trend analysis of activity generation attributes over time". It has also been published in the journal "Transportation".

Chapter 7, titled "Population segmentation based on similarity in activity patterns", presents a methodology to segment the population to develop activity generation attributes distributions; this research will be submitted for publication in the near future.

Chapter 8 provides general discussions on research findings.

Chapter 9 summarises the most important conclusions of the thesis and presents research contributions, limitations and directions for future research in this area.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

As the main goals of this research are to enhance the current approach of modelling of travel demand using an activity-based approach in the GMA as well as to validate and enhance the activity-based travel demand modelling framework, an extensive literature review on different travel demand modelling approaches is provided in this chapter. The discussion focuses on traditional four-stage models, disaggregate trip-based models using discrete choice methods, activity-based models and transport models used by the MTQ in Montreal. It identifies the benefits and limitations of these modelling approaches. Then, a detailed discussion of how activity-based models have been validated is presented. Next, an overview of the TASHA modelling framework with focus on its major features, conceptual framework, assumptions, and validation procedure is provided. Then, a comparison between the traditional models, the transport models used by the MTQ and the activity-based model TASHA is presented. Next, a detailed discussion on the methods, namely sequence alignment method as a clustering approach and multinomial logit model is provided. To propose a methodology of segmenting the population to develop the activity attributes distributions, these methods are used. Finally, a summary of existing research gaps and limitations is presented.

2.2 Overview of travel demand models

Travel demand forecasting is an essential component of transportation planning and policy analysis. A travel-demand model is defined as

> "*a mathematical relationship between travel-demand flows and their characteristics on the one hand, and given activity and transportation supply systems and their characteristics*".
>
> (Cascetta, 2009, p. 169)

A large number of travel demand models have been developed based on different assumptions and using different approaches to forecast the future demand since the 1950s. Cascetta (2009) has classified travel demand models based on different criteria. According to the level of detail, travel demand models have been classified into two broad categories:

a) aggregate travel demand models and
b) disaggregate travel demand models

The first generation travel demand models developed in 1950s are aggregate in nature. In aggregate models, the attributes are compiled for a group of users (for instance, the average travel times or costs of all trips between two zones). In disaggregate travel demand models, the attributes are considered at the individual level (for instance, the travel times or costs of between actual origin and destination points of a trip).

Based on the sequence of choices, Cascetta (2009) has also classified the travel demand models into three broad classes:

a) trip-based models
b) trip-chaining models and
c) activity-based models

Trip-based models assume that each origin-destination trip decision is taken independently without considering the interrelationship between the choice attributes (for instance time, destination, and mode) of different trips. Alternatively, trip-chaining models assume that each origin-destination trip decision of a set of consecutive trips, i.e. trip chain (which begin and end at an individual's home (Primerano, Taylor, Pitaksringkarn, & Tisato, 2008)) affects each other. Activity-based models estimate the travel demand as a derived demand from participating in activities at different points in space and time. The following sections describe the sequence of choices in aggregate and disaggregate travel demand models in detail.

2.2.1 Aggregate travel demand models

Aggregate travel demand models have been developed in 1950s and this trip-based approach mainly focuses on the zonal system. Several attributes are aggregated within a geographical zone, which serves as a unit of analysis for estimating travel demand. These first generation aggregate travel demand models are known as the traditional four-stage models or classical sequential procedure. Comprehensive research and several software packages make traditional models convenient to apply in any region. Traditional models have been enhanced and modified since their first development, but in this thesis, we have only focused on the original framework of the traditional model explained in the literature. Figure 2.1 shows the general framework of the traditional models. The model is presented as a sequence of four sub models, namely trip generation, trip distribution, trip mode choice, and trip assignment.

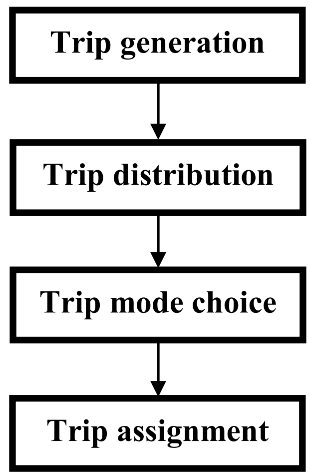


Figure 2.1: General framework of four-stage travel demand models

The four stages are applied sequentially, each with a different mathematical model to predict future travel demand. A brief description of the four different stages is presented below:

- **Trip generation**

Trip generation is the first stage of the traditional four-stage models. In this stage, based on the zonal information from land use, population and economic forecasts of an area, the models estimate the total number of trips produced and attracted in each zone of that area for different trip purposes, such as home-based work trips (work trips that begin or end at home), home-based non-work trips (shopping trips, school trips and so on that begin or end at home), and non-home-based trips (trips that neither begin or end at home). The produced and attracted trips are estimated using the characteristics of the zones. The trip production is typically estimated based on household characteristics, namely household size, and the number of vehicles available in a household. Trip attractions are generally estimated based on the level of employment in a zone. The total numbers of trips produced and attracted are estimated by different mathematical models, namely linear regression models, cross-classification models or trip rate models (Bonnel, 2004; Ortúzar & Willumsen, 2011; Roorda & Miller, 2006). The output of this first stage is the number of trips produced and attracted by each zone which is used in the trip distribution stage to make a trip matrix.

- **Trip distribution**

The second stage of the traditional four-stage model is trip distribution. This stage makes a linkage between the produced and attracted trips to form an origin-destination trip pattern. That means, trip distribution models allocate the generated trips from each zone to various other destination zones in the study area and produce a trip matrix denoting the trips from each zone to every other zone. The gravity model is the most commonly used mathematical model for estimating trip distribution. The gravity model allocates the produced trips of a zone to other destination zones based on their trip attractions and the distance to other zones (Bonnel, 2004; McNally, 2007b). Other mathematical models, such as growth factor, proportional flow, biproportional updating, and opportunity models are also used for estimating trip distribution (Bonnel, 2004; Ortúzar & Willumsen, 2011; Roorda & Miller, 2006).

- **Trip mode choice**

Mode choice is the third stage of the traditional four-stage model. This stage allocates different transport modes for the generated trips based on the modal and individual attributes and produces mode specific trip matrices. Multinomial or nested logit models are the most commonly used models to estimate the mode choice models (Bonnel, 2004; McNally, 2007b; Ortúzar & Willumsen, 2011).

- **Trip assignment**

The last stage of the traditional four-stage model is trip assignment. This stage assigns trips on the transportation network for each trip matrix estimated in the previous stage. For example, considering two modes, i.e. auto and public transit, the final output of the estimation is the volume of vehicles on each road and the number of passengers for transit on each link. Trip assignment models usually assume that each individual tries to minimize his/her generalized cost of travel for each O-D pair and a given mode (Bonnel, 2004). Different methods are used for trip assignment in the network such as "all-or-nothing" assignment, equilibrium assignment and so on (Bonnel, 2004; McNally, 2007b; Ortúzar & Willumsen, 2011).

Traditional models have originally been developed to forecast expected highway demands in 1950s; however these models are not well adapted to today's complex travel behaviours and have been widely criticized in the literature (Beinborn, 1995; McNally, 2007a; Shiftan & Ben-Akiva, 2011). These trip-based models are static and aggregate in nature. Within this framework, the attributes of the individuals are aggregated at the zonal level. Thus, these traditional models have limited ability to represent underlying travel behaviours of the individuals and predict changes in demand with the changes in different policies and strategies (such as TDM, ITS technology, HOV lanes and so on) in an area. Additionally, these models are not capable of capturing behavioural understanding of interactions among household members (intra-household interactions). Furthermore, this trip-based approach has no ability to address the interrelationship between trips and activities, temporal constraints and dependencies of activity scheduling, or the underlying activity behaviour.

2.2.2 Disaggregate trip-based models using discrete choice method

Travel surveys in different urban areas provide detail information on travel behaviour and socio-demographic attributes of an individual. This allows researchers and practitioners to estimate models directly without first aggregating into zones. These travel demand models are known as disaggregate travel demand models. This disaggregate approach assumes that aggregate behaviours are the result of numerous individual decisions. The approach models individual choices as a function of the characteristics of the alternatives available and the socio-demographic attributes of each individual.

One of the major innovations in the travel demand analysis was the development of disaggregate models using discrete choice methods in the early 1970s, however, in general, these models have focused on modelling individual trips made during the day (Roorda, 2005). Discrete choice models mainly rely on the microeconomic utility maximization theory where the concept of utility denotes as the benefit or satisfaction which a person can derive from consumption of a good or a service. This theory assumes that each individual seeks to maximize his/her utility for a good or a service. In a transportation context, the theory implies that each individual selects the alternatives available for travel and activities that maximize its utility. Discrete choice models have been widely used in the transportation field, detailed description of the theory and application of these models can be found in (Ben-Akiva & Lerman, 1985). The utility of a particular alternative is generally measured by a linear combination of individual's characteristics and attributes of the alternative.

Disaggregate discrete choice modelling is extensively used in travel demand analysis because of its behavioural response and the associated advantages over aggregate approach (Tye, Sherman, Kinnucan, Nelson, & Tardiff, 1982). Discrete choice models can be used for either trip-based, tour-based or activity-based models. However, disaggregate trip-based models have several limitations, for instance no adequate consideration of the individual's social context, no representation of intra-household interactions, and no ability to understand the interrelationship between other trips made during the day and the underlying participation of activities which actually lead to make the trips (Roorda, 2005).

2.2.3 Modelling approach used by the MTQ in Montreal

The Ministère des transports du Québec (MTQ) has developed several transport modelling tools with the help from the multi-platform software involving many data sets (for instance, O-D travel surveys) since the late 1970s (Tremblay, 2007). In Montreal, the O-D travel surveys are conducted by a telephone interview approximately every five years in the regions to collect detail travel and socio-economic information of the 5% population of the region (Chapleau, Allard, Trépanier, & Morency, 2001; Trépanier, Chapleau, & Morency, 2008). The modelling tools are regularly updated with new O-D travel surveys in the region.

The following Figure 2.2 presents the major modelling tools and the interrelationships between them. The models are used in five major cities in Quebec, namely Montreal, Quebec, Gatineau, Sherbrooke and Trois-Rivières. In this thesis, we have mainly focused on the travel demand modelling tools used in the Montreal region by the MTQ, more details on these tools can be found in other documents (Ministère des transports du Québec [MTQ], 2014; Tremblay, 2007).

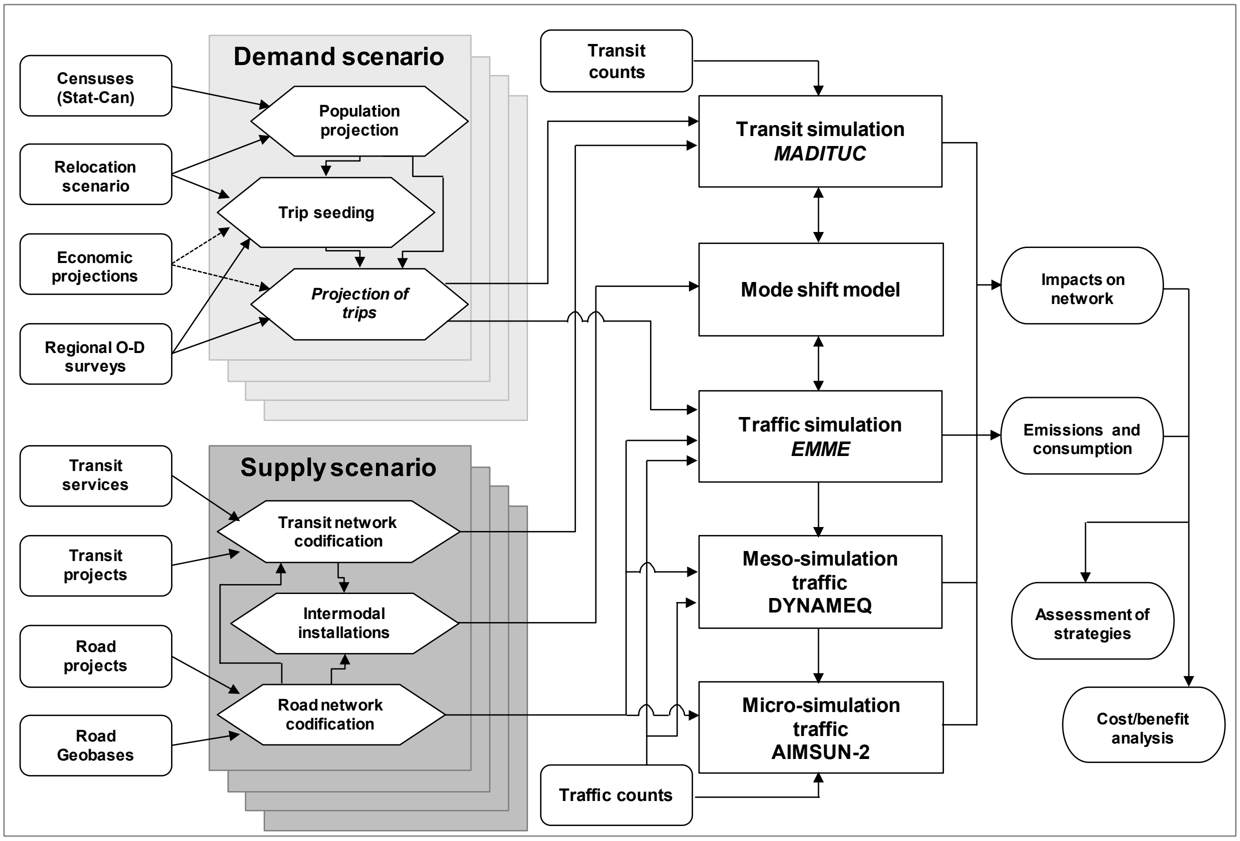


Figure 2.2: Modelling components of the Montreal Area and their interrelationships to each other (adapted from (Tremblay, 2007))

The travel demand modelling tools are discussed below:

2.2.3.1 Projection of future travel demand

The travel demand forecasting model has been developed under a broad project in the mid 1990s by the research group, MADITUC at Polytechnique Montreal, Canada. The forecasting model, embodied in the application known as MADEDE (Modèle d'Analyse Désagrégée de la Demande), has been developed based on the O-D travel surveys which allow developing this disaggregate modelling framework (Pimparé & Thiffault, 2000; Tremblay, 2007).

The basic concept of this model is to adjust the weighting factor of each trip observed in the O-D travel survey for expected changes, revealed from the observed trends using historical O-D travel surveys, in characteristics associated with the profile of an individual and the movements he/she conducts. The characteristics are examined across variables such as age, gender, and territory of residence under the five broad dimensions, namely the population size, the activity status of the individual (in terms of main occupation), the motorization of the individual (in terms of access to automobiles), the growth of transit uses and the employment forecasts. The forecasting model to project travel demand proceeds as below (Ministère des transports du Québec [MTQ], 2014; Tremblay, 2007):

- Apply first three adjustment factors successively to the individual of the O-D survey based on the expected changes in the population size, activity status, and motorization of the individual.
- Apply another adjustment for the "year-effect" on the modes.
- Redistribute work trips based on the expected changes in the employment attractions of each destination.

Adjustment related to population size, activity status and motorization

The mathematical equation to modify the basic expansion factor of the O-D survey based on the expected changes in the population size, activity status, and motorization of the individual is expressed as below:

$$F^{H}{}_{tsg} = F^{B}{}_{tsg} * \frac{P^{H}{}_{tsg}}{P^{B}{}_{tsg}} * \frac{S^{H}{}_{tsg}}{S^{B}{}_{tsg}} * \frac{M^{H}{}_{tsg}}{M^{B}{}_{tsg}} \qquad \text{.......................................} (2.1)$$

where,

H = Future year	M = Motorization
B = Base year	t = territory of residence
P =Population	s = gender
S = Individual status (worker, student, other, or non-mobile)	g = age group

$F^{B}{}_{tsg}$ = the expansion factor of the trip conducted by a person having the characteristic *tsg* in base year

F^{H}_{tsg} = the expansion factor of the trip in future year, calculated by multiplying the expansion factor of the base year with the expected evolution of the population with the characteristics tsg, the activity rate for a given status and the motorization rate

To forecast the future population, the MTQ uses the population projection model, ES-3 (Entrants - Sortants à 3 niveaux) since the mid 1990s. The population forecast for each 5-year horizon is based on a detailed calculation (by age-sex) of the mortality, loss migration (outgoing), inflows population (incoming), and fertility. The MTQ updates the population projection with each new census of the Statistics Canada. The future evolution of other characteristics, namely activities, and motorization are gathered from a comparative analysis of trends using all O-D surveys available in this region.

Adjustment related to year effect on the modes

Another adjustment related to year-effect on the modes, newly introduced in 2008 due to the growth in public transit uses observed between 2003 and 2008 is applied by estimating logistic regression models for three modes (car, transit and active transportation) for each of five aggregated trip purpose (work, school, leisure and visits, shopping, and other) (Ministère des transports du Québec [MTQ], 2014).

Adjustment related to employment forecasts

The model then applies a final adjustment for the work trips to redistribute them in to destinations, based on the trend of the relative employment attractiveness of different destinations across the region.

The weighting factor of each trip for the future horizons is estimated based on the most recent O-D travel survey. The final forecasting result provides the expansion factors for future horizons for each trip of the O-D survey. This forecasting approach is comprehensive as it includes the whole region and all types of travel in the modelling framework. This approach is totally disaggregate in nature; thus it differs with the traditional approach of travel demand forecasting in terms of level of aggregation as well as of method. However, as a trip-based approach, like the traditional models, it also does not capture the interrelationship between trips and activities made during the day and the underlying activity behaviours that mainly generates the trips, and consider temporal constraints and dependencies of activity scheduling. In addition, this approach does not include intra-household interactions among the household members.

2.2.3.2 Trip assignment

- **Public transit assignment**

For public transit assignment in Montreal, the MTQ uses the MADITUC (Modèle d'Analyse Désagrégée des Itinéraires de Transport Urbain Collectif) model, developed in the mid 1980s by the MADITUC group at Polytechnique Montreal (Chapleau, 1986). MADITUC is a totally disaggregate approach based on O-D travel survey data and geographically refined transit infrastructures. Various transport agencies are using the MADITUC software extensively in the Montreal Area. The codification of transit networks includes nodes, the route of each line, length, commercial speeds and periods of services. The codification window within the MADITUC software allows to validate the routes reported by the respondents in the O-D survey. The model assigns the route to the transit service which minimizes the total time (generalized cost) of travel. MADITUC allows two kinds of analysis:

a. First, the model can analyze the routes reported by the respondents in the O-D survey. In this case, the model estimates the descriptive variables of the public transit services (e.g. access, waiting, and travel time) and provides different statistics results of the transit operation which provides opportunity to conduct modal shift analysis.
b. Second, the model can evaluate different scenarios (e.g. introduction of a new service) by simulating public transit trips in the network.

- **Traffic assignment**

Traffic assignment of other transport modes, such as auto, truck and van are conducted using the Static User Equilibrium Assignment (more specifically SOLA algorithm (Second-Order Linear Approximation (Florian & Morosan, 2014))) using EMME software. The traffic assignment is based on the aggregate demand using an O-D matrix at the traffic analysis zone (TAZ) level of a study area. The output results of the assignment are presented in terms of volumes on the links of the network. In the context of Montreal, the road network is codified using the real geographic coordinates including all roads accessible to large vehicles, urban and regional arteries and collector streets. Generally, local streets are not codified explicitly, they are represented by the connectors to centroids of the zones.

2.2.3.3 Incremental mode shift model with threshold

A mode shift model, known as incremental mode shift model with threshold, is used in Montreal to predict the modal shift based on the relative changes in generalized cost of travel by available alternative modes using the results obtained from traffic and public transit assignments estimated in EMME and MADITUC, respectively. The model is applied on the total future demand to predict the modal shift of the users and then their path in transportation network. The mathematical equation of the shift model with two travel alternatives i.e. auto and public transit is given below (Tremblay, 2007):

$$\Delta\left[\frac{Trips_{public\ transit}}{Trips_{total}}\right] = f\left(\Delta\left[\frac{Time_{public\ transit}}{Time_{auto}}\right]\right) \quad \text{.......................................} (2.2)$$

A threshold of three minutes improvement in travel times is considered to eliminate the choices with small improvements. Then, the probability of transfer is calculated using the diversion curves calibrated based on the results obtained from the O-D travel surveys. The diversion curves are calibrated in six broad dimensions (trip purposes, destination attractions, combination of different transit modes, the main mode in the combination of different transit modes, number of price increments, and presence or absence of congestion on the auto trajectory). The following Figure 2.3 shows an example of the diversion curves.

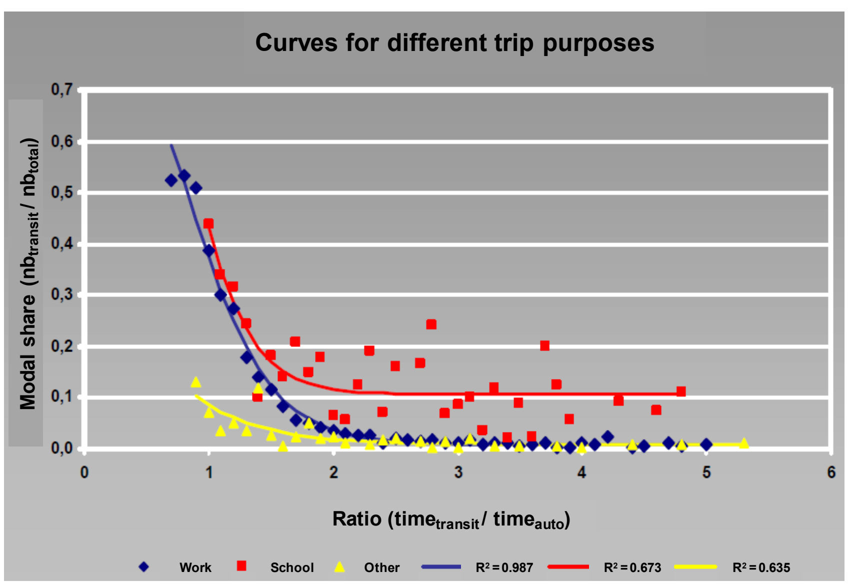


Figure 2.3: Diversion curves for different trip purposes
(adapted from (Tremblay, 2007))

2.2.4 Activity-based travel demand models

The activity-based approach has emerged in the literature since the 1970s to overcome the limitations observed in the traditional four-stage models. In addition, the motivation to shift from trip-based to activity-based models is that individuals' travel decisions are components of a broader set of activity scheduling decisions. The activity-based approach treats travel as derived demand from different activities distributed in space and time (Jones, Koppelman, & Orfeuil, 1990; Kitamura & Fujii, 1998; Recker, 1995).

The root of activity-based modelling is often referred to the contributions from Hagerstrand (1970), Chapin (1974) and Fried, Havens, and Thall (1977). Hagerstrand (1970) has proposed a conceptual framework, known as space-time prism which implies that individual can perform different activities in different locations at different points in time. He also explains an individuals' actions in space and time considering three kinds of constraints, such as capability constraints (limit the movement of individual because of human factors i.e. physical or biological), coupling constraints (require presence of another member of a family (i.e. joint household activities) or other resources (i.e. automobile access to participate in an activity)), and authority constraints (restrictions set by individuals or institutions such as office or store hour, and regulations such as noise restrictions). Another influencing study has been conducted by Chapin (1974), who has observed patterns of behaviour in time and space by different population segments. He has hypothesized a theory of motivations, subject to society imposed constraints, of basic human desires to participate in activities. Fried et al. (1977) have proposed a theory of adaptational change within an urban structure which motivates the behaviour and also constrains it by physical and social resources. Jones, Dix, Clarke, and Heggie (1983) have conducted the first comprehensive study of activities and travel behaviour combining these theories. More specifically, they have first defined the activity-based approach and attempted to model complex activity behaviour with empirical testing.

Activity-based models are interested in activity participation including activity generation, spatial distribution and temporal programming (start time and duration). They consider relationships among different trips made by the same individual during a day (or sometimes during a week) and often consider joint participation in activities by the individuals from same household (i.e. intra-household interactions). They are capable of understanding the underlying activity behaviours that mainly generate the trips. Recent advances in technology allow developing activity-based microsimulation models which permits explicit representation of activities and trip making decisions of a large number of individuals and their households.

Literature reviews on recent progress on activity-based modelling frameworks to the analysis of travel behaviour can be found in other papers (Axhausen & Gärling, 1992; Gärling, Kwan, & Golledge, 1994; Mohammadian, Auld, & Yagi, 2009; Shiftan & Ben-Akiva, 2011). Activity-based modelling frameworks, mainly developed based on the utility maximization principle (e.g. CARLA, STARCHILD) have been reviewed by Axhausen and Gärling (1992); whereas the models, developed based on computational-process models (CPMs) have been reviewed by Gärling et al. (1994). A detailed literature review on the recent activity-based models with the overview of statistical procedures used in the model components can be found in other paper (Mohammadian et al., 2009). Discussions of the development of activity-based models have also been found in some other literature (Bhat & Koppelman, 1999b; Bowman & Ben-Akiva, 2000). Moreover, Shiftan and Ben-Akiva (2011) have reviewed a model developed for Tel Aviv and some existing US models and indicated that improved behavioural realism and policy sensitivity could be attained using the more complex behavioural models.

2.2.4.1 Model components and their modelling approaches

Activity-based models generally consist of two major components namely activity generation and activity scheduling (Bhat & Koppelman, 1993; Habib & Miller, 2009).The first component (i.e. activity generation) mainly deals with the desire or need to participate in activities (i.e. activity demand). The second component (i.e. scheduling) deals with spatial-temporal opportunities and constraints which lead to an executed activity pattern. It is evident that the interactions between the generation of activities and their scheduling are two-way. A detailed discussion on the relationship between these two components can be found in (Bhat & Koppelman, 1993). Modelling approaches used in these two components are discussed as below:

- **Activity generation**

Although a substantial progress has been seen in the development of activity-based travel demand modelling, the activity generation component is an under-researched area compared to the scheduling component (Habib, 2007; Yasmin, Morency, & Roorda, 2012). Thus, it is still a challenge to adequately represent the activity generation choice process, including the number of activities, their sequence, temporal decisions (start time and duration), their spatial locations, and so on.

Some activity-based models (for example, CARLA (Jones et al., 1983), AMOS (Kitamura, Lula, & Pas, 1993), and SMASH (Ettema, Bergers, & Timmermans, 1993)) have used activity patterns as given input in the modelling framework directly drawn from the observed data without any modelling effort to specify how the activity agenda is developed. Another activity-based model, STARCHILD (Recker, McNally, & Root, 1986a, 1986b) has separate activity generation and scheduling components; but feedback and updating between these two components are missing in the framework. Other activity-based models have utilized Monte-Carlo simulation to generate activity patterns based on the observed distributions of activity attributes (for example, FAMOS (Pendyala, Kitamura, Kikuchi, Yamamoto, & Fujii, 2005), and TASHA (Miller & Roorda, 2003; Roorda et al., 2008)), which is another way of dealing with the activity generation component as given input.

CEMDAP has employed an activity generation approach using a range of econometric models including regression, binary logit, multinomial logit, and hazard-duration (Bhat, Guo, Srinivasan, & Sivakumar, 2004; Pinjari & Bhat, 2011). Both generation and scheduling components of ALBATROSS have been modelled comprehensively (Arentze & Timmermans, 2004). In this framework, individual activity generation has been developed from the household activity calendar which consists of different activities related to long-, medium- and short-term household decisions. This generation process has employed CHAID (Chi-squared automatic interaction detector) decision tree induction method.

Habib and Miller (2008) have argued that the generation components, dealt with activity patterns as given input or developed based on static rules could not address the behavioural trade-offs in the activity demand generation; as they did not consider any activity utility/benefits. They have developed a conceptual framework based on random utility maximization theory with the aim of providing a theoretical foundation for the activity generation components of activity-based models. This generation framework provides opportunity for the dynamic interrelationship with the activity scheduling process. However, the model does not include intra-household interaction explicitly within the modelling framework. Arentze and Timmermans (2009) have proposed a dynamic activity generation model for multiday planning period based on a needs-based approach which implies that an individual conducts activities to satisfy particular needs.

The activity generation component indicates the activity demand which also varies with changes in transportation policies; thus, this component must receive substantial attention in the activity-based modelling framework.

- **Activity scheduling**

Activity scheduling has been conducted by either using econometric models or following a rule-based procedure. Based on these two modelling approaches, activity-based models can be classified into two broad categories, econometric activity-based models and rule-based activity scheduling models.

The decision making process in econometric models is a mathematical function, mostly utility maximization-based equations, that capture relationships among activity and travel attributes, and predict the probability of decision outcomes (Bhat et al., 2004; Kitamura & Fujii, 1998; Pinjari & Bhat, 2011). Econometric activity-based models have the capability to examine alternative hypotheses in terms of the causal relationships among activity-travel patterns, land use and socio-demographic characteristics of individuals, but have limitations in terms of capturing the actual decision making processes and behavioural mechanisms that lead to observed activity-travel decisions. Examples of econometric activity-based models are Bowman and Ben-Akiva model (Bowman & Ben-Akiva, 2000), the MORPC model (Vovsha, Petersen, & Donnelly, 2004), the CEMDAP model (Bhat et al., 2004), and the Jakarta model (Yagi & Mohammadian, 2008a, 2008b).

An advanced modelling framework has been developed for activity scheduling to address some behavioural assumptions found in the econometric models (Gärling et al., 1994). This framework is known as computational process model (CPM) or rule-based approach. A CPM is a computer program implementation of a production system model consisting of a set of condition-action rules (IF-THEN) that specify an action to be executed when a condition is met (Gärling et al., 1994). CPM has the ability to model the interdependent decisions, which is not possible by other means as well as to incorporate behavioural principles (Golledge, Kwan, & Gärling, 1994). However, activity generation components get less attention than activity scheduling components in most CPMs (Pinjari & Bhat, 2011; Yasmin et al., 2012).

The first model built using the CPM framework is the Hayes-Roth model. The SCHEDULER model, an update of the earlier Hayes-Roth model, has tried to reduce some limitations in the decision making process of Hayes-Roth model (Golledge et al., 1994).

Arentze and Timmermans (2004) have also defined CPM by weak and strong definition. A CPM which applies a heuristic consisting of some sequential or partially sequential decision making process, but still assumes utility maximization or some other form of unbounded rationality at the level of individual decision steps, is called weak CPM. Examples of some activity-based models that meet the weak definition are PCATS (Kitamura & Fujii, 1998), the model system proposed by Bhat (Bhat, 1999), STARCHILD (Recker et al., 1986a, 1986b), and SMASH (Ettema et al., 1993). In addition, a CPM which employs a production system or some other rule-based formalism at the level of individual choice facets is called strong CPM. The examples are SCHEDULER (Golledge et al., 1994), GISICAS (Kwan, 1997), ALBATROSS (Arentze & Timmermans, 2004; Arentze & Timmermans, 2008), TASHA (Miller & Roorda, 2003; Roorda, 2005), and ADAPTS (Auld & Mohammadian, 2009).

2.3 Model application and validation

Activity-based models are still less used in practice than the traditional travel demand models (Arentze & Timmermans, 2008; Mohammadian et al., 2009; Shiftan et al., 2003). Some recent applications of the activity-based models are Portland (METRO), San Francisco (SFCTA), New York (NYMTC), Columbus (MORPC), Atlanta (ARC), Sacramento (SACOG), Bay Area (MTC), Denver (DRCOG) (Bradley & Bowman, 2006; Shiftan & Ben-Akiva, 2011), Dallas/Forth-Worth (CEMDAP) (Bhat et al., 2004), Southeast Florida (FAMOS) (Pendyala et al., 2005), Netherlands (ALBATROSS) (Arentze & Timmermans, 2008), and Toronto (TASHA) (Roorda et al., 2008). It is evident that any kind of model must be validated extensively before reaching the application stage. Activity-based models have been validated in several ways.

The most common validity testing of activity-based model is the application of the model using the base year dataset from which the model is originally derived. Some models have been validated using such a method; examples include STARCHILD (Recker et al., 1986b), SMASH (Ettema, Borgers, & Timmermans, 1996), Bowman and Ben-Akiva model (Ben-Akiva & Bowman, 1998; Bowman, 1998), PCATS (Kitamura & Fujii, 1998), ALBATROSS (Arentze &

Timmermans, 2004), the Jakarta model (Yagi & Mohammadian, 2007a, 2007b), TASHA (Miller & Roorda, 2003; Roorda et al., 2008), DRCOG (Childress, Sabina, Kurth, Rossi, & Malm, 2010; Kurth, Childress, Sabina, & Rossi, 2006), and CEMDAP (Bhat, Paleti, Pendyala, & Goulias, 2013). However, these model validations were only at the macro-level; very few models have been validated at different levels such as by population segments (PCATS (Pendyala et al., 2005)).

Models are also validated by comparing the estimated forecasts of daily travel behaviour of a future year with observed survey data of the same year (e.g. DRCOG and TASHA); however these model validations were also at the macro-level (Childress et al., 2010; Kurth et al., 2006; Roorda et al., 2008).

The assessment of spatial transferability of the model is another kind of validation test; however to date, validation of the activity-based models using such a test is still rare (Arentze et al., 2002; Auld & Mohammadian, 2012; Bowman et al., 2014).

2.3.1 Model transferability

The main concept of model transferability is the application of previously estimated model parameters into a new context (Karasmaa, 2001). A clear definition of transfer and transferability has been provided by Koppelman and Wilmot (1982). Transfer is defined as "*the application of a model, information, or theory about behaviour developed in one context to describe the corresponding behaviour in another context*", whereas transferability is defined as "*the usefulness of the transferred model, information, or theory in the new context*".

Sikder, Pinjari, Srinivasan, and Nowrouzian (2013) have reviewed the literature on spatial transferability of travel demand models. In travel demand analysis, mostly trip generation and mode choice models have been validated using such test (Yasmin, Morency, & Roorda, 2013; Yasmin et al., 2015a). Gangrade, Kasturirangan, and Pendyala (2000) have demonstrated that the variability of activity behaviours across geographic areas is low, which indicates potential for geographic transferability. Also, Sikder et al. (2013) have indicated that activity-based models, developed with the greater theoretical basis and behavioural realism are more transferable than the traditional trip-based models. However, the assessment of spatial transferability of activity-based models is infrequent (Arentze et al., 2002; Auld & Mohammadian, 2012; Bowman et al., 2014).

The evaluation of the spatial transferability of ALBATROSS model system at both individual and aggregate levels have been done by simulating activity patterns for two municipalities in the Netherlands (Arentze et al., 2002). The results indicate that the components of the ALBATROSS model are quite transferable except for transportation mode choice; however this research has indicated that transferring a model to an area by simulating activity patterns using a larger sample improves transferability.

There is another study on the spatial transferability of the planning order model (which determines when and in what order the activity attributes are planned), a component of the ADAPTS activity-based modelling framework (Auld & Mohammadian, 2009). The planning order model has potential for having good transferability properties; however the authors have indicated a need for further evaluation (Auld & Mohammadian, 2012).

Bowman et al. (2014) have empirically assessed the transferability of DaySim, an activity-based model for six regions in California and Florida in the US. But, they could not confirm considerable comparability between the six regions due to the small sample sizes available in some regions. However, they have indicated that activity generation and scheduling models are more transferable than mode choice and location choice models. They have also showed that larger survey samples (for instance 6,000 households compared to 2,500 households or less) considerably improve model estimation, which suggests that transferring a model developed based on a large sample from a comparable region is better than estimating a new model using a smaller local sample.

Developing a new model for a context requires a large amount of time, cost and expertise; but if an existing model can be transferred to that context and if it can successfully reproduce travel behaviours of that region, this process can significantly reduce these efforts. Thus, future study is required in this research area to get more evidence towards the application of activity-based models. This research selects the activity-based model, TASHA (developed for a Canadian context - Toronto) to test its spatial transferability and explore the possibility of implementing the model in another Canadian context (i.e. Montreal). Next section provides a detailed discussion on the TASHA modelling framework.

2.4 TASHA modelling framework

The activity-based model, TASHA (Travel Activity Scheduler for Household Agents) has been developed by Miller and Roorda (2003) as a module of an integrated land use and transportation model, ILUTE (Integrated Land Use, Transportation, Environment) (Salvini & Miller, 2005). ILUTE is a microsimulation framework in which TASHA has been developed as travel scheduling module; however it can be used as a standalone software to generate daily schedules of activities for individuals and households to estimate the travel demand of a region. TASHA has been mainly developed with the aim of improving travel demand modelling framework in the Greater Toronto Area (GTA), particularly the behavioural representation of human decision making, the spatial and temporal precision of outputs, and the sensitivity to demand oriented policies. The empirical application indicates that it improves behavioural modelling of travel demand in the GTA (Roorda et al., 2008). The key characteristics of the TASHA modelling framework are presented below:

- **Activity-based:** TASHA is an activity-based travel demand model, which recognizes travel as a derived demand from the demand to participate in out-of-home activities. Thus, to model trip-making, it is crucial to understand how people organize their daily activities in space and time.
- **Household-based:** TASHA includes intra-household interactions (i.e. interactions between household members) within the modelling framework, as individual's daily activity (thus travel) decisions are significantly influenced by the household-level interactions, constraints and needs.
- **Agent-based microsimulation model:** TASHA represents persons and households as "intelligent objects" or "agents" who are able to observe their own environment, make decisions and act into their environment. Microsimulation provides that opportunity to develop the fully disaggregated activity-based approach.

Miller and Roorda (2003) have presented the full conceptual design and methodology of this activity scheduling microsimulation model. It is a fully disaggregate model which estimates activity schedules and travel patterns for a 24-hour typical weekday for all persons in a household. The model has been developed based on the trip diary data from the 1996

Transportation Tomorrow Survey (TTS) for the GTA. There is an opportunity to implement this model in any city where this kind of dataset is available.

2.4.1 Major features of the TASHA model

The major features of the TASHA model are presented below:

2.4.1.1 Use of the concept of the project

The TASHA model employs the concept of project which has a specific project agenda. To fulfill the agenda, the project organizes the activity episodes into person's schedule in a household. The concept of project, project agenda and activity episode are defined below:

Project: A project includes a group of logically connected set of activity episodes, which together achieve a common objective. The project may include different activity types and one or more household members. Using the concept of project, it is possible to represent household interactions (i.e. joint participation, for example, dine out together) between household members realistically.

Project agenda: Projects have project agendas, which contain a list of activity episodes associated with each project.

Activity episodes: To make a person's schedule, activity episodes are the basic unit of analysis, they include a start time, a duration, a location and an activity type.

2.4.1.2 Interactive household agents

The model has been developed such that the interaction between family members can be represented realistically. Therefore, the model allows simultaneous scheduling of the individuals within the household, which facilitates interaction between family members that normally happens within a household.

2.4.1.3 Microsimulation of a 5% sample of households in the Greater Toronto Area

From the 1996 TTS in the GTA, approximately 89,000 households including 243,000 individuals are incorporated as individual agents/entities in the computer model. This microsimulation model allows generating activity/travel schedules for each person individually.

2.4.1.4 Designed using an object oriented programming technique

The TASHA model has been developed using an object oriented programming technique. Object orientation is defined as *"a modelling paradigm that attempts to mirror real life objects relevant to the scheduling process directly as "classes" in the program code"* (Roorda, 2005, p. 70).

2.4.1.5 Assumption of broad project and episode types

Four different broad project types at person and household level are assigned; details are presented in Figure 2.4 and 2.5.

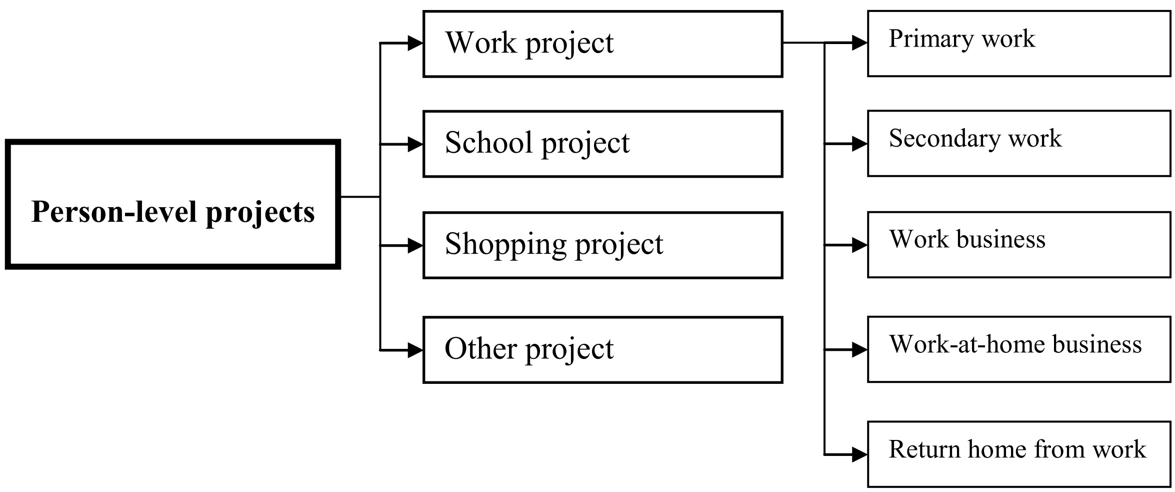


Figure 2.4: Person-level projects in the TASHA model

(adapted from (Roorda, 2005))

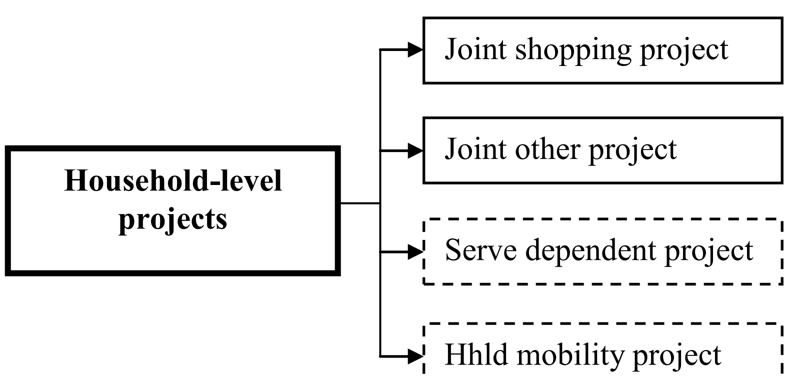


Figure 2.5: Household-level projects in the TASHA model

(adapted from (Roorda, 2005))

The work project includes different episode types as shown below:

- **Primary work -** work episodes occurring at the usual place of work that are part of the primary work event. The primary work event is defined as the sequence of work episodes beginning with the first work episode of the day plus any work episodes from subsequent work chains that begin before 3:00 p.m.

- **Secondary work -** work episodes occurring at the usual place of work that are part of the secondary work event. The secondary work event is defined as the sequence of work episodes in a chain that starts after 3:00 p.m., given that a primary work event has occurred.

- **Work business -** work episodes that occur at a location other than the usual place of work for a person who normally works at a location other than his or her home.

- **Work-at-home business -** work episodes that occur at a location other than home for a person who normally works at home.

- **Return home from work -** at-home episodes that are embedded within the primary work event. These episodes can be thought of as lunch trips but may include other at-home activities.

(Miller & Roorda, 2003, p.116)

2.4.1.6 The model assumes sequential household decisions

As many decisions of household and household members are taken simultaneously for different aspects, the TASHA model assumes that several household decisions are also made sequentially. In addition, it assumes home and work location and auto ownership as exogenous inputs into the model.

2.4.2 Assumptions of the TASHA modelling framework

The behavioural assumptions made within the TASHA modelling framework are listed below (Roorda, Doherty, & Miller, 2002):

- Mode choice is assumed to occur in sequence after formation of the schedule.
- Allocation of activity location is only based on household, person and zone characteristics, and may not be possible to modify during the scheduling procedure.

- Joint household scheduling decisions only include the generation of joint activities that include more than one household member.
- Schedule formation decision rules are the same for all people.
- Schedules are developed for a 24-hour period.
- Projects consider broad activity types (such as work, school, shopping and other) and only include activities of that type.
- A single fine-tuning algorithm to “clean-up” short work episodes is used in the “execution” stage.
- The scheduling of activities is done following an order of precedence based only on activity type; it does not consider the time dimension of a planned activity.
- The model simulates a single schedule for one point in time without considering an individual’s history of how the schedule is formed.
- There is no incorporation of learning or habit formation.
- Generation of activity attributes are initially based on the observed activity schedules from the TTS data.
- Conflicts, which come up during the insertion of activities into project agendas and person schedules, are handled using a limited set of resolution strategies.

2.4.3 Conceptual design of the TASHA modelling framework

TASHA is a rule-based model, mainly developed following a ‘‘bottom up’’ approach, i.e. the activities are generated first and then scheduled (Miller & Roorda, 2003; Roorda et al., 2008). The model estimates activity schedules and travel patterns for a 24-hours typical weekday for all individuals in a household (Miller & Roorda, 2003). Figure 2.6 presents the conceptual framework of the TASHA model including five components, namely activity generation, location choice, activity scheduling, mode choice, and trip assignment. Detailed descriptions of each component are given below:

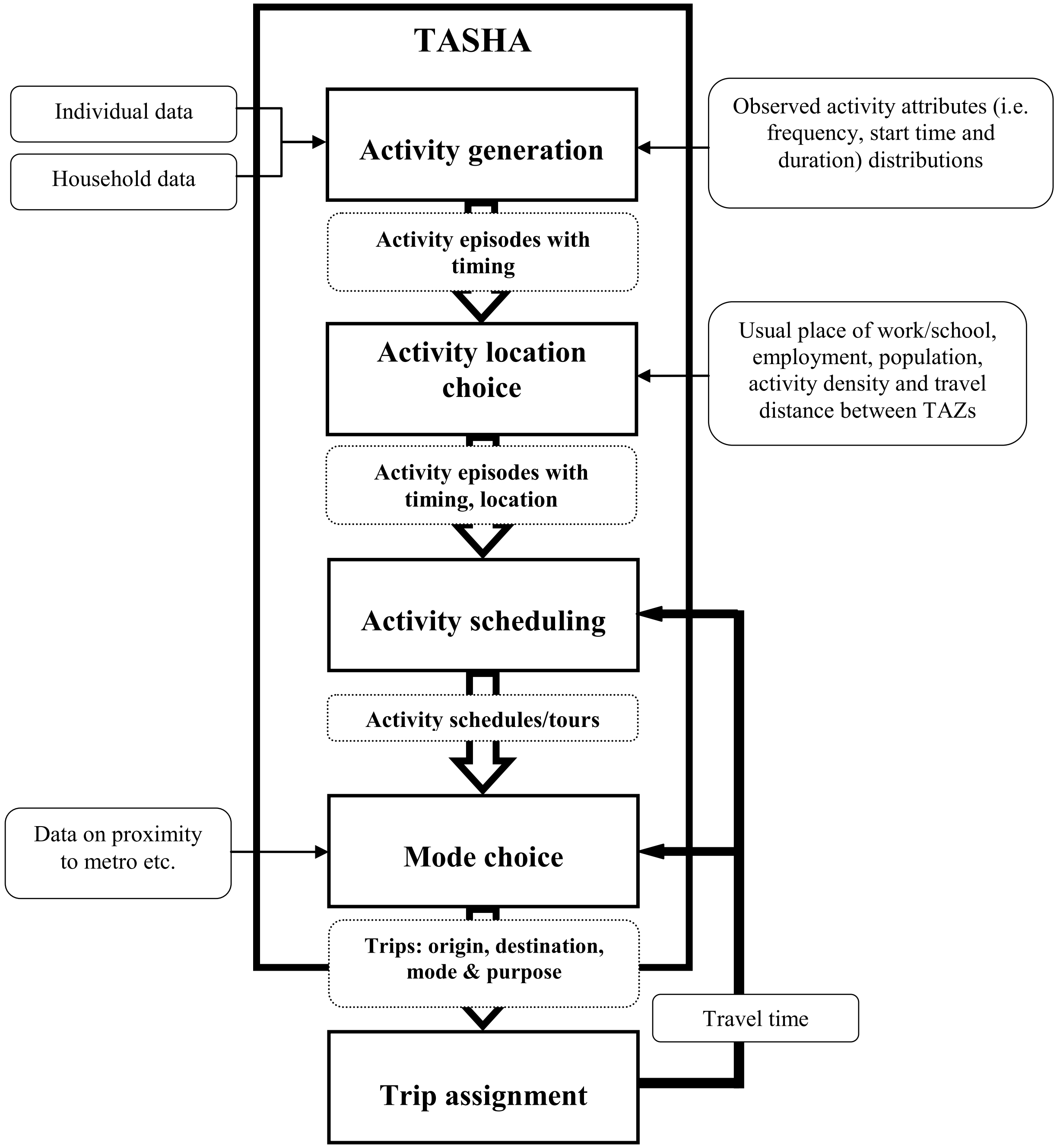


Figure 2.6: Conceptual design of the TASHA modelling framework

(adapted from (Roorda et al., 2008))

2.4.3.1 Activity generation

The first component of the TASHA modelling framework is activity generation, which generates both individual and joint activities (with other household members) for each person. Work, school, and return to home activities are generated first and then, based on the duration and start time of work and school activities, other types of activities are generated. The inputs in this stage are person and household data and series of activity generation behaviours related to frequency distribution of different types of activities, duration and start time of activities. In TASHA, the activity generation is done using Monte-Carlo simulation to draw the number of activities and their start time and duration from observed frequency distributions from the 1996 TTS of the GTA (Miller & Roorda, 2003).

262 frequency distributions of these activity attributes were developed from the 1996 TTS of the GTA, cross-classified by activity type, individual, household and schedule attributes such as age, gender, occupation, employment status, student status, presence of children, number of adults, individual work-school project status, and household work-school project status (Roorda, 2005). Table 2.1 presents the explanatory variables and number of distributions for each activity type.

Table 2.1: Explanatory variables used for classification of each activity type

Activity type	**Explanatory variables used for classification**	**Number of distributions**
Primary work	Age, occupation, employment status	32
Secondary work	Occupation, employment status	8
Work-business	Age, occupation, employment status	32
Work-at-home business	Age, occupation, employment status	12
Return home from work	Occupation, employment status	8
School	Age, student status	10
Independent other	Age, gender, individual work-school project status	56
Joint other	Presence of children, number of adults, household work-school project status	24
Independent shopping	Age, gender, Individual work-school project status	56
Joint shopping	Presence of children, number of adults, household work-school project status	24
Total		**262**

TASHA assumes that these distributions of activity attributes remain constant over time (thus, temporally transferable). However, these distributions may change over time. Furthermore, the population segmentation for each activity type by different variables has been done based on intuition and testing; the selection of variables to generate activities was not through a systematic process and these segments are fixed over time.

2.4.3.2 Activity location choice

This component allocates locations to each activity episode for an individual. Home location and usual place of work/school are exogenous input. The activity location choices of other activities are estimated using a series of entropy models (Eberhard, 2002). The mathematical expression of the entropy model is presented below. If a person lives in zone i, then the probability of choosing location zone j by that person for an activity is defined as:

$$P_{j/i} = \frac{\exp\left(\sum_k \delta_{jk}\ [\alpha_k + \beta_k \log(E_j) + \phi_k \log(P_j) + \gamma_k d_{ij}]\right)}{\sum_f \exp\left(\sum_k \delta_{j'k}\ [\alpha_k + \beta_k \log(E_{j'}) + \phi_k \log(P_{j'}) + \gamma_k d_{ij'}]\right)} \quad \text{.......................................} (2.3)$$

where,

δ_{jk}	= 1, if zone j belongs to zone activity category k; 0 otherwise
E_j	= employment in zone j
P_j	= population in zone j
d_{ij}	= distance from zone i to zone j
α_k, β_k, ϕ_k, γ_k	= parameters to be estimated
K	= 1, if the zone is the city core
	= 2, if employment density > 3000workers/km^2 (work), shopping mall floor space > 100,000 sq. ft. (shopping), >50 retail stores/km^2 (other)
	=3, otherwise

2.4.3.3 Activity scheduling

The activity scheduling component of TASHA is rule-based. The method of activity scheduling is summarized here. The detail of this process can be found in other papers (Miller & Roorda, 2003; Roorda et al., 2008).

Step 1: Activity episodes are added into a project agenda with preliminary time sequence with other activity episodes to achieve a common purpose.

Step 2: Person schedules are formed by moving activity episodes into person schedule based on the order of precedence observed from an interactive computer survey of activity scheduling (Doherty, Nemeth, Roorda, & Miller, 2004).

Step 3: Fine-tuning is done in the final schedule just before /during execution of the schedule by using a “clean up” algorithm.

2.4.3.4 Mode choice

The scheduling component provides the activity schedules of the individuals for 24-hour of a typical weekday, and then the random utility mode choice model allocates the modes to perform the activities. It includes joint mode choice for joint activities, vehicle allocation and also rideshare opportunities within the household.

2.4.3.5 Trip assignment

The last component, trip assignment, is done using a static user equilibrium traffic assignment model, implemented in EMME. This model assigns the route taken for each trip. The model can be run iteratively to return travel time feedback from EMME into the activity scheduling process.

2.4.4 Application and validation of the TASHA model

The feasibility and practicability of the prototype version of the TASHA model have initially been tested using the base year trip diary data from the 1996 TTS in the GTA (Miller & Roorda, 2003). This initial investigation shows that the model reproduces the observed 1996 trip making characteristics within an acceptable limit for a prototype model with underestimation of approximately 311,000 daily trips (-3.3%).

Validation of the TASHA model has been conducted in two ways (Roorda et al., 2008). First, model verification has been done by comparing base year model outcomes to the base year dataset from the 1996 TTS from which the model is originally derived. This verification assesses the components of the TASHA model, activity generation (frequency, start time, and duration), activity location choice, and activity scheduling.

Secondly, model validation has been done by comparing forecasts of a future year with the observed data of the same year (i.e. 2001 TTS in the Greater Toronto Area). The validation results indicate that TASHA could be considered as a powerful alternative tool than the traditional travel demand modelling framework for the GTA, at least for short-term forecasts, as activity attributes distributions have remained quite stable. However, there must be further improvements in the modelling framework for the GTA (for instance - improvement in location choice model, development of an activity generation model, applying different rules of scheduling for different population segments and so on) (Roorda et al., 2008).

In addition, the TASHA model has been applied in an European context (i.e. London), assuming that all model components are transferable (Le Vine, Sivakumar, Roorda, & Polak, 2010). The study by (Le Vine et al., 2010) presents the preliminary results in the development of an activity-based travel demand model for London. However, it has several limitations. The TASHA model application in the context of London was a simple transfer with only one run and without undertaking any quantitative evaluation. Moreover, the validation of the TASHA model was not comprehensive in terms of validation criteria.

Furthermore, Roorda and Miller (2006) have compared the TASHA model with the traditional four-stage models based on three different policy instruments, including alternative hours, high occupancy vehicle (HOV) lanes, and intelligent transportation system (ITS) technology to demonstrate the benefits of TASHA. They have also compared the important features and methods used in different components of both TASHA and four-stage models. This comparison indicates that TASHA has stronger potential to assess the demand-oriented policy measures than the four-stage model. In principle, it is capable of providing a closer representation of the actual human behaviour in making travel and activity decisions. Since TASHA has been developed based on a conventional trip diary data, it could be applied in different urban settings where this kind of data is available. Roorda (2005) has argued that TASHA will be a “next generation” travel demand modelling tool with some additional validation, testing and software engineering in the GTA. The TASHA model has recently been implemented within the GTAModel V4.0 in the analysis of major rail transit investment alternatives for the City of Toronto (Miller et al., 2015). This recent application estimates and calibrates the parameters of the TASHA model by using the 2011 TTS of the GTA (Data Management Group, 2014).

2.5 Comparison between the traditional four-stage model, the transport models used by the MTQ and the TASHA model

The traditional four-stage model is trip-based and aggregate in nature, whereas the TASHA model is a fully disaggregate and activity-based microsimulation model. Here in Montreal, a trip-based disaggregate approach is used to project the future travel demand. These three models are similar in terms of data requirements, as all of them are developed based on conventional trip diary data. The following Table 2.2 and 2.3 show the comparison between the traditional four-stage travel demand model, the transport models used by the MTQ in Montreal and the TASHA modelling framework in terms of model features and model components (Roorda & Miller, 2006; Tremblay, 2007).

Table 2.2: Comparison between features of the four-stage model, models used by the MTQ in Montreal and the TASHA model

Model features	**Four-stage model***	**Transport model used by MTQ**	**Activity-based model (TASHA)***
Level of disaggregation	Household/zone	Person/zone	Person/household
Unit of analysis	Trip	Trip	Activity episode
Inputs requirements	Traditional trip diary data, population and employment forecasts	Traditional trip diary data, population and employment forecasts	Traditional trip diary data, population and employment forecasts
Outputs	Zonal origin-destination trip tables Assigned transit and traffic flows	Origin-destination trips at individual level Assigned transit and traffic flows	Trip chains, activity schedules Assigned transit and traffic flows
Period of analysis	Peak hour	AM peak / 24-hour period	24-hour period
Incorporation of intra-household interactions	Since there is no disaggregation to a level lower than the household, within household interactions cannot be represented	There is no consideration of household interactions	Explicitly represents joint activities, sharing of household vehicles, and within-household ridesharing. Does not incorporate task allocation of maintenance activities within the household
Incorporation of inter-household interactions	No inter-household interactions (aside from trip assignment) such as carpooling between households can be represented	No representation of inter-household interactions	No inter-household interactions are represented. However, the agent-based microsimulation framework provides a laboratory for testing interactions such as carpooling

* The information on the four-stage model and activity-based model (TASHA) has been adapted from (Roorda & Miller, 2006, pp. 17-18)

Table 2.3: Comparison between methods of the four-stage model, models used by the MTQ in Montreal and the TASHA model

Model component	**Four-stage model method***	**Transport model used by MTQ**	**TASHA method***
Trip/activity generation	Trips are generated using linear regression, cross-classification or trip rate models.	The MTQ does not follow the traditional approach. Use MADEDE model to project the future travel demand using origin-destination of the individual trips from the O-D surveys, thus there is no separate trip generation or trip distribution stage.	Activities are simulated based on observed distributions available from TTS data. Tours are an outcome of the activity scheduling process.
Trip/activity distribution	Trip distribution is done using gravity model, a proportional flow model, or biproportional updating. These methods are all at the zonal level of analysis.		Home location and usual place of work/school are exogenous inputs. Location choices of other activities are simulated using a series of entropy models.
Mode choice	Multinomial or nested logit models are used to predict mode choice. These models are at the zone or the individual level of analysis. Choices are typically made for individual trips, without a tour logic, and without enforcing an overall household logic.	Incremental mode shift model with threshold is used which is developed based on the elasticity of demand with respect to travel times (generalized cost).	A tour-based mode choice model operates at the person and the household level. Mode choices are made at the individual level, but are subject to household vehicle constraints via an explicit vehicle allocation model. The mode choice algorithm treats ridesharing as a utility-based negotiation between household members.
Trip assignment	Trip assignment is done for total O-D peak hour trips using EMME, TransCAD, or other static equilibrium models.	Public transit assignment is done using the MADITUC model. Auto trip assignment is done using EMME software.	Trip assignment is done using EMME. However, because individual tours are generated by TASHA with a 5-minute temporal precision, a micro- or meso-scopic traffic simulation with dynamic traffic assignment is feasible.

* The information on the four-stage model and activity-based model (TASHA) has been adapted from (Roorda & Miller, 2006, p. 18; Eberhard, 2002)

2.6 Population segmentation methods

This research proposes a methodology to segment the population based on the similarities in individuals' daily activity patterns and their socio-demographic characteristics by combining two methods, the multiple sequence alignment method (SAM) as a clustering approach and the multinomial logit model (MNL). A detailed description of these two methods is presented below.

2.6.1 Sequence alignment method as a clustering approach

Clustering, a data mining technique, classifies a set of physical or abstract objects into a number of different clusters with similar characteristics. Clustering is a great tool to better understand the similarity of peoples' needs and choices. Many clustering algorithms are available; commonly used clustering algorithms are partitioning and hierarchical (Berkhin, 2006; Xu & Wunsch II, 2005). These algorithms group similar objects based on some similarity (or dissimilarity) measures. Commonly used measures are Euclidean, squared Euclidean distance functions and so on (Xu & Wunsch II, 2005). These functions measure the similarity of variables of corresponding positions among two objects and sum the results calculated for all the variables to get a measure of similarity between two objects (Joh, Arentze, & Timmermans, 2001). However, activity pattern is represented as a sequential order of activities, the similarities between the activity patterns cannot be adequately captured by using such functions (Joh et al., 2001; Schlich, 2001). In activity patterns, there could be a simple offset which could lead to a very poor similarity measure using such function; however it could just be a little difference in the activity schedules. To overcome this limitation, sequence alignment method seems quite promising in clustering of activity patterns, as this method allows researchers to use the whole activity pattern as a unit of analysis (Joh et al., 2001).

Sequence alignment method provides a quantitative measure of distance or similarity between two character sequences. This method, also known as optimal matching, is widely used in molecular biology since 1970 (Needleman & Wunsch, 1970). Still, it was only introduced in the field of activity scheduling analysis in 1998 (Wilson, 1998). Wilson (1998) has highlighted the need for a powerful quantitative approach to summarize the information gathered in activity diaries. Further research of Wilson and others have provided significant theoretical improvements in sequence alignment procedure (Joh et al., 2001; Wilson, 1999).

Among others, multidimensional alignment is a useful theoretical improvement (Joh, Arentze, Hofman, & Timmermans, 2002). Wilson (1999) has also developed a new software package ClustalG, an improved version of Clustal software series (i.e. ClustalX - general format (Thompson, Gibson, Plewniak, Jeanmougin, & Higgins, 1997) and Clustal W - DOS format (Thompson, Higgins, & Gibson, 1994)), which provides an opportunity to apply the sequence alignment method in other fields including activity pattern analysis (Saneinejad & Roorda, 2009; Schlich, 2001; Wilson, 1999).

2.6.1.1 Terms used in sequence alignment method

Basic operations

The basic operations of sequence alignment are identical matches, substitutions (inexact matches), and insertions or deletions. It can be explained by an example with two sequences [aim] and [am] (Wilson, 1998). These sequences can be aligned by inserting an *i* in the middle of [am] or by deleting the *i* from [aim]. In both ways, the alignment can be shown as follows:

a-m

aim

Insertions and deletions must take place in pairs, which is defined as indel. It means that when there is an insertion of a character in one sequence, there must be a deletion of it from the other. In the above example, the hyphen is used to represent the gap made by the indel.

Another basic operation in sequence alignment is substitution, which can be demonstrated by a mismatch or obtained by a pair of insertion and deletion. The sequence [am] can be transferred to [as] by substituting *s* for *m* or by inserting or deleting first *m* and then *s*. These alignments are shown as below:

Substitution	*Insertion/Deletion*
am	am-
as	a-s

Similarity, distance, and weights

Weights or scores are given to the operations to measure the quality of the alignment by similarity or distance measures. The alignment analysis is conducted through an algorithm which either maximizes similarity or minimizes distance. An analyst can use different weights to reflect the importance of operations, the similarity of particular elements, the position of elements in the sequence, or the number or type of neighbouring elements or gaps (Wilson, 1998). Also, the weights could vary based on the subject area (Schlich, 2003).

Global and local alignment

Global and local alignments refer to the analysis of complete and partial sequences, respectively. Global alignment employs the Needleman-Wunsch global algorithm that finds the best match of complete sequences (Needleman & Wunsch, 1970). Local alignment employs the Smith-Waterman local alignment algorithm that finds the best partial sequence match (Smith & Waterman, 1981). Both algorithms are dynamic programming methods.

Pairwise and multiple sequence alignment

Alignment of two character sequences (complete (global alignment) or partial (local alignment)) is defined as pairwise alignment. To measure the similarity of two sequences, first it generates a comparison table with elements of the two sequences on the two axes of a matrix. Then, an algorithm finds the best path from top-left to bottom-right in the comparison table obtained by the least cost (or highest score) using only eligible operations, such as insertion, deletion, or substitution of a single character. The operation of alignment algorithms can be explained by the dot matrix comparison method. For example, there are two character sequences [thought] and [thorough]. First, to construct a dot matrix plot, these sequences are written on the two axes of a matrix, as shown below:

	t	h	o	u	g	h	t
t	●						●
h		●				●	
o			●				
r							
o			●				
u				●			
g					●		
h						●	

Then, dots are placed at the cells where the characters in the appropriate column match. For highly similar sequences, the dot plots will appear as a single line along the matrix's main diagonal. In this example, two sequences are quite similar, with a few dissimilarities. Two sequences can be aligned as shown below:

```
thorough-
tho--ught
```

The characters [r o] are deleted from [thorough] or inserted in [thought] at the same time, which is called indel. Another indel with [t] ends the alignment procedure. The weighting scheme, which evaluates matches, mismatches, near matches, and gaps, determines the exact patterns of the optimum path through the comparison table and the resulting alignment.

Alignment of more than two character sequences is described as multiple sequence alignment. However, to align multiple activity sequences, N (where N is greater than two), pairwise alignment can be discussed as multiple sequence alignment by using comparison tables in N dimensions. However, the process with multiple sequences (when the sequences are more than ten) requires a significant amount of time and computer memory. Thus, heuristic methods are applied in multiple sequence alignments using pairwise measures. ClustalG is also developed based on such heuristic methods (for instance, neighbor joining algorithm (Saitou & Nei, 1987)).

2.6.1.2 Multiple sequence alignment procedure in ClustalG

Multiple sequence alignment in ClustalG proceeds as below:

- First, it conducts pairwise alignment of all sequences to create a matrix of similarity scores using Needleman-Wunsch global algorithm (Needleman & Wunsch, 1970).
- Then, it conducts the multiple alignment by using the branching pattern of a tree constructed from the pairwise similarity scores and certain other adjustments using neighbor joining algorithm (Saitou & Nei, 1987).

Finally, clusters are identified by observing and analyzing several elements. These are:

- Tree structure - Tree structure suggests the grouping of sequences. This structure can be visualized using a tree-viewing software (for instance, Treeview, or Archaeopteryx beta).
- Output log file - ClustalG provides an output log file, which gives the information on the procedure and order in which sequences, and groups of sequences are joined with one another. It also provides the score calculated for each grouping step.
- Aligned sequences - Color coded aligned sequences can also be visualized.

2.6.2 Multinomial logit model

Logit model has been widely used in travel demand analysis in the context of modelling discrete choices, especially when it has involved more than two alternatives (i.e. multinomial logit models (MNL)) (Bhat, 1995; Train, 2003). The simplicity, easy estimation and interpretation, and ability to add or remove choice alternatives have made the MNL model widely used (Bhat, 1995; Koppelman & Bhat, 2006).

In deriving the functional form of the MNL, it has been assumed that the random components of the utilities of the different alternatives are (Koppelman & Bhat, 2006):

- extreme-value (or Gumbel) distributed,
- identically and independently distributed across alternatives, and
- identically and independently distributed across observations/individuals.

Specifying the random components of the utility function in this way, McFadden (1974) has derived the simple closed-form MNL model, which provides the choice probabilities of each alternative as a function of the deterministic component of the utility of all the alternatives. The structural model can be written as:

$$P(i) = \frac{e^{(V_i)}}{\sum_{j=1}^{J} e^{(V_j)}} \quad \text{.. (2.4)}$$

where,

$P(i)$ = probability of choosing an alternative i *(i = 1, 2,....., J)* by an individual from a set of J alternatives

V_j = the deterministic component of the utility of alternative choices

The utility function includes attributes of alternatives and characteristics of individuals that give individual's utility valuation for each alternative.

Among others, the independence from irrelevant alternatives (IIA) property of the MNL has been widely discussed (Koppelman & Bhat, 2006). This property means that the ratio of the probabilities of choosing two alternatives is independent over all other alternatives. It implies that the relative probability of choosing one alternative over another is not affected by the presence or absence of any other alternatives in the choice set. Therefore, it is possible to add or remove an alternative from the choice set without affecting the structure or parameters of the model.

This property is advantageous for model estimation and application for different members of the population with different sets of alternatives and also for prediction of the choice probabilities for a new alternative (Koppelman & Bhat, 2006). However, the property has also been criticized when the characteristics of two or more alternatives of the choice set are close to each other. In this case, the relative probability of choosing one alternative over another may be affected by the presence or absence of any other alternatives in the choice set. This may provide predictions of choice probabilities with errors. To better understand this limitation, the classic example of red bus/blue bus can be discussed here (Arminger, 1994). For example, a commuter has two alternative choices, auto or red bus to go to work and the probability of choosing auto is two-thirds and red bus is one-third. Thus, the ratio of market shares between auto and red bus is 2:1. Now, a blue bus service which is totally identical to the red bus service except colour is introduced. It is expected that the new market shares will be the same, i.e. two-thirds for the car and one-third for both buses (i.e. one sixth for each bus (red or blue)). But, due to the IIA property, the MNL model will estimate the market share of the car as only twice of the red bus, it will not estimate the car as four times of it.

The assumption of the independence of random components in the utility of the alternatives leads to this limitation in the MNL model. To address this limitation, other models, developed with different assumptions related to the random components distributions in the utilities of the alternatives, which are free of the IIA property can be used (for instance, Nested logit model) (Domencich & McFadden, 1975; Koppelman & Bhat, 2006).

2.7 Synthesis

In travel demand modelling, the Ministère des transports du Québec (MTQ) does not follow the framework of the traditional four-stage model, they use their own modelling tools (combination of aggregate and disaggregate approach) developed using the O-D travel surveys (Ministère des transports du Québec [MTQ], 2014; Tremblay, 2007); however these models are trip-based. The trip-based approach has widely been criticized for many reasons and the activity-based approach has been recognized as a more powerful framework in travel demand analysis over traditional trip-based approach (Bhat & Lawton, 2000; Kitamura, 1996; Kitamura et al., 1996; Roorda, 2005; Shiftan et al., 2003). However, the practical application of these models is still relatively infrequent in the real world contexts (Arentze & Timmermans, 2008; Mohammadian et al., 2009; Shiftan et al., 2003).

To increase the practical applications of activity-based models, there must be more evidence that these models are capable of reproducing actual travel behaviour in an area. Thus, extensive validation and improvements within the modelling framework are crucial. The literature review points out the following research gaps and limitations in past research, which will be addressed in this thesis in the following chapters:

- There has been no attempt to develop or apply an activity-based travel demand modelling framework in the Montreal region.
- Spatial transferability, as a validation test, has not been frequently examined for the activity-based models.

- Validations of activity-based models are mostly conducted at the macro-level. Few models have been validated at different levels such as by population segments.
- Like some other activity-based models, in TASHA, less effort has been applied to modelling activity generation than scheduling.
    - Activity generation is based on the observed distributions of activity attributes from the 1996 TTS of the GTA and these distributions are assumed to remain constant over time (thus, temporally transferable). However, this assumption of temporal stability may not be the case for long term forecast in the GTA as well as for other regions, if the model is applied there.
    - Population segmentation for each activity type by different variables has been done based on intuition and testing. The selection of variables to generate activities was not through a systematic process and these segments are fixed over time.

The next chapter explains the methodology followed to fulfill the research objectives of this thesis.

CHAPTER 3 METHODOLOGICAL FRAMEWORK

3.1 Introduction

The objectives of this research are to enhance the current process of modelling of travel demand of the GMA using an activity-based approach (i.e. TASHA) and to contribute to the validation and enhancement of the activity-based modelling framework. The limitations in the current modelling tools used by the MTQ in the GMA and the opportunity of improvement in the current process of modelling using an activity-based approach in this region are identified at the end of the Chapter 2 (Section 2.7). This section also summarizes the research gaps and limitations that exist in the field of activity-based travel demand modelling and in the TASHA modelling framework. These research gaps and limitations justify the significance of this research and its objectives. This thesis addresses several issues presented below:

3.2 Application of TASHA and evaluation of its spatial transferability

The activity-based approach has already been recognized as a more powerful framework than the traditional four-stage travel demand models in travel demand analysis. However, the traditional travel demand models are the majority in practice; practical application of the activity-based models is still relatively rare. Similar to other models, activity-based models must also be validated before the practical applications in the real world contexts. Though spatial transferability has been recognized as a useful validation test for travel demand models, to date, activity-based models have not been frequently assessed using such a test.

This limitation is addressed in the first article entitled "Assessment of spatial transferability of an activity-based model, TASHA" which has been published in the journal "Transportation Research Part A: Policy and Practice" in 2015, and presented in Chapter 4 of this thesis (Yasmin et al., 2015a). In summary, this paper transfers three of the five components of the TASHA model to the context of the Island of Montreal (activity generation, activity location choice and activity scheduling) and examines its spatial transferability as a validation test. It employs the Toronto distributions of activity attributes (such as frequency, start time, and duration of different types of activity) and location choice parameters, estimated using the 1996 TTS of the GTA without any adjustment. Other input data are gathered from the Montreal 2003 O-D travel survey and the 2001 Canadian Census. The trip diary data from the 1996 TTS of the GTA and the O-D travel survey of Montreal are quite similar; however some input variables are not available directly from the O-D dataset. We have to impute them for this research, a detailed description of the data preparation effort can be found in (Morency & Trépanier, 2013). TASHA generates daily schedules of activities (individual and joint) for each individual of the Island of Montreal. Then, we compare the modelled activity attributes (frequency, start time, duration and distance) from TASHA and the observed attributes from the 2003 O-D travel survey at the aggregate level for five different activities (i.e. work, school, shopping, other, and return to home).

The results, at the aggregate level, indicate that TASHA provides reasonable outcomes (in some cases - better results than for the Toronto Area) for all four attributes for work, school and return to home activities with few exceptions (for instance, school start time). It also demonstrates promising results for shopping frequency and start times; however, TASHA provides larger differences for average shopping durations and distances. Only the modelling outputs for all four attributes for the ‘other’ activity type vary largely with the observed attributes for the Montreal Island.

In general, we propose to re-estimate the model parameters and to use the local activity attribute distributions (frequency, start time and duration) if available, while transferring the TASHA model from one context to another. The implication of this first paper is twofold. First, it contributes to the current literature by assessing spatial transferability of an activity-based travel demand model, i.e. TASHA. Second, it aims to enhance the current travel demand modelling framework in Montreal.

3.3 Validation of activity-based travel demand models

Validation is a vital element of the model development process for practical applications in the real world context. Activity-based models have been validated using different methods, such as:

1. Verification by comparing estimated outputs with the observed from the base year dataset from which the model is originally derived.
2. Validation by comparing forecasts of daily travel behaviour for a future year with observed survey data from the same year.

3. Validation by applying model in different context and testing its spatial transferability.

Majority of validation efforts for activity-based models have followed first and/or second methods, however these validations were mostly at the macro-level; only few models have been validated at different levels such as by population segments. Additionally, few models have been tested by evaluating their spatial transferability and these validations were also mostly at the macro-level. TASHA has also been validated following first and second method but at the macro-level.

Our first article evaluates the spatial transferability of the TASHA model at the macro-level (Yasmin et al., 2015a), whereas our second article, which is an extension of the first article, focuses on the validation of the activity-based travel demand models at three different levels. This article titled "Macro-, meso-, and micro-level validation of an activity-based travel demand model", submitted to the Journal "Transportmetrica A: Transport Science" is presented in Chapter 5 (Yasmin, Morency, & Roorda, submitted). In particular, this paper assesses the spatial transferability as a validation test of the TASHA model by validating model outcomes at three different levels of aggregation (macro-level - aggregated across the entire population; meso-level - across population segments; and micro-level - individuals). It employs the 2003 O-D travel survey of Montreal and the 2001 Canadian Census. This paper compares the simulated activity attributes (activity frequency, start time, duration, and trip distance) with the observed attributes from the 2003 O-D survey for five activity types (work, school, shopping, other, and return to home) at macro-, and meso-level. It also compares the simulated outputs for each individual with the observed attributes from the 2003 O-D survey using some criteria for micro-level validation.

The validation results at macro- and meso-level are promising; TASHA is capable of reproducing activity behaviours of another context, at least for fixed activities (work, school) with few exceptions. But, in the case of some activity attributes of flexible activities (shopping, other), the TASHA model provides large differences. This research also discusses potential reasons behind the large differences found in some cases.

This paper shows how and at which level the validation of activity-based travel demand models must be performed.

3.4 Activity generation behaviours and their trends

In general, activity-based models have two major components, activity generation and activity scheduling. The activity generation component relatively gets less attention than the scheduling component in the modelling framework. This is also true for the TASHA model. Activity scheduling in TASHA follows a rule-based approach; however activity generation is based on the observed distributions of activity frequency, start time and duration from the 1996 TTS of the GTA.

This research has transferred TASHA from Toronto to Montreal without changing any parameters from the Toronto settings; the results have been presented in our first and second articles (Yasmin et al., 2015a; Yasmin et al., submitted). These papers use the observed distributions of activity attributes (frequency of different types of activity, start time and duration) from the 1996 TTS of the Toronto Area as input for the Montreal application. Like the Toronto Area application (Roorda et al., 2008), we have also assumed these attribute distributions to remain constant over time (thus, temporally transferable). In Toronto area, the validation results indicate that TASHA is successful to reproduce activity/travel patterns in the GTA, at least for short-term forecasts, as activity attributes distributions have remained quite temporally transferable. However, this assumption of temporal stability over time may not be the case for the Montreal region.

Chapter 6 empirically examines whether observed distributions of activity attributes in the Montreal region have changed over time. It also investigates how these attributes vary with peoples' socio-demographic characteristics. The results are presented in the article entitled "Trend analysis of activity generation attributes over time" published in the journal "Transportation" in 2015 (Yasmin, Morency, & Roorda, 2015b). This paper compares three activity generation attributes for different types of activity (work, school, shopping, and other) across the three travel surveys (i.e. the 1998, 2003 and 2008 O-D surveys of the GMA). In addition, it examines these attributes with different socio-demographic groups (by age, gender, occupation, employment status, student status, *et cetera*).

The results indicate that distributions of activity attributes are changing over time in the Montreal region. The paper also provides insights into possible reasons for these changes in terms of demographic, socio-economic, land use, technology, and policy changes in this region. It indicates preparing the activity attributes for the application of an activity-based model, TASHA, such that they reflect temporal changes in travel behaviour of the GMA.

3.5 Population segmentation to develop activity generation attributes

Activity generation in TASHA is based on observed distributions of activity attributes (frequency, start time and duration) from the 1996 TTS of the Toronto Area. A total of 262 distributions were developed from the observed data, cross-classified by activity type, person, household and schedule attributes. The population segmentation for each activity type was based on intuition and testing; the selection of attributes to generate activities was not through a systematic process and those segments are fixed over time.

This limitation is addressed in Chapter 7 entitled "Population segmentation based on similarity in activity patterns" which will be submitted for publication in the near future. This research focuses on how to systematically obtain segments of population based on similarity in activity patterns (represented by activity frequency, start time and duration) and socio-demographic characteristics to develop activity attributes distributions for an activity-based travel demand model, i.e. TASHA. More specifically, this research first utilizes a clustering approach, namely sequence alignment method to segment the individuals based on the similarities in individuals' daily activity patterns. Then, it estimates the multinomial logit models for the distinct segments of individuals using their socio-demographic characteristics. This research employs the 2003 O-D travel survey of Montreal.

After clustering the individuals based on the similarities in their daily activity patterns, we obtain eleven distinct segments (*Work 1, Work 2, Work 3, School 1, School 2, Shopping 1, Shopping 2, Other 1, Other 2, Home* and *Outlier*). Multinomial logit models are estimated for these segments of individuals (except *Outlier* - consists of a variety of uncommon activity sequences) using their socio-demographic characteristics. This research then selects a final population segmentation model which offers better data fit compared to other estimated models and provides useful insights into the determinants of belonging to different segments. The population segmentation model is quite simple in terms of data requirements (variables, age and gender), and estimation; thus it could be easier to use this model for prediction purposes.

The following three chapters present three articles published or submitted for publication in scientific journals. Therefore, Chapter 4, 5, and 6 have an independent structure consisting of an abstract, introduction, background, data and research method, analysis and results, a conclusion and references. Then, Chapter 7 presents a methodology to segment the population to develop activity attributes distributions, which will be submitted for publication in the near future. Finally, this thesis concludes with a general discussion (Chapter 8) and conclusions (Chapter 9).

CHAPTER 4 ARTICLE 1: ASSESSMENT OF SPATIAL TRANSFERABILITY OF AN ACTIVITY-BASED MODEL, TASHA

This chapter is presented in the form of a paper published in the *Transportation Research Part A: Policy and Practice,* 78 (2015), 200-213.

Farhana Yasmin[a,*], Catherine Morency[a,1], Matthew J. Roorda[b,2]

[a] *Department of Civil, Geological and Mining Engineering, Polytechnique Montreal, C.P. 6079, Station Centre-Ville, Montreal, Quebec, Canada H3C 3A7*
[b] *Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario, Canada M5S 1A4*
[*] *Corresponding author. E-mail: farhana.yasmin@polymtl.ca*
[1] *E-mail: cmorency@polymtl.ca*
[2] *E-mail: roordam@ecf.utoronto.ca*

Abstract

Spatial transferability has been recognized as a useful validation test for travel demand models. To date, however, transferability of activity-based models has not been frequently assessed. This paper assesses the spatial transferability of an activity-based model, TASHA (Travel Activity Scheduler for Household Agents), which has been developed for the Greater Toronto Area (GTA), Canada. TASHA has been transferred to the context of the Island of Montreal, Canada using the 2003 Origin-Destination (O-D) travel survey and the 2001 Canadian Census. It generates daily schedules of activities (individual and joint) for each individual in this region. The modelled activity attributes (frequency, start time, duration and distance) from TASHA and observed attributes from the 2003 O-D travel survey are compared for five different activities (i.e. work, school, shopping, other, and return to home). At the aggregate level, TASHA provides quite reasonable outcomes (in some cases - better results than for the Toronto Area) for all four attributes for work, school and return to home activities with few exceptions (for instance, school start time). The model outcomes are also promising for shopping frequency and start times; however, TASHA provides larger differences for average shopping durations and distances. Only the forecasts for all four attributes for the 'other' activity type differ greatly with the observed

attributes for the Montreal Island. These large differences most likely indicate the differences in behaviour between the Montreal Island and the Toronto Area. In general, we conclude that re-estimation of model parameters and the use of local activity attribute distributions (frequency, start time and duration) is a desirable step in the transfer of the TASHA model from one context to another.

Keywords: Spatial transferability; Activity-based model; Activity scheduling; TASHA; Montreal

4.1 Introduction

Spatial transferability has been recognized as a useful validation test for travel demand models. To date, however, transferability of activity-based models has not been frequently assessed. This paper examines the spatial transferability of an activity-based model, TASHA (Travel Activity Scheduler for Household Agents), which has been developed by Miller and Roorda (2003). TASHA is a fully disaggregate model that estimates activity schedules and travel patterns for a 24 hour typical weekday for all individuals in a household. The model has been developed based on trip diary data from the 1996 Transportation Tomorrow Survey (TTS) for the Greater Toronto Area (GTA), Canada. Consequently, it is possible to implement this model in any city where this kind of dataset is available. Montreal is renowned for large scale Origin-Destination (O-D) travel surveys conducted every five years since the early 1970s. This is very similar data to that of the TTS allowing an opportunity to apply TASHA in the context of Montreal. Therefore, the main focus of this paper is to examine to what extent the activity-based model, TASHA is transferable to the Montreal Island, Canada.

The remainder of this paper is organized as follows. First, a brief review of the related literature is provided. Then, an overview of the TASHA model is presented. Next, data and research method are described. Then, a comparison between simulated results by TASHA and observed data in the Montreal Island is presented. Also, a systematic comparison of the TASHA model performance in the Montreal Island and the Toronto Area is provided. The paper concludes with a discussion of the implications of the study.

4.2 Background

Travel demand models have been employed as decision support tools for transportation planning over the last several decades. However, policy instruments in the field of transportation planning such as Travel Demand Management (TDM), Intelligent Transportation Systems (ITS) technology and, High Occupancy Vehicle (HOV) lanes need more precise decision support tools than the traditional four-stage travel demand model (Roorda & Miller, 2006; Shiftan et al., 2003). The activity-based approach has emerged in the literature since the 1970s to overcome the limitations observed in the traditional four-stage travel demand model. The importance of activity-based models in travel demand analysis is also well recognized in the literature (Bhat & Lawton, 2000; Kitamura, 1996; Kitamura, Pas, Lula, Lawton, & Benson, 1996; Shiftan et al., 2003). However, the practical application of the activity-based modelling approach is still relatively rare; traditional travel demand models are the majority of models used in practice (Arentze & Timmermans, 2008; Mohammadian, Auld, & Yagi, 2009; Shiftan et al., 2003). There have been some excellent literature reviews on progress on activity-based modelling frameworks to the analysis of travel behaviour (Axhausen & Gärling, 1992; Gärling, Kwan, & Golledge, 1994; Mohammadian et al., 2009; Shiftan & Ben-Akiva, 2011). Axhausen and Gärling (1992) have reviewed some activity-based modelling frameworks which are mainly developed based on the utility maximization principle (e.g. CARLA, STARCHILD). Gärling et al. (1994) have reviewed several computational-process models (CPMs), whereas Mohammadian et al. (2009) have reviewed recent activity-based models with an overview of statistical procedures used in the model components. In addition, Shiftan and Ben-Akiva (2011) have reviewed some existing US models and a model developed for Tel Aviv and argued that improved behavioural realism and policy sensitivity could be attained using the more complex behavioural models.

4.2.1 Classification of activity-based models

Activity-based models can be classified into two broad categories, econometric activity-based models and rule-based activity scheduling models. The decision making process in these two models are different; it is a mathematical function in econometric models, whereas it is a computational process in rule-based models (Kitamura & Fujii, 1998).

Econometric activity-based models employ the econometric systems of equations (mostly utility maximization-based equations) that capture relationships among activity and travel attributes, and predict the probability of decision outcomes (Bhat, Guo, Srinivasan, & Sivakumar, 2004; Pinjari & Bhat, 2011). The strength of this approach is the ability to examine alternative hypotheses in terms of the causal relationships among activity-travel patterns, land use and socio-demographic characteristics of individuals; however, it has lacking in terms of capturing the actual decision making processes and behavioural mechanisms that lead to observed activity-travel decisions. Econometric activity-based models include Bowman and Ben-Akiva model (Bowman & Ben-Akiva, 2000), the MORPC model (Vovsha, Petersen, & Donnelly, 2004), the CEMDAP model (Bhat et al., 2004), and the Jakarta model (Yagi & Mohammadian, 2008a, 2008b).

Computational process models (CPM), also known as rule-based approaches, were developed to address some behavioural assumptions found in the econometric models (Gärling et al., 1994). A CPM is a computer program implementation of a production system model consisting of a set of condition-action rules (IF-THEN) that specify an action to be executed when a condition is met (Gärling et al., 1994). This approach deals with how a decision maker formulates and executes the schedules and has the ability to capture schedule constraints explicitly. The strength of the CPM is the ability to model the interdependent decisions, which is not possible by other means as well as to incorporate behavioural principles (Golledge, Kwan, & Gärling, 1994). However in most CPMs, activity scheduling components get more attention than activity generation components (Pinjari & Bhat, 2011; Yasmin, Morency, & Roorda, 2012). Arentze and Timmermans (2004) have further categorized CPM as either weak or strong. A CPM which applies a heuristic consisting of some sequential or partially sequential decision making process, but still assumes utility maximization or some other form of unbounded rationality at the level of individual decision steps, is called weak CPM. Examples include PCATS (Kitamura & Fujii, 1998), the model system proposed by Bhat (Bhat, 1999), STARCHILD (Recker, McNally, & Root, 1986a, 1986b), and SMASH (Ettema, Bergers, & Timmermans, 1993). A CPM that employs a production system or some other rule-based formalism at the level of individual choice facets is called strong CPM. Examples include SCHEDULER (Golledge et al., 1994), GISICAS (Kwan, 1997), ALBATROSS (Arentze & Timmermans, 2004; Arentze & Timmermans, 2008), TASHA (Miller & Roorda, 2003; Roorda, 2005), and ADAPTS (Auld & Mohammadian, 2009).

4.2.2 Model application and validation

The practical application of activity-based modelling approach is still relatively rare. Recent applications include Portland (METRO), San Francisco (SFCTA), New York (NYMTC), Columbus (MORPC), Atlanta (ARC), Sacramento (SACOG), Bay Area (MTC), Denver (DRCOG) (Bradley & Bowman, 2006; Shiftan & Ben-Akiva, 2011), Dallas/Forth-Worth (CEMDAP) (Bhat et al., 2004), Southeast Florida (FAMOS) (Pendyala, Kitamura, Kikuchi, Yamamoto, & Fujii, 2005), Netherlands (ALBATROSS) (Arentze & Timmermans, 2008), and Toronto (TASHA) (Roorda, Miller, & Habib, 2008). Any travel demand model should be subjected to validity testing before reaching the application stage. Activity-based models have been validated in several ways. The most common validity testing of activity-based travel demand model is the application of the model using the base year dataset from which the model is originally derived. Examples include STARCHILD (Recker et al., 1986b), SMASH (Ettema, Borgers, & Timmermans, 1996), Bowman and Ben-Akiva model (Ben-Akiva & Bowman, 1998; Bowman, 1998), PCATS (Kitamura & Fujii, 1998), ALBATROSS (Arentze & Timmermans, 2004), the Jakarta model (Yagi & Mohammadian, 2007a, 2007b), TASHA (Miller & Roorda, 2003; Roorda, Miller, et al., 2008), DRCOG (Childress, Sabina, Kurth, Rossi, & Malm, 2010; Kurth, Childress, Sabina, & Rossi, 2006), and CEMDAP (Bhat, Paleti, Pendyala, & Goulias, 2013). Models have also been validated by comparing forecasts of daily travel behaviour for a future year with observed survey data from the same year. TASHA (Roorda, Miller, et al., 2008), and DRCOG (Childress et al., 2010; Kurth et al., 2006) have been validated using such methods. Spatial transferability has also been recognized as a useful validation test for the models, however, to date, this kind of test has not been applied extensively to activity-based travel demand models (Arentze, Hofman, Van Mourik, & Timmermans, 2002; Auld & Mohammadian, 2012; Bowman, Bradley, Castiglione, & Yoder, 2014). Also from a practical point of view, if the model is transferable, it will save time, cost and expertise needed to develop a model for a new context.

4.2.3 Model transferability

Model transferability is the application of previously estimated model parameters into a new context (Karasmaa, 2001). Koppelman and Wilmot (1982) distinguish transfer and transferability. Transfer is defined as "*the application of a model, information, or theory about* *behaviour developed in one context to describe the corresponding behaviour in another context*". Transferability is defined as "*the usefulness of the transferred model, information, or theory in the new context*". A literature review on the spatial transferability of travel demand models can be found in (Sikder, Pinjari, Srinivasan, & Nowrouzian, 2013). To date, model transferability research has mostly focused on trip generation and mode choice models. The transferability discussion has included the spatial and temporal contexts as well as model specification and level of aggregation (Cotrus, Prashker, & Shiftan, 2005). Researchers have found mixed results in their investigations of spatial transferability of trip generation and mode choice models. A number of studies have found acceptable transferability of trip generation and mode choice models (Agyemang-Duah & Hall, 1997; Karasmaa, 2001; Rose & Koppelman, 1984), while others have reported poor transferability of the models (Daor, 1981). Wilmot (1995) has indicated that disaggregated models of trip generation tend to show better transferability than aggregate models since the parameters used in the disaggregate models are not dependent on the zone system. Also, the quality of the model specification of the transferred model has great impact on transferability. Wilmot and Stopher (2001) and Wilmot (1995) have proven that partial transfers (i.e. some model parameters are estimated locally and the rest of the parameters are transferred) improve the transferability of the trip generation model.

Gangrade, Kasturirangan, and Pendyala (2000) have indicated that the variability of activity behaviours across geographic areas is low indicating potential for geographic transferability. Also, it is assumed that activity-based models, developed with the greater theoretical basis and the behavioral realism are more transferable than the traditional trip-based models (Sikder et al., 2013). However, few studies have examined the spatial transferability of activity-based models. Arentze et al. (2002) have tested the spatial transferability of ALBATROSS model system at both individual and aggregate levels by simulating activity patterns for two municipalities in the Netherlands. They have reported acceptable transferability of ALBATROSS model except for transportation mode choice, however indicated that transferring a model to an area by simulating activity patterns using larger sample improves transferability. Another study on the spatial transferability of the planning order model, which is one of the components of the ADAPTS activity-based modelling framework (Auld & Mohammadian, 2009), has shown potential for having good transferability properties, however the study indicates a need for further evaluation (Auld & Mohammadian, 2012). A recent study (Bowman et al., 2014), has empirically tested the transferability of an activity-based model (DaySim) for six regions in California and Florida in the US; however they could not confirm substantial comparability between the six regions because of the small sample sizes available in some regions. But, they have showed that activity generation and scheduling models are more transferable than mode choice and location choice models. They have also indicated that larger survey samples (for instance 6,000 households compared to 2,500 households or less) substantially improve the model estimability and thus suggested that transferring a model developed based on a large sample from a comparable region is better than estimating a new model using a smaller local sample. The TASHA model, estimated using quite a large dataset (4.5% sample of GTA households i.e. 81,554 households) is being transferred to the Montreal Island using also a large dataset (i.e. 26,960 households). As prior studies have suggested that using larger survey samples for model estimation and for model transfer improves model estimability and transferability, respectively, the study assumes that the Toronto TASHA model has potential for good transferability in the Montreal Island. In addition, verification and validation results of the TASHA model indicate that TASHA could successfully replicate the activity/travel patterns in the geographic context (i.e. GTA) for which the model was developed (Roorda, Miller, et al., 2008); however the question arises of how well the model would perform if applied in another context. This research is an effort to examine the spatial transferability of the activity-based model, TASHA with the application in the context of the Montreal Island.

4.3 The TASHA model

The activity-based model, TASHA (Travel Activity Scheduler for Household Agents) has been developed by researchers at the University of Toronto as part of the broader development of the integrated urban model, ILUTE (Integrated Land Use Transportation Environment). However, TASHA can also be used as a stand-alone travel demand model. It is a fully disaggregate microsimulation model which estimates activity schedules and travel patterns for a twenty-four hour typical weekday for all individuals in a household. The model has been developed based on a large-scale traditional trip diary data from the 1996 Transportation Tomorrow Survey (TTS) for the Greater Toronto Area (GTA). The advantages of using traditional trip-based travel survey are twofold. First, it shows the possibility of developing improved travel demand models for the urban areas where the activity-based travel surveys are not available for many reasons. Second, it provides opportunity to apply the model with reasonably low expense in other large metropolitan areas where such datasets are available. The major features of the operational model are as follows (Miller & Roorda, 2003):

- The model makes use of the concept of the project to organize activity episodes into the schedules of persons in a household.
- The model features interactive household agents.
- The model is a microsimulation of a 5% sample of Greater Toronto Area households.
- The model was designed by using an object oriented programming technique.
- The model assumes broad project and episode types.

The model assigns broad project types to each household and to each person. Household-level projects include joint shopping and joint other, and person-level projects include work, school, shopping, and other. Each project may contain one or more activity episode types. In the current version of the TASHA model, each project other than work project contains only episodes of the same type as the project (e.g. the shopping project contains only shopping episodes). The work project consists of several activity episode types, namely primary work, secondary work, work business, work-at-home business, and return home from work. Detailed definitions of the episode types can be found in Miller and Roorda (2003).

The full conceptual design and methodology of this prototype activity scheduling model can be found in Miller and Roorda (2003). This modelling framework includes five components, activity generation, location choice, activity scheduling, mode choice, and trip assignment. This research limits the application to the first three components i.e. activity generation, activity location choice, and activity scheduling. A brief overview of these three methods is provided here.

1. Activity generation: TASHA follows a ‘‘bottom up’’ approach, i.e. activities are generated first and then scheduled. This feature supports dynamic scheduling so that schedules can change constantly with new opportunities and constraints, an individual faces before execution of his or her schedule. The activity generation component generates activities (individual and joint with other household members) for each individual. Activities such as work and school are generated first and then, considering the duration and start time of work and school activities, other types of activities are generated. This stage requires person and household attributes, and the frequency distributions of number, start time and duration of activities, by activity type as input. Frequency distributions of these activity attributes were developed from the 1996 TTS of the GTA, cross-classified by activity type, individual, household and schedule attributes such as age, gender, occupation, employment status, student status, presence of children, number of adults, individual work-school project status, and household work-school project status. In TASHA, these observed distributions of activity attributes are assumed to remain constant over time. The activity generation component employs Monte-Carlo simulation to draw the number of activities and their start time and duration from these observed frequency distributions.

2. Activity location choice: Home location and usual place of work/school are exogenous inputs into the model. The location choices of other activities are simulated using a series of entropy models based on employment, population, activity density, and distance variables (Eberhard, 2002).

3. Activity scheduling: The activity scheduling component is a rule-based procedure that first organizes the generated activities into work, school, shopping (joint and individual) and other (joint and individual) projects. Then consistent and feasible activity-travel schedules are developed for interacting household members. The method of activity scheduling is described here briefly. The detail of this process can be found in other papers (Miller & Roorda, 2003; Roorda, 2005).

Step 1: Activity episodes are inserted into a project agenda along with other activity episodes with a common purpose. A preliminary time sequence is formed.

Step 2: A person schedule is formed by taking activity episodes from the project agenda and adding them into the person schedule based on the order of precedence observed from an interactive computer survey of activity scheduling (Doherty, Nemeth, Roorda, & Miller, 2004).

Step 3: A “clean up” algorithm is applied to fine tune the person schedule just before /during execution of the schedule.

TASHA has been validated in two ways (Roorda, Miller, et al., 2008). First, the model is verified by comparing base year model outcomes to the base year dataset (i.e. 1996 TTS from which the model is originally derived). This base year verification tests the activity generation (frequency, start time, and duration), activity location choice, and activity scheduling model components of TASHA. Then, the model is further validated by comparing forecasts of a future year with the observed data of the same year (i.e. 2001 TTS in the Greater Toronto Area). The validation results indicate that TASHA is capable of reproducing activity/travel patterns in the GTA, at least for short-term forecasts, as activity attributes distributions have remained quite stable. However it needs further improvements in the modelling framework for the GTA (for instance - improvement in location choice model, development of activity generation model, applying different rules of scheduling for different population segments and so on) (Roorda, Miller, et al., 2008).

4.4 Data and research method

Three of the five components of the TASHA model have been transferred from Toronto to Montreal (activity generation, activity location choice and activity scheduling). A detailed methodological framework of transferring the model is presented in Figure 4.1.

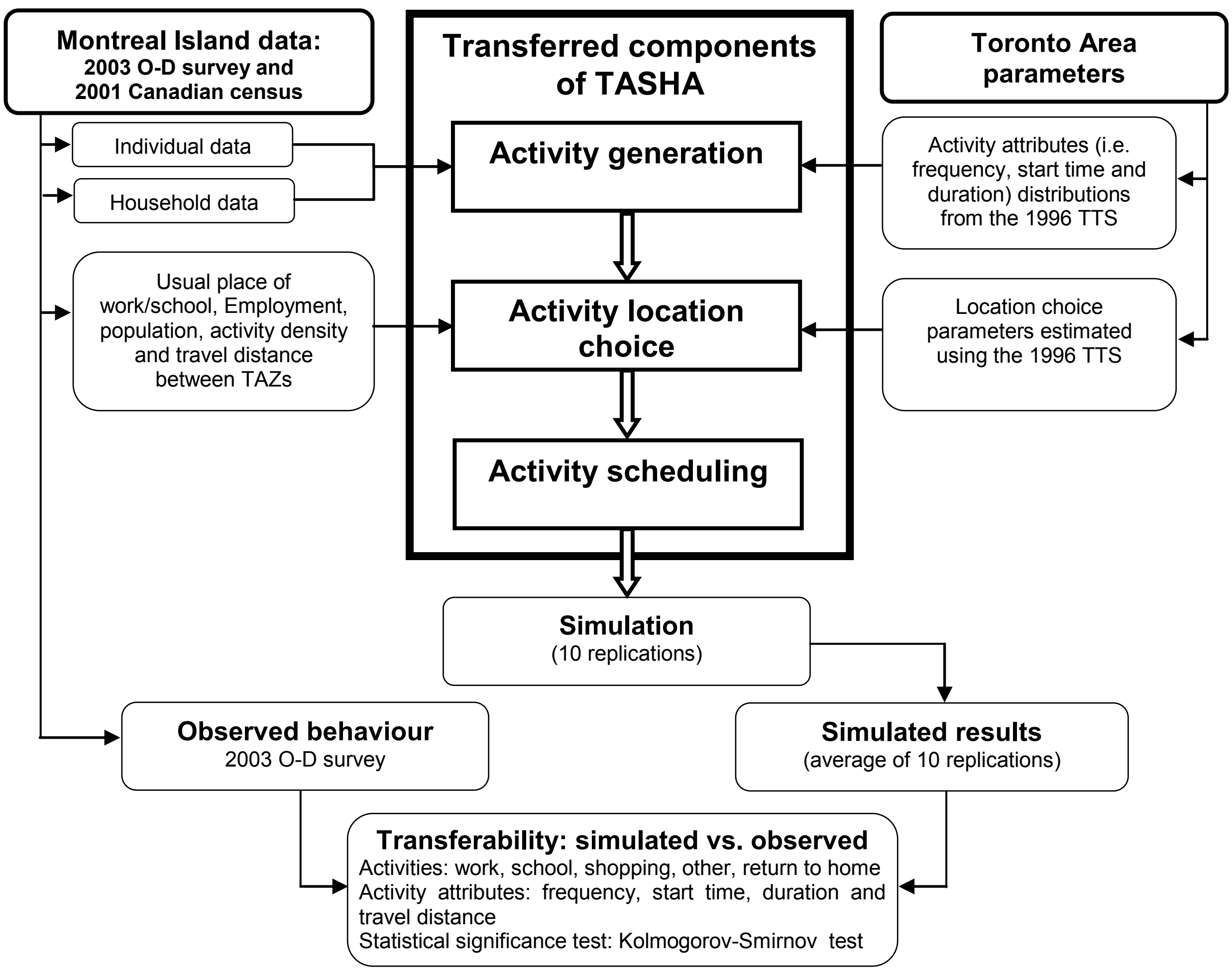


Figure 4.1: Methodological framework of transferring the TASHA model

This research employs the Toronto parameters, estimated using the 1996 TTS of the GTA without any adjustment to simulate the activity schedules of the individuals of the Montreal Island. We hypothesize that activity attributes distributions from the 1996 TTS of the GTA and location choice parameters estimated using the 1996 TTS are similar for the Montreal Island. Also like Toronto, we assume that these observed activity attributes distributions remain constant over time. Thus, observed distributions of activity attributes (such as frequency, start time, and duration of different types of activity) and location choice model parameters from the 1996 TTS are used as inputs in the activity generation and activity location choice components of TASHA, respectively. Other input data such as individual and household attributes for activity generation, as well as home and work/school location, employment, population, activity density, and travel distance between traffic analysis zones for the activity location choice model are gathered from the Montreal 2003 O-D travel survey and the 2001 Canadian Census.

Montreal has undertaken large scale Origin-Destination (O-D) travel surveys every five years since the early 1970s. The survey collects detailed travel and socio-economic information of approximately 5% of the total Greater Montreal Area (GMA) population aged 5 years and older. The socio-economic data of individuals and households and travel data of a specific weekday of the fall period (September to December) of all household members are collected by a telephone interview. Detailed information on the O-D travel surveys can be found on the AMT website (Agence Métropolitaine de Transport [AMT], 2014).

Some individuals in the 2003 O-D travel survey made open chains (trip chains that did not start and end at home). Households with individuals that made open trip chains were excluded. Thus, after data preparation, a sample of 59,624 individuals (26,960 households) is used for evaluating the spatial transferability of the TASHA model.

Since TASHA uses a stochastic approach for simulating activity scheduling, the model is run for ten replications following the same number of runs as the Toronto Area to compare the TASHA model performance in both cities (Roorda, Miller, et al., 2008). However, few recent studies have investigated the minimum number of runs needed to obtain stable outcomes and reported mixed results. Castiglione, Freedman, and Bradley (2003) have showed that the stable outcomes at the aggregate level could be obtained after only one simulation; however others have indicated that multiple runs are required to attain the stable outcomes. Different studies have suggested different number of runs (for instance 20 runs by (Ziems, Sana, Plotz, & Pendyala, 2011), 25 - 30 runs by (Rasouli, Arentze, & Timmermans, 2012), or even 200 runs by (Cools, Kochan, Bellemans, Janssens, & Wets, 2011)).

TASHA outputs and the 2003 O-D travel survey are then processed to prepare activity attributes for comparison. TASHA simulates activity schedules that include eleven types of activities, which are aggregated into five broad activity classes according to the commonly used patterns in Montreal (work, school, shopping, other, and return to home). For both TASHA output and the O-D travel survey, activity start time corresponds to the travel start time. The activity duration also includes travel time to the activity location. Trip distances (Euclidean distances) are calculated using the coordinates of the centroids of traffic analysis zones (TAZs) from an origin to a destination point. Average values of activity attributes over ten replications are computed. Finally, modelled average values computed over ten replications are compared with the 2003 O-D travel survey for different activity attributes, activity frequency, start time, duration, and travel distance. Kolmogorov-Smirnov tests (K-S tests) are performed to examine whether the modelled and observed distributions come from the same distribution. In addition, a systematic comparison of the TASHA model performance in the Montreal Island and the Toronto Area has been conducted for all activity attributes to better understand the model transferability. The Toronto Area validation (using the 2001 TTS) results are gathered from (Roorda, Miller, et al., 2008).

Though the results in this paper are produced using the model outcomes of ten replications of the TASHA model, an additional ten replications have also been conducted to examine whether ten runs are sufficient to obtain the stable outcomes. The start time distributions of work and shopping activities over the day have been reproduced compiling all twenty replications and compared with the same distributions produced using the model outcomes of ten replications. The comparison indicates that there is no significant variability between these two distributions (also confirmed by the K-S tests (P = 1.000 for both activities)), thus we conclude that ten replications are sufficient to obtain the stable outcomes for the TASHA model.

As of 2012, the Montreal Island is composed of an area of 500 square km and a population of 1,886,000, whereas the GTA is composed of an area of 7,125 square km with a population of 6,054,000 (Statistics Canada, 2011). This indicates that the GTA is approximately 14 and 3 times higher in terms of area and population, respectively, than the Montreal Island.

Table 4.1: Some key socio-demographic characteristics - Montreal Island vs. Toronto Area

Variables	Montreal Island (%)	Greater Toronto Area (%)	(Δ %)
Age groups			
Ages 24 years old and less	28.24%	33.10%	-4.86% *
Ages 25-54 years old	46.35%	47.40%	-1.05% *
Ages 55 years old and more	25.41%	19.50%	5.91% *
Gender			
Male	48.89%	48.90%	-0.01% NS
Female	51.11%	51.00%	0.11% NS
Employment status			
Full-time worker	39.43%	37.90%	1.53% *
Part-time worker	5.92%	8.50%	-2.58% *
Full-time work at home	0.07%	1.70%	-1.63% *
Part-time work at home	0.0%	0.50%	-0.50% *
Not employed	46.44%	51.40%	-4.96% *
Student status			
Full-time student	22.47%	21.80%	0.67% *
Part-time student	0.95%	2.70%	-1.75% *
Not a student	76.58%	75.40%	1.18% *
Possession of driver's license			
Yes	59.27%	63.00%	-3.73% *
No	40.73%	37.00%	3.73% *
Possession of a transit pass			
Yes	16.59%	5.00%	11.59% *
No	83.41%	94.80%	-11.39% *
Number of vehicles			
0	32.89%	15.30%	17.59% *
1	46.71%	41.40%	5.31% *
2	17.42%	34.90%	-17.48% *
3+	2.98%	8.40%	-5.42% *
Number of children			
0	77.41%	63.10%	14.31% *
1	11.11%	15.40%	-4.29% *
2	8.53%	15.40%	-6.87% *
3+	2.95%	6.10%	-3.15% *
Number of adults			
0	0.06%	0.00%	0.06% *
1	42.15%	25.50%	16.65% *
2	46.00%	52.90%	-6.90% *
3	8.32%	13.60%	-5.28% *
4	2.83%	6.00%	-3.17% *
5+	0.64%	1.50%	-0.86% *

* indicates statistical significance at the 99% confidence level and NS indicates not significant

A comparative analysis of key socio-demographic characteristics between the Montreal Island and the Toronto Area has been presented in Table 4.1 (Data Management Group, 1996). The Toronto Area has a greater proportion of individuals aged 24 years old and less, and a lower proportion over 55 years old. However, gender composition is quite similar in both cities. The unemployed population is significantly higher and the proportion of students is slightly higher in the Toronto Area. More people possess a driver's license in the Toronto Area; whereas more people have a transit pass in the Montreal Island, which could indicate their preferences in travel choices. Also, the proportion of the households that do not own a car is significantly higher on the Montreal Island, whereas the proportion with more than one car is significantly higher in the Toronto Area. Households with children are more common in the Toronto Area and one adult households are more common on the Montreal Island.

4.5 Spatial transferability of the TASHA model

4.5.1 Activity frequency

Table 4.2 compares the average activity frequencies simulated by TASHA with the observed frequencies of O-D travel survey by activity type in the Montreal Island. It also presents the Toronto Area validation (using 2001 TTS) results of activity frequency, gathered from (Roorda, Miller, et al., 2008) to systematically compare the model performance in both cities. It demonstrates that in the Montreal Island, TASHA undersimulates the total number of observed activities by 10.3% (13,570 activities). In the Toronto Area, TASHA undersimulates the total number of observed activities by 0.2% in the verification test using the same year dataset (i.e.1996 TTS) on which TASHA has been developed, and undersimulates the total number of observed activities by 3.2% in the validation test using a future year dataset (i.e. using the 2001 TTS), as also shown in Table 4.2 (Roorda, Miller, et al., 2008). Though the Montreal Island result is not as good as the result found for the Toronto Area, considering spatial (using Toronto parameters) and temporal differences (using survey parameters from an earlier year) the result is not unexpected, but seems quite promising.

For the Montreal Island, all activity types are undersimulated except for the work activity. The reason could be behavioural differences between the Toronto Area and the Montreal Island or could be in part due to conflicts in activity scheduling which cause the rejection of lower priority activity episodes. Only the work activity is oversimulated (+5.4%). As also shown in the Toronto Area validation (Table 4.2), TASHA simulates the work activity more closely than school, shopping and other activities. This is expected as the work activity is considered to be the highest priority activity among others in TASHA, thus it is scheduled first. On the other hand, school and return to home activities are undersimulated by 6.4%, and 8.7%, respectively. The undersimulation of shopping activities is greater (13.9%) than work, school and return to home activities; however the simulation result is quite similar as the Toronto Area (-13.2%), as shown in Table 4.2. The largest undersimulation is associated with other activities (30.6%). This difference indicates that in the urban area of the Montreal Island, individuals are making a larger number of other household maintenance trips, potentially because they have greater accessibility to opportunities at a shorter distance.

4.5.2 Activity start time

Modelled and observed travel/activity start time distributions by activity type in the Montreal Island are presented in Figure 4.2. K-S tests are applied to examine whether modelled and observed distributions are similar. K-S tests results (D statistic[1] and P-value[2]) of activity start time distributions of both cities are presented in Table 4.3 to systematically examine the TASHA model performance in both cities.

Figure 4.2 (a) shows that work activity start times are closely simulated over the day within 2% for all but 1 hour (7:00 - 7:59 AM). Like the Toronto Area as shown in Table 4.3, the K-S test result (P = 0.893) for work activities provides strong evidence that the simulated and observed distributions are similar. We perceive, however, that the simulated distributions by TASHA are shifted in time (undersimulation of early starting activities at 7:00 AM and oversimulation of starting activities at 9:00 AM). The comparison indicates that people in the Montreal Island leave

---

[1] Test based on the largest difference between the two cumulative distributions.

[2] P-value, equal or smaller than the assumed significance level suggests that two distributions are sampled from populations with different distributions.

earlier for work activities than residents in the Toronto Area. However, some people could start early because of the congestion as comparison shows the travel start times.

Figure 4.2 (b) shows the same distributions for school activities. These start times are also simulated over the day within 2% for all but 2 morning hours of the day (7:00 - 7:59 AM and 8:00 - 8:59 AM). There is a significant undersimulation of trips in start time between 7:00 - 7:59 AM and oversimulation of those occurring between 8:00 - 8:59 AM. As shown in Table 4.3, the K-S test (P = 0.441) result is inconclusive for these two distributions. We observe that in the Toronto Area, the majority of observed school activities is concentrated in one morning hour (i.e. 8:00 - 8:59 AM) in both base year and future year datasets (Roorda, Miller, et al., 2008), but in the Montreal Island the large portion of observed school activities (total 71.44%) is distributed in two morning hours (7:00 - 7:59 AM and 8:00 - 8:59 AM) and the percentage distributions at these two hours are close to each other. This indicates differences in school start times between the Montreal Island and the Toronto Area.

Shopping activity start times are simulated within -3.3% to 4.3% (Figure 4.2 (c)). The K-S test result (P = 0.992), shown in Table 4.3 indicates a very high probability that the modelled and observed shopping distributions are from the same distributions. However, TASHA slightly undersimulates the proportion of shopping activities starting in early hours as well as in late afternoon. The K-S test for the other activity type (P = 0.031) rejects the null hypothesis that both distributions are from the same distribution (Table 4.3). For return to home activities (Figure 4.2 (e)), despite an oversimulation from 12:00 PM to 3:59 PM and undersimulation from 4:00 PM to 6:59 PM, the K-S test (P = 0.992) suggests that the distributions are similar (Table 4.3).

It is useful to compare these results to similar tests (i.e. K-S tests) undertaken for model validation in the Toronto Area. There, the modelled and observed activity start time distributions for all five activities were similar as indicated by the K-S tests (P > 0.97 for all activity types in verification test and P > 0.86 for all activity types in validation test, as also shown in Table 4.3 (Roorda, Miller, et al., 2008). In the Montreal application, the comparison shows that overall TASHA captures the temporal aspects of travel behaviours almost as well for work, shopping, and return to home activities, whereas it shows large differences only for other activity. For school activity, TASHA simulated results are quite good for 24 hours except two morning hours.

Table 4.2: Activity frequency and average distance comparison in the Montreal Island and the Toronto Area, TASHA vs. observed

Activity type	Montreal Island (TASHA vs. 2003 O-D survey)						Toronto Area[a] (TASHA vs. 2001 TTS)					
	Work	**School**	**Shopping**	**Other**	**Return to home**	**Total**	**Work**	**School**	**Shopping**	**Other**	**Return to home**	**Total**
Model average total activities (TASHA)[b]	26027	12170	9329	16862	53946	118333	143990	41987	46844	84577	265031	582429
Model Standard deviation total activities (TASHA)[b]	131	41	138	194	172	382	329	62	357	360	364	1131
Observed total activities	24693	13004	10829	24309	59068	131903	145123	43930	53989	93771	264588	601401
Model ± activities (#)	1334	-834	-1500	-7447	-5122	-13570	-1133	-1943	-7145	-9194	443	-18972
Model ± activities (%)	5.4	-6.4	-13.9	-30.6	-8.7	-10.3	-0.8	-4.4	-13.2	-9.8	0.2	-3.2
Model average distance (km)[b]	7.94	4.43	5.11	6.91	6.12	6.38	13.20	5.58	5.80	7.88	8.92	9.34
Model standard deviation distance (km)[b]	0.03	0.02	0.05	0.03	0.02	0.02	0.02	0.02	0.03	0.03	0.01	0.01
Observed average distance (km)	7.94	4.21	3.63	5.01	5.61	5.63	12.78	5.30	5.40	8.02	9.22	9.26
Model ± distance (km)	0.00	0.22	1.48	1.90	0.51	0.75	0.42	0.28	0.40	-0.14	-0.30	0.07
Model ± distance (%)	0.01	5.3	40.9	37.9	9.1	13.3	3.2	5.3	7.4	-1.8	-3.2	0.8

[a] The Toronto Area validation results are gathered from (Roorda, Miller, et al., 2008).
[b] Model results are the average value (standard deviation) of ten replications.

Table 4.3: K-S test results for activity start time and average duration distributions in the Montreal Island and the Toronto Area, TASHA vs. observed

Activity type	Montreal Island (TASHA vs. 2003 O-D survey)					Toronto Area[a] (TASHA vs. 2001 TTS)				
	Work	**School**	**Shopping**	**Other**	**Return to home**	**Work**	**School**	**Shopping**	**Other**	**Return to home**
Activity start time	D= 0.167 P = 0.893	D= 0.250 P = 0.441	D= 0.125 P = 0.992	D= 0.417 P = 0.031	D= 0.125 P = 0.992	D= 0.167 P = 0.861	D= 0.135 P = 0.993	D= 0.087 P = 1.000	D= 0.083 P = 1.000	D= 0.120 P = 0.990
Activity duration	D= 0.333 P = 0.139	D= 0.083 P = 1.000	D= 0.292 P = 0.259	D= 0.333 P = 0.139	------	D= 0.286 P = 0.304	D= 0.263 P = 0.462	D= 0.217 P = 0.593	D= 0.417 P= 0.021	------

[a] The Toronto Area validation results are gathered from (Roorda, Miller, et al., 2008).

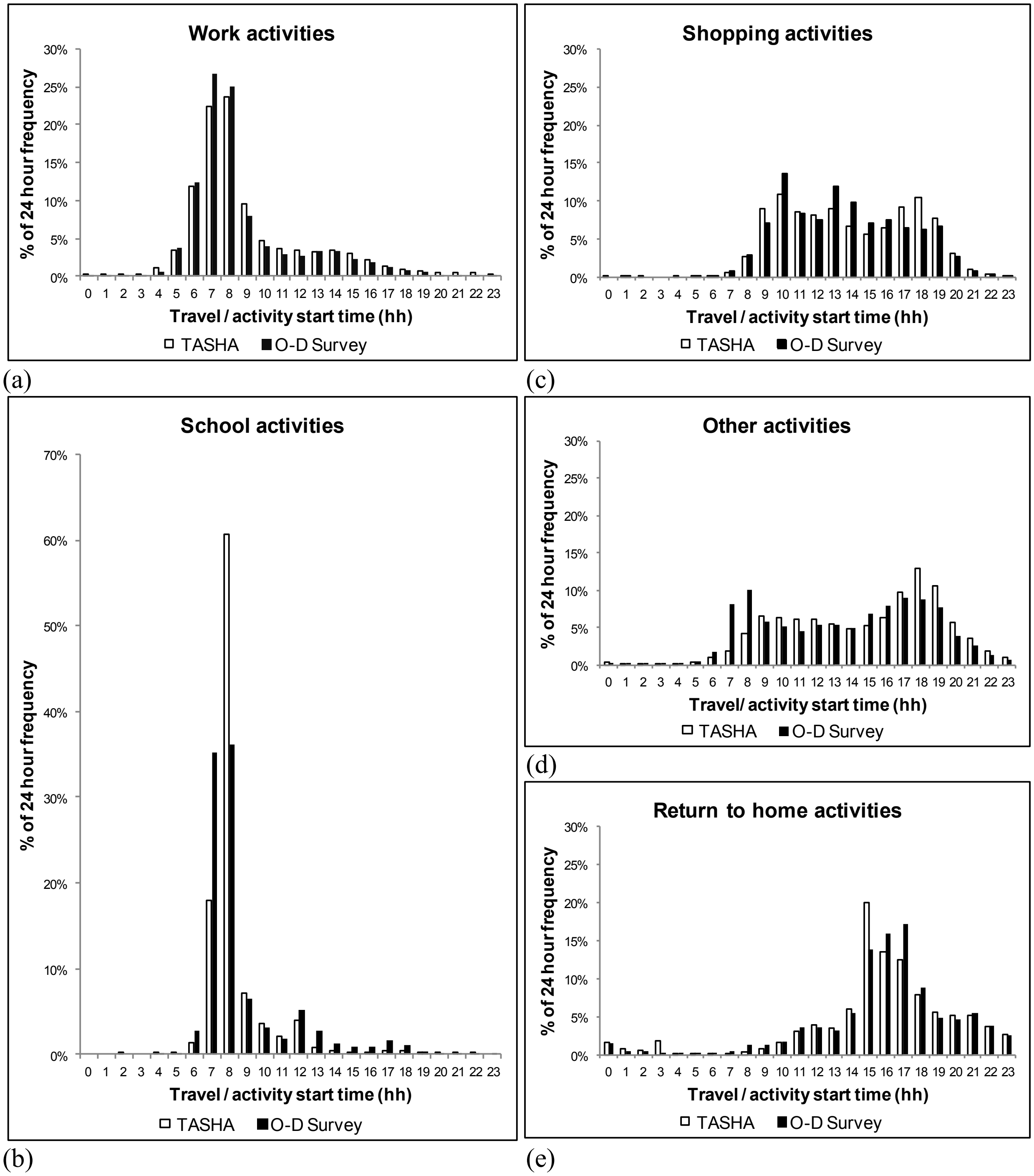


Figure 4.2: Travel / activity start time distributions in the Montreal Island: TASHA vs. O-D Survey (a) Work activities, (b) School activities, (c) Shopping activities, (d) Other activities, and (e) Return to home activities

People in the Montreal Island also leave home earlier and arrive home later from all activities than the model based on the Toronto Area data would predict.

4.5.3 Activity duration

Figure 4.3 presents average activity durations for each activity type by travel/activity start times in the Montreal Island. As expected, average activity durations for all activity types are lower for activities that started later in the day. K-S tests are also applied to examine significant differences between modelled and observed distributions for all activities. K-S test results of average activity duration distributions of both cities have been presented in Table 4.3.

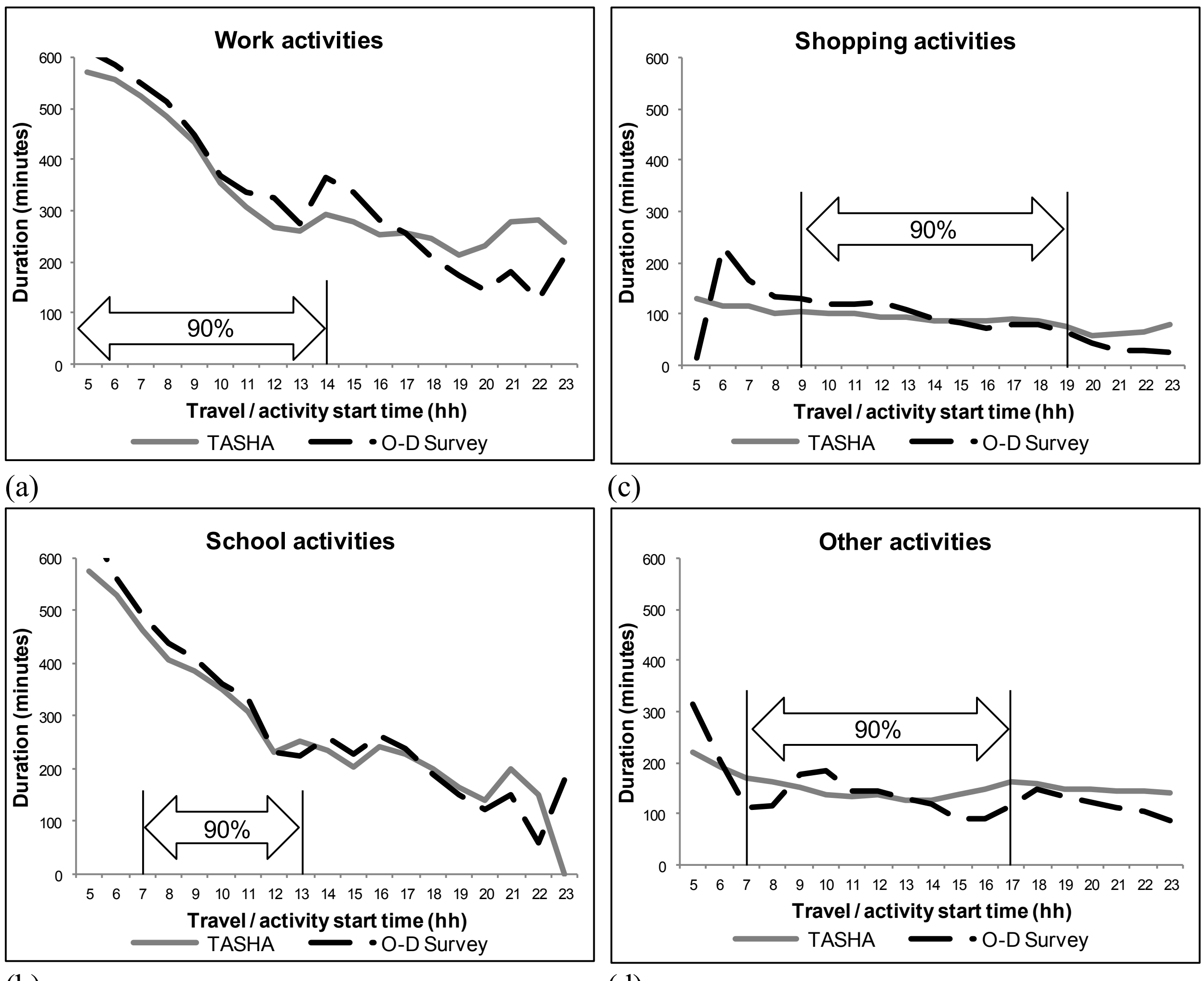


Figure 4.3: Activity duration distributions in the Montreal Island: TASHA vs. O-D Survey (a) Work activities, (b) School activities, (c) Shopping activities, and (d) Other activities

For the Montreal Island as shown in Table 4.3, the K-S tests reject the null hypotheses for work (P = 0.139), shopping (P = 0.259) and other activities (P = 0.139) and it accepts the null hypothesis for school activities (P = 1.000) that modelled and observed distributions are similar.

However, we observe that the higher differences between modelled and observed average durations are seen at those hours of the day when there are low activity frequencies. The bands on Figure 4.3 indicate the hours of the day when the majority (90%) of the trips takes place. The K-S test results for the Montreal Island application are similar (except school activities) to those found in the Toronto Area verification and validation tests (Table 4.3), in which the hypothesis of similar duration distributions are either rejected (P ranges from 0.021 to 0.109) or inconclusive (P ranges from 0.304 to 0.593) for all activity types (Roorda, Miller, et al., 2008). For work activities, average durations are undersimulated for activities starting before 5:00 PM (Figure 4.3(a)). As observed, the major portion (90%) of work activity start between 5:00 AM -2:00 PM which is shown by the band. Within the band, the simulated average durations are within 10% of the observed average durations at every hour except 12:00 PM and 2:00 PM (for which the differences are within 20%). Also, a K-S test (P = 0.759) for the portion of the distribution within the 90% band hours provides strong evidence that the simulated and observed distributions are similar.

School activity durations are also undersimulated for almost every time except 12:00 PM to 1:59 PM, and after 6:00 PM (Figure 4.3(b)). For the Montreal Island, activity duration estimates for the school activity provides a better result than both verification and validation tests (Table 4.3) in the Toronto Area (Roorda, Miller, et al., 2008). 90% of school trips take place between 7:00 AM and 1:00 PM shown by the band in Figure 4.3(b). Within this time period, the simulated average durations are almost always within 10% of observed durations.

For shopping activities (Figure 4.3(c)) average durations are undersimulated for activities starting before 3:00 PM, and oversimulated after 3:00 PM (though the differences are smaller). The band shows that 90% of shopping activity starts between 9:00 AM and 7:00 PM and within this time period the differences are within 20%.

For other activities (Figure 4.3(d)), the differences between simulated and observed are much more variable throughout the day; TASHA undersimulates the durations of activities which started early in the day and oversimulates the durations for activities in the late afternoon.

4.5.4 Activity location

Table 4.2 compares the TASHA simulated average distances with the observed distances by activity type in the Montreal Island and the Toronto Area (Roorda, Miller, et al., 2008). For the Montreal Island, TASHA oversimulates the observed average distance by 13.3%. The average travel distance reported in the O-D survey is 5.63 km, whereas the simulated average distance is 6.38 km. The oversimulation of distance is greater than the minor differences found in verification (0.9%) and validation tests (0.8%), as also shown in Table 4.2, conducted in the Toronto Area (Roorda, Miller, et al., 2008).

Oversimulation of distance occurs for all activity types. This may happen because of using location choice parameters estimated using the 1996 household travel survey data from the Toronto Area without any change for the Montreal Island. The Toronto Area, for which the model has been developed, is approximately 14 times larger than the Montreal Island. This scale difference could be a possible source of error. Interestingly, the simulation provides better results for work and school activities than the Toronto Area. Average distance for the work activity is not oversimulated (i.e. within 0.01%) in the Montreal Island, while work travel distances are oversimulated by 2.7% in the verification test and by 3.2% in the validation test (Table 4.2) in the Toronto Area (Roorda, Miller, et al., 2008). In addition, average distance for school activities is oversimulated by 5.3% whereas in the Toronto Area the distances were oversimulated by 10.5% in verification test and by 5.3% (similar to the Montreal Island) in the validation test, as also shown in Table 4.2. It is important to mention here that usual place of work/school is directly input in to the modelling framework which, of course, greatly reduces uncertainty for work and school. The largest differences between modelled and observed distances are for shopping (40.9%) and other (37.9%) activity types. This may happen because the location choice model (within the TASHA framework) for activities which are not taking place at the usual place of work/school is a spatial interaction model which is either home-based or work-based (Roorda, Miller, et al., 2008). For example, locations for shopping and other activities are chosen based only on the distances from home, other interactions (for instance people could do these activities close to work place) are not taken in to account while choosing the locations.

4.5.5 Summary of the transferability results and the TASHA model performance

This research also systematically investigates the TASHA model performance in the Montreal Island and the Toronto Area for all activity attributes, as shown in Table 4.2 and Table 4.3. Overall similar to the Toronto Area, TASHA undersimulates the total number of observed activities in the Montreal Island, but with higher difference (10.3% compared to 3.2% in the Toronto Area). In addition, the simulations for work, school, and return to home activities are quite satisfactory and the simulation for shopping activity are also reasonably accurate in both cities. Only for other activities, the difference between the simulated and observed activities is significantly larger in the Montreal Island than the Toronto Area. Also, TASHA captures the temporal aspects of travel behaviour quite well for work, shopping, and return to home activities in both cities, however it shows larger differences for the other activity type in the Montreal Island than the Toronto Area. For school activities, TASHA simulated results are good throughout the day except two morning hours in the Montreal Island. The study reveals that people in the Montreal Island engage in activities (for instance, work, school and other activities) earlier than predicted by the model based on behaviour in the Toronto Area. Activity duration simulation for school activities provides a very good result, even better than the validation test in the Toronto Area. Similar to the Toronto Area, TASHA also provides reasonable results for work activity durations for those hours when the majority of work activities take place. However simulated durations are not good for shopping and other activities; these results are similar in both cities. TASHA oversimulates the observed average distance to all activities by 13.3% in the Montreal Island, which is larger than the slight difference found in the Toronto Area. Interestingly, the simulation provides better results for work and school activities than the Toronto Area. The result is also quite reasonable for return to home activities. The largest differences are observed for shopping and other activity types.

The systematic comparison of the model performance in both cities demonstrates that TASHA provides quite similar, in some cases even better results in the Montreal Island for most of the activity attributes (except start time for school activities, average distance for shopping activities and, frequency, start time and average distance for other activities).

4.6 Conclusions

This paper presents an empirical assessment of the spatial transferability of an activity-based travel demand model, TASHA (Travel Activity Scheduler for Household Agents). To assess the spatial transferability, simulated activity attributes by TASHA and observed attributes from the 2003 O-D travel survey in the Montreal Island have been compared. Activity attributes include activity frequency, start time, average duration, and average travel distance.

At the aggregate level of analysis, TASHA provides reasonable outcomes for the Montreal Island (in some cases - better results than the Toronto Area) for all four attributes for work, school and return to home activities with few exceptions (for instance, school start time). Also, the model outcomes are promising for shopping frequency and start times, however TASHA provides larger differences for average shopping durations (similar to the Toronto Area result (Roorda, Miller, et al., 2008)) and distances. Only the outcomes for all four attributes for the other activity type differ greatly from the observed attributes in the Montreal Island, however the Montreal Island result for average activity duration is quite similar to the Toronto Area validation result.

The large differences observed in some cases most likely indicate real differences in behaviour between the Montreal Island and the Toronto Area. Two cities are distinctly different in terms of socio-demographic, cultural, economic development, employment pattern, commuting pattern and so on (Balakrishnan, Maxim, & Jurdi, 2005; Heisz, 2006; Roorda, Morency, & Woo, 2008; Rose, 1999; Shearmur, 2006). Also the model, developed for the Toronto Area has been applied on the highly urbanized setting of the Montreal Island, which is approximately 14 times smaller than the Toronto Area. This scale difference could also contribute to those large differences found in this research. In addition, TASHA has been transferred to the Montreal Island without any adjustment in the parameters from Toronto settings. The model parameters, developed from the 1996 TTS of the GTA have been used to simulate the activity schedules of the Montreal Island individuals for the year 2003. As it shows, there is a significant time gap between these two datasets of the Montreal Island and the Toronto Area. However TASHA is able to replicate activity/travel patterns with precision, at least for short term forecasts (Roorda, Miller, et al., 2008), this time gap might also contribute to the large differences found in this research. It is plausible that parameters, estimated from a previous survey year (i.e. the 1996 TTS) may not remain constant over time. It would be useful to examine the changes in activity attributes for different activities over time. Yasmin et al. (2012) have confirmed that activity attributes for work, school, shopping and other activities have changed over a period of 10 years (1998-2008) in the Greater Montreal Area. Thus, an improved activity generation model within the TASHA modelling framework should be developed so that it would be sensitive to activity scheduling constraints, could capture behavioural trade-offs involved in time allocation decisions as well as could reflect temporal changes in travel behaviour (Habib & Miller, 2009; Yasmin et al., 2012).

Considering these facts, transferability results seem quite promising. This application provides substantial evidence for the spatial transferability of TASHA, at least in case of fixed activities (i.e. work, school); however the model is less successful in replicating some activity attributes for flexible activities (i.e. shopping, other). This also indicates that scheduling behaviour of fixed activities in two geographic areas are quite similar and more stable than those of flexible activities in spite of having spatial and temporal differences between two areas.

To validate whether large differences indicate real behavioural differences between the Montreal Island and the Toronto Area, future research could analyze differences in activity-travel behaviour using travel survey data from both cities. It could also compare observed activity attribute distributions from travel surveys of both cities. This validation test (spatial transferability) has been done at the macro-level (aggregation of the entire population), future research could address the validation at the meso-level (for instance, aggregation by population segments by age group and gender, and by home location), and the micro-level (for instance, households/individuals), which is customary in Montreal while evaluating the performance of a model (Yasmin, Morency, & Roorda, 2014). In general, we conclude that the TASHA model could be transferred to a new developing area where the dataset is not still available. However, re-estimation of model parameters and the use of local activity attribute distributions (frequency, start time and duration) is a desirable step in the transfer of the TASHA model from one context to another.

Activity behaviours are becoming complex due to several reasons such as socio-economic changes, growing congestion, innovative policy instruments and so on (Shiftan & Ben-Akiva, 2011). Activity-based travel demand models are more efficient in capturing individuals' activity behaviours than the traditional four-stage travel demand models. Along with other advantageous features, policy sensitivity towards emerging policies was one of the key motivations in development of these more behaviourally realistic activity-based models. However to increase practical application of activity-based models, policy sensitivity of such models (such as TASHA) need to be assessed, which is of great interest to practitioners. Thus, future research could test some policy scenarios with the adapted TASHA with parameters calibrated using the local datasets from the Montreal Island. Some policy scenarios could include:

- Demographic scenarios (for instance impact of ageing population, and different composition of labour force participation),
- Land-use scenarios (for instance introducing new shopping mall, concentrated development in a centre, mixed development), and
- Demand-oriented scenarios (for instance alternative working hours (such as flexible working hours), HOV lanes (such as ridesharing), parking restrictions, and congestion pricing in AM period).

Acknowledgements

The authors would like to thank the Ministère des transports du Québec (Service de la modélisation des systèmes de transport) for their support and contributions. They also wish to thank the technical committee on O-D travel surveys in the Montreal region for providing access to data for research purposes.

References

Agence Métropolitaine de Transport. (2014). Portrait de la mobilité. Retrieved December 10, 2014, from https://www.amt.qc.ca/fr/a-propos/portrait-mobilite

Agyemang-Duah, K., & Hall, F. L. (1997). Spatial transferability of an ordered response model of trip generation. *Transportation Research Part A: Policy and Practice, 31*(5), 389-402.

Arentze, T., Hofman, F., Van Mourik, H., & Timmermans, H. (2002). Spatial transferability of the Albatross model system: Empirical evidence from two case studies. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 1-7.

Arentze, T., & Timmermans, H. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological, 38*(7), 613-633. doi: 10.1016/j.trb.2002.10.001

Arentze, T., & Timmermans, H. (2008). *ALBATROSS: Overview of the model, application and experiences*. Paper presented at the Innovations in Travel Modeling 2008 Conference, Portland, Oregon, USA.

Auld, J., & Mohammadian, A. K. (2012). Activity planning processes in the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Research Part A: Policy and Practice, 46*(8), 1386-1403. doi: http://dx.doi.org/10.1016/j.tra.2012.05.017

Auld, J., & Mohammadian, A. K. (2009). Framework for the development of the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Letters: The International Journal of Transportation Research, 1*(3), 243-253.

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

Balakrishnan, T. R., Maxim, P., & Jurdi, R. (2005). Social class versus cultural identity as factors in the residential segregation of ethnic groups in Toronto, Montreal and Vancouver for 2001. *Canadian Studies in Population, 32*(2), 203-227.

Ben-Akiva, M. E., & Bowman, J. (1998). Activity-based travel demand model systems. In P. Marcotte & S. Nguyen (Eds.), *Equilibrium and Advanced Transportation Modeling* (pp. 27-46). Boston/ Dordrecht/London: Kluwer Academic Publishers.

Bhat, C. R. (1999). *A comprehensive and operational analysis framework for generating the daily activity travel profiles of workers.* Paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC, USA.

Bhat, C. R., Guo, J. Y., Srinivasan, S., & Sivakumar, A. (2004). Comprehensive econometric microsimulator for daily activity-travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1894*, 57-66. doi: 10.3141/1894-07

Bhat, C. R., & Lawton, K. T. (2000). Passenger travel demand forecasting *Millennium Paper*. Washington D.C., USA: Transportation Research Board, National Research Council.

Bhat, C. R., Paleti, R., Pendyala, R. M., & Goulias, K. G. (2013). SimAGENT activity-based travel demand analysis *SimAGENT Core Models* (Vol. 4): GeoTrans Laboratory, University of California, Santa Barbara.

Bowman, J. L. (1998). *The day activity schedule approach to travel demand analysis.* (Doctor of Philosophy), Massachusetts Institute of Technology, Cambridge, Massachusetts.

Bowman, J. L., & Ben-Akiva, M. E. (2000). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice, 35*(1), 1-28.

Bowman, J. L., Bradley, M., Castiglione, J., & Yoder, S. L. (2014). *Making advanced travel forecasting models affordable through model transferability.* Paper presented at the 93rd Transportation Research Board Annual Meeting, Washington D.C., USA.

Bradley, M. A., & Bowman, J. L. (2006). *A summary of design features of activity-based microsimulation models for U.S. MPOs.* Paper presented at the TRB Conference on Innovations in Travel Demand Modeling, Austin, Texas

Castiglione, J., Freedman, J., & Bradley, M. (2003). Systematic investigation of variability due to random simulation error in an activity-based microsimulation forecasting model. *Transportation Research Record: Journal of the Transportation Research Board, 1831*, 76-88. doi: 10.3141/1831-09

Childress, S., Sabina, E., Kurth, D., Rossi, T., & Malm, J. (2010). *DRCOG activity-based model calibration/ validation.* Paper presented at the Innovations in Travel Modeling Conference, Transportation Research Board, Tempe, Arizona.

Cools, M., Kochan, B., Bellemans, T., Janssens, D., & Wets, G. (2011). Assessment of the effect of micro-simulation error on key travel indices: Evidence from the activity-based model Feathers. *Transportation Research Record, 1558*.

Cotrus, A. V., Prashker, J. N., & Shiftan, Y. (2005). Spatial and temporal transferability of trip generation demand models in Israel. *Journal of Transportation and Statistics, 8*(1), 37-56.

Daor, E. (1981). *The transferability of independent variables in trip generation models.* Paper presented at the Transportation Analysis and Models: Proceedings of Seminar N, PTRC Education and Research Services, Ltd., London, England.

Data Management Group. (1996). Transportation Tomorrow Survey: Verson 2.1 Data Guide: Joint Program in Transportation, University of Toronto, Toronto.

Doherty, S. T., Nemeth, E., Roorda, M. J., & Miller, E. J. (2004). Design and assessment of the Toronto area computerized household activity scheduling survey. *Transportation Research Record: Journal of the Transportation Research Board, 1894*, 140-149.

Eberhard, L. K. (2002). *A 24-h household-level activity based travel demand model for the GTA.* (Masters Thesis), University of Toronto, Toronrto, Canada.

Ettema, D., Bergers, A., & Timmermans, H. (1993). Simulation model of activity scheduling behaviour. *Transportation Research Record: Journal of the Transportation Research Board, 1413*, 1-11.

Ettema, D., Borgers, A., & Timmermans, H. (1996). SMASH (Simulation Model of Activity Scheduling Heuristics): Some simulations. *Transportation Research Record: Journal of the Transportation Research Board, 1551*, 88-94.

Gangrade, S., Kasturirangan, K., & Pendyala, R. M. (2000). Coast-to-coast comparison of time use and activity patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1718*, 34-42.

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

Golledge, R. G., Kwan, M.-P., & Gärling, T. (1994). Computational process modeling of household travel decisions using a geographic information system. *Papers in Regional Science, 73*(2), 99-117.

Habib, K. M. N., & Miller, E. J. (2009). Modelling activity generation: A utility based model for activity agenda formation. *Transportmetrica, 5*(1), 3-23. doi: 10.1080/18128600802591384

Heisz, A. (2006). Canada's global cities: Socio-economic conditions in Montréal, Toronto and Vancouver *Trends and Conditions in Census Metropolitan Areas*. Ottawa: Statistics Canada.

Karasmaa, N. (2001). *The spatial transferability of the Helsinki Metropolitan Area mode choice models.* Paper presented at the 5 th workshop of the Nordic Research Network on Modeling Transport, Land-use, and the Environment, Nynashamn, Sweden.

Kitamura, R. (1996). *Applications of models of activity behavior for activity-based demand forecasting*. Paper presented at the Activity-Based Travel Forecasting Conference, New Orleans, Louisiana.

Kitamura, R., & Fujii, S. (1998). Two computational process models of activity-travel behavior. In T. Garling, T. Laitila & K. Westin (Eds.), *Theoretical Foundations of Travel Choice Modeling* (pp. 251-279). Bingley, UK: Emerald Group Publishing.

Kitamura, R., Pas, E. I., Lula, C. V., Lawton, T. K., & Benson, P. E. (1996). The sequenced activity mobility simulator (SAMS): An integrated approach to modeling transportation, land use and air quality. *Transportation, 23*(3), 267-291. doi: 10.1007/BF00165705

Koppelman, F. S., & Wilmot, C. G. (1982). Transferability analysis of disaggregate choice models. *Transportation Research Record: Journal of the Transportation Research Board, 895*, 18-24.

Kurth, D. L., Childress, S., Sabina, E. E., & Rossi, T. (2006). *Proposed validation and sensitivity testing of Denver region activity-based models.* Paper presented at the Innovations in Travel Demand Modeling Conference, Transportation Research Board, Austin, TX.

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

Miller, E. J., & Roorda, M. J. (2003). Prototype model of household activity-travel scheduling. *Transportation Research Record: Journal of the Transportation Research Board, 1831*, 114-121. doi: 10.3141/1831-13

Mohammadian, A. K., Auld, J., & Yagi, S. (2009). Recent progress on activity-based microsimulation models of travel demand and future prospects. In B. W. Sloboda (Ed.), *Transportation Statistics* (pp. 151-171): J. Ross Publishing, Inc.

Pendyala, R. M., Kitamura, R., Kikuchi, A., Yamamoto, T., & Fujii, S. (2005). Florida Activity Mobility Simulator: Overview and preliminary validation results. *Transportation Research Record: Journal of the Transportation Research Board, 1921*, 123-130. doi: 10.3141/1921-14

Pinjari, A. R., & Bhat, C. R. (2011). Activity-based travel demand analysis. *A Handbook of Transport Economics, 10*, 213-248.

Rasouli, S., Arentze, T., & Timmermans, H. (2012). *Analysis of uncertainty in performance indicators of a complex activity-based model: The case of the Albatross model system.* Paper presented at the Innovations in Travel Demand Modelling Conference, Tampa, USA.

Recker, W. W., McNally, M. G., & Root, G. S. (1986a). A model of complex travel behavior: Part I--Theoretical development. *Transportation Research Part A: General, 20*(4), 307-318. doi: 10.1016/0191-2607(86)90089-0

Recker, W. W., McNally, M. G., & Root, G. S. (1986b). A model of complex travel behavior: Part II--An operational model. *Transportation Research Part A: General, 20*(4), 319-330. doi: 10.1016/0191-2607(86)90090-7

Roorda, M. J. (2005). *Activity-based modelling of household travel.* (Doctor of Philosophy), University of Toronto, Toronto, Canada.

Roorda, M. J., & Miller, E. J. (2006). Past Presidents' Award for merit in transportation engineering: Assessing transportation policy using an activity-based microsimulation model of travel demand. *ITE Journal (Institute of Transportation Engineers), 76*, 16-21.

Roorda, M. J., Miller, E. J., & Habib, K. M. N. (2008). Validation of TASHA: A 24-h activity scheduling microsimulation model. *Transportation Research Part A: Policy and Practice, 42*, 360-375. doi: 10.1016/j.tra.2007.10.004

Roorda, M. J., Morency, C., & Woo, K. (2008). *Two cities, two realities? A closer look at the evolution of trip rates in Toronto and Montreal*. Paper presented at the 87th Transportation Research Board (TRB) Annual Meeting, Washington D.C., USA.

Rose, D. (1999). Urban hierarchies and the changing characteristics of 'urban professionals' in Toronto and Montreal. *Canadian Journal of Regional Science, XXII-1,2*, 133-142.

Rose, G., & Koppelman, F. S. (1984). Transferability of disaggregate trip generation models *Proceedings of the 9th International Symposium on Transportation and Traffic Theory* (pp. 471-491). Utrecht, Netherlands.

Shearmur, R. (2006). Travel from home: An economic geography of commuting distances in montreal. *Urban Geography, 27*(4), 330-359.

Shiftan, Y., & Ben-Akiva, M. (2011). A practical policy-sensitive, activity-based, travel-demand model. *The Annals of Regional Science, 47*(3), 517-541.

Shiftan, Y., Ben-Akiva, M., Proussaloglou, K., Jong, G. d., Popuri, Y., Kasturirangan, K., & Bekhor, S. (2003). *Activity-based modeling as a tool for better understanding travel behaviour*. Paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Sikder, S., Pinjari, A., Srinivasan, S., & Nowrouzian, R. (2013). Spatial transferability of travel forecasting models: A review and synthesis. *International Journal of Advances in Engineering Sciences and Applied Mathematics, 5*(2-3), 104-128. doi: 10.1007/s12572-013-0090-6

Statistics Canada. (2011). Population and dwelling counts, for Canada and census divisions. *Canada 2011 Census.* Retrieved July 30, 2013, from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=701

Vovsha, P., Petersen, E., & Donnelly, R. (2004). Impact of intra-household interactions on individual daily activity travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1898*, 87-97.

Wilmot, C. G. (1995). Evidence on transferability of trip-generation models. *Journal of Transportation Engineering, 121*(5), 405-410.

Wilmot, C. G., & Stopher, P. R. (2001). Transferability of transportation planning data. *Transportation Research Record: Journal of the Transportation Research Board, 1768*, 36-43.

Yagi, S., & Mohammadian, A. K. (2007a). *Microsimulation and policy applications of an activity-based modeling system of travel demand.* Paper presented at the 86th Transportation Research Board Annual Meeting, Washington DC, United States.

Yagi, S., & Mohammadian, A. K. (2007b). *Validation and policy applications of an activity-based microsimulation model of travel demand.* Paper presented at the 11th World Conference on Transportation Research, Berkeley, California.

Yagi, S., & Mohammadian, A. K. (2008a). Joint models of home-based tour mode and destination choices: Applications to a developing country. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 29-40.

Yagi, S., & Mohammadian, A. K. (2008b). Modeling daily activity-travel tour patterns incorporating activity scheduling decision rules. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 123-131.

Yasmin, F., Morency, C., & Roorda, M. J. (2012). *An empirical analysis of change in activity attribute distributions over time*. Paper presented at the 13th International Conference on Travel Behaviour Research organized by the International Association for Travel Behavior Research (IATBR), Toronto, Canada.

Yasmin, F., Morency, C., & Roorda, M. J. (2014). *Macro, meso, and micro-level validation of an activity-based travel demand model*. Paper presented at the 93rd Transportation Research Board (TRB) Annual Meeting, Washington D.C., USA.

Ziems, S., Sana, B., Plotz, J., & Pendyala, R. M. (2011). *Stochastic variability in microsimulation modeling results and convergence of corridor-level characteristics*. Paper presented at the 90th Transportation Research Board (TRB) Annual Meeting, Washington D.C., USA.

CHAPTER 5 ARTICLE 2: MACRO-, MESO-, AND MICRO-LEVEL VALIDATION OF AN ACTIVITY-BASED TRAVEL DEMAND MODEL

This chapter is presented in the form of a paper submitted to the *Transportmetrica A: Transport Science* on June 15, 2015.

Farhana Yasmin[a,*], Catherine Morency[a,1], Matthew J. Roorda[b,2]

[a] *Department of Civil, Geological and Mining Engineering, Polytechnique Montreal, C.P. 6079, Station Centre-Ville, Montreal, Quebec, Canada H3C 3A7*
[b] *Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario, Canada M5S 1A4*
[*] *Corresponding author. E-mail: farhana.yasmin@polymtl.ca*
[1] *E-mail: cmorency@polymtl.ca*
[2] *E-mail: roordam@ecf.utoronto.ca*

Abstract

Validation is an essential part of the model development process for practical applications in the real world context. This paper focuses on how and at which level the validation of activity-based travel demand models must be performed. It then examines the spatial transferability, as a validation test, of an activity-based model, TASHA (Travel Activity Scheduler for Household Agents). This paper applies the TASHA model to the Island of Montreal, Canada, using the 2003 Origin-Destination (O-D) travel survey and the 2001 Canadian Census, and validates the transfer by comparing modelled and observed activity attributes at three different levels of aggregation. Validation results at different levels (specially at macro- and meso-level) seem quite promising. TASHA can successfully reproduce activity behaviours of another context, at least for fixed activities (work, school) with few exceptions.

Key words: travel demand modelling; activity-based model; TASHA; model validation; spatial transferability; Montreal

5.1 Introduction

Validation is an essential part of the model development process for practical application in the real world context. Spatial transferability has been utilized as a useful validation test for travel demand models. To date, however, activity-based travel demand models have not been frequently assessed using this test. Accordingly, this paper focuses on how and at which level the validation of activity-based travel demand models must be performed. More specifically, it considers spatial transferability as a validation test of an activity-based model, TASHA (Travel Activity Scheduler for Household Agents) which has been developed by Miller and Roorda (2003) for the Greater Toronto Area (GTA), Canada. TASHA is a fully disaggregate microsimulation model, which estimates activity schedules and travel patterns for a 24 hour typical weekday for all individuals in a household. The trip diary data from the 1996 Transportation Tomorrow Survey (TTS) for the GTA has been employed to develop the TASHA model. Therefore, the application of the model is possible in any city where a similar dataset is available. In the Montreal Area, large scale Origin-Destination (O-D) travel surveys have been conducted every five years since the early 1970s. These surveys' outputs are very similar to those provided by the TTS, as the latter were inspired by the former ones, allowing application of TASHA in the context of Montreal. Therefore, the TASHA model has been applied to the Island of Montreal, Canada and this paper examines to what extent this activity-based model is transferable to this region by validating model outcomes aggregated across the entire population (macro-level), across population segments (meso-level) and for individuals (micro-level).

The remainder of this paper is organized as follows. First, a brief review of validation procedures for activity-based travel demand models is provided. Next, a brief overview of the TASHA model is presented. Then, data and research method are described. Next, the validation results at three different levels of aggregation (macro-, meso-, and micro-levels) are presented. Then, the paper summarizes the validation results and provides a discussion. A conclusion follows.

5.2 Background

Travel demand models are important decision-making tools in the analysis of transportation planning. Recent advances of innovative transportation related policy instruments as well as growing complexity in activity-patterns urge for more efficient decision support tools than traditional four-stage travel demand models (Roorda & Miller, 2006; Shiftan & Ben-Akiva, 2011; Shiftan et al., 2003). As a consequence, researchers and practitioners have sought a new approach; they have tried to enhance traditional trip-based, aggregate travel demand models with the disaggregate approach based on the discrete choice theory since the 1970s. But, the notion that travel demand is derived from decisions to participate in out-of-home activities changes the focus of travel demand modelling from trip-based to activity-based. The activity-based modelling approach has emerged in the transportation planning literature to overcome the limitations observed in the traditional four-stage travel demand model since the 1980s, though the concept was discussed in the literature since the 1970s.

5.2.1 Activity-based models and their applications

Some excellent literature reviews on activity-based modelling frameworks to the analysis of travel behaviour can be found in (Axhausen & Gärling, 1992; Gärling, Kwan, & Golledge, 1994; Mohammadian, Auld, & Yagi, 2009; Shiftan & Ben-Akiva, 2011). In recent years, advances in technology allowed the development of activity-based microsimulation models, which can be classified in two broad categories, econometric activity-based models and rule-based activity scheduling models. Econometric activity-based models are Bowman and Ben-Akiva model (Bowman & Ben-Akiva, 2000), the MORPC model (Vovsha, Petersen, & Donnelly, 2004), the CEMDEP model (Bhat, Guo, Srinivasan, & Sivakumar, 2004), and the Jakarta model (Yagi & Mohammadian, 2008a, 2008b). Rule-based activity scheduling models include STARCHILD (Recker, McNally, & Root, 1986a, 1986b), SMASH (Ettema, Bergers, & Timmermans, 1993), SCHEDULER (Golledge, Kwan, & Gärling, 1994), GISICAS (Kwan, 1997), PCATS (Kitamura & Fujii, 1998), the model system proposed by Bhat (Bhat, 1999), TASHA (Miller & Roorda, 2003; Roorda, 2005), ALBATROSS (Arentze & Timmermans, 2004; Arentze & Timmermans, 2008), and ADAPTS (Auld & Mohammadian, 2009).

The activity-based approach has already been recognized as a powerful framework for travel demand analysis; however traditional travel demand models are still more often used in practice (Arentze & Timmermans, 2008; Bhat & Lawton, 2000; Mohammadian et al., 2009; Shiftan et al., 2003). Some recent applications of activity-based modelling approach include Portland (METRO), San Francisco (SFCTA), New York (NYMTC), Columbus (MORPC), Atlanta (ARC), Sacramento (SACOG), Bay Area (MTC), Denver (DRCOG) (Bradley & Bowman, 2006; Shiftan & Ben-Akiva, 2011), Dallas/Forth-Worth (CEMDAP) (Bhat et al., 2004), Southeast Florida (FAMOS) (Pendyala, Kitamura, Kikuchi, Yamamoto, & Fujii, 2005), Netherlands (ALBATROSS) (Arentze & Timmermans, 2008), and Toronto (TASHA) (Roorda, Miller, & Habib, 2008).

5.2.2 Activity-based model validation methods

In light of the shortage of applied experience, activity-based models require extensive validation prior to application. Activity-based models have been validated using different methods.

The most common validity test is the verification of the model using the base year estimation dataset with analysis at the macro-level. Models validated using such a method include STARCHILD (Recker et al., 1986b), SMASH (Ettema, Borgers, & Timmermans, 1996), Bowman and Ben-Akiva model (Ben-Akiva & Bowman, 1998; Bowman, 1998), PCATS (Kitamura & Fujii, 1998), ALBATROSS (Arentze & Timmermans, 2004), the Jakarta model (Yagi & Mohammadian, 2007a, 2007b), TASHA (Miller & Roorda, 2003; Roorda, Miller, et al., 2008), DRCOG (Childress, Sabina, Kurth, Rossi, & Malm, 2010; Kurth, Childress, Sabina, & Rossi, 2006) and CEMDAP (Bhat, Paleti, Pendyala, & Goulias, 2013). Pendyala et al. (2005) argue for further validation tests on specific population segments, modes, and geographical areas. However, only few models have been validated at different levels such as by population segments (PCATS (Pendyala et al., 2005)).

Another validation test is the comparison of the estimated forecasts of daily travel behaviour of a later year with the observed survey data of the same year. DRCOG and TASHA have been assessed using this method; however these validations were also at the macro-level (Childress et al., 2010; Kurth et al., 2006; Roorda, Miller, et al., 2008).

Spatial transferability can be a useful validation test of activity-based models; however to date, activity-based models have not been extensively validated using this test (Arentze, Hofman, Van Mourik, & Timmermans, 2002; Bowman, Bradley, Castiglione, & Yoder, 2014; Yasmin, Morency, & Roorda, 2015a). The focus of model transferability research has been mostly on trip generation and mode choice models (Yasmin, Morency, & Roorda, 2013). It is assumed that the activity-based models, developed with greater theoretical basis and behavioral realism, are more transferable than the traditional trip-based models (Gangrade, Kasturirangan, & Pendyala, 2000; Sikder, Pinjari, Srinivasan, & Nowrouzian, 2013). Arentze et al. (2002) have tested the spatial transferability of ALBATROSS model system and reported satisfactory transferability at both individual and aggregate levels except for transportation mode choice. Auld and Mohammadian (2012) have shown that the planning order model of the ADAPTS activity-based modelling framework (Auld & Mohammadian, 2009) is quite transferable, however they indicate a need for further evaluation. After empirically testing the transferability of an activity-based model (DaySim), Bowman et al. (2014) have recommended that transferring a model developed based on a large sample from a comparable region is better than estimating a new model using a smaller local sample. Yasmin et al. (2015a) have also examined the spatial transferability of the TASHA model by transferring it to the Montreal Island, however the examination was also only at the aggregate level. As evident from the literature review, very few research efforts have focused on spatial transferability as a validation test of the activity-based models, and most of them have only focused on the macro-level validation. Thus, this research addresses the question of how well the activity-based model, TASHA would perform if applied in another context and if validated at different levels of aggregation (macro-, meso-, and micro-levels).

5.3 The TASHA model

The TASHA (Travel Activity Scheduler for Household Agents) model has been developed by Miller and Roorda (2003) based on trip diary data from the 1996 Transportation Tomorrow Survey (TTS) for the Greater Toronto Area (GTA). The model is a fully disaggregate microsimulation model. It estimates activity schedules and travel patterns for a twenty-four hour typical weekday for all individuals in a household. Miller and Roorda (2003) present the full conceptual design and methodology of this activity scheduling model. The TASHA model consists of five components, namely activity generation, location choice, activity scheduling, mode choice, and trip assignment. This paper focuses on the spatial transferability of the first three components i.e. activity generation, location choice, and activity scheduling. A brief description of the methodology of these three components is presented here.

Activity generation

In TASHA, activities are generated first, and then scheduled. Work and school activities are generated first and then, based on the duration and start time of these activities, other types of activities are generated. The input data required at this stage are person and household attributes, observed distributions of frequency, start time and duration of activities, by activity type. The activity generation component generates both individual and joint activities for each individual using Monte-Carlo simulation from the 1996 TTS observed distributions of activity attributes (Roorda, Miller, et al., 2008).

Activity location choice

The locations of home and usual place of work/school are exogenously input into the model. Other than these, the locations of other activities are simulated using a series of entropy models (Eberhard, 2002). The input data required at this stage are employment, population, activity density, and travel distance between traffic analysis zones.

Activity scheduling

In TASHA, the generated activities are scheduled following a rule-based procedure. First, it organizes the generated activities into work, school, shopping (joint and individual) and other (joint and individual) projects. Next, consistent and feasible activity-travel schedules are developed for interacting household members. A brief description of the activity scheduling method is discussed below, detailed process can be found in other papers (Miller & Roorda, 2003; Roorda, Miller, et al., 2008).

Step 1: Insert activity episodes into a project agenda with other activity episodes that achieve a common purpose (for instance work) resulting a preliminary sequence of activities in time.

Step 2: Form person schedules by taking activity episodes from the project agenda and adding them into person schedule following the order of precedence of activity episodes (i.e. work-business, work-at-home-business, primary work, secondary work, return home from work, school, joint other, joint shopping, individual other, and individual shopping) observed from an interactive computer survey of activity scheduling (Doherty, Nemeth, Roorda, & Miller, 2004). Definitions of the episode types can be found in other papers (Miller & Roorda, 2003; Roorda, Miller, et al., 2008).

Step 3: Apply a 'clean up' algorithm to fine tune the schedule for execution. A single algorithm is applied to rearrange activities in the schedule to remove unrealistically short work episodes with duration of 30 minutes or less.

The components of the TASHA modelling framework, namely activity generation, activity location choice, and activity scheduling have been tested following the most common validation method (i.e. base-year verification). Thus, the base year model outcomes at the aggregate level have been compared with the base year observed dataset (i.e. 1996 TTS) from which the model is originally derived for the GTA. In addition, another validation of the TASHA model has been done by comparing forecasts of a future year also at the aggregate level with the observed data of the same year (i.e. 2001 TTS). Both validation results at the aggregate level show that TASHA could reproduce the activity/travel patterns in the GTA (Roorda, Miller, et al., 2008). In addition, TASHA has also been validated using a spatial transferability test at the aggregate level (Yasmin et al., 2015a), the key result of this research is also reported in the macro-level validation section of this paper.

5.4 Data and research method

The Toronto Area model, TASHA has been transferred to the Montreal Island using the large scale Origin-Destination (O-D) travel survey of Montreal. This survey has been conducted every five years since the early 1970s to collect detailed travel and socio-economic information of approximately 5% of the total Greater Montreal Area (GMA) population aged 5 years and older. Detailed information on these large scale O-D travel surveys can be found on the AMT website (Agence Métropolitaine de Transport [AMT], 2015). As of 2012, the Montreal Island is composed of an area of 500 square km with a population of 1,886,000 (Statistics Canada, 2011b). On the other hand, the GTA is composed of an area of 7,125 square km with a population of 6,054,000 (Statistics Canada, 2011a).

TASHA application in the context of the Montreal Island was based on the model parameters (observed activity attributes distributions and location choice parameters) estimated using the Toronto data. Thus, observed distributions of activity attributes (activity frequency, start time, and duration of different types of activity) from the 1996 TTS are used as inputs in the activity generation component of TASHA. Location choice model parameters estimated using the 1996 TTS for the GTA are also used for this application. Other required data such as individual and household attributes (including home and work location), employment, population, activity density, and travel distance between traffic analysis zones are gathered from the 2003 O-D travel survey of Montreal and the 2001 Canadian Census. During data preparation, we have excluded the households of which an individual made an open trip chain (i.e. his/her travel did not start and/or end at home). After cleaning, a sample of 59,624 individuals from 26,960 households has been used in the application of TASHA to generate their daily activity schedules.

As TASHA employs a stochastic approach for simulating activity scheduling, the model is run for ten replications to generate activity schedules for Montreal residents. Several activity attributes have been prepared from both the simulated outputs by TASHA and observed data from the 2003 O-D travel survey for comparison. Typically, TASHA generates eleven types of activities, which are aggregated into five broad activity classes (work, school, shopping, other, and return to home). Here, activity start time stands for the travel start time for both TASHA output and the O-D travel survey. Consequently, the duration of an activity includes travel time to the activity location as it was calculated using the start times of two successive trip observations. The distances are the Euclidean distances between an origin and a destination point which are calculated using the coordinates of the centroids of traffic analysis zones (TAZs).

The validation has been conducted at three different levels of aggregation, macro-level (aggregation of the entire population), meso-level (aggregation by population segments by age group and gender, and by home location), and micro-level (individuals). We compute the average values of activity attributes over ten replications for both macro- and meso-levels. Four activity attributes (activity frequency, start time, duration, and trip distance) have been compared for five activity types (work, school, shopping, other, and return to home). The validation results are reported, here, as percentage differences between simulated results and observed from the 2003 O-D survey for macro- and meso-levels. In addition, the Kolmogorov-Smirnov tests (K-S tests) have been performed for these two levels to examine whether the modelled and observed distributions are drawn from the same distribution, with at least 90% confidence. For micro-level validation, simulated outputs for each individual are compared with the observed attributes from the 2003 O-D survey using criteria which are discussed in the micro-level validation section.

5.5 Model validation

5.5.1 Macro-level validation

5.5.1.1 Activity frequency

Table 5.1 presents the observed values as well as the percentage differences between TASHA simulated outputs and observed from the 2003 O-D travel survey for activity frequency by activity type. In the Montreal Island at the macro-level, overall TASHA undersimulates the total number of observed activities by 10.3%. In the Toronto Area, TASHA undersimulates the total number of observed activities by 0.2% while verifying using the base year dataset (i.e.1996 TTS) and, undersimulates the number of observed activities by 3.2% while validating using a future year dataset (i.e. using the 2001 TTS) (Roorda, Miller, et al., 2008).

TASHA undersimulates the observed total activities for all but work activity. These undersimulations may happen if there are conflicts in activity scheduling which cause the rejection of activity episodes. Like the Toronto Area, frequency of the work activity (+5.4%) is more closely simulated than other types of activities. This is expected as work activity is scheduled first giving highest priority among others. The simulations of school and return to home activities are also quite good (within 10%). However, the percentage difference of simulated and observed shopping activities is slightly higher (-13.9%) than work, school and return to home activities, but the simulation result is quite similar to the validation result of the Toronto Area (Roorda, Miller, et al., 2008). The other activity shows the largest undersimulation (-30.6%). The individuals of the Montreal Island might be conducting more other household maintenance activities than the individuals of the Toronto Area because of having more accessibility to the opportunities at a shorter distance.

5.5.1.2 Activity start time

Percentage differences between modelled and observed travel/activity start time distributions for a typical weekday for all five activities are presented in Figure 5.1 (a). K-S tests are performed to examine whether modelled and observed distributions are the same.

TASHA simulation for work activity start times is quite promising, as start times are closely simulated over the day within ±2% for all except one hour (7:00 - 7:59 AM). The K-S test (P = 0.893) also indicates that the simulated and observed distributions are similar. However, the study observes a shift in time (i.e. undersimulation of early starting activities at 7:00 AM and oversimulation of starting activities at 9:00 AM) of the simulated distributions by TASHA. This indicates that people in the Montreal Island leave earlier for work activities than in the Toronto Area. However, congestion could also be the reason for the early travel start in Montreal.

Start times simulation of the school activity is also good (within ±2%) for the entire day except for two morning hours (7:00 - 7:59 AM and 8:00 - 8:59 AM). In the Montreal Island, these two morning hours consist of the most of the observed school activities (around 71.4%). But in the

Toronto Area, both observed data of base year and future year show that the most of the school activities start between 8:00 - 8:59 AM (Roorda, Miller, et al., 2008). It is clear that the simulation result reflects the behavioural differences between the two cities. Figure 5.1 (a) shows a significant undersimulation of trips starting between 7:00 - 7:59 AM and oversimulation of those occurring between 8:00 - 8:59 AM. As expected, the K-S test result (P = 0.441) is inconclusive for these two distributions.

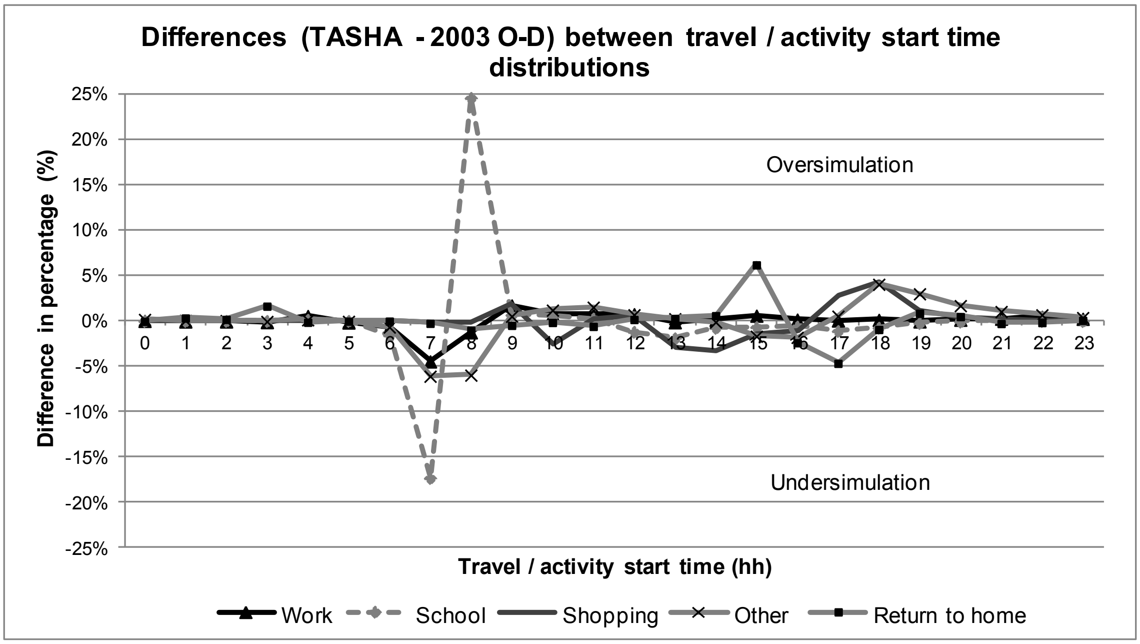
TASHA simulation of shopping activity start times is also satisfactory (between -3.3% and 4.3%). There is an undersimulation of shopping activities starting in early hours of the morning and in late afternoon. The K-S test for shopping distributions accepts the null hypothesis (P = 0.992) that activity start times are from the same distribution. For other activities, there is an undersimulation of activities starting before 9:00 AM and between 2:00 PM and 4:59 PM. The K-S test provides a strong evidence (P = 0.031) that the modelled and observed distributions are drawn from different distributions. For return to home activities, there is an oversimulation of activities starting between 12:00 PM and 3:59 PM and undersimulation of activities starting between 4:00 PM and 6:59 PM. The K-S test indicates a strong evidence (P = 0.992) that the modelled and observed distributions are from the same distribution.

In the Toronto Area, activity start times simulation by TASHA for all activities in both base year and future year are quite good; the K-S tests in both verification (P > 0.97 for all activity types) and validation (P > 0.86 for all activity types) analysis provides strong evidence that modelled and observed activity start time distributions are similar for all activities (Roorda, Miller, et al., 2008). From this analysis, we observe that people in the Montreal Island engage in these activities at different times than people in the Toronto Area. However in the Montreal Island, the start time simulations by TASHA are fairly promising for work, shopping, and return to home activities; the simulation is also good for school activities over the day except two morning hours. The distributions are significantly different for other activities.

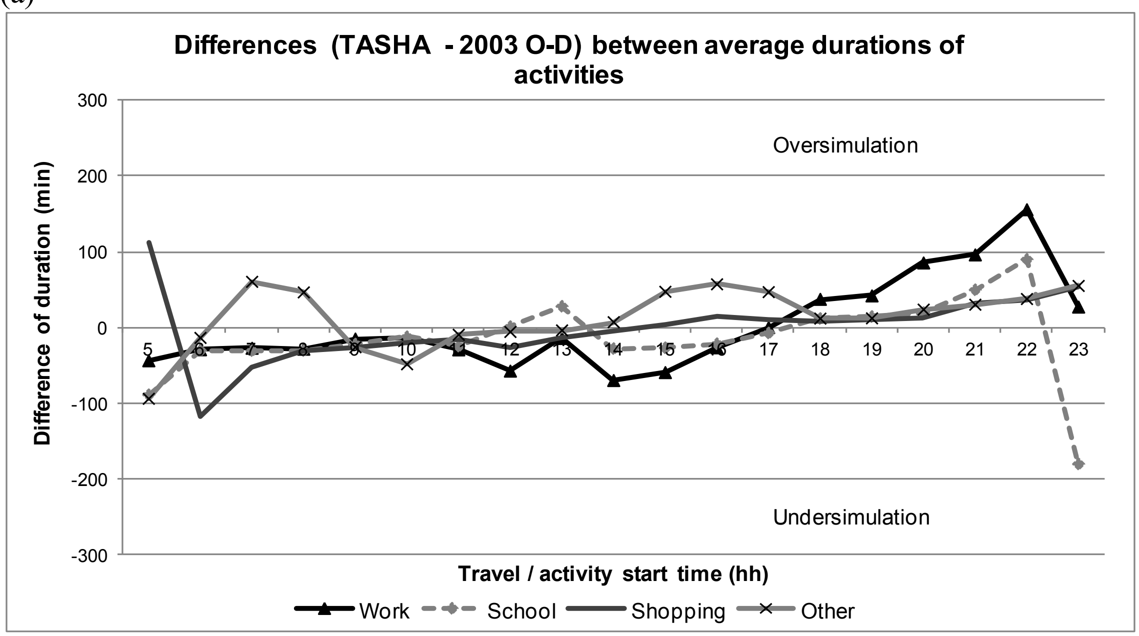
Table 5.1: Observed values at macro-level and percentage differences between TASHA simulated outputs and observed from the 2003 O-D survey at macro-, and meso-level: Activity frequency and average trip distance

Activities	**Macro-level**		**Meso- level**									
			Population segments by age groups and gender						**Population segments by home location regions**			
			Ages 24 years old and younger		**Ages 25 - 54 years**		**Ages 55 years old and older**		**Region 1** (Montreal Downtown)	**Region 2** (Montreal Centre)	**Region 3** (Montreal East)	**Region 4** (Montreal West)
			Men	**Women**	**Men**	**Women**	**Men**	**Women**				
Activity frequency												
	Observed values[a]	% diff[b]	% diff[b]									
Work	24693	5.4%	-8.5%	-9.6%	6.8%	10.2%	-3.8%	-0.1%	5.4%	5.4%	5.9%	5.1%
School	13004	-6.4%	-6.2%	-5.7%	-7.7%	-8.0%	-59.0%	-47.2%	2.2%	-7.1%	-9.6%	-4.9%
Shopping	10829	-13.9%	-9.6%	-19.8%	-15.2%	-7.7%	-15.8%	-18.1%	-26.4%	-14.9%	-6.4%	-14.0%
Other	24309	-30.6%	-23.6%	-27.2%	-36.2%	-41.1%	-19.8%	-10.0%	-25.2%	-30.9%	-24.6%	-33.9%
Return to home	59068	-8.7%	-7.4%	-6.9%	-7.8%	-9.0%	-13.6%	-9.7%	-5.8%	-8.9%	-6.6%	-9.7%
Total	**131903**	**-10.3%**	**-8.7%**	**-9.2%**	**-9.3%**	**-11.1%**	**-14.0%**	**-10.5%**	**-8.1%**	**-10.6%**	**-7.8%**	**-11.4%**
Activity location												
Work	7.94 km	0.0%	-0.3%	7.1%	-2.4%	3.2%	-7.6%	5.4%	-8.1%	1.2%	0.8%	-1.0%
School	4.21 km	5.3%	6.7%	6.1%	-1.3%	3.4%	-2.1%	3.3%	-4.3%	5.0%	6.4%	5.6%
Shopping	3.63 km	40.9%	26.9%	35.7%	33.3%	37.5%	44.4%	53.7%	19.3%	48.1%	49.2%	29.4%
Other	5.01 km	37.9%	46.1%	46.8%	28.6%	54.1%	18.4%	37.4%	24.5%	46.8%	38.4%	30.1%
Return to home	5.61 km	9.1%	7.8%	9.7%	1.8%	13.6%	6.9%	27.8%	-2.7%	12.6%	10.6%	5.8%
Total	**5.63 km**	**13.3%**	**10.1%**	**12.0%**	**6.4%**	**18.8%**	**10.2%**	**31.3%**	**1.7%**	**16.9%**	**14.6%**	**9.9%**

[a] values indicate observed total activities for attribute, activity frequency and average trip distance in km for attribute, activity location from the 2003 O-D survey.
[b] values indicate percentage differences between TASHA simulated outputs and observed from the 2003 O-D survey, and positive and negative values indicate oversimulation and undersimulation, respectively.



(a)



(b)

Figure 5.1: Differences (TASHA - 2003 O-D) between (a) travel / activity start time distributions, (b) average durations of activities

5.5.1.3 Activity duration

Figure 5.1 (b) presents the differences between modelled and observed average activity durations by travel/activity start times for four activities. K-S tests have also been performed to assess if modelled and observed distributions are similar. K-S tests results indicate that the simulation results in the Montreal Island are quite similar to the results in the Toronto Area for all activities except school; here the K-S test results reject the null hypotheses that modelled and observed distributions are the same (P ranges from 0.139 to 0.259). In the Toronto Area validation test, the K-S tests also either reject the null hypothesis (P = 0.021) or shows inconclusive results (P ranges from 0.304 to 0.593) for all activity types (Roorda, Miller, et al., 2008). However, the higher differences between modelled and observed average durations are observed in the Montreal Island at those hours of the day when there are low activity frequencies. The simulation of average duration of school activity over the day (P = 1.000) is better than the Toronto Area simulation.

The result shows undersimulation of average durations for all work activities starting before 5:00 PM. By this time the majority of work trips have taken place. For the school activity, it also shows undersimulation of average durations for almost every hour except activities starting between 12:00 PM and 1:59 PM, and after 6:00 PM. There is undersimulation of shopping activities starting before 3:00 PM, and oversimulation of the ones after 3:00 PM (though the differences are smaller). For the other activity, there is underestimation of durations of activities which started early in the day and oversimulation for activities taking place in the late afternoon.

5.5.1.4 Activity location

Observed average trip distances from the 2003 O-D survey and percentage differences between modelled and observed average distances to all activities are shown in Table 5.1. In the Montreal Island, at the macro-level, the result shows an oversimulation of the average observed distances by 13.3%. The result can be compared to the Toronto Area; overall TASHA oversimulates the average observed distances by 0.9% while verifying TASHA using the base year dataset (i.e.1996 TTS) and, by 0.8% while validating using the future year dataset (i.e. using the 2001 TTS) (Roorda, Miller, et al., 2008).

TASHA oversimulates the average distances for all activities but work. The simulation of average distance of school and return to home activities are also quite good (within 10%). Interestingly, the simulation of average distance of work and school activities is either better than or similar to the results found in the verification and validation tests in the Toronto Area (Roorda, Miller, et al., 2008). The usual place of work/school is directly input into the modelling framework which, of course, greatly reduces uncertainty for these activities. The largest differences are observed for the shopping and other activity types.

5.5.2 Meso-level validation

Meso-level validation allows us to examine how TASHA reproduces observed activity behaviour for different population segments. Two meso-level validations have been performed for two population segmentations: 1) by age group and gender, and 2) by home location. Accordingly, the entire population (59,624 individuals) has first been divided by age groups and gender resulting in six groups, ages 24 years old and younger (men and women), ages 25 - 54 years old (men and women), and ages 55 years old and older (men and women). Then, the population is sub-divided by their home location regions. The Montreal Island is divided into four regions, as shown in Figure 5.2. As of the 2003 O-D survey, Region 1 (Montreal-Downtown), Region 2 (Montreal-Centre), Region 3 (Montreal-East), and Region 4 (Montreal-West) have a population (density) of 71,350 (5,544 persons/km^2), 994,938 (6,386 persons/km^2), 305,715 (3,277 persons/km^2) and 500,387 (2,093 persons/km^2), respectively.

5.5.2.1 Activity frequency

Table 5.1 presents meso-level validation results for both population segmentations for activity frequency by activity type. Similar to the macro-level, TASHA undersimulates the total number of observed activities for population segmentations by age group and gender, as well as home location. Comparatively for different segments by age group and gender, TASHA simulates the observed frequencies closely (mostly within 10%) for work, school (with a few exceptions), and return to home activities, but simulates other activities with high differences similar to the macro-level. The simulation of shopping activity frequency shows mixed results with low and high differences, which ranges from -7.7% to -19.8% for different population segments.

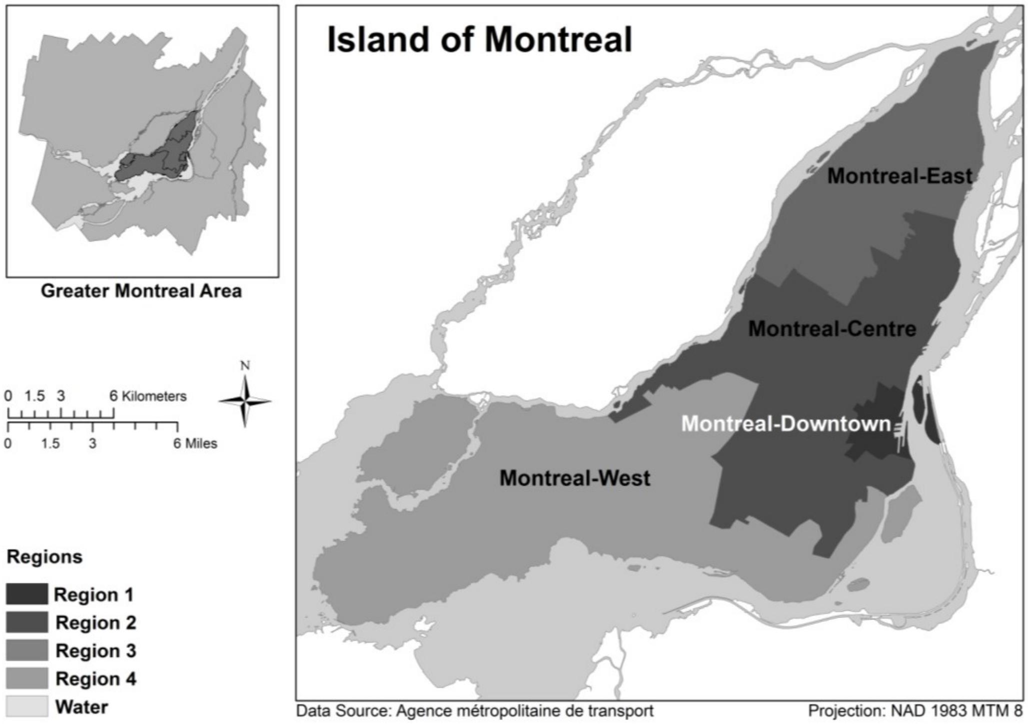


Figure 5.2: Island of Montreal by regions

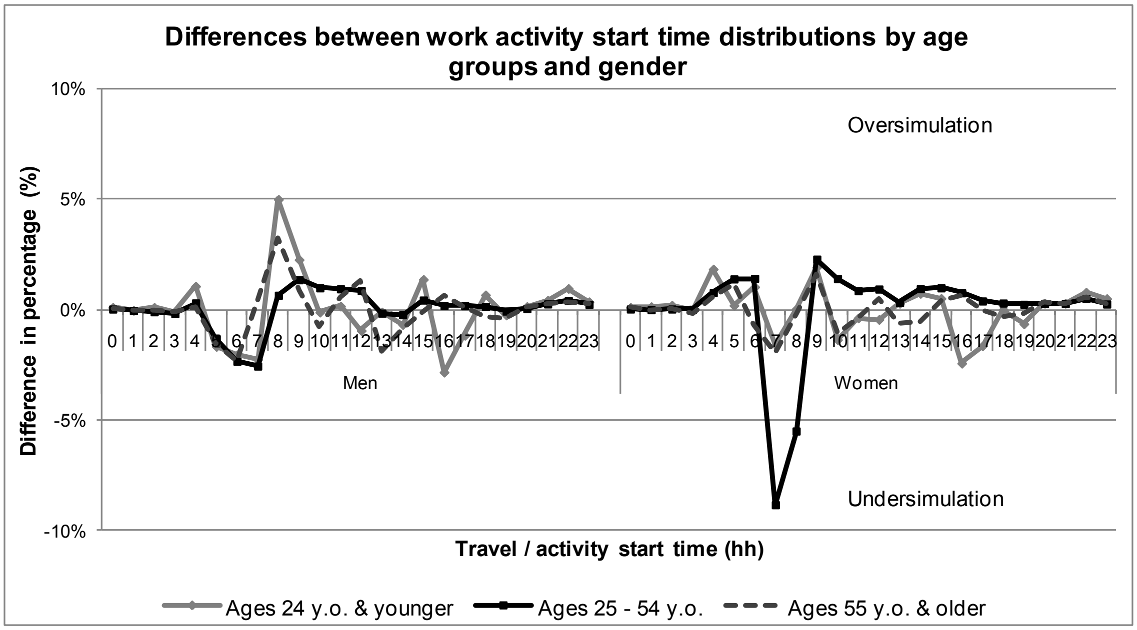
TASHA undersimulates the frequency of activities for all activity types for the population segment of ages 24 years old and younger. It also shows a similar variability between men and women for all activities except shopping for this age group. However, school activities are more frequent for this population segment (86.6% of total school activities) and TASHA simulates the observed frequencies more closely (within 6%) for this activity for both men and women for this segment. For population segment with ages 25-54 years old, TASHA undersimulates all types of activities except work for both men and women, similar to the macro-level. The majority (77.7%) of work activities are conducted by this population segment and the TASHA simulation of work activity frequency for both men and women is quite good (within 10%). For individuals aged 55 years old and older, TASHA underestimates all types of activities for both men and women. The variation is higher for school activities, however these activities are very infrequent for this population segment (0.5% of total school activities).

Similar to the macro-level as well as the population segmentation by age group and gender, TASHA simulation of activity frequency is also quite good for work, school, and return to home activities (within 10%), but it shows high differences for other activities for different population segments by home location region. Also, shopping activity shows mixed results with low and high values, which ranges from -6.4% to -26.4% for these population segments, similar to the other population segments (by age group and gender). The frequencies of work activities are more closely simulated (within 6%) than other types of activities for all population segments by home location regions except Region 1 (Montreal-Downtown); however the differences are close to each other for the four regions. TASHA slightly oversimulates the observed school activities for Region 1, however undersimulates for the other three regions with larger differences than Region 1.

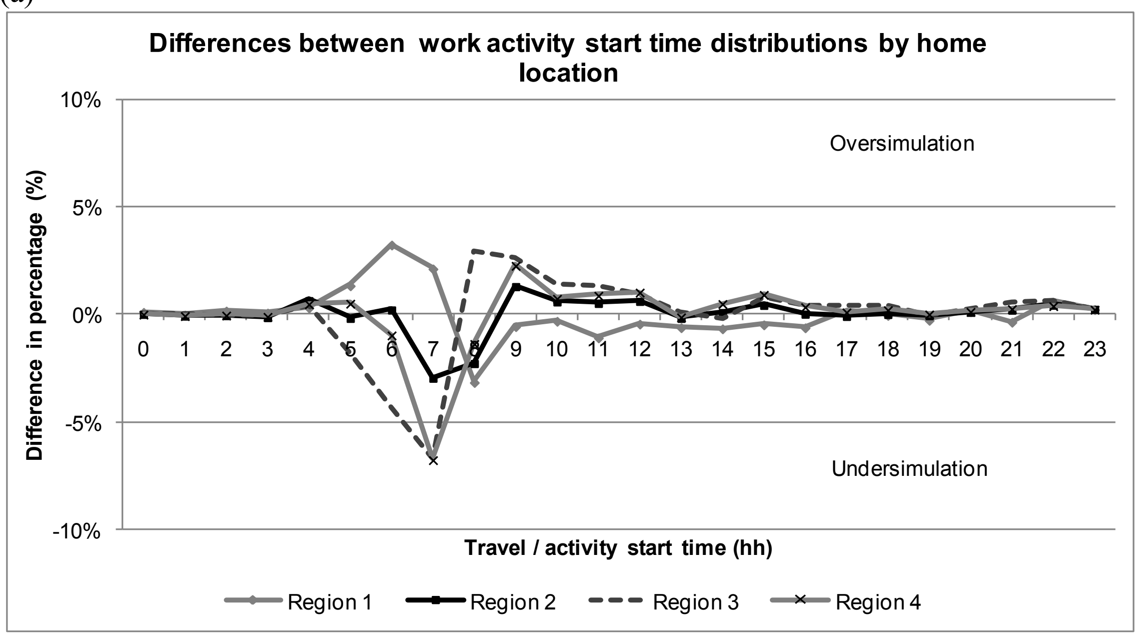
5.5.2.2 Activity start time

Meso-level validations for both population segmentations have also been conducted for travel/activity start time distributions, for all activities. Here, we only report the results of the work activity start time distributions for both population segmentations. Figure 5.3 (a) presents the percentage differences between work activity start time distributions by age groups and gender. Results show that the differences are within 5% for all age groups and both gender, except for women aged 25 - 54 years old for two hours. For this group, TASHA undersimulates the activities starting between 7:00 - 7:59 AM by 8.8% and between 8:00 - 8:59 AM by 5.5%. For men younger than 24 years old or older than 55 years old, the largest oversimulations (5% and 3%, respectively) are observed for activities starting between 8:00 - 8:59 AM.

Figure 5.3 (b) presents the percentage differences between work activity start time distributions by home location region. This result also shows that the difference is within 5% over the day for all regions except for 1 hour for Region 3 (-6.8%) and Region 4 (-6.7%). These largest variations are found for Region 3 (Montreal-East) and Region 4 (Montreal-West) for the activities starting between 7:00 - 7:59 AM. As the segments are based on the home locations of the individuals, the observation is logical considering the different levels of congestion experienced by the residents of these regions (congestion is higher for the residents of the west and the east of the Montreal Island than for people residing in the central regions). The largest undersimulations for both population segmentations again indicate that individuals in the Montreal Island start their work trips earlier than in the Toronto Area, which has already been observed at the macro-level.



(a)



(b)

Figure 5.3: Differences (TASHA - 2003 O-D) between work activity start time distributions: (a) by age groups and gender, and (b) by home locations

5.5.2.3 Activity duration

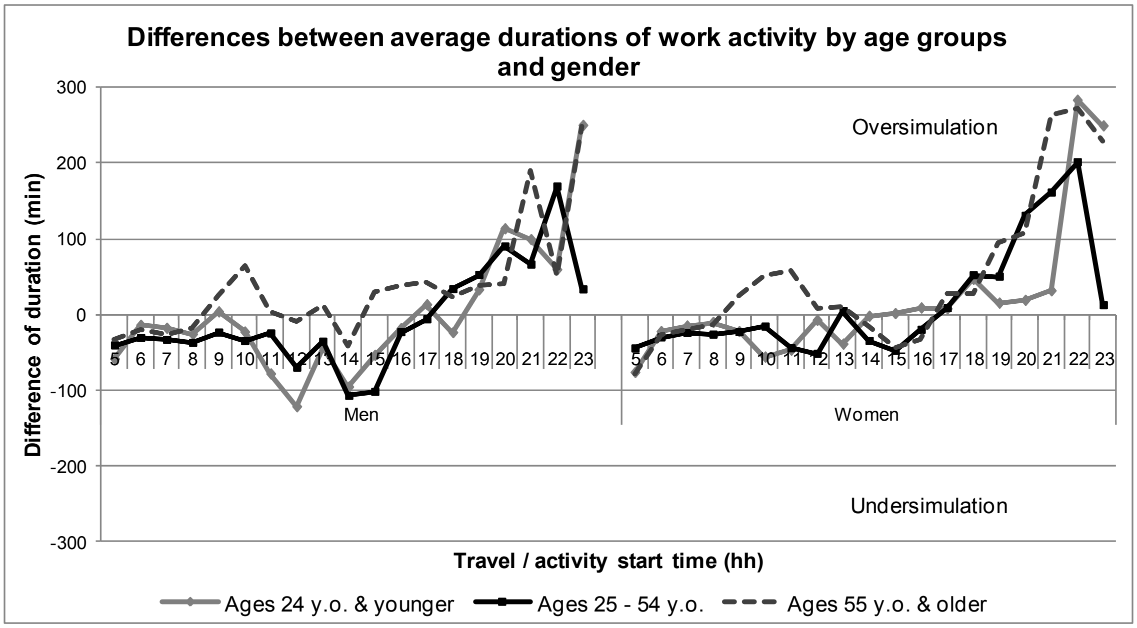
For each population segmentation, meso-level validations for average activity durations have also been conducted for all activities. Here, we only present the results of the average durations of work activity for both population segmentations. The differences of work activity average durations between TASHA simulated outputs and the 2003 O-D survey for population segments by age groups and gender are shown in Figure 5.4 (a). TASHA oversimulates the average durations of work activities for individuals aged 55 years old and older, men and women, but undersimulates for the other two segments for the same hours which are the most common for this type of activity departure.

Figure 5.4 (b) presents the differences of average durations between TASHA outputs and the 2003 O-D survey for work activities by home location. Undersimulations with small differences are observed for work activities starting early in the day, when most of the work trips begin in typical weekday. The findings are relatively similar for all regions.

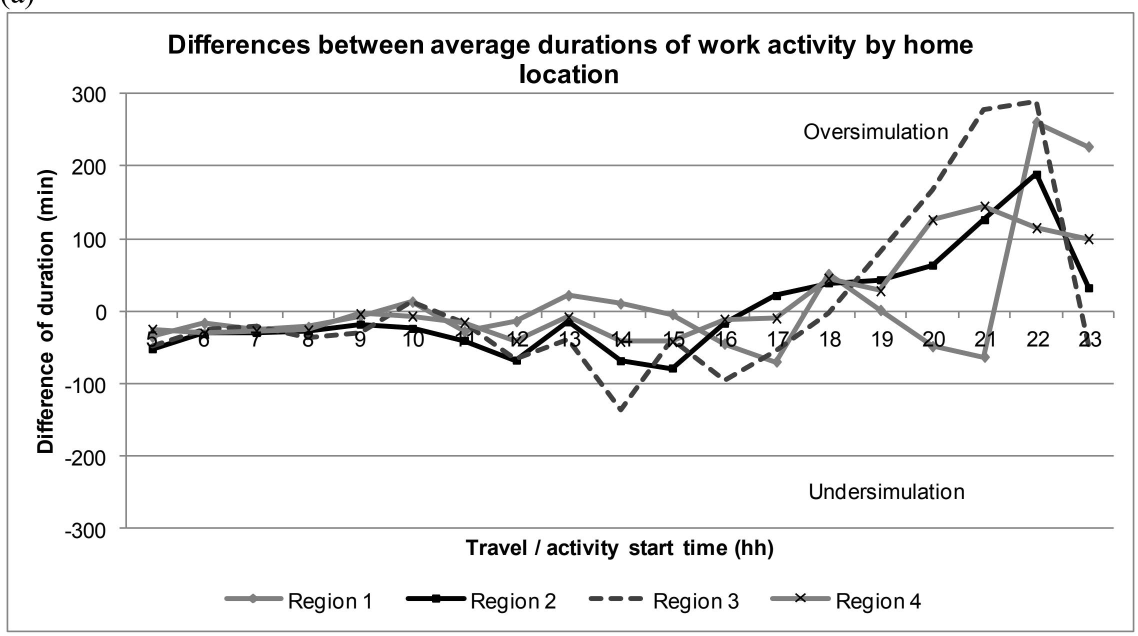
5.5.2.4 Activity location

Table 5.1 presents meso-level validation results of average distances by activity type for both population segmentations. Overall, similar to the macro-level, the results show lower variability for work, school, and return to home activities (within 10% with a few exceptions) but higher variability for shopping and other activities. Also, as for macro-level, TASHA oversimulates the average distances of all trips made by all segments of population.

The simulation of average distances are quite good for all segments of the population by age group and gender (within 12%, close to the macro-level simulation results) except for women aged 25 - 54 years old (+18.8%) and aged 55 years old and older (+31.3%). For these segments of population, the largest differences are also observed for shopping and other activities. Overall, the better result is found for men aged between 25 - 54 years old (+6.4%) among all segments of population. This population segment has performed most work activities and the TASHA simulation of work activity for both men and women of this segment is quite good (within ±3%). The population segment aged 24 years old and younger school activities, and the TASHA simulation is quite reasonable for school activities for both men and women of this segment (within ±7%).



(a)



(b)

Figure 5.4: Differences (TASHA - 2003 O-D) between average durations: (a) work activity by age groups and gender, and (b) work activity by home location

Table 5.1 also presents percentage difference in average distance between TASHA outputs and the 2003 O-D travel survey by home location. Overall, the lowest difference in average travel distance is observed for the Montreal Downtown. Similar to the macro- and other meso-level comparisons (population segments by age group and gender), there are lower differences (within ±10%) in average distance to work, school, and return to home (except Region 2) and higher differences in average distance to shopping and other activities. The differences are small for work and school activities for all regions since, among other things, usual place of work and school are model input. It is hence logical that predicting destination for these purposes is much easier.

5.5.3 Micro-level validation

Micro-level validation has examined the number of individuals whose activity schedules are accurately reproduced by TASHA. This type of validation is customary in Montreal. In this paper, the micro-validation is based on the following criteria estimated at the person level:

- Total number of trips;
- Number of trips for each activity type;
- Total durations of activities per day (± 60 minutes);
- Total durations of activities by activity type per day (± 60 minutes);
- Number of trips in the morning peak;
- Number of trips in the afternoon peak

The micro-level validation identifies the percentage of individuals that are correctly simulated according to the above criteria applied cumulatively in sequence. Table 5.2 results can be interpreted as follows. Based on the first criterion alone, TASHA accurately simulates the total number of trips of 43% of individuals (including zero trip) and 32% of individuals (excluding zero trip). If the first and second criteria are applied, TASHA accurately simulates the total number of trips and the number of trips for each activity type for 25% of individuals. When the third criterion (the total durations of activities of the individual) is added, TASHA replicates approximately 9% of individuals' behaviour accurately, and so on. 6% of individuals' activity schedules are correctly simulated when all the criteria listed above are applied.

Table 5.2: Micro-level validation results

Criteria		Matched percentage (%)
Number of trips	Including zero trip	43.4%
	Excluding zero trip	31.9%
Number of trips by activity type	Work	27.9%
	School	26.8%
	Return to home	26.6%
	Shopping	34.8%
	Other	24.5%
Duration of activities per day		8.7%
Duration by activity type	Work	8.6%
	School	8.6%
	Shopping	8.6 %
	Other	8.6 %
Number of trips in the morning peak		7.1%
Number of trips in the afternoon peak		6.3%

5.6 Summary of validation results and discussions

At the macro-level, overall TASHA undersimulates the total number of observed activities by 10.3%. TASHA replicates work, school, and return to home activities reasonably well (within 10%) and also provides promising results for shopping activities (slightly greater than 10%, however similar to the Toronto Area). TASHA provides larger differences only for other activity type. In addition, TASHA captures the temporal aspects of travel behaviour quite well for work, shopping, and return to home activities, however it shows large differences for other activity type. For school activities, TASHA simulated results are quite good throughout the day except two morning hours. The study reveals that people in the Montreal Island engage in activities (for instance, work, school and other activities) earlier than predicted by the model based on behaviour in the Toronto Area. Activity duration simulation for school activities provides a very good result, even better than both verification and validation tests in the Toronto Area. TASHA also provides reasonable results for work activity durations for those hours when the majority of work activities take place. However, similar to the Toronto Area, the simulated durations are not good for shopping and other activities. Like the Toronto Area, TASHA oversimulates the average distance to all activities by 13.3%. Interestingly, the simulation provides better results for work and school activities than for the Toronto Area. The result is also quite reasonable for return to home activities. The largest differences are observed for shopping and other activity types.

At the meso-level, TASHA simulation is also quite good (mostly within 10%) for activity frequency of work, school (with a few exceptions), and return to home activities for both population segmentations by age group and gender, and by home location, however TASHA provides large differences for other activities, similar to the macro-level. TASHA simulation for shopping activity shows mixed results with low and high differences (ranges from -6.4% to -26.4%) for both population segmentations. The simulation of work activity start time over the day is also promising (within 5%) for both population segmentations with a few exceptions (for instance for two hours for women aged 25 - 54 years old (however within 10%), and for 1 hour for Regions 3 and 4 (within 7%). Undersimulations with small differences are observed for average durations of work activities starting early in the day, when most of the work trips begin in typical weekday for both population segmentations except individuals aged 55 years old and older, men and women. Overall similar to the macro-level, the simulation results for average distance show lower differences for work, school, and return to home activities (within 10%) with a few exceptions, but higher differences for shopping and other activities.

At the micro-level, TASHA reproduces approximately 43% of individuals' behaviour accurately based on the first criterion (total number of trips, including zero trip, performed by the individual). This percentage decreases to 6% when other criteria (number of trips and activity duration by activity type, and time of day) are also applied sequentially.

Validation results show large differences between the simulated and the observed attributes in some cases at different levels (mostly for shopping and other activities). The following section provides a discussion on the potential reasons behind these large differences.

- **Behavioural differences**

The large variations observed in some cases for different activities at different levels most likely indicate real differences in behaviour exist between residents of the Montreal Island and the Toronto Area. These two cities are distinctly different in terms of socio-demographic, cultural, economic development, employment pattern, commuting pattern and so on (Balakrishnan, Maxim, & Jurdi, 2005; Heisz, 2006; Roorda, Morency, & Woo, 2008; Rose, 1999; Shearmur, 2006). Yasmin et al. (2015a) have compared the Montreal Island (using 2003 O-D survey) and the Toronto Area (using 1996 TTS) based on some key socio-demographic characteristics. Average age is increasing in both cities with an increasing proportion of elderly people (Roorda, Morency, et al., 2008), however the proportion of people aged 55 years old and older is significantly higher in the Montreal Island. Else, there is no significant difference in gender composition between these cities. There is a higher proportion of full-time workers in the Montreal Island. Average household size is higher in the Toronto Area (2.73 people per household in 2001) than the Montreal Area (2.31 people per household in 2003) (Roorda, Morency, et al., 2008), because households with no children and one adult households are more common on the Montreal Island, whereas two adults' households are more common on the Toronto Area. On the other hand, the proportion of students is slightly higher in the Toronto Area. There, more people possess a driver's license and a lower proportion of people has a transit pass. Also, car ownership rate is higher in the Toronto Area (1.41 in 2001) compared to the Montreal Area (1.18 in 2003), however it is increasing more dramatically in the Montreal Area (Roorda, Morency, et al., 2008). These socio-demographic and other differences between the two cities are a likely cause of differences in activity-travel behaviours.

- **Temporal issue**

The TASHA application in the Montreal Island uses model parameters (such as activity attribute distributions) from Toronto settings from an earlier year (i.e. 1996) to simulate activity schedules of the individuals for the year of 2003. However these activity attribute distributions used as input in TASHA may not remain constant over time. A recent study shows that activity attributes for work, school, shopping and other activities have changed over three time points in a 10-year period (1998, 2003 and 2008) in Montreal (Yasmin, Morency, & Roorda, 2015b).

- **Scale difference**

The TASHA model, developed for the Toronto Area has been transferred to the highly urbanized setting of the Montreal Island. The Toronto Area (7,125 square km) is approximately 14 times larger than the Montreal Island (500 square km). This scale difference could also contribute to the large differences found in the validation results. Thus, it is plausible that individuals in the Montreal Island has different activity-travel behaviours than in the Toronto Area. It is noted here that the Montreal Island covers 11% of the area of the Greater Montreal Area (GMA); however 49% of total population of the GMA resides in the Montreal Island.

- **Differences in activity attribute distributions**

This research assumes that the activity attributes distributions in the Toronto Area and the Montreal Island are similar and thus employs the observed activity attributes distributions from the 1996 TTS of the GTA while transferring the TASHA model to the Montreal Island. A total of 262 distributions for ten activity types are developed based on particular properties of individuals, households, and schedules (Roorda, Miller, et al., 2008). Table 5.3 presents the explanatory variables and number of distributions for each activity type (Roorda, 2005). Detailed definition of activity type can be found in (Roorda, Miller, et al., 2008). However, it is quite plausible that all 262 distributions might not be similar in both cities which might lead to the large differences found in some cases in the comparative analyses of this research. To examine this, the study further prepares these 262 distributions of three activity attributes for the Montreal Island using the 2003 O-D travel survey and then compares them with the similar distributions of the Toronto Area. K-S tests have also been performed to examine whether these activity attributes distributions between the two cities are the same, with at least 90% confidence.

This study finds that the TASHA simulation provides large differences mostly for flexible activities (shopping and other activities) at different levels. Thus, we have focused the discussion here on the distributions of these activities; more specifically, we have observed those distributions which cover approximately 80% of total population of the Montreal Island. Table 5.4 indicates whether these distributions of shopping and other activities (both independent and joint) are from the same distribution based on the K-S tests results (*D* statistic[3] and *P*-value[4]).

---

[3] Test based on the largest difference between the two cumulative distributions.

[4] P-value, equal or smaller than the assumed significance level suggests that two distributions are sampled from populations with different distributions.

Table 5.3: Explanatory variables for different activity

Activity type	Explanatory variables used for classification	Number of distributions (*)
Primary work	Age, occupation, employment status	32 (5)
Secondary work	Occupation, employment status	8 (4)
Work-business	Age, occupation, employment status	32 (5)
Work-at-home business	Age, occupation, employment status	12 (3)
Return home from work	Occupation, employment status	8 (4)
School	Age, student status	10 (3)
Independent other	Age, gender, individual work-school project status	56 (10)
Joint other	Presence of children, number of adults, household work-school project status	24 (8)
Independent shopping	Age, gender, Individual work-school project status	56 (10)
Joint shopping	Presence of children, number of adults, household work-school project status	24 (8)
Total		**262**

* number in the parentheses indicates the number of distributions covering approximately 80% of population in each activity type of the Montreal Island

Among 56 distributions of activity type 'independent other', 10 distributions cover a significant proportion of total population (approximately 80%) of the Montreal Island. Among them, the K-S tests indicate that the frequency distributions of 8 distributions between two cities are quite similar, whereas it indicates that the hypothesis of similar start time and duration distributions of these 10 distributions are either rejected or inconclusive. In case of activity type 'joint other', 8 distributions cover most of the population (approximately 80%) of the Montreal Island and the K-S tests provide strong evidence that the frequency distributions of 5 distributions are similar in both cities. However, the start time distributions of these 8 distributions between two cities are mostly different. But, the K-S tests indicate that the hypothesis of similar duration distributions of 4 distributions are either rejected or inconclusive. For activity types 'independent and joint shopping', the K-S tests indicate a very high probability that the frequency distributions of the distributions, which cover most of the population in both activity types (10 distributions for independent and 8 distributions for joint shopping), are mainly the same. However, the K-S tests indicate that the hypothesis of similar start time distributions of most of the distributions between the two cities are either rejected or inconclusive. On the other hand, the K-S tests show strong evidence that some duration distributions are similar in the two cities, whereas it indicates that the hypothesis of similar duration distributions of rest of the distributions are inconclusive.

Table 5.4: K-S test results[a] for the distributions (cover approximately 80% of population) of shopping and other activities

Number of distributions[b]	Activity attributes					
	Frequency	Start time	Duration	Frequency	Start time	Duration
	Independent other			Independent shopping		
10 distributions	*S*	*D*	*D*	*In*	*D*	*In*
	S	*In*	*D*	*S*	*D*	*In*
	S	*In*	*D*	*S*	*D*	*S*
	In	*D*	*D*	*S*	*S*	*S*
	S	*D*	*D*	*S*	*S*	*S*
	S	*D*	*In*	*S*	*In*	*S*
	S	*D*	*In*	*S*	*D*	*S*
	S	*D*	*D*	*S*	*D*	*In*
	In	*D*	*D*	*S*	*D*	*In*
	S	*D*	*D*	*S*	*S*	*In*
	Joint other			Joint shopping		
8 distributions	*S*	*In*	*S*	*S*	*In*	*S*
	In	*In*	*In*	*S*	*In*	*S*
	S	*S*	*S*	*S*	*S*	*S*
	In	*D*	*D*	*S*	*In*	*In*
	S	*D*	*S*	*S*	*D*	*S*
	In	*D*	*D*	*S*	*D*	*In*
	S	*S*	*S*	*In*	*D*	*S*
	S	*D*	*In*	*In*	*D*	*S*

a	For the *D* statistic and *P*-value, contact the authors.
b	indicates the number of distributions covering approximately 80% of population in each activity type of the Montreal Island
D	indicates that K-S test rejects the null hypothesis that two sample distributions come from the same distribution, with at least 90% confidence (*P*-value ≤ 0.10).
S	indicates that K-S test provides strong evidence (*P*-value ≥ 0.90) that two sample distributions come from the same distribution.
In	indicates that K-S test result is inconclusive (P-value ranges from 0.20 to 0.81).

The comparison of activity attributes distributions for shopping and other activities between two cities reveals that activity frequency distributions of these activities are mostly similar in both cities, whereas start time and duration distributions are mostly different. This finding clearly indicates behavioural differences in activity-travel patterns between the Montreal Island and the Toronto Area.

In TASHA, the population segmentation in each activity type, shown in Table 5.3, has been done based on intuition. Thus, there could be further improvement in population segmentation by employing segmentation process (also known as clustering) based on their activity generation behaviours (frequency, start time and duration) to develop activity attribute distributions.

- **Scheduling rules in TASHA**

TASHA is a rule-based model, which schedules the generated activities following a rule-based procedure. In TASHA, a single set of strategies of activity scheduling is applied for all individuals in a household (Miller & Roorda, 2003). However, different people may have different strategies for scheduling their activities. Though this research finds that the validation results at macro- and meso-level are quite similar (for instance close simulation of fixed activities); at the meso-level, TASHA simulation of the activity attributes for some population segments provides larger differences than the simulation results at the macro-level. Thus, there could be further improvement in the TASHA modelling framework by applying different rules of scheduling for different population segments.

5.7 Conclusions

This paper has focused on the spatial transferability as a validation test of an activity-based model, TASHA. The validation has been performed at three different levels, macro-level (aggregation of the entire population), meso-level (aggregation by population segments by age group and gender, and by home location), and micro-level (individuals).

Validation results at macro- and meso-level clearly demonstrate that TASHA can successfully reproduce activity behaviours of another context, at least for fixed activities (work, school) with few exceptions. However, TASHA provides large differences for some activity attributes of flexible activities (shopping, other). Although micro-level validation of a model is customary in Montreal, the need for validation of a model at this level depends on the purpose of using the model. Considering spatial, temporal and other differences between the Montreal Island and the Toronto Area, the transferability results at different levels (especially at macro- and meso-level) seem quite promising. However in general, we recommend to re-estimate model parameters and to use local activity attribute distributions (frequency, start time and duration) if available, when transferring the TASHA model from one context to another.

Acknowledgements

The authors would like to thank the Ministère des transports du Québec (Service de la modélisation des systèmes de transport) for their support and contributions. They also wish to thank the technical committee on O-D travel surveys in the Montreal region for providing access to data for research purposes. They would also like to thank Mr. Hubert Verreault for micro-level validation analysis. They also wish to thank Mr. Pierre-Leo Bourbonnais for travel time calculation of transit trips.

References

Agence Métropolitaine de Transport. (2015). Portrait de la mobilité. Retrieved January 14, 2015, from https://www.amt.qc.ca/fr/a-propos/portrait-mobilite

Arentze, T., Hofman, F., Van Mourik, H., & Timmermans, H. (2002). Spatial transferability of the Albatross model system: Empirical evidence from two case studies. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 1-7.

Arentze, T., & Timmermans, H. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological, 38*(7), 613-633. doi: 10.1016/j.trb.2002.10.001

Arentze, T., & Timmermans, H. (2008). *ALBATROSS: Overview of the model, application and experiences*. Paper presented at the Innovations in Travel Modeling 2008 Conference, Portland, Oregon, USA.

Auld, J., & Mohammadian, A. K. (2012). Activity planning processes in the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Research Part A: Policy and Practice, 46*(8), 1386-1403. doi: http://dx.doi.org/10.1016/j.tra.2012.05.017

Auld, J., & Mohammadian, A. K. (2009). Framework for the development of the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Letters: The International Journal of Transportation Research, 1*(3), 243-253.

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

Balakrishnan, T. R., Maxim, P., & Jurdi, R. (2005). Social class versus cultural identity as factors in the residential segregation of ethnic groups in Toronto, Montreal and Vancouver for 2001. *Canadian Studies in Population, 32*(2), 203-227.

Ben-Akiva, M. E., & Bowman, J. (1998). Activity-based travel demand model systems. In P. Marcotte & S. Nguyen (Eds.), *Equilibrium and Advanced Transportation Modeling* (pp. 27-46). Boston/ Dordrecht/London: Kluwer Academic Publishers.

Bhat, C. R. (1999). *A comprehensive and operational analysis framework for generating the daily activity travel profiles of workers.* Paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC, USA.

Bhat, C. R., Guo, J. Y., Srinivasan, S., & Sivakumar, A. (2004). Comprehensive econometric microsimulator for daily activity-travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1894*, 57-66. doi: 10.3141/1894-07

Bhat, C. R., & Lawton, K. T. (2000). Passenger travel demand forecasting *Millennium Paper*. Washington D.C., USA: Transportation Research Board, National Research Council.

Bhat, C. R., Paleti, R., Pendyala, R. M., & Goulias, K. G. (2013). SimAGENT activity-based travel demand analysis *SimAGENT Core Models* (Vol. 4): GeoTrans Laboratory, University of California, Santa Barbara.

Bowman, J. L. (1998). *The day activity schedule approach to travel demand analysis.* (Doctor of Philosophy), Massachusetts Institute of Technology, Cambridge, Massachusetts.

Bowman, J. L., & Ben-Akiva, M. E. (2000). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice, 35*(1), 1-28.

Bowman, J. L., Bradley, M., Castiglione, J., & Yoder, S. L. (2014). *Making advanced travel forecasting models affordable through model transferability.* Paper presented at the 93rd Transportation Research Board Annual Meeting, Washington D.C., USA.

Bradley, M. A., & Bowman, J. L. (2006). *A summary of design features of activity-based microsimulation models for U.S. MPOs.* Paper presented at the TRB Conference on Innovations in Travel Demand Modeling, Austin, Texas

Childress, S., Sabina, E., Kurth, D., Rossi, T., & Malm, J. (2010). *DRCOG activity-based model calibration/ validation.* Paper presented at the Innovations in Travel Modeling Conference, Transportation Research Board, Tempe, Arizona.

Doherty, S. T., Nemeth, E., Roorda, M. J., & Miller, E. J. (2004). Design and assessment of the Toronto area computerized household activity scheduling survey. *Transportation Research Record: Journal of the Transportation Research Board, 1894*, 140-149.

Eberhard, L. K. (2002). *A 24-h household-level activity based travel demand model for the GTA.* (Masters Thesis), University of Toronto, Toronrto, Canada.

Ettema, D., Bergers, A., & Timmermans, H. (1993). Simulation model of activity scheduling behaviour. *Transportation Research Record: Journal of the Transportation Research Board, 1413*, 1-11.

Ettema, D., Borgers, A., & Timmermans, H. (1996). SMASH (Simulation Model of Activity Scheduling Heuristics): Some simulations. *Transportation Research Record: Journal of the Transportation Research Board, 1551*, 88-94.

Gangrade, S., Kasturirangan, K., & Pendyala, R. M. (2000). Coast-to-coast comparison of time use and activity patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1718*, 34-42.

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

Golledge, R. G., Kwan, M.-P., & Gärling, T. (1994). Computational process modeling of household travel decisions using a geographic information system. *Papers in Regional Science, 73*(2), 99-117.

Heisz, A. (2006). Canada's global cities: Socio-economic conditions in Montréal, Toronto and Vancouver *Trends and Conditions in Census Metropolitan Areas*. Ottawa: Statistics Canada.

Kitamura, R., & Fujii, S. (1998). Two computational process models of activity-travel behavior. In T. Garling, T. Laitila & K. Westin (Eds.), *Theoretical Foundations of Travel Choice Modeling* (pp. 251-279). Bingley, UK: Emerald Group Publishing.

Kurth, D. L., Childress, S., Sabina, E. E., & Rossi, T. (2006). *Proposed validation and sensitivity testing of Denver region activity-based models.* Paper presented at the Innovations in Travel Demand Modeling Conference, Transportation Research Board, Austin, TX.

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

Miller, E. J., & Roorda, M. J. (2003). Prototype model of household activity-travel scheduling. *Transportation Research Record: Journal of the Transportation Research Board, 1831*, 114-121. doi: 10.3141/1831-13

Mohammadian, A. K., Auld, J., & Yagi, S. (2009). Recent progress on activity-based microsimulation models of travel demand and future prospects. In B. W. Sloboda (Ed.), *Transportation Statistics* (pp. 151-171): J. Ross Publishing, Inc.

Pendyala, R. M., Kitamura, R., Kikuchi, A., Yamamoto, T., & Fujii, S. (2005). Florida Activity Mobility Simulator: Overview and preliminary validation results. *Transportation Research Record: Journal of the Transportation Research Board, 1921*, 123-130. doi: 10.3141/1921-14

Recker, W. W., McNally, M. G., & Root, G. S. (1986a). A model of complex travel behavior: Part I--Theoretical development. *Transportation Research Part A: General, 20*(4), 307-318. doi: 10.1016/0191-2607(86)90089-0

Recker, W. W., McNally, M. G., & Root, G. S. (1986b). A model of complex travel behavior: Part II--An operational model. *Transportation Research Part A: General, 20*(4), 319-330. doi: 10.1016/0191-2607(86)90090-7

Roorda, M. J. (2005). *Activity-based modelling of household travel.* (Doctor of Philosophy), University of Toronto, Toronto, Canada.

Roorda, M. J., & Miller, E. J. (2006). Past Presidents' Award for merit in transportation engineering: Assessing transportation policy using an activity-based microsimulation model of travel demand. *ITE Journal (Institute of Transportation Engineers), 76*, 16-21.

Roorda, M. J., Miller, E. J., & Habib, K. M. N. (2008). Validation of TASHA: A 24-h activity scheduling microsimulation model. *Transportation Research Part A: Policy and Practice, 42*, 360-375. doi: 10.1016/j.tra.2007.10.004

Roorda, M. J., Morency, C., & Woo, K. (2008). *Two cities, two realities? A closer look at the evolution of trip rates in Toronto and Montreal*. Paper presented at the 87th Transportation Research Board (TRB) Annual Meeting, Washington D.C., USA.

Rose, D. (1999). Urban hierarchies and the changing characteristics of 'urban professionals' in Toronto and Montreal. *Canadian Journal of Regional Science, XXII-1,2*, 133-142.

Shearmur, R. (2006). Travel from home: An economic geography of commuting distances in montreal. *Urban Geography, 27*(4), 330-359.

Shiftan, Y., & Ben-Akiva, M. (2011). A practical policy-sensitive, activity-based, travel-demand model. *The Annals of Regional Science, 47*(3), 517-541.

Shiftan, Y., Ben-Akiva, M., Proussaloglou, K., Jong, G. d., Popuri, Y., Kasturirangan, K., & Bekhor, S. (2003). *Activity-based modeling as a tool for better understanding travel behaviour*. Paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Sikder, S., Pinjari, A., Srinivasan, S., & Nowrouzian, R. (2013). Spatial transferability of travel forecasting models: A review and synthesis. *International Journal of Advances in Engineering Sciences and Applied Mathematics, 5*(2-3), 104-128. doi: 10.1007/s12572-013-0090-6

Statistics Canada. (2011a). Population and dwelling counts, for Canada and census divisions. *Canada 2011 Census.* Retrieved July 30, 2013, from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=701

Statistics Canada. (2011b). Population and dwelling counts, for census metropolitan areas. *Canada 2011 Census.* Retrieved February 27, 2014, from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=205&S=3&RPP=50

Vovsha, P., Petersen, E., & Donnelly, R. (2004). Impact of intra-household interactions on individual daily activity travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1898*, 87-97.

Yagi, S., & Mohammadian, A. K. (2007a). *Microsimulation and policy applications of an activity-based modeling system of travel demand.* Paper presented at the 86th Transportation Research Board Annual Meeting, Washington DC, United States.

Yagi, S., & Mohammadian, A. K. (2007b). *Validation and policy applications of an activity-based microsimulation model of travel demand.* Paper presented at the 11th World Conference on Transportation Research, Berkeley, California.

Yagi, S., & Mohammadian, A. K. (2008a). Joint models of home-based tour mode and destination choices: Applications to a developing country. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 29-40.

Yagi, S., & Mohammadian, A. K. (2008b). Modeling daily activity-travel tour patterns incorporating activity scheduling decision rules. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 123-131.

Yasmin, F., Morency, C., & Roorda, M. J. (2013). *Spatial transferability of activity-based models: Moving TASHA from Toronto to Montreal*. Paper presented at the 13th World Conference on Transport Research, Rio de Janeiro, Brazil.

Yasmin, F., Morency, C., & Roorda, M. J. (2015a). Assessment of spatial transferability of an activity-based model, TASHA. *Transportation Research Part A: Policy and Practice, 78*, 200-213. doi: 10.1016/j.tra.2015.05.008

Yasmin, F., Morency, C., & Roorda, M. J. (2015b). Trend analysis of activity generation attributes over time. *Transportation*. doi: 10.1007/s11116-015-9624-z

CHAPTER 6 ARTICLE 3: TREND ANALYSIS OF ACTIVITY GENERATION ATTRIBUTES OVER TIME

This chapter is presented in the form of a paper published in the *Transportation* (2015), 1-21.

Farhana Yasmin[a,*], Catherine Morency[a,1], Matthew J. Roorda[b,2]

[a] *Department of Civil, Geological and Mining Engineering, Polytechnique Montreal, C.P. 6079, Station Centre-Ville, Montreal, Quebec, Canada H3C 3A7*
[b] *Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario, Canada M5S 1A4*
[*] *Corresponding author. E-mail: farhana.yasmin@polymtl.ca*
[1] *E-mail: cmorency@polymtl.ca*
[2] *E-mail: roordam@ecf.utoronto.ca*

Abstract

Activity generation models are relatively poorly developed in activity-based travel demand modelling frameworks. This research investigates whether observed distributions of activity attributes (activity frequency, start time and duration) used as inputs in the activity generation component of an activity-based travel demand model have changed over time. This research empirically examines changes in the distributions of activity generation attributes over time in the Greater Montreal Area (GMA), Quebec, Canada. It also focuses on how these attributes vary with peoples' socio-demographic characteristics. This research relies on the 1998, 2003 and 2008 Origin-Destination (O-D) household travel surveys of the GMA. The comparative analysis at three time points in a 10-year period clearly reveals that distributions of activity attributes are significantly changing over time. Work and school activities show similar trends; frequency "1" has increased and frequency "2+" has decreased over time. The occurrence of shopping activity on weekdays is decreasing over time. Start time and duration distributions for each activity have also changed significantly over time. The research allows preparing activity attributes for the application of an activity-based model, TASHA, such that they reflect temporal changes in travel behaviour of the GMA.

Keywords: Travel demand modelling; Activity generation; Activity attribute distribution; TASHA; Montreal

6.1 Introduction

Activity-based travel demand models consist of two major components, namely activity generation and activity scheduling (Bhat & Koppelman, 1993; Habib & Miller, 2009). Unlike the activity scheduling components, modelling efforts are still relatively rare for the activity generation components in activity-based modelling frameworks. The activity-based model TASHA (Travel Activity Scheduler for Household Agents), developed for the Greater Toronto Area (GTA), Canada by Miller and Roorda (2003), has been spatially transferred to Montreal, Canada (Yasmin, Morency, & Roorda, 2014, 2015). In TASHA, observed distributions of activity attributes (frequency of different types of activity, start time and duration) are used as inputs in the activity generation component and these distributions are assumed to remain constant over time. However it is quite possible that activity generation behaviour of people and households may change over time. Therefore, this research empirically examines changes in the distributions of activity generation attributes over time in the Greater Montreal Area (GMA). It also focuses on how these attributes vary with peoples' socio-demographic characteristics. Some of the key results of this research have been presented in (Yasmin, Morency, & Roorda, 2012).

The remainder of this paper is organized as follows. First, a brief review of the related literature is provided. Next, data sources and research methodology are defined; this section also includes an analysis of the main changes in demography, socio-economic characteristics, and land-uses in the GMA over time. Then, the trends of activity generation behaviour over time are presented. Next, the important changes of activity attributes are discussed with respect to different socio-demographic characteristics. Finally, important findings are summarized and implications of this research are identified in the conclusion.

6.2 Background

As transportation significantly influences mobility, economic and land-use development, environmental quality, government finance, and quality of life, careful planning is required in this field. Transportation planning involves forecasts of travel behaviour necessary to provide adequate facilities and services to meet future demand. For effective planning, travel demand models must be sensitive to changes in the attributes of the transportation system and the behaviour of people who use the system, as a result of different policies and strategies. To date, it is well recognized that travel behaviour can be best understood within a framework of activity participation (Kitamura, 1996; Kitamura, Pas, Lula, Lawton, & Benson, 1996; Shiftan et al., 2003). Accordingly, since the 1970s the activity-based approach has been examined or implemented in several major cities in North America and Europe. The activity-based approach treats travel as a demand derived from different activities distributed in space and time (Kitamura & Fujii, 1998).

Activity-based travel demand models consist of two major components, namely activity generation and activity scheduling (Bhat & Koppelman, 1993; Habib & Miller, 2009). The activity generation component mainly deals with the desire or need to participate in activities (i.e. activity demand), whereas the scheduling component deals with spatial-temporal opportunities and constraints which lead to an executed activity pattern. In several activity-based modelling frameworks, activity generation components get less attention than scheduling components. Some models have used activity patterns as given input in the modelling framework directly drawn from the observed data without any modelling effort to specify how the activity agenda is developed (e.g. CARLA (Jones, Dix, Clarke, & Heggie, 1983), AMOS (Kitamura, Lula, & Pas), and SMASH (Ettema, Bergers, & Timmermans, 1993)). The activity-based model, STARCHILD (Recker, McNally, & Root, 1986a, 1986b) has separate activity generation and scheduling components; however limitations exist in terms of feedback and updating within the framework. The Bowman and Ben-Akiva model simulates daily activity patterns and generate travel tours for each household member using nested logit model structures (Bowman & Ben-Akiva, 2000). Activity-based models developed based on a similar approach to the Bowman and Ben-Akiva model, but with some modifications, have been applied in several US cities and regions (e.g. CT-RAMP (Davidson, Vovsha, Freedman, & R., 2010; Vovsha, Petersen, & Donnelly, 2004), DaySim (Bradley, Bowman, & Griesenbeck, 2010), and TourCast (Meeks et al., 2013)). Some models have employed Monte-Carlo simulation to generate activity patterns based on the observed distributions of activity attributes (e.g. FAMOS (Pendyala, Kitamura, Kikuchi, Yamamoto, & Fujii, 2005), and TASHA (Miller & Roorda, 2003; Roorda, Miller, & Habib, 2008)). This stochastic technique produces random error from the Monte Carlo draws. However, Veldhuisen, Timmermans, and Kapoen (2000) have shown that the effect of simulation error is negligible for aggregate outcomes, but, is significant for disaggregate outcomes. This problem can be handled by running the model multiple times and averaging the outcomes, running the simulation with a supersample, or coordinating random seed for comparison of scenarios (Bowman, 2008). CEMDAP (Bhat, Guo, Srinivasan, & Sivakumar, 2004) employs an activity generation approach similar to trip generation models of the traditional four-stage travel demand modelling framework. ALBATROSS (Arentze & Timmermans, 2004) models both generation and scheduling components comprehensively. Habib and Miller (2008) have developed an activity generation modelling framework for a week-long time span based on random utility maximization theory. Arentze and Timmermans (2009) have proposed a dynamic activity generation model for multiday planning period based on a needs-based approach which implies that an individual conducts activities to satisfy particular needs. Given the importance of activity generation as a driver of travel, the activity generation component should receive considerably more attention in the field of activity-based travel demand modelling.

As part of a broad research project with an aim to enrich the current process of modelling of travel demand of the Greater Montreal Area using an activity-based approach, the activity-based model TASHA has been applied in Montreal (Yasmin et al., 2014, 2015). TASHA is a fully disaggregate microsimulation model which estimates activity schedules and travel patterns for a twenty-four hour typical weekday for all individuals in a household. The full conceptual design and methodology of this activity scheduling model can be found in (Miller & Roorda, 2003). The model has been developed based on trip diary data from the 1996 Transportation Tomorrow Survey (TTS) for the Toronto Area. As of 2012, the GMA is comprised of an area of 4,258 square km with a population of 3,824,221, whereas the GTA is composed of an area of 7,125 square km with a population of 6,054,000 (Statistics Canada, 2011a, 2011b). Montreal was founded in 1642; whereas Toronto was founded in 1793 as York. Montreal experienced significant urban growth during the 1950s and 60s and was the uncontested metropolis of Canada during the first half of the twentieth century before being surpassed by Toronto in 1971.

Like other activity-based models, TASHA includes both activity generation and activity scheduling components. The activity generation component generates activities (individual and joint) for each person using Monte-Carlo simulation from observed distributions of activity frequency, start time and duration (Miller & Roorda, 2003). Yasmin et al. (2015) have transferred TASHA from Toronto to Montreal without changing any parameters from the Toronto settings. They have used observed distributions of activity attributes (frequency of different types of activity, start time and duration) from the 1996 TTS of the Toronto Area as input for the

Montreal application. These distributions are assumed to remain constant over time; however it is quite possible that activity generation attributes may change over time and may need some modelling efforts to prepare the distributions of attributes such that they reflect temporal changes in travel behaviour.

This research empirically examines changes in the distributions of activity generation attributes over time in the GMA and provides insights into the possible reasons for these changes. Our analysis of the reasons for these changes considers demographic, socio-economic, land use, technology, and policy changes in the GMA. This research specifically addresses research questions including: are activity generation behaviours of the GMA changing over time?, and if yes, then which activities and socio-demographic cohorts are experiencing the important changes? Four different types of activity are examined: work, school, shopping, and other.

6.3 Data and research method

In the Greater Montreal Area, large-scale traditional Origin-Destination (O-D) household travel surveys have been conducted every five years since the early 1970's. Detailed information on the O-D travel surveys can be found on the AMT website (Agence Métropolitaine de Transport [AMT], 2015). The survey collects detailed travel and socio-economic information for approximately 5% of the total GMA population aged 5 years and older. Socio-economic information on households (home location, size, car ownership), individuals (age, gender, employment status, driving license, mobility) and travel information of all household members (trip purpose, trip start time, sequence, mode, transit path, highways/bridges used, type of parking spaces, and so on) of a specific weekday in the fall period (September to December) are collected. The survey gathers precise spatio-temporal details on each trip. The data are collected by a telephone interview, however, in recent years, this traditional survey method is facing several challenges (for instance response rate decreases, sampling frame limitations, high costs and so on) (Bonnel, Lee-Gosselin, Madre, & Zmud, 2009). Advancement in technologies (for instance availability of smart phones, global positioning satellite (GPS), and easy access to internet) allows the development of new, improved methods, such as smart phone-based, GPS-based, and web-based surveys (Bourbonnais & Morency, 2013; Cottrill et al., 2013; Stopher, Prasad, Wargelin, & Minser, 2013). However, these methods are still evolving and require more research to understand their benefits, problems as well as challenges (Zmud, Lee-Gosselin, Munizaga, & Carrasco, 2013).

This research relies on the 1998, 2003 and 2008 O-D travel surveys of the GMA. During data preparation, open chains (i.e. trip chains that did not start and end at home) have been observed for some individuals in each survey and we have excluded those individuals from the analysis. Therefore, the trend analysis of activity generation behaviours includes trip observations of 145,083 (1998), 126,982 (2003) and 141,761 (2008) individuals. However, socio-demographic analyses include all individuals in a household.

The analysis focuses on three activity attributes, frequency of different types of activity (work, school, shopping, and other excluding return to home) and their temporal organization within a day (start time and duration). The "other" activity type includes en route, leisure, visit friends/family, health, drop someone off, pick someone up, and other. Trip information from three travel surveys is processed to prepare activity attributes data for the analysis. Activity type has been defined from the destination purpose of a trip observation. Start time of activity is taken directly from the start time of the trip observation. Activity duration has been calculated using successive trip start times, thus, it includes travel duration from the previous activity.

Activity generation attributes have been compared across the three travel surveys. Also, the attributes are examined for socio-demographic groups (by age, gender, occupation, employment status, student status, *et cetera*). Kolmogorov-Smirnov tests (K-S tests) have been performed to determine the statistical differences between activity attributes distributions for different groups of people within each survey and between surveys over different years.

Table 6.1: Some key socio-demographic characteristics and land uses in the GMA

Variables		1998	2003	2008
Average age		36	38[a]	39[a]
Age groups (%)				
Ages 24 years old and younger		32.5%	30.3%[a]	29.7%[a]
Ages 25 - 54 years old		46.8%	46.6%	45.0%[a]
Ages 55 years old and older		20.7%	23.1%[a]	25.3%[a]
Gender (%)				
Men		49.1%	49.6%[a]	48.6%[a]
Women		50.9%	50.4%[a]	51.4% [a]
Employment status by gender (%)				
Full-time worker	Men	23.4%	23.8%[a]	23.5%[a]
	Women	16.9%	18.2%[a]	18.7%[a]
Part-time worker	Men	1.5%	1.5%	1.5%
	Women	3.2%	3.0%[a]	3.0%
Student	Men	11.8%	11.7%	11.4%[a]
	Women	11.6%	11.1%[a]	11.5%[a]
Retired	Men	6.6%	7.1%[a]	7.3%
	Women	8.5%	9.6%[a]	10.3%[a]
Other status	Men	2.6%	2.6%	1.2%[a]
	Women	7.9%	4.8%[a]	1.2%[a]
Possession of driver's license (%)				
Yes		62.1%	64.3%[a]	65.7%[a]
No		18.0%	17.4%[a]	16.1%[a]
Household car ownership rate		1.18	1.22 [a]	1.28 [a]
Average household size		2.41	2.32 [a]	2.33
Number of children (%)				
0		68.9%	71.9%[a]	71.3%[a]
1		13.8%	12.7%[a]	13.0%
2		12.6%	11.5%[a]	11.6%
3+		4.7%	4.0%[a]	4.1%
Number of adults (%)				
1		33.3%	34.2%[a]	34.8%[a]
2		51.7%	52.3%[a]	51.3%[a]
3		10.6%	9.7%[a]	10.0%[a]
4		3.5%	3.1%	3.1%
5+		0.8%	0.6%	0.7%
Population density[b] (%)				
>5,000 persons/km^2	(high-density urban)	29.0%	28.3%[a]	27.6%[a]
>2,000 and ≤5,000 persons/km^2	(medium-density urban)	25.4%	27.1%[a]	26.6%[a]
>1,000 and ≤2,000 persons/km^2	(medium-density suburban)	12.4%	14.3%[a]	14.5%
≤1000 persons/km^2	(low-density suburban)	33.3%	30.3%[a]	31.3%[a]

[a]Value is significantly different at the 95% confidence level than that observed in the previous survey.
[b]Classification of density areas has been taken from (Roorda, Morency, & Woo, 2008).

A comparative analysis of key socio-demographic characteristics and land-uses in the GMA over three time points (1998, 2003 and 2008) in a 10-year period has been presented in Table 6.1. As a land-use indicator, the proportion of population living in four types of density areas (defined by Roorda, Morency, et al. (2008)) across three travel surveys has been compared. The population density has been measured for all municipal sectors using the commonly surveyed area from the 1998 travel survey for the GMA. Then, the municipal sectors are aggregated in four types of density areas, shown in Table 6.1. This comparative analysis reveals some important changes in socio-demographic characteristics and land uses over 10-year period (for instance ageing of the population, increasing female participation in workforce, increasing motorization rate, decreasing household size, and suburbanization) must affect the overall trends in activity generation behaviours in the GMA.

6.4 Trends of activity generation behaviour

This section presents the trend analysis of activity frequency, start time and duration and reveals whether distributions of these activity attributes in the GMA are changing over time. Table 6.2 indicates whether the activity attributes distributions among different years are from the same distribution based on K-S test results (*D* statistic[5] and *P*-value[6]).

6.4.1 Frequency of activities

This analysis presents the frequency comparison of four activities, i.e. work, school, shopping and other over a 10-year period (1998 to 2008). K-S tests (Table 6.2) reject the null hypothesis that frequency distributions of each activity among three years are the same except for the distribution of shopping activity from the year 2003 to 2008. The analysis shown in Table 6.3

[5] Test based on the largest difference between the two cumulative distributions.

[6] P-value, equal or smaller than the assumed significance level suggests that two distributions are sampled from populations with different distributions.

Table 6.2: K-S test results* for trends of three activity generation attributes

			Frequency				Start time				Duration			
		Year	Work	School	Shopping	Other	Work	School	Shopping	Other	Work	School	Shopping	Other
K-S test results for trends for the entire population														
		1998-2003	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
		2003-2008	*D*	*D*	*In*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
Variables	**Groups**	**K-S test results for trends by socio-demographic characteristics**												
Age groups	24 years old or younger	1998-2003	*S*	*D*	*D*	*D*	*S*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
		2003-2008	*D*	*D*	*D*	*D*	*In*	*D*	*In*	*D*	*D*	*D*	*D*	*D*
	25 - 54 years old	1998-2003	*D*	*S*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
		2003-2008	*D*	*In*	*D*	*D*	*D*	*D*	*In*	*D*	*D*	*D*	*In*	*D*
	55 years old and older	1998-2003	*D*	*S*	*D*	*D*	*D*	*In*	*D*	*D*	*D*	*S*	*D*	*In*
		2003-2008	*In*	*S*	*D*	*D*	*D*	*In*	*D*	*In*	*D*	*In*	*D*	*D*
Gender	Men	1998-2003	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
		2003-2008	*D*	*D*	*S*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*In*
	Women	1998-2003	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
		2003-2008	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
Occupation	Full-time worker	1998-2003	*D*	*In*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*In*	*D*	*D*
		2003-2008	*D*	*S*	*D*	*D*	*D*	*In*	*D*	*D*	*D*	*D*	*D*	*D*
	Part-time worker	1998-2003	*In*	*In*	*D*	*In*	*In*	*S*	*D*	*D*	*D*	*S*	*D*	*D*
		2003-2008	*D*	*S*	*S*	*D*	*D*	*D*	*In*	*D*	*D*	*D*	*In*	*D*
	Student	1998-2003	*S*	*D*	*D*	*D*	*In*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
		2003-2008	*In*	*D*	*In*	*D*	*S*	*D*	*S*	*D*	*In*	*D*	*D*	*D*
	Retired	1998-2003	*S*	*S*	*D*	*In*	*S*	*S*	*D*	*D*	*D*	*In*	*D*	*In*
		2003-2008	*S*	*S*	*D*	*D*	*In*	*S*	*In*	*In*	*In*	*S*	*In*	*In*

* For the *D* statistic and *P*-value, contact the authors.

D indicates that K-S test rejects the null hypothesis that two sample distributions come from the same distribution, with at least 90% confidence (*P*-value ≤ 0.10).

S indicates that K-S test provides strong evidence (*P*-value ≥ 0.90) that two sample distributions come from the same distribution.

In indicates that K-S test result is inconclusive (*P*-value ranges from 0.15 to 0.79).

indicates that work activity frequency “0 times” is slightly decreasing from 1998 to 2003, but again increasing a little from 2003 to 2008. Thus, overall the occurrence of work activities is slightly increasing (0.4%) over time. This might be happening because the number of jobs are increasing; people are engaging more in work activity in this area (Communauté Métropolitaine de Montréal, 2010), especially female participation in the workforce is growing over time (Fortin, Godbout, & St-cerny, 2012), as also observed in this research (Table 6.1).

Table 6.3: Comparison of frequency of activities

Activity type	Frequency (activities per day)	1998 (%)	2003 (%)	Diff 1998-2003 (Δ %)	2008 (%)	Diff 2003-2008 (Δ %)
Work	0	60.53%	59.53%	-1.00%	60.10%	0.57%
	1	35.38%	37.56%	2.18%	37.59%	0.03%
	2+	4.09%	2.90%	-1.19%	2.31%	-0.60%
School	0	77.72%	78.35%	0.63%	78.68%	0.32%
	1	19.55%	20.23%	0.68%	20.32%	0.08%
	2+	2.72%	1.41%	-1.31%	1.01%	-0.40%
Shopping	0	80.73%	84.45%	3.72%	85.59%	1.14%
	1	15.83%	13.32%	-2.51%	12.48%	-0.85%
	2+	3.44%	2.23%	-1.21%	1.94%	-0.29%
Other	0	69.36%	71.98%	2.62%	73.97%	1.99%
	1	20.92%	19.32%	-1.60%	18.46%	-0.86%
	2+	9.72%	8.69%	-1.03%	7.57%	-1.12%

Another observed trend is that the proportion of people doing at least one work activity per day is increasing, but the proportion of people conducting more than one work activity is decreasing over time. People who are working are less often returning home for lunch and/or doing other activities during the lunch hour. 5.1% of people (who did a work activity) went home for lunch and/or did other activities during the lunch hour in 1998, which decreased to 3.3% in 2003 and further decreased to 2.4% in 2008. Overall population is increasing in both urban and suburban areas in the GMA; however the population growth rate is significantly higher in lower-density suburban areas in the GMA (such as North Shore and South Shore). Hence, employment is mostly concentrated in the central areas (Montreal Downtown and Montreal Centre) in the GMA. Also, the household car ownership increased over time (Table 6.1), which is related to suburbanization, the increase in average commuting distance, and urban congestion. These are potential reasons that people are less often returning home during the lunch hour.

The analysis also reveals that school activity occurrence has decreased (1.0%) from 1998 to 2008. Similar to the work activity trend; the proportion of people conducting at least one school activity is increasing while more than one school activity is becoming less common. The proportion of return home trips during the lunch hour is also decreasing over time; fewer children are going home for lunch perhaps because no one is at home at that time which might be related to the increasing presence of women in the workforce. The study demonstrates that 9.7% of children went home for lunch/conducting other activities in 1998, which decreased to 3.2% in 2008. Also, the analysis indicates that shopping activity occurrence has decreased (4.8%) from 1998 to 2008. This is happening because may be people are changing their shopping habits by doing more shopping in weekends than weekdays (Zhong, Hunt, & Lu, 2008). Online shopping may also be contributing to this reduction of shopping activity (McKeown & Brocca, 2009). Other activity occurrence has also significantly decreased over the 10-year period; however it is expected that a decrease in household size might increase trips related to household maintenance activities (Roorda, Morency, et al., 2008). Although Table 6.1 reveals that average household size is decreasing in the GMA, it also indicates that multi-adult households are still common, thus we assume that people in the household are sharing their household maintenance activities.

6.4.2 Start time of activities

This section shows the changes in weekday start time distributions of different activities over the 10-year period in the GMA. It is noted here that activity start time is defined directly from the start time of travel for participating in an activity. Figure 6.1 (a), (b), (c) and (d) present the activity start time distribution of work, school, shopping, and other activities, respectively. K-S tests reject the hypothesis of similar start time distributions of each activity among three years (Table 6.2).

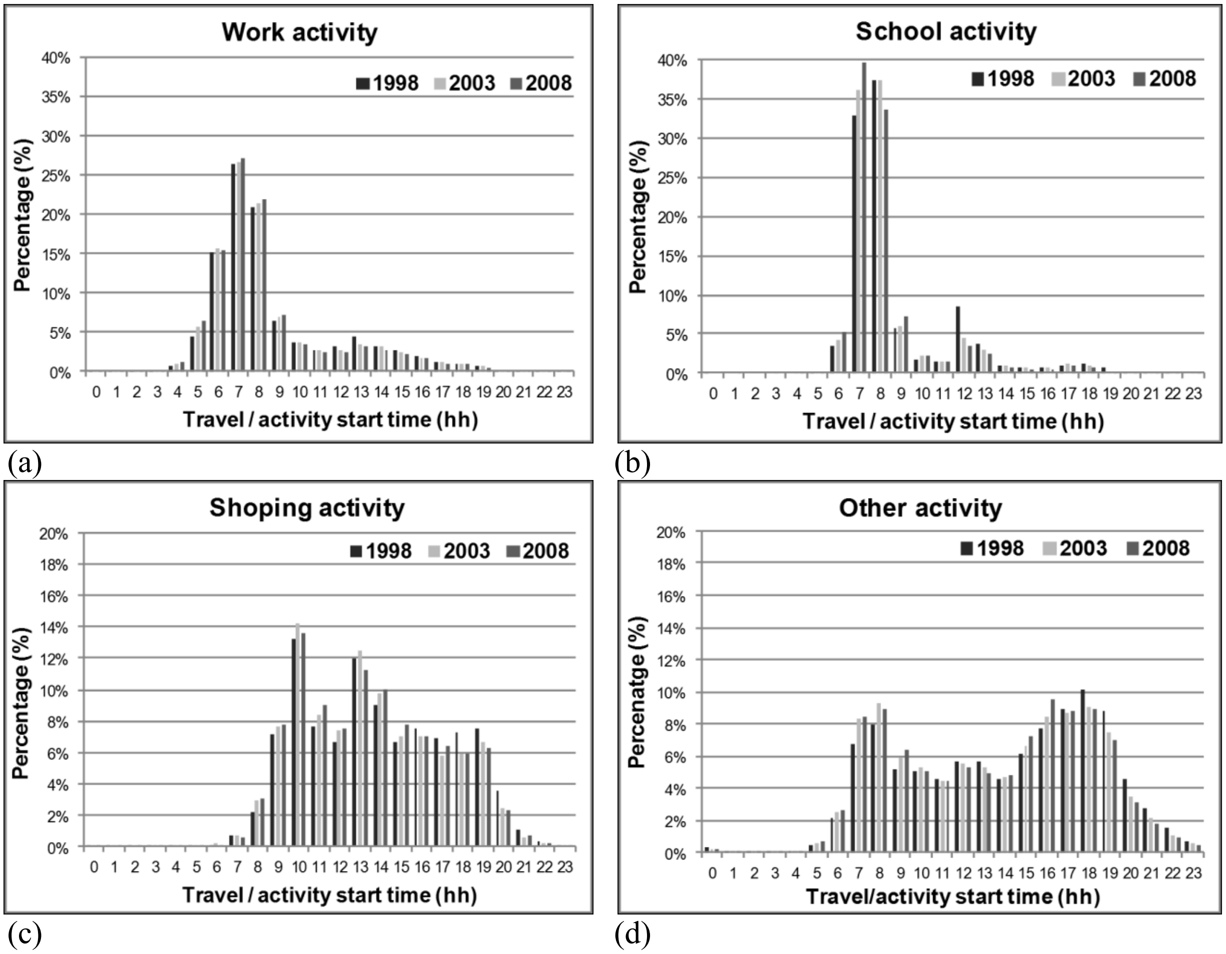


Figure 6.1: Travel/activity start time: a. Work activity, b. School activity, c. Shopping activity, and d. Other activity

Figure 6.1 (a) and (b) illustrate, as expected, that the morning hours are most common for work and school activities, whereas shopping activities are mostly associated with early afternoon to mid-day (Figure 6.1 (c)). Also, Figure 6.1 (d) presents the start time of other activity, which shows morning and evening peaks in the three surveys. It is interesting to note that work and school activities are showing a trend toward earlier trip start times. This fact indicates that travel/activity start times of work and school activity are being adjusted over time. In addition, fewer people (who are doing work/school activities) are returning home for lunch and/or doing other activities during lunch hour. This phenomenon leads to the change in the distributions of start times of both activities; there is a large reduction in trips starting just after lunch hour and an increase in the proportion of trips starting in the morning.

6.4.3 Duration of activities

This part of the analysis presents the average duration of work, school, shopping, and other activities for a typical weekday in 1998, 2003 and 2008. The null hypotheses that the duration distributions for each activity among the survey years are the same are rejected (Table 6.2). Figure 6.2 clearly shows that the average durations of work, school and shopping activities are increasing over time. This may be happening because people are spending longer hours at their activity locations (i.e. work, school and shopping) and/or their travel times have increased because of increasing urban congestion (Les Conseillers ADEC inc., 2009, 2014) and/or faster growing of suburban areas than the urban core. This study finds that returning home trips during lunch hour from work/school activities have decreased significantly over time; more people are doing one long work/school activity instead of 2 or more short activities. Population in low-density sub-urban areas in the GMA are also rapidly increasing over time, which increase travel distances (thus, travel times). The study confirms that in Montreal, average commuting distance is increasing over time; the average distance was 10.97 km in 1998 which has increased to 11.51 km in 2008.

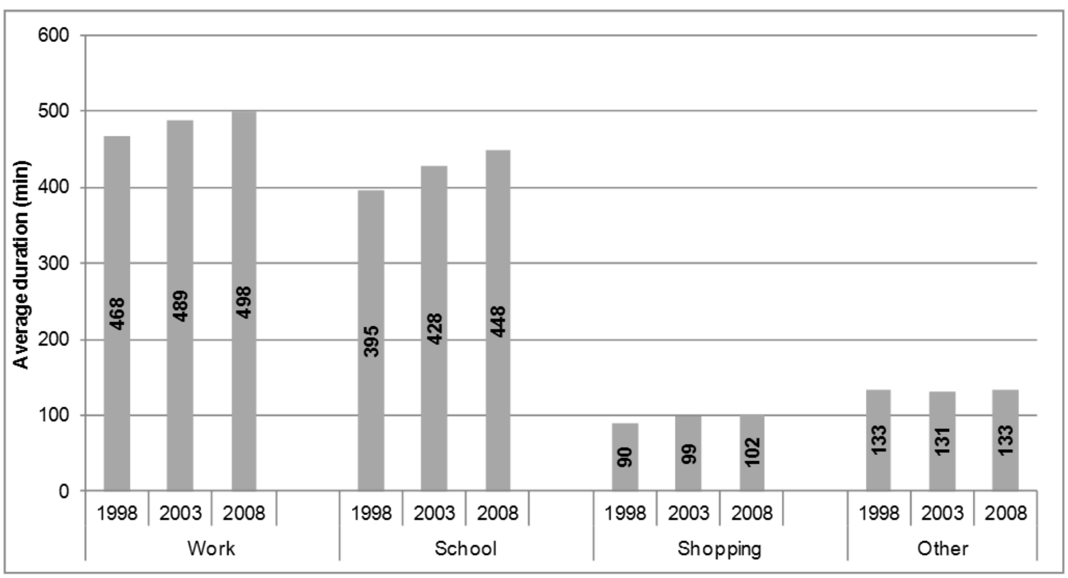


Figure 6.2: Average duration of activities

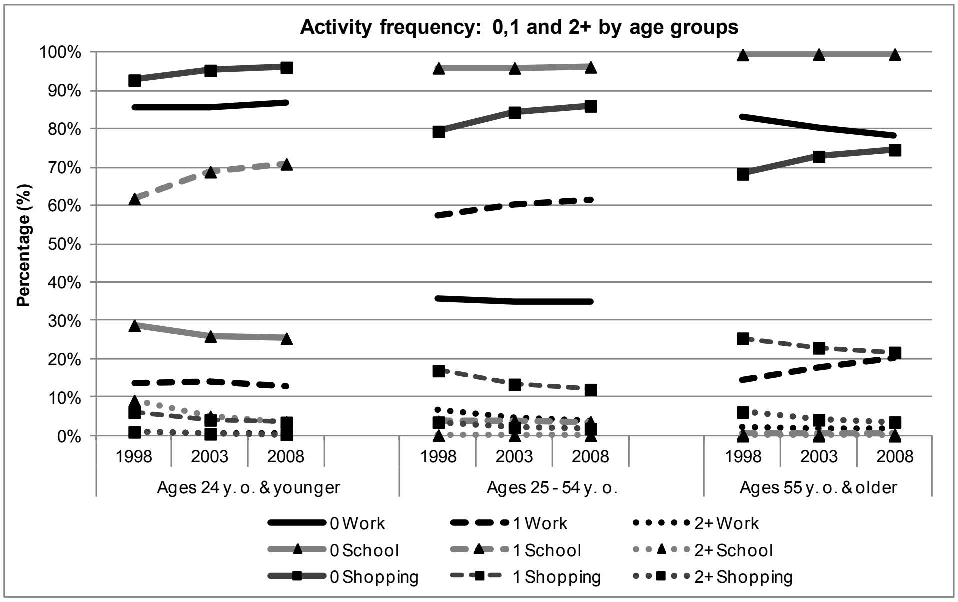
6.5 Major trends with socio-demographic characteristics

The previous analyses in this paper reveal some interesting trends of activity generation behaviour in the Greater Montreal Area. The subsequent sections investigate which socio-demographic cohorts are contributing to these trends. The analysis includes socio-demographic variables, such as age, gender, employment status, and student status. Table 6.2 indicates whether three activity attribute distributions between different groups of people among the surveys come from an identical distribution based on the statistical test (i.e. K-S test) results.

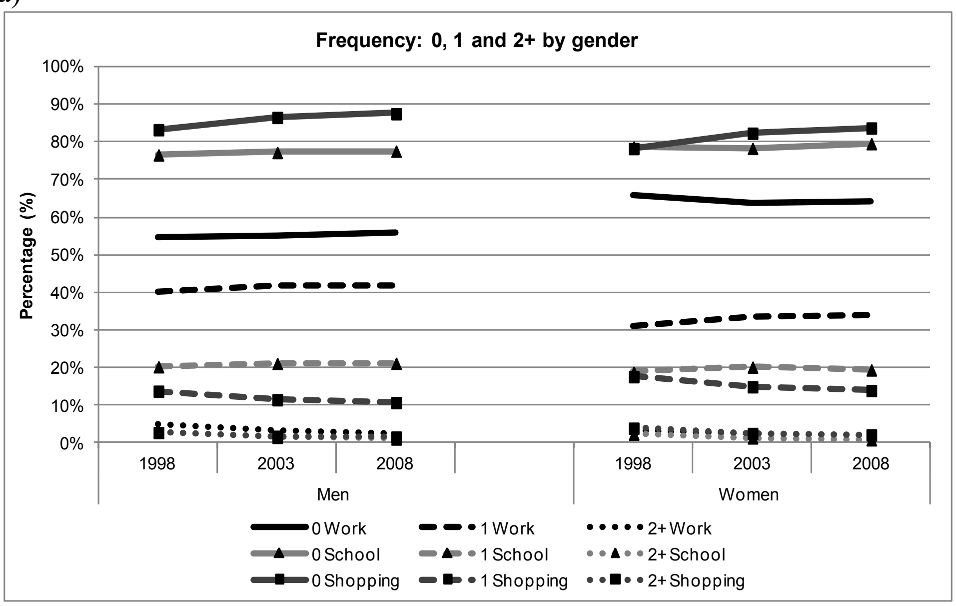
6.5.1 The pattern of work activity is changing over time

Analysis of activity frequency reveals that work activity patterns are changing over time (Table 6.3). Figure 6.3 (a) presents the frequency (0, 1 and 2+) of three activities (work, school, and shopping) by three age groups. It indicates that people who are 24 years old or younger are mostly performing school activities, whereas those between 25 and 54 years old are mostly conducting work activities in each survey. On the other hand, people aged 55 years old and older are most often performing shopping as well as other activities in each survey. The trend (i.e. frequency "0" decrease, "1" increase and "2+" decrease) has also been observed for people aged between 25 - 54 years old for work activities. The K-S test confirms that the frequency distributions of work activity of this age group between survey years come from different distributions (Table 6.2). The same trend occurs for people aged 55 years old and older; however the K-S test indicates that the distributions are different from year 1998 to 2003, but the test result is inconclusive for the distributions from year 2003 to 2008.

Figure 6.3 (b) shows the activity frequency (0, 1 and 2+) of activities by gender. K-S tests reveal that frequency distributions are different from one survey to another for both men and women except for shopping activity for men from year 2003 to 2008 (Table 6.2). The analysis illustrates that, as expected, male participation in work activities is still higher than female participation in each survey. However, female participation in work activities is rapidly increasing, while male participation shows a slight decrease over the 10-year period. Thus, we can conclude that women are contributing to the increasing overall work activity occurrence over time. This finding is



(a)



(b)

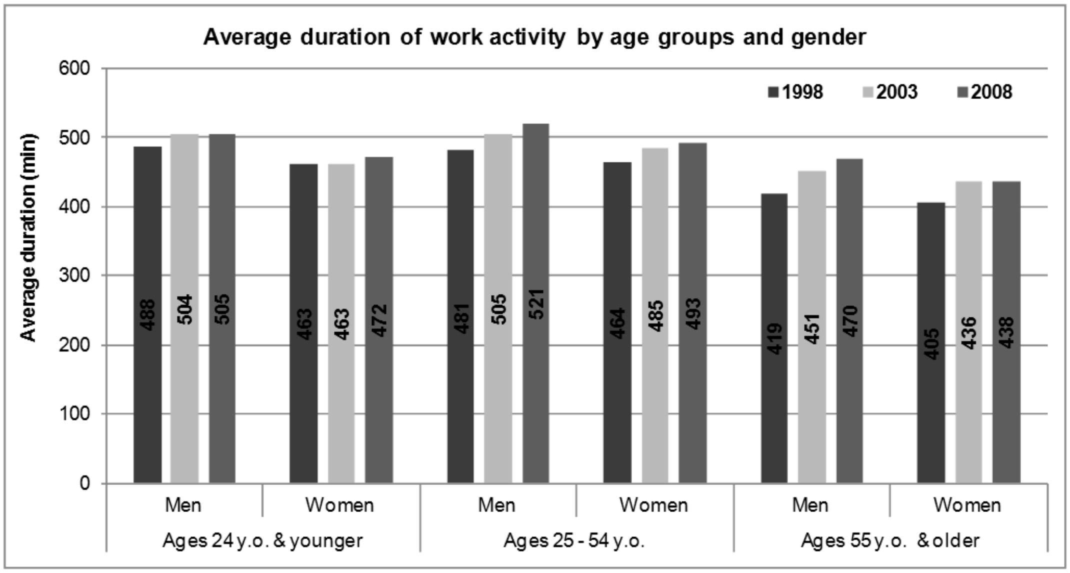
Figure 6.3: Activity frequency (0, 1 and 2+) by a. Age groups, b. Gender

consistent with a study that shows that female participation in the workforce has rapidly grown (from 63% in 1996 to 75% in 2011 for women aged 15-64 years old) since the launch of the universal low fee childcare program in Quebec in 1997 (Fortin et al., 2012; Statistics Canada, 2014). This low fee child care program initially cost a parent $5 per day, which was increased to $7 in 2004 and further increased to $7.5 in 2014 (Fortin et al., 2012; Ministère de la Famille, 2014). The program initially targeted only 4 year old children, but over the years after several changes in children age groups, it was opened to all preschool-age children (0-4 years old) in 2000.

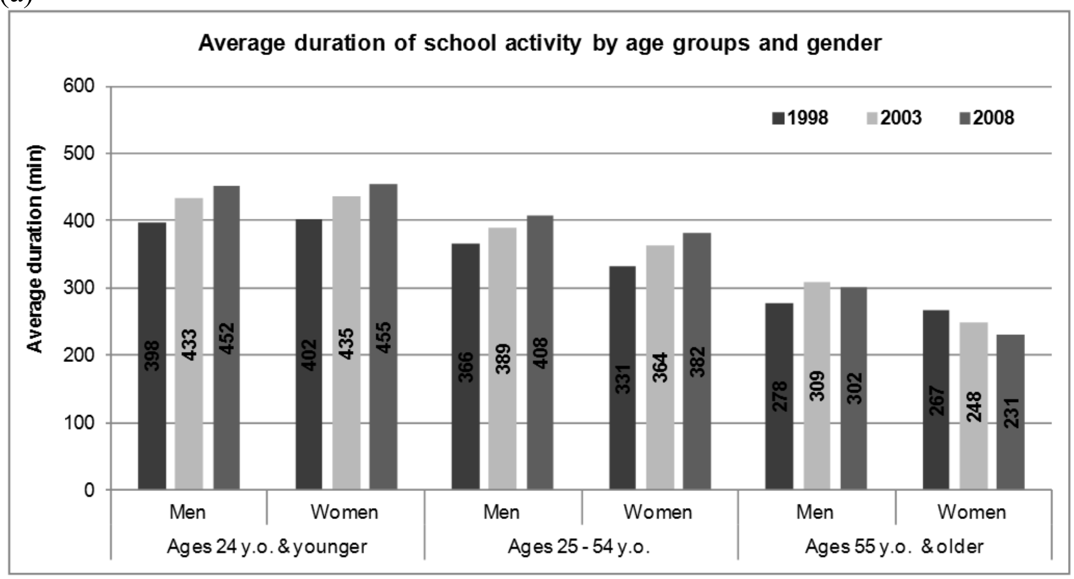
The proportions of both men and women doing at least one work activity are increasing while the proportion doing this activity more than once is decreasing over time. The increased proportion of doing at least one work activity is higher for women than men, however the decreased proportion of conducting this activity more than once in a day is higher for men than women. The study confirms that more men (1.6%) have stopped returning home and/or doing other activities during the lunch hour than women (1.1%). Also, it finds that average commute distances travelled by both men and women are increasing; however the distance travelled by men is higher than women in each survey. Women may be living closer to their workplace than men. Also, women may have more household responsibility (for instance taking care of children) than men.

6.5.2 Workers and students are spending longer hours travelling to and at work

We have seen earlier that people between 25 and 54 years old have the lowest share of work activity frequency "0 times" (Figure 6.3 (a)), and the average duration of work activity by this age group is increasing over time for both men and women (Figure 6.4 (a)). Also, for people aged 54 years old and older, the average work duration is increasing for both men and women over time. Figure 6.4 (a) also reveals that the average duration of work activity by men is increasing more than women, for all age groups. Figure 6.4 (b) presents the average duration of school activities by different age groups over the three time periods. People who are 24 years old and younger are mostly performing school activities (Figure 6.3 (a)), and Figure 6.4 (b) indicates that average duration of school activities by this age group is increasing over time for both men and women.



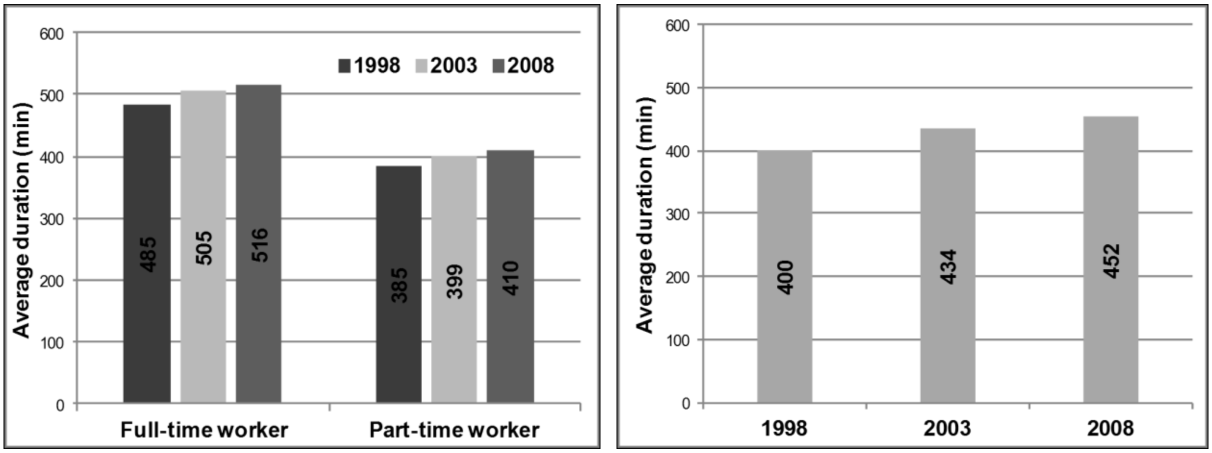
(a)



(b)

Figure 6.4: Average duration of a. Work activity, b. School activity by age groups and gender

Here, we also analyze work and school activity durations with employment and student status, which indicates that the average duration of work activity is increasing over time for both full-time and part-time workers (Figure 6.5 (a)). It also demonstrates that students are spending longer hours travelling to and at study location than before (Figure 6.5 (b)).



(a) (b)

Figure 6.5: Average duration of a. Work activity by employment status, b. School activity by student

Also, a portion of increasing duration of work and school activity is related to the increasing travel time due to urban congestion (Les Conseillers ADEC inc., 2009, 2014) and/or due to increasing commuting distance (for work activity) which is also related to the rapidly growing sub-urban population. In addition, fewer people are returning home during the lunch hour resulting in more single long work/school activity instead of two or more short activities.

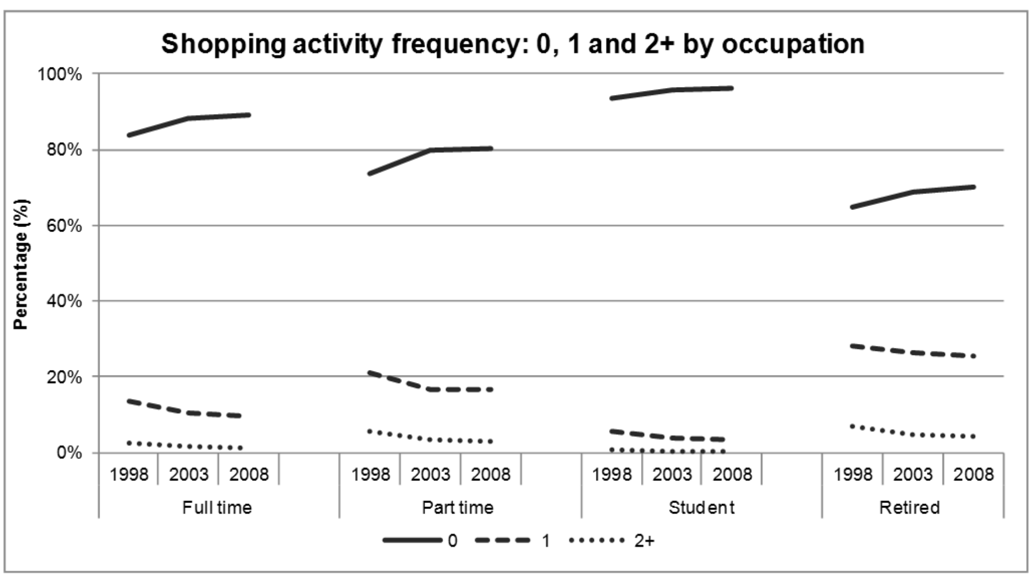


Figure 6.6: Shopping activity frequency: 0, 1 and 2+ by occupation

In addition, Figure 6.6 illustrates that the occurrence of shopping activities by full-time and part-time workers and students is decreasing over time in a typical weekday which may be a result of spending longer hours at work and study, respectively.

6.5.3 The nature of shopping activity is changing over time

Previous analyses in this paper reveal that overall the nature of the shopping activity is changing over time; the proportion of people conducting shopping activities on weekdays is decreasing over time (Table 6.3). Figure 6.3 (a) shows that this phenomenon of decreasing shopping activity is happening for all three age groups, however people aged 55 years old and older are participating in more shopping activities than others. Analysis illustrates that workers and students are spending longer hours at their activity locations (Figure 6.5 (a) and (b)), and their participation in shopping activity is decreasing in a typical weekday over time (Figure 6.6). It is possible that they are shopping less often in the weekdays and going shopping during weekends (Zhong et al., 2008). In addition, though the magnitude of online shopping in Canada is still relatively smaller than that of traditional shopping, it is growing over time (McKeown & Brocca, 2009).

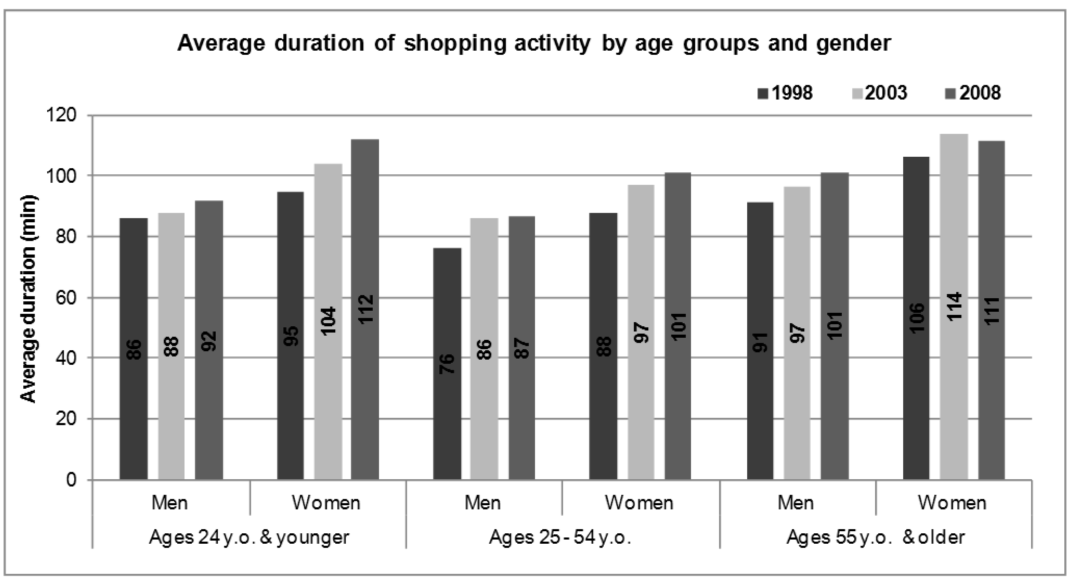
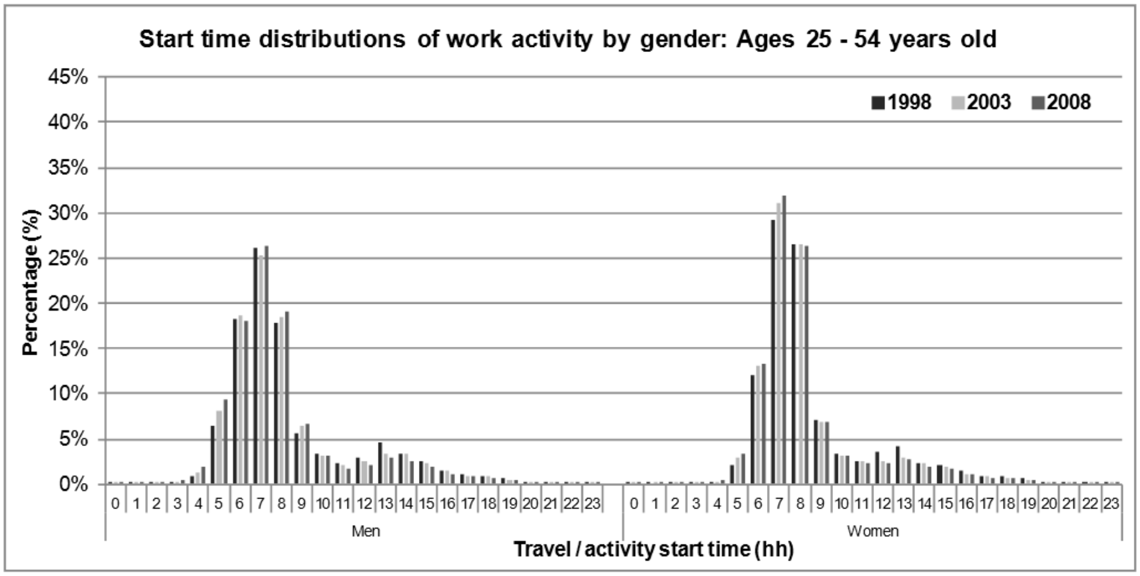


Figure 6.7: Average duration of shopping activity by age groups and gender

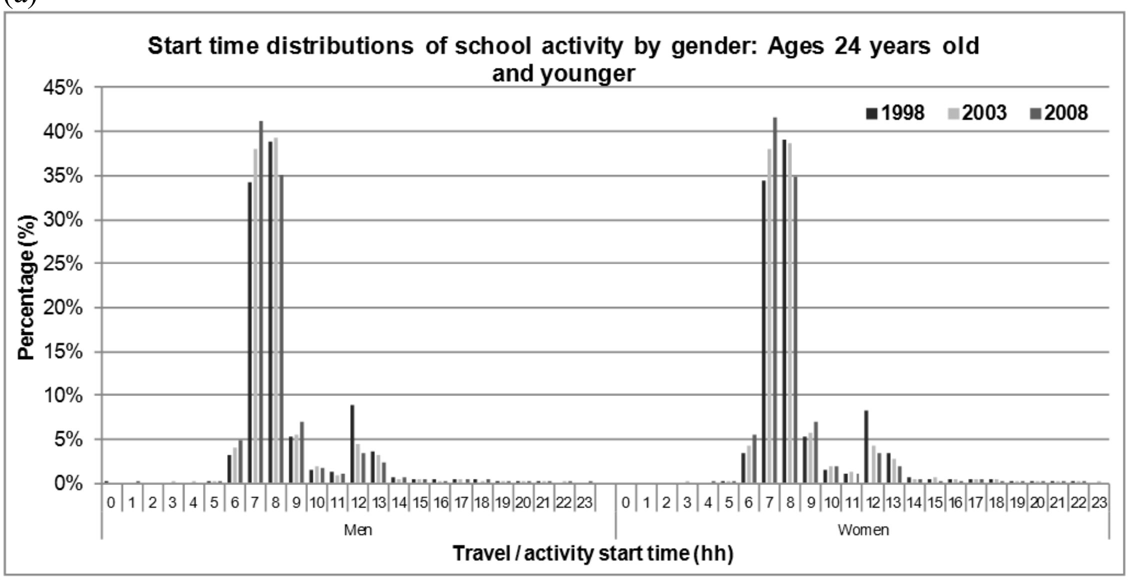
In addition, Figure 6.7 reveals that the average duration of shopping activity is increasing over time for all age groups and both genders. Because travel duration is included within the activity duration, a portion of increasing duration for shopping activity may be due to urban congestion or longer trip distances but is definitely less important than for work and study trips since shopping trips are typically not conducted during the peak periods. This study reveals that overall average trip distance for shopping activity has increased over this 10-year period in Montreal; however the average distance for this activity has increased from 4.20 km in 1998 to 4.63 km in 2003 but again decreased to 4.52 km in 2008. Hence, shopping frequency has decreased while duration has increased. This could be happening because shopping is consolidating into fewer places at fewer times. People may be shopping at power centres or big box retailers which decreases frequency but increases duration at the same time.

6.5.4 Travel/activity start times are being adjusted as a result of urban congestion

Previous analyses in this paper indicate a trend towards earlier start times of work and school activities. The reason behind this trend could be that people are starting their activities earlier to avoid urban congestion. Another reason could be that increasing travel time due to congestion (Les Conseillers ADEC inc., 2009, 2014) and/or faster growth of suburban areas than the urban core requires people to start their travel for work and school activities early. The study finds that the largest proportion of the employment is still agglomerated in the Montreal central areas (Montreal Downtown and Montreal Centre); however a significant proportion of population resides in medium- and low-density suburban areas.



(a)



(b)

Figure 6.8: Start time distributions of a. Work activity by gender (ages 25 - 54 years old), b. School activity by gender (ages 24 years old and younger)

Figure 6.8 (a) presents the start time distributions of work activity by people between 25 and 54 years old. While both men and women exhibit morning peak work activities, women start their work activities slightly later than men. Previous analysis in this research also reveals that both men and women are less often returning home and/or doing other activities during the lunch hour over time. Figure 6.8 (b) shows the start time distributions of school activity for the people aged 24 years old and younger, by gender. It illustrates a trend towards earlier start times for school activity by both men and women. This could alternatively be due to a greater effect of children on work start times for women 25 years and older, or could simply be due to reduced gender differences in the workforce for younger generations.

6.5.5 Gender differences are decreasing over time

Gender differences are decreasing for the work activity over the 10 year-period. Work activity occurrence for women has increased to become closer to that of men. Figure 6.3 (b) shows a large decrease of frequency “0 times” for women over the period of 10 years, but slight increase of frequency "0 times" for men. However, the average duration of work activities by men has increased more than that of women’s (Figure 6.4 (a)). Figure 6.4 (a) also reveals that the average duration of work activities by men is longer than women, while for shopping activity it shows the opposite result in each survey (Figure 6.7). For both men and women who are between 25 and 54 years old, most work activities are associated with morning hours, though women are starting their work activities slightly later than men (Figure 6.8 (a)).

6.6 Conclusions

This paper empirically examines changes in the distributions of activity generation attributes over three time points in a 10-year period (1998, 2003 and 2008) in the Greater Montreal Area (GMA). Also, it examines which activities and socio-demographic cohorts are experiencing the important changes. The analysis demonstrates that distributions of work, school, shopping and other activity frequency are significantly changing over this 10-year period. Specifically, work activity patterns are changing over time; more women are working, thus gender differences have decreased over time. Individuals aged 25 years and older are contributing to the changing pattern of work activity. The study finds an interesting trend of work and school activity i.e. frequency "1" has increased and frequency "2+" decreased over time. Workers, (especially more men than women) less often return home or do other activities during lunch hour. In the GMA, travel time has significantly increased due to urban congestion and/or increase in commuting distance for both men and women, resulting from rapidly growing sub-urban population over time; these might be the reasons for not coming home during lunch hour. Children are also less often returning home for lunch as there may be no adult at home, which might be related to the higher presence of women in the workforce. Shopping activity occurrence on weekdays is decreasing over time; it is decreasing for all three age groups, both for men and women. People might be shopping more often in weekends than weekdays; online shopping might also be another contributing factor in this reduction.

Start time and duration distributions for each activity have also changed significantly over time. Work and school activities starting in the morning hours are generally increasing. Travel/activity start times of work and school activity at morning hours are changing over time maybe to avoid urban congestion. Also, the reduction in returning home and/or doing other activities during the lunch hour has brought changes in the distributions of start times of both activities; it shows a large reduction in trips starting just after the lunch hour and increase in the proportion of trips starting in the morning. Average durations of work, school and shopping activity are also increasing over time, while durations have not changed for other activities. The longer hours that workers and students are spending travelling and working are accompanied by a reduced occurrence of shopping in a typical weekday. A portion of increasing durations of work and school activities are linked to increasing travel time due to urban congestion and/or increasing commuting distance which is also associated with rapidly growing suburban population. Also, as people are less often returning home and/or doing other activities during the lunch hour they are doing one long work/school activity instead of two or more short activities. Though occurrence of shopping has decreased, shopping duration has increased for all three age groups for both men and women. This may be due to the choice of shopping at larger consolidated shopping destinations (for example, big box shopping centres).

The scope of this research was to examine the trends of activity attributes distributions over time and which activities and socio-demographic cohorts are experiencing the important changes. However, to better understand some phenomena identified in this research, future research could investigate further, as follows:

- The work activity pattern in a typical weekday is changing over time. Thus, it would be useful to investigate differences in work activity attributes for each of the five days of the week, and how they are changing over time.
- The research reveals that workers/children are less often returning to home from work/school activities during lunch hour. Future research could investigate the reasons for this trend in detail.

- Changes have been found in shopping activity attributes in a typical weekday. Future research could examine differences in these attributes for each day of the week with specific attention to changing patterns of weekday vs. weekend shopping. It could also investigate shopping location choices (for instance big box shopping centres vs. small retailers).
- The activity type "other" includes seven types of activities; future research could analyze some of these activities (for instance leisure activity) separately.
- Future research could also observe intra-household interactions to investigate how the individuals of a household interact with each other to generate activities (for instance household maintenance activities).
- Activity attributes could be further investigated with respect to other variables including socio-demographic (for instance number of children and adults in a household), cultural (for instance country of birth), socio-economic (for instance income), and locational (for instance sub-urban vs. urban core) variables.

The study clearly reveals that the distributions of activity attributes are changing over time in the GMA; therefore attributes must be prepared for the activity-based modelling framework (i.e. TASHA) such that they reflect temporal changes in travel behaviour resulting from other changes in the GMA. It also demonstrates some interesting trends which could be beneficial to researchers, planners, and policy makers in transportation planning.

Acknowledgements

The authors would like to thank the Ministère des transports du Québec (Service de la modélisation des systèmes de transport) for their support and contributions. They also wish to thank the technical committee on O-D travel surveys in the Montreal region for providing access to data for research purposes.

References

Agence Métropolitaine de Transport. (2015). Portrait de la mobilité. Retrieved January 14, 2015, from https://www.amt.qc.ca/fr/a-propos/portrait-mobilite

Arentze, T., & Timmermans, H. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological, 38*(7), 613-633. doi: 10.1016/j.trb.2002.10.001

Arentze, T., & Timmermans, H. (2009). A need-based model of multi-day, multi-person activity generation. *Transportation Research Part B: Methodological, 43*(2), 251-265. doi: 10.1016/j.trb.2008.05.007

Bhat, C. R., Guo, J. Y., Srinivasan, S., & Sivakumar, A. (2004). Comprehensive econometric microsimulator for daily activity-travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1894*, 57-66. doi: 10.3141/1894-07

Bhat, C. R., & Koppelman, F. S. (1993). A conceptual framework of individual activity program generation. *Transportation Research Part A: Policy and Practice, 27*(6), 433-446.

Bonnel, P., Lee-Gosselin, M., Madre, J.-L., & Zmud, J. (2009). Keeping up with a changing world: Challenges in the design of transport survey methods *Transport Survey Methods: Keeping Up with a Changing World* (pp. 3-13): Emerald.

Bourbonnais, P.-L., & Morency, C. (2013). Web-Based Travel Survey: A Demo. In J. Zmud, M. Lee-Gosselin & J. A. Carrasco (Eds.), *Transport Survey Methods: Best practice for decision making* (pp. 207-223). United Kingdom Emerald Group Publishing Limited, Bingley.

Bowman, J. L. (2008). *The day activity schedule approach of Bowman, Ben-Akiva and Bradley: 1994-2008*. Paper presented at the Innovations in Travel Modeling 2008: A Transportation Research Board Conference, Portland, Oregon, USA.

Bowman, J. L., & Ben-Akiva, M. E. (2000). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice, 35*(1), 1-28.

Bradley, M., Bowman, J. L., & Griesenbeck, B. (2010). SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution. *Journal of Choice Modelling, 3*(1), 5-31. doi: http://dx.doi.org/10.1016/S1755-5345(13)70027-7

Communauté Métropolitaine de Montréal. (2010). Portrait of Greater Montreal, 2010 edition (Vol. 1). Quebec, Canada: Policy and Interventions Department, Communauté Métropolitaine de Montréal.

Cottrill, C. D., Pereira, F. C., Zhao, F., Dias, I. F., Lim, H. B., Ben-Akiva, M. E., & Zegras, P. C. (2013). Future mobility survey: Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board, 2354*, 59-67.

Davidson, W., Vovsha, P., Freedman, J., & R., D. (2010, September 29 – October 1, 2010). *CT-RAMP family of activity-based models.* Paper presented at the Australasian Transport Research Forum, Canberra, Australia.

Ettema, D., Bergers, A., & Timmermans, H. (1993). Simulation model of activity scheduling behaviour. *Transportation Research Record: Journal of the Transportation Research Board, 1413*, 1-11.

Fortin, P., Godbout, L., & St-cerny, S. (2012). Impact of Quebec's universal low fee childcare program on female labour force participation, domestic income, and government budgets. Quebec, Canada: The Research Chair in Taxation and Public Finance at the University of Sherbrooke.

Habib, K. M. N., & Miller, E. J. (2008). Modeling daily activity program generation considering within-day and day-to-day dynamics in activity-travel behaviour. *Transportation, 35*(4), 467-484. doi: 10.1007/s11116-008-9166-8

Habib, K. M. N., & Miller, E. J. (2009). Modelling activity generation: A utility based model for activity agenda formation. *Transportmetrica, 5*(1), 3-23. doi: 10.1080/18128600802591384

Jones, P. M., Dix, M. C., Clarke, M. I., & Heggie, I. G. (1983). *Understanding travel behaviour*. Aldershot, UK: Gower.

Kitamura, R. (1996). *Applications of models of activity behavior for activity-based demand forecasting*. Paper presented at the Activity-Based Travel Forecasting Conference, New Orleans, Louisiana.

Kitamura, R., & Fujii, S. (1998). Two computational process models of activity-travel behavior. In T. Garling, T. Laitila & K. Westin (Eds.), *Theoretical Foundations of Travel Choice Modeling* (pp. 251-279). Bingley, UK: Emerald Group Publishing.

Kitamura, R., Lula, C. V., & Pas, E. I. *AMOS: An activity-based, flexible and truly behavioral tool for evaluation of TDM measures.* Paper presented at the Proceedings of the 21st Summer Annual Meeting: Transportation Planning Methods, PTRC Education and Research Services, Ltd., England.

Kitamura, R., Pas, E. I., Lula, C. V., Lawton, T. K., & Benson, P. E. (1996). The sequenced activity mobility simulator (SAMS): An integrated approach to modeling transportation, land use and air quality. *Transportation, 23*(3), 267-291. doi: 10.1007/BF00165705

Les Conseillers ADEC inc. (2009). Évaluation des coûts de la congestion routière dans la région de Montréal pour les conditions de référence de 2003. Montréal, Québec: Ministère des Transports du Québec.

Les Conseillers ADEC inc. (2014). Évaluation des coûts de la congestion routière dans la région de Montréal pour les conditions de référence de 2008. Montréal, Québec: Ministère des Transports du Québec.

McKeown, L., & Brocca, J. (2009). Internet shopping in Canada: An examination of data, trends and patterns *Business Special Surveys and Technology Statistics Division Working Papers*. Ottawa, Canada: Business Special Surveys and Technology Statistics Division, Statistics Canada.

Meeks, W. S., Sabina, E., Childress, S., Van Slyke, C., Mullins III, J. A., Ziering, E., & Rossi, T. F. (2013). *TourCast: Trade-offs in Activity-Based Model Implementations*. Paper presented at the 92nd Transportation Research Board (TRB) Annual Meeting, Washington D.C., USA.

Miller, E. J., & Roorda, M. J. (2003). Prototype model of household activity-travel scheduling. *Transportation Research Record: Journal of the Transportation Research Board, 1831*, 114-121. doi: 10.3141/1831-13

Ministère de la Famille. (2014). Programme de places à contribution réduite. Retrieved Februay 11, 2015, from http://www.mfa.gouv.qc.ca/fr/services-de-garde/parents/programme-contribution-reduite/pages/index.aspx

Pendyala, R. M., Kitamura, R., Kikuchi, A., Yamamoto, T., & Fujii, S. (2005). Florida Activity Mobility Simulator: Overview and preliminary validation results. *Transportation Research Record: Journal of the Transportation Research Board, 1921*, 123-130. doi: 10.3141/1921-14

Recker, W. W., McNally, M. G., & Root, G. S. (1986a). A model of complex travel behavior: Part I--Theoretical development. *Transportation Research Part A: General, 20*(4), 307-318. doi: 10.1016/0191-2607(86)90089-0

Recker, W. W., McNally, M. G., & Root, G. S. (1986b). A model of complex travel behavior: Part II--An operational model. *Transportation Research Part A: General, 20*(4), 319-330. doi: 10.1016/0191-2607(86)90090-7

Roorda, M. J., Miller, E. J., & Habib, K. M. N. (2008). Validation of TASHA: A 24-h activity scheduling microsimulation model. *Transportation Research Part A: Policy and Practice, 42*, 360-375. doi: 10.1016/j.tra.2007.10.004

Roorda, M. J., Morency, C., & Woo, K. (2008). *Two cities, two realities? A closer look at the evolution of trip rates in Toronto and Montreal*. Paper presented at the 87th Transportation Research Board (TRB) Annual Meeting, Washington D.C., USA.

Shiftan, Y., Ben-Akiva, M., Proussaloglou, K., Jong, G. d., Popuri, Y., Kasturirangan, K., & Bekhor, S. (2003). *Activity-based modeling as a tool for better understanding travel behaviour*. Paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Statistics Canada. (2011a). Population and dwelling counts, for Canada and census divisions. *Canada 2011 Census.* Retrieved July 30, 2013, from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=701

Statistics Canada. (2011b). Population and dwelling counts, for census metropolitan areas. *Canada 2011 Census.* Retrieved February 27, 2014, from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=205&S=3&RPP=50

Statistics Canada. (2014). CANSIM Table 282-0002. Labour force survey estimates (LFS), by sex and detailed age group, annual (Persons), 1976 to 2014. *CANSIM (database).* Retrieved February 10, 2015, from http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=2820002&tabMode=dataTable&srchLan=-1&p1=-1&p2=9

Stopher, P. R., Prasad, C., Wargelin, L., & Minser, J. (2013). Conducting a GPS-only household travel survey In J. Zmud, M. Lee-Gosselin & J. A. Carrasco (Eds.), *Transport Survey Methods: Best Practice for Decision Making* (pp. 91-113). United Kingdom: Emerald Group Publishing Limited, Bingley.

Veldhuisen, J., Timmermans, H., & Kapoen, L. (2000). Microsimulation model of activity-travel patterns and traffic flows: Specification, validation tests, and Monte Carlo error. *Transportation Research Record: Journal of the Transportation Research Board, 1706*, 126-135.

Vovsha, P., Petersen, E., & Donnelly, R. (2004). Impact of intra-household interactions on individual daily activity travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1898*, 87-97.

Yasmin, F., Morency, C., & Roorda, M. J. (2012). *An empirical analysis of change in activity attribute distributions over time*. Paper presented at the 13th International Conference on Travel Behaviour Research organized by the International Association for Travel Behavior Research (IATBR), Toronto, Canada.

Yasmin, F., Morency, C., & Roorda, M. J. (2014). *Macro, meso, and micro-level validation of an activity-based travel demand model*. Paper presented at the 93rd Transportation Research Board (TRB) Annual Meeting, Washington D.C., USA.

Yasmin, F., Morency, C., & Roorda, M. J. (2015). Assessment of spatial transferability of an activity-based model, TASHA. *forthcoming Transportation Research Part A: Policy and Practice, 78C*. doi: 10.1016/j.tra.2015.05.008

Zhong, M., Hunt, J. D., & Lu, X. (2008). Studying differences of household weekday and weekend activities: A duration perspective. *Transportation Research Record: Journal of the Transportation Research Board, 2054*, 28-36. doi: 10.3141/2054-04

Zmud, J., Lee-Gosselin, M., Munizaga, M. A., & Carrasco, J. A. (2013). Transport surveys: Considerations for decision makers and decision making. In J. Zmud, M. Lee-Gosselin & J. Carrasco (Eds.), *Transport Survey Methods: Best practice for decision making* (pp. 3-19). United Kingdom Emerald Group Publishing Limited, Bingley.

CHAPTER 7 POPULATION SEGMENTATION BASED ON SIMILARITY IN ACTIVITY PATTERNS

7.1 Overview

Population segmentation to develop activity attributes distributions in TASHA, an activity-based model, was based on intuition and testing, it was not based on any systematic segmentation process. This chapter focuses on how to systematically obtain population segments based on similarity in individuals' activity patterns (represented by three activity attributes, activity frequency, start time, and duration) to develop activity attribute distributions for an activity-based model. More specifically, it proposes a methodology combining a clustering approach (multiple sequence alignment method (SAM)) and discrete choice method (multinomial logit model (MNL)) to segment the population of the Island of Montreal, Canada based on the individual's daily activity patterns and their socio-demographic characteristics. This research uses the 2003 Origin-Destination (O-D) travel survey of Montreal.

The remainder of this chapter is organized as follows. First, a brief review of the related literature is provided. Next, the proposed methodology of population segmentation is presented. Then, data and research method are described. Next, cluster analysis and modelling results are presented. The chapter concludes with a summary of the research and a discussion of the implications of the study.

7.2 Background

Activity-based models are recognized as a more powerful framework in capturing individuals' activity behaviours than the traditional four-stage models (Bhat & Lawton, 2000; Kitamura, 1996; Kitamura et al., 1996; Roorda, 2005; Shiftan et al., 2003). However, the application of the activity-based models is still relatively less common in practice than traditional models (Arentze & Timmermans, 2008; Mohammadian et al., 2009; Shiftan et al., 2003). But the activity-based models are also coming into practice in various parts of North America in recent years (Bradley & Bowman, 2006; Davidson et al., 2010; Miller et al., 2015; Shiftan & Ben-Akiva, 2011; Vovsha, Bradley, & Bowman, 2004). To increase the real world application of such models, there must be extensive validation and testing prior to application to examine their capability of reproducing activity behaviours. Validation of activity-based models has been done using different methods; a review of the validation methods can be found in other papers (Yasmin et al., 2015a; Yasmin et al., submitted).

Among others, assessment of spatial transferability of the models has been recognized as a useful validation test; however to date, validation of activity-based models using such assessment is still infrequent (Arentze et al., 2002; Auld & Mohammadian, 2012; Bowman et al., 2014; Yasmin et al., 2015a; Yasmin et al., submitted). There are two advantages of examining the spatial transferability of a model by applying the model in different contexts. First, it provides opportunity to validate the model by assessing its spatial transferability. Second, if the model is spatially transferable to another context; it provides opportunity to implement the model in that context which will save time, cost and expertise needed to develop a new model for that context. Thus, this approach may facilitate increasing the practical application of the activity-based models.

As part of a broad research project focusing on the enhancement of current process of modelling of travel demand of Montreal using an activity-based approach, the activity-based model TASHA (Travel Activity Scheduler for Household Agents (Miller & Roorda, 2003)) has been transferred to the Montreal Island using the 2003 O-D travel survey (Yasmin et al., 2015a; Yasmin et al., submitted). We have examined the spatial transferability of the TASHA model and indicated that the TASHA model was quite transferable, at least in case of fixed activities (for instance work and school). However, it was less successful in reproducing some activity attributes for flexible activities (for instance shopping, and other). A detailed discussion of the potential reasons behind the large differences found between the simulated and the observed attributes in some cases for flexible activities can be found in other papers (Yasmin et al., 2015a; Yasmin et al., submitted).

Activity generation in TASHA was based on observed distributions of activity attributes (frequency, start time and duration) from the travel survey data (i.e. Transportation Tomorrow Survey (TTS) of the Toronto Area) of the year 1996 (Miller & Roorda, 2003; Roorda et al., 2008). A total of 262 distributions were developed from the observed data, cross-classified by activity type, person, household and schedule attributes (Roorda, 2005). For example, the segment of full-time workers, aged between 26-64 years old has a fixed distributions of activity frequency, start time and duration. The population segmentation for each activity type was based on intuition and testing; the selection of attributes to generate activities was not through a systematic process and those segments are fixed over time. Therefore, this research proposes a methodology to segment the population based on the similarities in individuals' daily activity patterns to develop the activity attributes distributions for an activity-based model (such as, TASHA).

7.2.1 Earlier research on population segmentation in travel demand modelling

The importance of population segmentation in travel demand modelling has already been well recognized (Badoe & Miller, 1998; Páez, 2006). Thus, travel demand analysis by different segments has become a common practice. However, *a priori* segmentation and the latent class models (LCM) are common (Badoe & Miller, 1998; Ishaq, Bekhor, & Shiftan, 2012a).

Using *a priori* segmentation process, there can only be a fixed, finite number of mutually exclusive segments. It is assumed that individuals in each segment have similar choice preferences. This segmentation process is easy to implement (Bhat & Koppelman, 1999a), but it has its limitations (Badoe & Miller, 1998).

The latent class model, (LCM) is another way of obtaining the segments, which integrates the discrete choice models and segmentation process. LCM groups the population who tend to have similar preferences and then estimates the utility for each segment. This modelling approach shows a significant improvement over other discrete choice models (Ishaq, Bekhor, & Shiftan, 2012b).

Traditional four-stage models generally segment the population based on one variable, i.e. trip-purpose assuming that different types of trips have different motivations (McNally, 2007b; Ortúzar & Willumsen, 2011). However, there is no theoretical evidence that segmenting the population based on trip purpose provides maximum similarity and dissimilarity within and among the segments, respectively (Ishaq et al., 2012b). In modelling travel demand, *a priori* segmentation is also done based on one or two key socio-demographic variables (for instance, gender, income, automobile ownership) (Bhat & Koppelman, 1999a), and sometimes using several other variables, geographical location, other trip characteristics, individual's attitudinal statements, or life cycle features (Golob, 2001; Lieberman, Schumacher, Hoffman, & Wornum, 2001; Mandel, 1998; Ryley, 2006; Shiftan, Outwater, & Zhou, 2008).

Activity-based models are also often estimated for the entire population or pre-defined segments (Ishaq et al., 2012a). Some examples of such activity-based models are the Portland model (Bowman, Bradley, Shiftan, Lawton, & Ben-Akiva, 1999), the San Francisco model (Jonnalagadda, Freedman, Davidson, & Hunt, 2001), the Florida model (Pendyala et al., 2005), the Jakarta model (Yagi & Mohammadian, 2010), the Atlanta model (Bradley & Vovsha, 2005), and the Tel-Aviv model (Shiftan & Ben-Akiva, 2011). However, in activity-based modelling, a systematic segmentation of population, other than LCM, using the wide range of existing segmentation methods (thus clustering) is still rare (Ishaq et al., 2012a).

A recent study (Ishaq et al., 2012a) has developed an integrated methodological framework (defined as the flexible model structure (FMS)),which relies on an optimization algorithm that segment given data and searches for the best model structure for each segment simultaneously. Three models have been estimated, two with a priori segmentation and nested logit model (NL) structure, and one with fuzzy segmentation method and more than one model structure (i.e. FMS). This experimentation has indicated a significant improvement in the estimation results for the flexible model structure compared to the other two approaches. However, the FMS estimation was based on fewer explanatory variables (gender, number of cars per household, number of persons per household, car and bus travel time). Also, this model structure has only focused on the two-dimensional travel choices (destination and mode choice) of a multidimensional activity-based model.

In the activity-based model TASHA, population segmentation was done by cross-classification of a set of variables such as activity type, person, household and schedule attributes. This segmentation process was based on intuition and testing. It may improve model performance if we systematically segment the population using similarity in multidimensional activity patterns (frequency, start time, duration) and individual's socio-demographic characteristics (such as age groups, and gender) to develop activity attributes distributions for an activity-based model.

Clustering based on similarities in activity patterns has gained attention in activity behaviour analysis. However, activity patterns are typically compared using conventional Euclidean distance measures. But, an activity pattern, represented as a sequential order of activities, requires a measure that can capture the sequential information and interdependency. The sequence alignment method (SAM) seems quite promising in clustering of activity patterns (Joh et al., 2001; Wilson, 1998). SAM is useful, because it allows using the whole activity pattern as a unit of analysis (Joh et al., 2001).

SAM, also known as optimal matching, provides a quantitative measure of distance or similarity between two character sequences. This method is widely used in molecular biology since the 1970 (Needleman & Wunsch, 1970). However, Wilson (1998) has introduced this approach into activity pattern analysis in 1998 and also improved the sequence alignment procedure (Wilson, 1999). Multidimensional alignment is also a useful development (Joh et al., 2002; Joh et al., 2001). Wilson (1999) has developed a new software package ClustalG, an improved version of Clustal software series (i.e. general format ClustalX), which can define an activity episode with up to twelve activity attributes (such as activity type, location, with whom and so on). However, Schlich (2003) has indicated that inclusion of more attributes or dimensions in similarity measuring procedure reduces the observed similarities. Several empirical applications of this method in activity pattern analysis can be found in other papers (Saneinejad & Roorda, 2009; Schlich, 2001; Wilson, 1999). In these papers, the similarity measurement of activity patterns is based on attributes, such as activity type, activity start and end time, trip or activity duration, activity location, and transport mode.

Fundamentally, activity behaviour analysis assumes that spatial and/or socio-demographic characteristics of individuals and households are systematically related to their activity patterns (Hanson, 1982; Kuppam & Pendyala; Pas, 1984). Thus, the goal of this proposed methodology is to provide a systematic procedure to segment the population based on both activity patterns and socio-demographic characteristics of the individuals and households.

7.3 Methodology

This research uses a methodology to systematically segment the population by combining two methods, the multiple sequence alignment method (SAM) as a clustering approach and the multinomial logit model (MNL). It first applies the multiple SAM to segment the population based on the similarities in individuals' daily activity patterns (represented by three activity attributes, frequency, start time and duration) and then estimate the multinomial logit models for the distinct segments of individuals based on their socio-demographic characteristics. A detailed description of the multiple SAM and MNL model can be found in Section 2.6.1 and Section 2.6.2 (Chapter 2 - Literature review), respectively.

7.4 Data and research method

This research demonstrates the methodology using the 2003 O-D travel survey of Montreal. Large scale O-D travel surveys are conducted every five years since the early 1970s in this region. The survey collects detailed travel and socio-economic information of approximately 5% of the total Greater Montreal Area (GMA) population aged 5 years and older. The travel data of all household members are collected for a specific weekday of the fall period (September to December) by a telephone interview (Trépanier et al., 2008). Detailed information on the O-D travel surveys can be found on the AMT website (Agence Métropolitaine de Transport [AMT], 2015).

This research applies the methodology to segment the population of the Island of Montreal, which is composed of an area of 500 square km and a population of 1,886,000 (Statistics Canada, 2011). A detailed methodological framework of this research is presented in Figure 7.1.

During data preparation, open chains (i.e. trip chains that did not start and end at home) have been observed for some individuals in the 2003 O-D travel survey. Households with individuals that made open trip chains have been excluded. The cluster analysis and model estimation have been done based on individuals' daily activity patterns represented by three activity attributes (activity frequency, start time and duration) and socio-demographic characteristics. The O-D survey supplies a trip database, thus it has been processed to transform them in to an activity database.

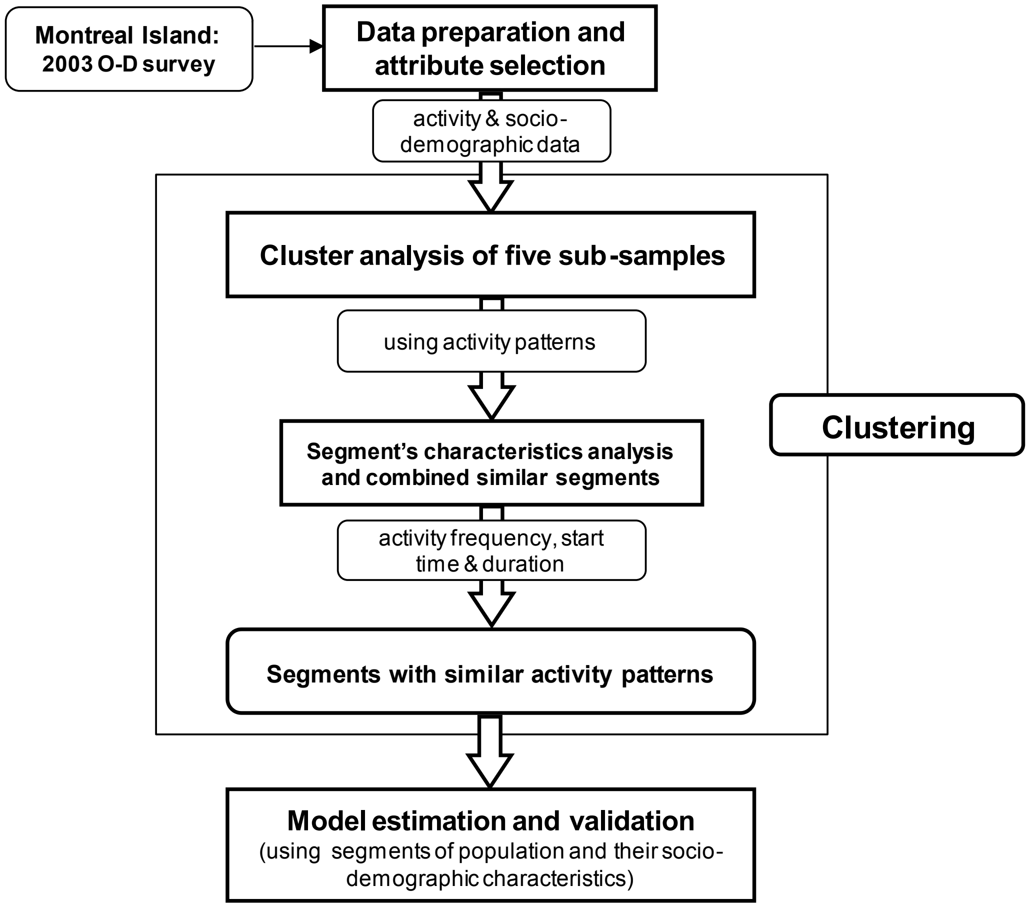


Figure 7.1: Methodological framework of population segmentation

Activity type has been defined from the destination purpose of a trip observation; the surveys collect trip data of thirteen types of activities, which have been aggregated into five broad activity classes according to the commonly used patterns in Montreal (work, school, shopping, other, and return to home). The "other" activity type includes en route, leisure, visit friends/family, health, drop someone off, pick someone up, and other. Activity start time has been computed by using trip start time and travel time information for each trip. However, travel time information is not available in the survey. Travel times for auto trips are collected from the Ministère des transports du Québec (MTQ). Travel times for trips conducted by other modes (for instance, public transit, cycle, walk and so on) are imputed using different methods. For example, travel times for transit trips are imputed based on the travel times estimated using the current network, and the ratio between the travel times for peak period observations for the year 2003 (estimated by the MTQ using an assignment model) and travel times estimated using the current network (by using the Connection Scan Algorithm (CSA) (Dibbelt, Pajor, Strasser, & Wagner, 2013)) for the same O-D pairs. Total duration (including travel and activity duration) has been calculated using successive trip start times, and then activity duration has been calculated by subtracting travel duration from total duration.

After data preparation and cleaning, there were in total 58052 individuals in the dataset of the Island of Montreal. The analysis in this research proceeds in two steps. First it conducts a cluster analysis using the sequence alignment method based on individual's activity patterns. Second, it estimates the multinomial logit models for the distinct segments, obtained from the cluster analysis, based on individual's socio-demographic characteristics.

First, in order to conduct the cluster analysis, five sub-samples with 2000 individuals in each sub-sample (total 10000 individuals) have been randomly selected from the dataset. In this research, sequence alignment method is applied in one dimension with activity type (work - W, school - S, shopping - M, other - O and home - H). Activity start time and duration are also represented by dividing the 24-hour typical weekday into 30 minutes time intervals. Table 7.1 presents a sample dataset showing the daily activity patterns (thus, character sequences) of ten individuals. Cluster analysis has been done for each sub-sample (i.e. 2000 sequences) using multiple sequence alignment method in ClustalG (Wilson, 1999). This research considers the default sequence alignment parameters as recommended by Wilson, Harvey, and Thompson (2005). The parameters are 10 points for matches, 0 points for mismatches, 8 points for gap insertion and 3 points for gap extension. Population segments in each sub-sample are identified by observing several elements, groups suggested by the tree, group score from the log file, and aligned sequences. Then, segments of each sub-sample have been analyzed based on activity attributes (frequency, start time and average durations). Next, segments from five sub-samples have been combined based on their similar activity attributes. Final segments are also analyzed with their activity attributes. A descriptive analysis of different socio-demographic characterictis of each segment has also been conducted. The socio-demographic variables include age, gender, employment status, possession of driver's license, vehicle ownership, presence of children, household size, and home location regions.

Finally, distinct segments are modelled. Among a total sample of 9292 (excluding *Outlier*), 7434 (80% of the sample) and 1858 (20% of the sample) activity patterns have been randomly selected for multinomial logit model estimation and validation, respectively. Model estimation has been performed using the BIOGEME software (Bierlaire, 2015).

Table 7.1: A sample data of daily activity patterns for cluster analysis

Individuals	Daily activity patterns (sequences)
1	HHHHHHWWWWWWWWWWWWWWWWWWWWWWWHHHHHHHHHHHHHHHHHHHH
2	HHHHHHHHHHHWWWWWWWWWWWWWWWHHHSSSSSSSHHHHHHHHHHHHH
3	HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
4	HHHHHHHHHHHHHMMMMMMMHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
5	HHHHHHHHHHHHHOOOOOOOOHHHHHHHHHHHHHHHHHHHHHHHHHHHH
6	HHHHHHHHHWWWWWWWWWOOWWWWWWWWOOOOOOOOOOOOOOOHHHHHH
7	HHHHHHHHHHSSSSSSSSSHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
8	HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
9	HHHHHWWWWWWWWWWWWWWWWWWWWWWWWHOOOOOOHHHHHHHHHHHHH
10	HHHHHHHHHHSSSSSSSSSSSSSSSSSSSSSSSSSSSHHHHHHHHHHHH

7.5 Empirical application

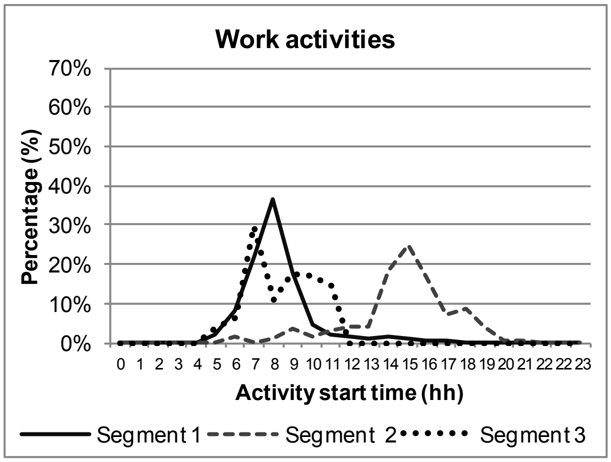
7.5.1 Clustering population based on similarities in activity patterns

Individuals are grouped based on the similarities in their daily activity patterns represented by the activity attributes (activity frequency, start time, and duration) for five broad activity types (work, school, shopping, other, and home). Cluster analysis using the multiple sequence alignment method has been done for five sub-samples separately due to limitation in computer capacity. After combining the segments from five sub-samples based on similar characteristics in each sub-sample, eleven distinct segments (Segment 1 - 11) are identified for 10000 activity sequences. The activity attributes (frequency, start time and duration), which represented the daily activity patterns are analyzed for each segment, as shown in Table 7.2 and Figure 7.2.

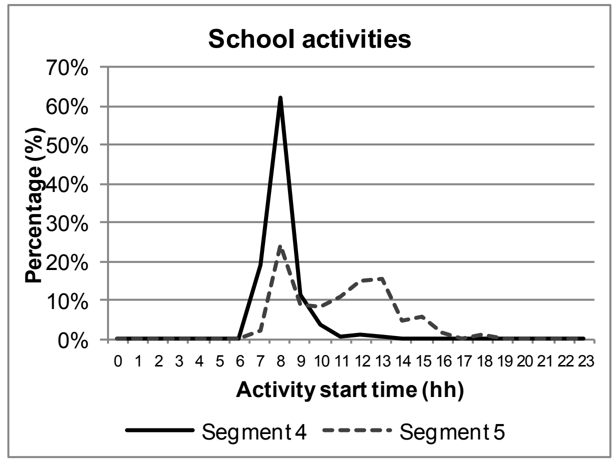
Table 7.2: Summary of activity attributes in each segment - proportion of individuals conducting activities and average durations

Segment no.	% of individuals conducting activities				Average durations (min)			
	Work	**School**	**Shopping**	**Other**	**Work**	**School**	**Shopping**	**Other**
Segment 1	100%	1%	8%	23%	535	2	5	25
Segment 2	100%	6%	14%	27%	423	13	7	31
Segment 3	100%	0%	12%	43%	208	0	5	36
Segment 4	1%	100%	4%	13%	2	456	3	16
Segment 5	13%	100%	6%	32%	25	358	5	50
Segment 6	1%	0%	100%	18%	1	0	164	11
Segment 7	1%	1%	100%	15%	0	0	85	5
Segment 8	8%	2%	21%	100%	11	4	15	268
Segment 9	0%	0%	11%	100%	0	0	5	170
Segment 10	0%	0%	0%	0%	0	0	0	0
Segment 11	22%	10%	40%	45%	46	18	29	27

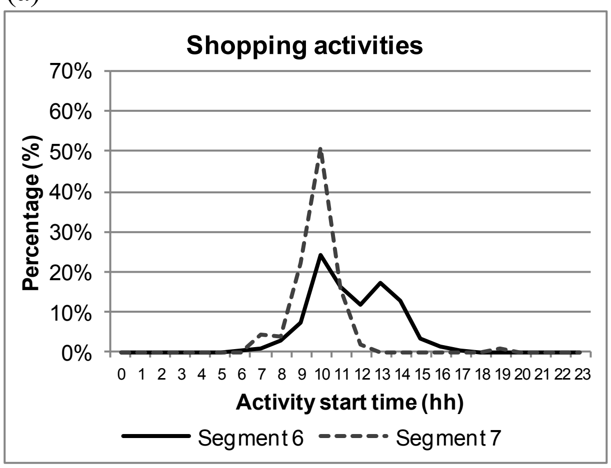
Each segment consists of activity sequences that are similar to each other except Segment 11. Segment 11 includes a variety of uncommon activity sequences and is labelled as "*Outlier*" (Table 7.3). Table 7.2 shows the proportions of individuals conducting activities and the average durations of activities in each segment. These attributes clearly indicate which activities are dominant activities in each segment. For example, in Segment 1, all individuals (100%) conduct at least one work activity with a long average duration of 535 minutes; however a small proportion of individuals also conduct other types of activities with lower average durations (school - 1% (average duration 2 minutes), shopping - 8% (average duration 5 minutes), and other - 23% (average duration 25 minutes)), as shown in Table 7.2. Figure 7.2 then shows the start time distributions of the dominant activity in each segment. The distributions of three activity attributes in each segment are characterized and labelled in Table 7.3. Table 7.3 also shows the proportion of the individuals in each segment. A comparative analysis of key socio-demographic characteristics of each segment has also been presented in Table 7.4.



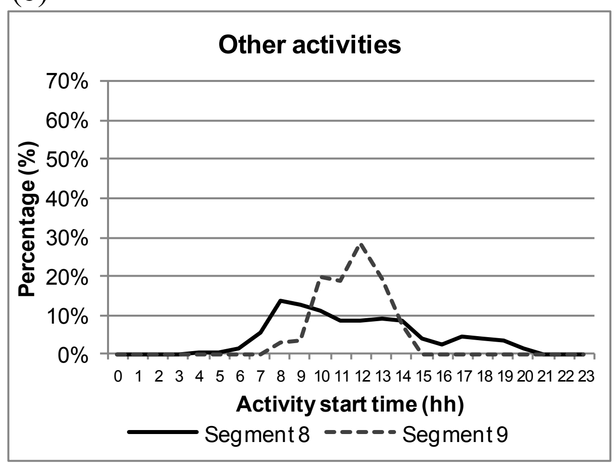
(a)



(b)



(c)



(d)

Figure 7.2: Summary of activity attributes in each segment - start time distributions of different activities (a) Segment 1, Segment 2 and Segment 3, (b) Segment 4 and Segment 5, (c) Segment 6 and Segment 7, and (d) Segment 8 and Segment 9

Overall, 32% of the population belongs to Segment 1, 17% to Segment 4, 5% to Segment 6, 10% to Segment 8 and 23% to Segment 10. Other segments include a small proportion of the population.

Segment 1, Segment 2 and Segment 3 are work-dominated segments, labelled as *Work 1*, *Work 2* and *Work 3*, respectively, as shown in Table 7.3. However, these segments differ in terms of start time distributions and activity durations.

The first segment, *Work 1* consists of a large number of individuals (32%) whose daily activity schedule follows a general pattern of home, a long duration of work and home. However, some individuals also report conducting school (1%), shopping (8%) and other (23%) activities with lower average durations during a typical weekday. The individuals in this segment start their work activities in the morning hours, which is common for these activities in the Montreal area (Yasmin et al., 2015b) and do the work activities for the longest average durations (535 minutes). Considering the daily time budget, such a long work-day might influence the participation in other types of activities (for instance, less participation in shopping activities) (Bhat, 1998b).

In segment *Work 2*, a small number of individuals (2%) is included. Like *Work 1*, their activity schedules follow a general pattern of home, work and home. However these individuals have a broad range of start times of work activities, which vary between afternoon to evening. In addition, they conduct work activities for a shorter duration (423 minutes) than *Work 1* (535 minutes). Thus, it is assumed that the individuals can participate more in other types of activities, school (6%), shopping (14%) and other (27%) than *Work 1*. The third work-dominated segment, *Work 3* consists of only 0.2% of individuals (the lowest proportion), who start their work activities between morning and afternoon. However, they conduct their work activities for a significantly lower average duration (208 minutes) compared to other work-dominated segments.

Work 1 (54%) and *Work 2* (59%) have a higher proportion of men with average age of 40 and 39 years, respectively, whereas gender composition is quite similar in *Work 3* which has a slightly higher average age of 45 years than *Work 1*and *Work 2*. A large majority (91%) of individuals of *Work 1* are full-time workers. It is not surprising as the analysis reveals that this is the segment of a long duration of work. In *Work 2*, there are less full-time (68%) and more part-time (16%) workers compared to *Work 1*; which may be one of the reasons for lower average durations in *Work 2* compared to *Work 1* (full-time work segment). Furthermore, *Work 3* consists of fewer full-time (46%) and more part-time (30%) workers than *Work 1* and *Work 2*. This seems reasonable because the patterns with the lowest average durations of work activities are included in *Work 3*. The majority of individuals of these three segments possess a driver's license (*Work 1* - 84%, *Work 2* - 78%, and *Work 3* - 82%). Household car ownership is also higher; however car ownership in *Work 1* is the largest (81%). The highest full-time work status and possession of driver's license in *Work 1* might be related to this fact, as also shown in other research (Potoglou & Kanaroglow, 2008). Though the households with three or more persons are higher; presence of children is lower. But, in segment *Work 3*, the proportion of the households with children is comparatively higher (54%) than the other two work-dominated segments. Home location distributions of *Work 1* and *Work 2* are quite similar with the Montreal Island; however *Work 3* shows some significant differences (for instance, lower individuals live in Montreal Centre and higher individuals live in Montreal West).

Table 7.3: Segment descriptions and proportions of individuals in each segment (N = 10000)

Segment no.	**Characterization of segments (dominant activity, start time, duration)**	**Segment label**	**Proportion of individuals (%)**
Segment 1	Work, morning, long duration	*Work 1*	32%
Segment 2	Work, afternoon to evening, less duration	*Work 2*	2%
Segment 3	Work, morning to afternoon, short duration	*Work 3*	0.2%
Segment 4	School, morning, long duration	*School 1*	17%
Segment 5	School, morning to evening, less duration	*School 2*	1%
Segment 6	Shopping, morning to afternoon, long duration	*Shopping 1*	5%
Segment 7	Shopping, morning to late afternoon, less duration	*Shopping 2*	1%
Segment 8	Other, morning to evening, long duration	*Other 1*	10%
Segment 9	Other, morning to late afternoon, less duration	*Other 2*	1%
Segment 10	No out-of-home activity	*Home*	23%
Segment 11	Outliers	*Outlier*	7%

Segment 4 (*School 1*) and Segment 5 (*School 2*) are the school-dominated segments with 17% and 1% of individuals, respectively. The daily activity schedules of the individuals in these segments consist of a general pattern of home, a long duration of school, and home; however more individuals in *School 2* participate in other types of activities (work - 13%, shopping - 6%, and other - 32%) than the individuals in *School 1* (work - 1%, shopping - 4%, and other - 13%). The average durations of participating in school activities in *School 1* is higher (456 minutes) than *School 2* (358 minutes); thus, it is assumed that the individuals in *School 2* may have more time to participate in other types of activities considering the daily time budget. The individuals in *School 1* start their school activities in the morning hours (typical for this activity type in the Montreal Area (Yasmin et al., 2015b)), whereas the individuals in *School 2* are more flexible with the start times of school activities, which vary from morning to evening.

The segments, *School 1* and *School 2* are the two youngest segments with the average age of 15 and 22 years, respectively. However, this suggests that *School 1* is a group of younger students (98%) and *School 2* consists of older students (95%). More men (52%) are in *School 1*, but more women (52%) are in *School 2*. The majority (81%) of the individuals in *School 1* do not have a driver's license; however more individuals (56%) possess a driver's license in *School 2*. These findings are consistent with their age distributions. Household car ownership is higher in both segments (*School 1* - 78%, and *School 2* - 63%); although the proportion is larger in *School 1*. It is assumed that the younger children live with their parents, which might be related to higher car ownership in their households. In addition, as expected for a youngest student segment, the presence of children and household size in *School 1* is higher. Home location distributions are quite similar as the Montreal Island in *School 1*; but the proportion (12%) is significantly higher for the Montreal Downtown for *School 2*. It may suggest that older students (for instance undergraduate and graduate students) tend to live in Montreal Downtown, where several universities are located.

There are two shopping-dominated segments, *Shopping 1* (Segment 6 - consists of 5% of individuals) and *Shopping 2* (Segment 7 - consists of 1% of individuals). However, they differ in terms of the start time distributions and average durations of participating in different activities. In *Shopping 1*, the start time distributions of shopping activities vary between morning to afternoon with average duration of 164 minutes, whereas in *Shopping 2*, the start time distributions of shopping activities vary between morning to late afternoon with less average duration of 85 minutes.

Table 7.4: Summary of key socio-demographic characteristics and their distributions in each segment (N = 10000)

Variables	**S1** Work 1	**S2** Work 2	**S3** Work 3	**S4** School 1	**S5** School 2	**S6** Shopping 1	**S7** Shopping 2	**S8** Other 1	**S9** Other 2	**S10** Home	**S11** Outlier	**Total**
Average age	40	39	45	15	22	57	57	51	59	54	47	42
Gender (%)												
Men	54%	59%	50%	52%	48%	39%	39%	46%	40%	41%	50%	49%
Women	46%	41%	50%	48%	52%	61%	61%	54%	60%	59%	50%	51%
Employment status (%)												
Full-time worker	91%	68%	46%	1%	2%	14%	14%	23%	12%	18%	28%	40%
Part-time worker	6%	16%	30%	0%	0%	6%	1%	6%	9%	5%	10%	5%
Student	2%	8%	7%	98%	95%	4%	2%	10%	3%	10%	14%	23%
Retired	1%	6%	13%	0%	1%	56%	50%	41%	65%	46%	30%	21%
Other status	1%	2%	4%	0%	1%	16%	25%	15%	11%	20%	14%	9%
At home	0%	0%	0%	0%	0%	3%	9%	5%	1%	2%	4%	2%
Possession of driver's license (%)												
Yes	84%	78%	82%	18%	56%	63%	77%	68%	60%	52%	72%	61%
No	16%	22%	18%	81%	43%	37%	23%	32%	40%	48%	28%	39%
Vehicle ownership (%)												
0 car	19%	23%	29%	22%	37%	32%	22%	29%	43%	34%	24%	27%
1+ car	81%	77%	71%	78%	63%	68%	78%	71%	57%	66%	76%	73%
Presence of children (%)												
Yes	33%	28%	46%	77%	38%	18%	27%	24%	8%	21%	29%	33%
No	67%	72%	54%	23%	62%	82%	73%	76%	92%	79%	71%	67%
Household size (%)												
1 person HH	19%	21%	15%	3%	10%	30%	21%	32%	44%	26%	21%	22%
2 persons HH	35%	34%	29%	16%	31%	45%	45%	35%	38%	38%	36%	35%
3 or more persons HH	47%	45%	56%	81%	59%	25%	34%	33%	18%	36%	42%	43%
Home location regions (%)												
Montreal Downtown	3%	6%	8%	3%	12%	5%	7%	5%	4%	4%	5%	4%
Montreal Centre	54%	58%	34%	51%	50%	52%	45%	56%	60%	54%	54%	54%
Montreal East	17%	17%	14%	17%	16%	17%	22%	16%	18%	18%	14%	17%
Montreal West	26%	18%	44%	29%	22%	27%	27%	23%	18%	23%	27%	25%

The average age (57 years) of the individuals is much higher in shopping-dominated segments than work- and school-dominated, implying that older individuals participate in more shopping activities in a typical weekday. In addition, the analysis reveals that more women are in these segments (61% in each segment). Table 7.4 also reveals that the majority of the individuals (*Shopping 1* - 72%, and *Shopping 2* - 75%) are either retired or individuals with other status. It is obvious that workers and students cannot engage in shopping activities for such a long durations in a typical weekday; however it has also been observed that some individuals, especially full-time workers (14% in each segment) have such daily activity schedule. These shopping-dominated segments have higher possession of driver's license (*Shopping 1* - 63% and *Shopping 2* - 77%) and car ownership (*Shopping 1* - 68% and *Shopping 2*- 78%), and lower presence of children (*Shopping 1*- 18%, and *Shopping 2* - 27%) .

Other-dominated patterns are included in two segments, *Other 1* (Segment 8) and *Other 2* (Segment 9). *Other 1* and *Other 2* consist of 10% and 1% of individuals, respectively. In general, the daily activity patterns of the individuals consist of home, a long duration of other activities, and home. Some individuals in *Other 1* also report participating in other out-of-home activities (work - 8%, school - 2%, and shopping - 21%); however few occurrences (11%) of shopping activities are only reported in segment *Other 2*. In *Other 1*, the start time distributions of other activities vary between morning to evening and the average duration of other activities is 268 minutes. In *Other 2*, the start time distributions of other activities vary between morning to late afternoon and the average duration of other activities is 170 minutes.

The individuals in *Other 1* and *Other 2* have an average age of 51 and 59 years, suggesting that *Other 2* is a group of the elderly individuals. More women (*Other 1* - 54%, and *Other 2* - 60%) are in these segments; however the proportion is higher in *Other 2*. The majority of the individuals are either retired or with other status (*Other 1* - 56%, and *Other 2* - 76%); however the proportion is significantly higher for *Other* 2, which is consistent with their age distributions. In addition, it is surprising to observe that some full-time and part-time workers are also included in these segments. More individuals have the driver's license (*Other 1* - 68% and *Other 2* - 60%). Household car ownership is higher (*Other 1* – 71%, and *Other 2* - 57%). More households have no children (*Other 1* - 76% and *Other 1* - 92%), this fact is significant in *Other 2* which is usual for a segment of elderly individuals.

Segment *Home* consists of the individuals (23%), who do not conduct any out-of-home activity in a typical weekday. The average age of the individuals in this segment is 54 years, implying that more older individuals are in this segment (similar result found in (Bhat, 1998b)). In addition, more women are in this segment. The majority of the individuals (66%) are either retired or with other status; but it is also interesting to observe that a significant proportion of full-time workers (18%) and student (10%) are also included in this segment. These individuals might conduct home-work or home-school or might be at home for some other reasons. In segment *Home*, the proportion of the individuals possessing driver's license is slightly higher (52%), which is comparatively lower than the other segments except *School 1* (the youngest segment). Household car ownership is higher (66%). Presence of children is lower; it is quite expected as it is a group of older individuals.

7.5.2 Modelling population segments

7.5.2.1 Modelling process

The potential variables have been chosen based on the previous research on population segmentation (Section 7.2.1). Initially, modelling data has been prepared for the independent variables, namely age groups (thirteen age groups with 5-year interval), gender (men and women), employment status (full-time worker, part-time worker, student, retired, other, at home), possession of driver's license, presence of children, household size (household with one, two, and three or more individuals), car ownership, distance from home to Central Business District (CBD) (home within 5 km from the CBD, home between 6 to 10 km from the CBD, home between 11 to 15 km from the CBD, home between 16 to 25 km from the CBD and home between 26 to 35 km from the CBD), and home location (Montreal Downtown, Montreal Centre, Montreal East and Montreal West). Additionally, the variables with gender-specific age groups (for instance, men aged 25-29 and women aged 25-29 years old) have also been prepared. Before estimating the models, tests for correlations between all pairs of explanatory variables have been performed. The variables with low correlation (<0.4) are considered in the model specifications. In addition, only one variable from the pair of variables with high correlation is considered in the same model specification. It can be illustrated by an example with the variables, household size and presence of children. Three variables have been prepared with household size (*hhsize1*, *hhsize2*, and *hhsize3+*). The variables, *hhsize2* and *hhsize3+* are highly correlated with the presence of children in a household. Thus, the variables, related to household size and presence of children are not included in the same model specification.

Multinomial logit models have been estimated for the distinctive segments using the individual's socio-demographic characteristics. The final model has been selected based on the model fit, agreement with prior hypotheses, and the statistical significance of explanatory variables.

The list of explanatory variables of the final population segmentation model, *Model 1,* is shown in Table 7.5. The variables, age groups with 5-year interval and gender are included in this model specification and in total 126 parameters have been estimated. However, *Model 1* assumes that the influence of gender remains the same across the population with different age groups for belonging to a segment. But, it is possible that the influence of gender might differ across the population with different age groups to belong to a segment. In order to include this potential difference another model, *Model 2*, has been estimated with gender-specific age groups and in total 234 parameters have been estimated. The explanatory variables for *Model 2* are also presented in Table 7.5.

Summary statistics of the characteristics of the estimation sample for *Model 1* are also presented in Figure 7.3. In the model estimation sample (i.e. 7434 individuals), the largest proportion (35%) of the individuals belong to the segment *Work 1*. Significant proportions of the individuals are also included in segment *School 1* (18%), *Shopping 1* (6%), *Other 1* (11%), and *Home* (25%). The rest of the segments include a small proportion of individuals (*Work 2* - 2%, *Work 3* - 0.3%, *School 2* - 2%, *Shopping 2* - 1%, and *Other 2* - 1%).

From the summary statistics of the final model variables (*Model 1*), shown in Figure 7.3, it is observed that a significant proportion (79%) of the individuals aged between 25-54 years old is in segment *Work 1*. *Work 2* shows a similar age distribution (71% of individuals aged between 25-54 years old) like *Work 1*; however this segment also includes younger individuals aged between 20-24 years old (12%). Though segment *Work 3* is also a work-dominated segment, it shows different age distributions compared to other work-dominated segments. The proportions of the individuals aged between 35-44 (41%) and 55-64 (24%) years old are significantly higher.

Table 7.5: Variables for estimating multinomial logit model

Variables	**Description**
Dependent variable	
Segments	Segment 1 - Segment 10
Independent variables (person attributes)	
Model 1	
Age groups	
age5-9	Aged between 5 and 9 years old
age10-14	Aged between 10 and 14 years old
age15-19	Aged between 15 and 19 years old
age20-24	Aged between 20 and 24 years old
age25-29	Aged between 25 and 29 years old
age30-34	Aged between 30 and 34 years old
age35-39	Aged between 35 and 39 years old
age40-44	Aged between 40 and 44 years old
age45-49	Aged between 45 and 49 years old
age50-54	Aged between 50 and 54 years old
age55-59	Aged between 55 and 59 years old
age60-64	Aged between 60 and 64 years old
age65more	Aged 65 years old or older
Gender	
men	Men
women	Women
Model 2	
Age groups by gender	
m5-9	Men aged between 5 and 9 years old
m10-14	Men aged between 10 and 14 years old
m15-19	Men aged between 15 and 19 years old
m20-24	Men aged between 20 and 24 years old
m25-29	Men aged between 25 and 29 years old
m30-34	Men aged between 30 and 34 years old
m35-39	Men aged between 35 and 39 years old

Table 7.5: Variables for estimating multinomial logit model (cont'd)

Variables	Description
m40-44	Men aged between 40 and 44 years old
m45-49	Men aged between 45 and 49 years old
m50-54	Men aged between 50 and 54 years old
m55-59	Men aged between 55 and 59 years old
m60-64	Men aged between 60 and 64 years old
m65more	Men aged 65 years old or older
w5-9	Women aged between 5 and 9 years old
w10-14	Women aged between 10 and 14 years old
w15-19	Women aged between 15 and 19 years old
w20-24	Women aged between 20 and 24 years old
w25-29	Women aged between 25 and 29 years old
w30-34	Women aged between 30 and 34 years old
w35-39	Women aged between 35 and 39 years old
w40-44	Women aged between 40 and 44 years old
w45-49	Women aged between 45 and 49 years old
w50-54	Women aged between 50 and 54 years old
w55-59	Women aged between 55 and 59 years old
w60-64	Women aged between 60 and 64 years old
w65more	Women aged 65 years old or older

Younger individuals (90%) with age less than 25 years old are the majority in segment *School 1*. In *School 2*, the majority (85%) of the individuals are between 15-34 years old. Older individuals are more common in the rest of the segments; however individuals aged 65 years old or older are the most highly represented in these segments (*Shopping 1* - 46%, *Shopping 2* - 44%, *Other 1* - 32 %, and *Other 2* - 52%). The age distributions in each 5-year cohort of segment *Home* are close to each other between 20-64 years old; however a significant proportion (39%) of individuals are 65 years old or older.

The analysis also reveals that more men are in segments *Work 1* (54%), *Work 2* (60%), and *School 1* (53%), whereas it shows that more women are in segments *School 2* (51%), *Shopping 1* (61%), *Shopping 2* (60%), *Other 1* (53%), *Other 2* (60%), and *Home* (58%). However, gender composition is quite similar in *Work 3.*

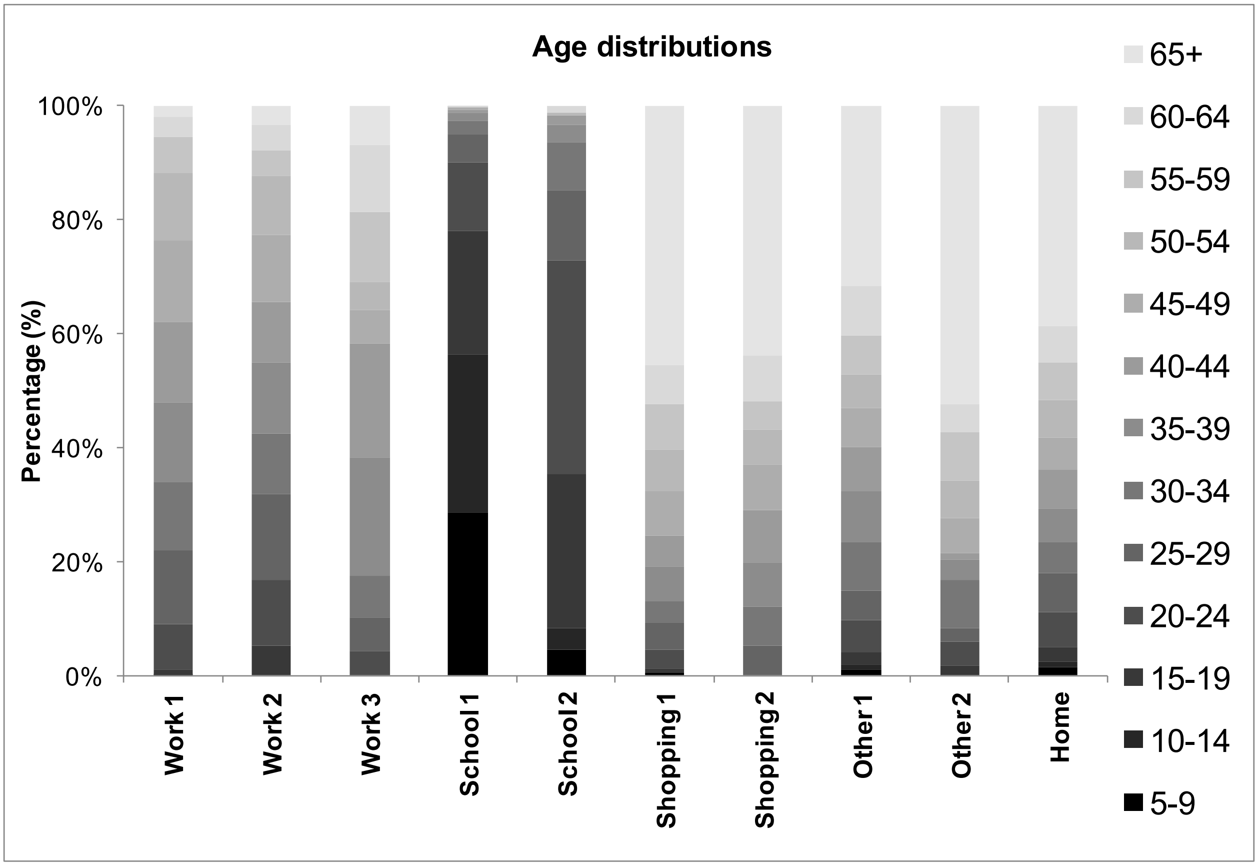


Figure 7.3: Summary statistics of variables for estimating models - age distributions for M*odel 1*

7.5.2.2 Discussion of the model results

Table 7.6 presents the parameter estimation results of the final population segmentation model (*Model 1*) for each segment. All independent variables, related to age groups and gender are kept in *Model 1*; however some of the variables show *t*-statistics less than the threshold value (1.96 for a two-tailed test at the 95% confidence interval). These variables are kept in the final model for completeness and to provide useful and suggestive insights. Also, with a larger data set these parameters might show statistical significance.

Table 7.6 shows that any of the parameters of the segments *School 2*, *Shopping* 2 and *Other 2* are not significant. The sample size of these segments is small; it is assumed that the coefficient values might show statistical significance if we would use a larger set of data. The model results for the rest of the segments are discussed below.

The result suggests that the key determinant of belonging to a segment is age group. For *Work 1*, most of the parameters related to age groups are significant except two (*age5-9* and *age10-14*).

These younger age groups exhibit negative relationships, whereas the older age groups (starting from 15 years old) show positive relationships, implying that older individuals are more likely to be in this work-dominated segment. This finding is consistent with prior expectation. The parameters of the older age groups (starting from 15 years old) also show an interesting trend that with the increase of age the magnitude of the positive influence significantly increases up to a certain age (49 years old) with an exception (*age40-44*) and then decreases. These results also comply with the observed behaviour, as shown in Table 7.4 and Figure 7.3 (a). Gender is also a significant determinant; more specifically, the parameter (0.44) of the variable *men* shows that men are more likely to be in this segment.

The parameters for *Work 2* show similar results to those of *Work 1.* All parameters related to age groups are significant except two younger age groups (*age5-9* and *age10-14*). Older age groups (starting from 15 years old) have positive influence and younger age groups (less than 15 years old) have negative influence; this implies that older individuals tend to be in this segment. However like *Work 1*, the parameters do not show a clear trend of increasing positive influence with the increase of age, but it shows a decreasing trend of positive influence after certain age (50 years old or older). It is noted here that the sample size of this segment is also low (2%) compared to *Work 1* (35%). The parameter value (0.64) for the variable *men* demonstrates that men are more likely to be in this segment.

Though *Work 3* is another work-dominated segment (sample size - 0.3%), the model results for this segment show different behaviours than other work-dominated segments, as also observed in descriptive analysis (Table 7.4 and Figure 7.3 (a)). For this segment, three younger age groups show negative relationship and the rest of the age groups show positive relationships. A significant proportion of individuals included in this segment is aged between 35-44 and 55-64 years old; the model replicates this behaviour. The parameters of these age groups are higher than others, implying that individuals at these ages are more likely to be in this segment.

For the largest school-dominated segment *School 1* (18%), the model result shows that most of the parameters are significant except two older age groups (*age55-59* and *age60-64*). All parameters of age groups exhibit a positive relationships except one (*age55-59*) and a clear trend of decreasing the magnitude of influence with the increase of the age with a few exceptions (for instance, *age10-14* has higher parameter value (9.37) than *age5-9* (9.14) and *age55-59* shows a negative relationship). These results indicate that younger individuals are more likely to be in this school-dominated segment, which comply with the prior expectation. The result also suggests that men are more likely to be in this segment; however the difference is very low.

A few number of parameters of segment *Shopping 1* is significant; however the result shows negative relationships with all age groups but two (*age45-49* and *age55-59*). The magnitude of negative influence is generally higher for the younger age groups, implying that older individuals are more likely to do shopping in a typical weekday for a longer duration. In addition, the negative parameter value (-0.14) of the variable *men* indicates that women are more likely to belong to this segment, implying that women tend to do more shopping. These results comply with the observed behaviours, as revealed in Table 7.4 and Figure 7.3 (a), and with previous research (Srinivasan & Bhat, 2005).

Other 1 is the largest other-dominated segment (11%). For this segment, the parameters for younger age groups (less than 30 years old) with an exception show negative relationship and for older age groups show positive relationship. These imply that older individuals are more likely to be in this segment of individuals that conduct other activities with a longer duration. The parameter of the variable *men* also indicates that men are more likely to be in this segment, which is observed in the descriptive analysis (Table 7.4).

Model 1 assumes that the influence of gender remains the same across the population with different age groups to belong to a segment. The results of *Model 1* indicate that men are more likely to belong to work-, school-, and largest other-dominated segments, whereas women are more likely to be in shopping-dominated segments. However, the gender effect might vary across the population with different age groups. *Model 2* examines this assumption by considering gender-specific age groups, as shown in Table 7.7.

Model 2 provides some useful insights on the gender differences across different age groups. Some of these insights for the segments with larger sample size are discussed below.

Table 7.6: *Model 1* (age groups and gender): Parameter estimates (N= 7434)

Variables	***Work 1***		***Work 2***		***Work 3***		***School 1***		***School 2***		***Shopping 1***		***Shopping 2***		***Other 1***		***Other 2***		***Home***	
	Parameters	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**
Constant	-2.83	-17.70	-5.34	-10.41	-6.52	-6.38	-6.51	-6.49	-10.00	-1.75	-1.18	-12.32	-2.87	-14.37	-1.04	-12.05	-3.06	-13.74	-	-
Age groups																				
age5-9	-6.44	-0.35	-4.09	-0.22	-2.76	-0.14	9.14	8.94	8.27	1.44	-1.28	-1.73	-6.15	-0.34	-0.06	-0.15	-5.96	-0.33	-	-
age10-14	-0.39	-0.37	-3.85	-0.21	-2.55	-0.13	9.37	9.12	8.50	1.48	-7.57	-0.41	-5.92	-0.32	-0.24	-0.50	-5.73	-0.31	-	-
age15-19	2.05	7.11	3.09	4.78	-3.37	-0.17	8.26	8.15	9.40	1.64	-1.22	-2.31	-6.78	-0.37	-0.05	-0.18	-0.77	-0.75	-	-
age20-24	3.16	16.20	2.98	5.21	1.66	1.17	6.77	6.71	8.87	1.55	-1.04	-3.31	-7.70	-0.42	0.06	0.34	-1.00	-1.35	-	-
age25-29	3.60	19.11	3.22	5.80	2.68	2.31	5.77	5.69	7.77	1.36	-0.44	-1.82	-0.55	-1.03	-0.14	-0.72	-1.05	-1.42	-	-
age30-34	3.71	19.15	3.02	5.22	2.47	2.00	5.17	5.06	7.33	1.28	-0.40	-1.55	0.04	0.09	0.57	3.28	0.23	0.50	-	-
age35-39	3.85	20.07	3.01	5.21	3.37	3.05	4.80	4.66	6.63	1.15	-0.05	-0.23	0.17	0.40	0.55	3.20	-0.89	-1.20	-	-
age40-44	3.70	19.78	2.80	4.87	3.18	2.88	3.51	3.26	5.74	1.00	-0.39	-1.64	0.12	0.29	0.32	1.85	-1.77	-1.73	-	-
age45-49	3.94	20.61	3.17	5.56	1.76	1.24	3.55	3.27	-3.55	-0.04	0.12	0.57	0.30	0.73	0.38	2.09	-0.20	-0.37	-	-
age50-54	3.57	18.96	2.68	4.57	1.58	1.11	2.67	2.30	5.06	0.87	-0.09	-0.40	-0.17	-0.38	0.16	0.89	-0.67	-1.09	-	-
age55-59	2.98	15.31	2.12	3.34	2.66	2.29	-4.86	-0.20	-4.14	-0.04	0.12	0.61	-0.36	-0.74	0.27	1.54	0.01	0.03	-	-
age60-64	2.44	11.79	2.03	3.11	2.72	2.34	1.62	1.14	5.10	0.88	-0.01	-0.03	-0.12	-0.27	0.44	2.63	-0.34	-0.62	-	-
age65more	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Gender																				
men	0.44	6.51	0.64	3.45	0.25	0.61	0.23	2.10	0.26	1.29	-0.14	-1.25	-0.03	-0.14	0.16	1.86	-0.03	-0.11	-	-
women	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Log-likelihood at zero coefficient															-17117.418					
Log-likelihood at sample shares															-12419.056					
Log-likelihood at convergence															-9214.641					
Rho-square															0.258024					

indicates statistical significance at the 95% confidence level.

Table 7.7: *Model 2* (age groups by gender): Parameter estimates (N= 7434)

Variables	*Work 1*		*Work 2*		*Work 3*		*School 1*		*School 2*		*Shopping 1*		*Shopping 2*		*Other 1*		*Other 2*		*Home*	
	Parameters	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**
Constant	-3.29	-12.08	-5.93	-5.91	-5.93	-5.91	-5.93	-5.91	-12.90	-0.39	-1.31	-11.70	-3.04	-12.58	-1.15	-10.91	-3.15	-12.36	-	-
m5-9	-9.32	-0.06	-6.64	-0.05	-6.78	-0.04	8.64	8.30	11.70	0.35	-0.63	-0.83	-9.59	-0.07	0.12	0.22	-9.49	-0.06	-	-
m10-14	-8.89	-0.06	-6.20	-0.04	-6.33	-0.04	9.03	8.52	10.70	0.32	-10.90	-0.07	-9.08	-0.06	0.34	0.55	-8.99	-0.06	-	-
m15-19	2.86	6.66	2.80	1.95	-7.29	-0.05	7.88	7.66	12.50	0.38	-1.13	-1.52	-10.00	-0.07	0.21	0.51	0.02	0.02	-	-
m20-24	3.85	12.08	3.97	3.70	-8.21	-0.05	6.22	6.10	11.60	0.35	-1.63	-2.71	-10.90	-0.08	0.00	-0.02	-10.80	-0.07	-	-
m25-29	4.23	13.75	4.48	4.31	-8.39	-0.05	5.35	5.22	11.00	0.33	-0.71	-1.91	-11.10	-0.08	0.06	0.24	-1.07	-1.03	-	-
m30-34	4.86	14.94	4.37	4.06	2.29	1.61	5.13	4.91	10.90	0.33	-0.94	-1.75	-0.60	-0.58	0.84	3.12	-0.48	-0.46	-	-
m35-39	4.62	14.58	4.26	3.99	3.45	3.05	4.26	3.99	9.74	0.30	-0.07	-0.22	0.78	1.47	0.57	2.18	-10.70	-0.07	-	-
m40-44	4.50	14.57	4.02	3.77	3.43	3.10	2.51	2.04	8.81	0.27	-1.01	-2.28	0.02	0.04	0.47	1.93	-10.90	-0.07	-	-
m45-49	4.95	15.11	4.65	4.37	-7.75	-0.05	3.04	2.45	-3.27	-0.01	0.29	0.86	0.15	0.19	0.56	1.88	-0.43	-0.41	-	-
m50-54	4.46	14.10	2.71	2.20	-8.08	-0.05	-7.95	-0.05	9.01	0.27	0.04	0.12	0.73	1.39	0.33	1.18	0.34	0.53	-	-
m55-59	3.95	12.19	3.65	3.29	2.04	1.43	-7.90	-0.05	-3.58	-0.01	-0.10	-0.28	-10.80	-0.08	0.81	3.31	0.36	0.56	-	-
m60-64	3.29	9.66	3.45	3.05	2.06	1.45	-7.87	-0.05	9.05	0.27	0.08	0.25	-0.14	-0.19	0.74	2.95	-0.02	-0.03	-	-
m65more	1.25	3.73	1.59	1.37	-9.61	-0.06	-9.44	-0.07	-5.69	-0.01	0.19	1.08	0.37	1.04	0.41	2.59	0.20	0.50	-	-

Table 7.7: *Model 2* (age groups by gender): Parameter estimates (N= 7434) (cont'd)

Variables	***Work 1***		***Work 2***		***Work 3***		***School 1***		***School 2***		***Shopping 1***		***Shopping 2***		***Other 1***		***Other 2***		***Home***	
	Parameters	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**	**Parameters**	**t-stat**
w5-9	-9.10	-0.06	-6.40	-0.04	-6.54	-0.04	8.74	8.32	10.50	0.32	-11.10	-0.08	-9.28	-0.07	0.14	0.23	-9.18	-0.06	-	-
w10-14	0.89	0.82	-6.40	-0.04	-6.53	-0.04	8.78	8.36	11.90	0.36	-11.00	-0.08	-9.33	-0.06	-0.56	-0.72	-9.24	-0.06	-	-
w15-19	2.60	5.82	4.55	4.12	-7.34	-0.05	7.71	7.50	12.40	0.38	-1.17	-1.58	-10.00	-0.07	0.05	0.12	-9.93	-0.07	-	-
w20-24	3.79	11.93	3.78	3.50	1.84	1.29	6.35	6.24	12.10	0.37	-0.59	-1.56	-11.00	-0.08	0.42	1.69	-0.25	-0.32	-	-
w25-29	4.36	13.91	3.52	3.18	3.00	2.58	5.24	5.09	10.50	0.32	-0.07	-0.23	0.40	0.70	0.01	0.04	-0.87	-0.84	-	-
w30-34	3.93	12.60	3.55	3.25	1.76	1.24	4.32	4.12	9.84	0.30	-0.09	-0.30	0.47	0.90	0.69	3.05	0.59	1.11	-	-
w35-39	4.42	14.15	3.50	3.16	1.89	1.33	4.37	4.15	9.57	0.29	0.10	0.34	-0.31	-0.41	0.84	3.68	-0.20	-0.26	-	-
w40-44	4.25	13.77	3.35	3.03	-8.35	-0.05	3.35	3.03	8.73	0.26	0.07	0.24	0.46	0.87	0.51	2.19	-1.04	-1.00	-	-
w45-49	4.33	14.16	3.47	3.19	1.68	1.18	3.07	2.72	-3.98	-0.01	0.15	0.57	0.58	1.19	0.54	2.36	0.00	0.01	-	-
w50-54	4.03	13.20	3.89	3.68	1.59	1.12	2.69	2.31	-4.09	-0.01	-0.04	-0.13	-1.31	-1.26	0.33	1.43	-11.20	-0.08	-	-
w55-59	3.38	10.80	2.26	1.83	2.26	1.83	-8.38	-0.06	-4.14	-0.01	0.34	1.42	0.28	0.53	0.11	0.45	-0.12	-0.18	-	-
w60-64	2.95	9.01	2.33	1.89	2.33	1.89	1.64	1.15	-4.04	-0.01	0.07	0.24	0.13	0.23	0.52	2.32	-0.44	-0.58	-	-
w65more	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Log-likelihood at convergence																	-9141.343			
Rho-square																	0.263926			

indicates statistical significance at the 95% confidence level.

Table 7.8: Model validation with a holdout sample (N= 1858)

Segments	**Actual shares (%)**	**Predictions**	
		Model 1	***Model 2***
Work 1	33%	35%	34%
Work 2	2%	2%	1%
Work 3	0.1%	0.3%	0.4%
School 1	20%	18%	18%
School 2	2%	1%	1%
Shopping 1	6%	6%	6%
Shopping 2	1%	1%	1%
Other 1	12%	12%	11%
Other 2	1%	1%	1%
Home	23%	24%	25%

Similar to *Model 1*, most of the parameters specific to age groups for both genders are significant for segment *Work 1.* The results also suggest that older individuals from both genders tend to be in this segment. *Model 1* has suggested that men are more likely to belong to this work-dominated segment; *Model 2* also supports that fact, as the parameters are always higher for men than women across different age groups except one (age group 25-29 years old). This implies that more women aged 25-29 years old tend to be in this segment.

For *School 1* (largest school-dominated segment), the number of significant parameters of different age groups for women is the same with *Model 1*; however, the number is less for different age groups for men. Similar to *Model 1*, *Model 2* also suggests that younger individuals from both genders tend to be in *School 1*, which is consistent with the priori expectation. *Model 1* also shows positive influence of being a man despite of his/her age, suggesting that men tend to be in this segment. However, *Model 2* indicates significant differences between men and women across different age groups; women from several age groups are more likely to be in this segment.

In *Model 2* for *Shopping 1*, only two parameters of different age groups (20-24 and 40-44 years old) for men are significant, while no parameters are significant for women. *Model 1* has suggested that older individuals are more likely to do shopping in a typical weekday. This fact is also observed for both men and women in *Model 2*; however the gender-specific parameters of different age groups suggest that it is also true for younger women but not men.

For *Other 1* overall, it is observed that men are more likely to be in this segment across all age groups with a few exceptions. Men from middle aged to elderly tend to belong in this segment, while middle-aged women are more likely to be in this other-dominated segment.

7.5.2.3 Overall model fit and model validation

The goodness of fit of the overall model is examined through the log-likelihood values and the rho-square value. The log-likelihood value at convergence for *Model 1* is -9214.6 with 126 parameters, whereas the log-likelihood at zero coefficient and at sample shares are -17117.4 and -12419.1, respectively. The rho-square value is 1 minus the ratio of log-likelihood value of the full model and the log-likelihood value of the null model (constant only model). The rho-square value is 0.258, which indicates a reasonable model fit.

Although the log-likelihood value at convergence (-9141.343 with 234 parameters) and rho-square value (0.264) for *Model 2* indicate a slight improvement over *Model 1*; considering the estimated and significant number of parameters and analyzing their useful insights, *Model 1* has been selected as the final model.

The performance of the estimated models has also been examined using a hold-out validation sample (1858 individuals). Using the parameters of both models, probabilities for the individuals to belong to a segment have been predicted. The probability estimation equations for both models (*Model 1* and *Model 2*) are presented in Appendix A. To evaluate the performance, the observed and predicted shares of each segment for both models have been compared, as shown in Table 7.8. The validation results indicate that both models provide quite reasonable prediction, however, *Model 1* provides slightly better predictions.

7.6 Summary

This chapter proposes a methodology to systematically segment the population based on both activity patterns and socio-demographic characteristics of the individuals. More specifically, this research first applies the multiple sequence alignment method (SAM) to segment the population based on the similarities in individuals' daily activity patterns (represented by three activity attributes, frequency, start time and duration) for five broad activity types (work, school, shopping, other, and home). It then estimates the multinomial logit models for the distinct segments of individuals based on their socio-demographic characteristics. This research demonstrates the methodology using the 2003 O-D travel survey of Montreal.

The cluster analysis using the multiple SAM provides eleven distinct segments (*Work 1, Work 2, Work 3, School 1, School 2, Shopping 1, Shopping 2, Other 1, Other 2, Home* and *Outlier*) based on the similarities in individuals' daily activity patterns. Multinomial logit models are estimated for these distinct segments of individuals (excluding *Outlier* which includes a variety of uncommon activity sequences) based on their socio-demographic characteristics. The population segmentation model with two basic demographic variables (age groups with 5-year interval and gender) offers better data fit compared to other estimated models and provides useful insights into the determinants of belonging to different segments. The validation exercise using the final model estimation results for a holdout sample further indicates the strength of this model to segment the population.

The approach of segmenting population demonstrated in this research provides several advantages. To segment the population, it considers multi dimensions of an activity pattern (frequency, start time, and duration), which leads to more behaviourally driven segments. This approach includes both activity patterns and socio-demographic characteristics of the individuals to systematically segment the population. Also, unlike other research, inclusion of less aggregated age groups (thirteen age groups with 5-year interval) provides better and detail understanding of their influences for belonging to a segment. Additionally, it will be easier to capture changes in behaviours (i.e. in the segments) when there is a switch in behaviours among some demographic groups over time (for instance switch from 15 to 24 years old or 55 to 64 years old). In addition, the population segmentation model is quite simple in terms of data requirements (variables, age and gender), and estimation, which makes it easier to use for prediction purposes. However, future research could further validate the model using a future year dataset (for instance, 2008 O-D survey). As we observe similarities between some segments (for instance same dominant activity for the segments *Work 1*, *Work 2* and *Work 3*), it is possible that the multinomial logit model may not be the best method. Nested logit models may be more appropriate; future research could estimate such models.

We assume that development of activity attributes distributions based on population segments by a systematic process may improve the TASHA model performance to simulate individuals' activity schedules. To better fit the TASHA model in another context (for instance, Montreal), we also recommend to use local activity attributes distributions (frequency, start time and duration) (see Chapter 4 and 5). Future research could integrate the population segmentation model within the TASHA modelling framework to segment the population to develop local activity attributes distributions, which would be used to simulate individuals' daily activity schedules. These future works will facilitate to further evaluate the TASHA model performances. The population segmentation model can also be utilized to identify a relevant group of interest to simulate their activity schedules and/or conduct detail investigation of their travel behaviours.

CHAPTER 8 GENERAL DISCUSSIONS

Activity-based travel demand models are recognized as a more powerful framework to better understand the behavioural basis for individuals' activity participation decisions in space and time (Bhat & Lawton, 2000; Kitamura, 1996; Kitamura et al., 1996; Roorda, 2005; Shiftan et al., 2003). This thesis focuses on the two main goals. The first goal is to enhance the current modelling approach of travel demand for the Greater Montreal Area (GMA) by applying an activity-based approach (i.e. TASHA) in the context of Montreal. The second goal is to contribute to the validation and enhancement in the activity-based modelling framework by demonstrating a validation procedure of activity-based models and proposing some improvements.

An extensive literature review on different modeling approaches allows us to identify several research gaps and limitations in the modelling tools used by the Ministère des transports du Québec (MTQ) in the GMA and in the current activity-based modelling framework, as presented in Chapter 2. Based on the perspectives gained through the review, this research attempts to contribute in several aspects. The objectives of this thesis are listed in Section 1.2 (Chapter 1), so in this chapter, it may be beneficial to assess the progress made through this research and highlight its achievements.

- The first objective is to apply the activity-based model TASHA (Travel Activity Scheduler for Household Agents) in the context of Montreal. In order to achieve this objective, the TASHA model has been transferred to the context of the Island of Montreal using the 2003 Origin-Destination (O-D) travel survey and the 2001 Canadian Census.

- The second objective is to compare TASHA simulated to observed activity attributes at three different levels of aggregation (macro-, meso-, and micro-level). This research compares TASHA simulated to observed activity attributes (activity frequency, start time, average duration, and average distance) from the 2003 O-D survey at three different levels, macro-level (aggregation of the entire population), meso-level (aggregation by population segments by age group and gender, and by home location), and micro-level (individuals).

- The third objective is to empirically examine daily activity generation behaviours and their evolutions over time in Montreal. To achieve this objective, this research examines changes in the distributions of activity generation attributes (activity frequency, start time and duration) over a 10-year period using the 1998, 2003 and 2008 O-D travel surveys in the GMA. It also investigates which activities and socio-demographic cohorts are experiencing the important changes in this region.
- The fourth objective is to develop a population segmentation model for segmenting the population to develop activity generation attributes distributions for an activity-based travel demand model (such as TASHA). To fulfill this objective, this research proposes a methodology combining two methods (multiple sequence alignment method (SAM) and multinomial logit model (MNL)) to segment the population of the Montreal Island based on the individual's daily activity patterns and their socio-demographic characteristics using the 2003 O-D travel survey.

Detailed discussions on the TASHA model application and its validation, activity generation behaviours and their trends analysis, and development of population segmentation model are presented in Chapter 4, 5, 6 and 7. Overall discussions and main conclusions of this research are presented below.

8.1 TASHA application and validation

This research applies the activity-based model TASHA in the context of the Montreal Island. TASHA has been developed based on the 1996 TTS (Transportation Tomorrow Survey) for the GTA (Greater Toronto Area). TASHA consists of five components, namely activity generation, location choice, activity scheduling, mode choice, and trip assignment. This research transfers the first three components i.e. activity generation, activity location choice, and activity scheduling to the context of the Montreal Island using the Toronto parameters, estimated using the 1996 TTS of the GTA without any adjustment. It assumes that the activity attributes distributions from the 1996 TTS of the GTA and location choice parameters estimated using the 1996 TTS are the same for the Montreal Island. Similar to Toronto, this research also assumes that these observed activity attributes distributions remain constant over time (thus, temporally transferable). Therefore, this research applies the TASHA model in the context of the Montreal Island using the observed distributions of activity attributes (such as frequency, start time, and duration of different types of activities) and location choice model parameters from the 1996 TTS. Other input data are gathered from the 2003 O-D travel survey, and the 2001 Canadian Census.

TASHA simulates daily schedules of activities (individual and joint) for a 24-hour typical weekday for all individuals in the Montreal Island. This research evaluates the spatial transferability of the TASHA model by comparing TASHA simulated to observed activity attributes (activity frequency, start time, duration, and distance) for five different activities (i.e. work, school, shopping, other, and return to home) at three different levels of aggregation (macro-, meso-, and micro-level). Few studies assess the spatial transferability of the activity-based models and validate such models at different levels; this research contributes to these under researched area.

This research also systematically compares the TASHA model performance in the Montreal Island and the Toronto Area for the activity attributes (i.e. activity frequency, start time, duration, and distance), which provides better understanding of the model performance.

Validation results at macro- and meso-level seem quite promising; TASHA can successfully replicate activity behaviours of another context, at least for fixed activities (work, school) with few exceptions. However, we observe large differences between the simulated and the observed attributes in some cases at different levels, mostly for flexible activities (shopping and other). The MTQ in Montreal generally conducts micro-level validation of a model. The utility function for the current transit assignment model in Montreal is calibrated using the individual observed transit trips, for which the itineraries are available in the O-D survey. However, the need for model validation at such level depends on what purpose the model would be used for.

This research recommends that the TASHA model could be implemented in a new developing area where the dataset is not still available. Prior studies (for instance (Bowman et al., 2014)) have suggested that transferring a model developed based on a large sample from a comparable region is better than estimating a new model using a smaller local sample. In a new developing area where the dataset is not still available, it requires time, cost and expertise to conduct a household survey as well as develop a new model. As this study provides substantial evidence for the spatial transferability of TASHA, at least in case of fixed activities (i.e. work, school), we suggest to apply such model in the new developing area. However, we suggest re-estimating the model parameters and using local activity attribute distributions (frequency, start time and duration) if available from a large sample, for the transfer of the TASHA model from one context to another. In addition, if there is only a small sample available in an area, local updating could be a useful option to spatially transfer the TASHA model to that context. Several researches have shown that local updating with a relatively cost-effective sample size significantly improves the quality of transferred data (Karasmaa, 2007; Rashidi, Auld, & Mohammadian, 2013; Zhang & Mohammadian, 2008). Among other techniques used in model transferability, Bayesian updating is the most popular updating technique (Zhang & Mohammadian, 2008). The transferred distributions (i.e. the Toronto distributions) of the TASHA model can be updated by applying this technique using the available small local sample of the area.

The TASHA model has been developed based on the trip diary data (i.e. 1996 TTS of the GTA) and the model application in this research also relies on the trip diary data (i.e. the 2003 O-D travel survey of Montreal). Both trip diary data are collected by a traditional survey method (i.e. telephone interview), which is facing several challenges (for instance response rate decreases, sampling frame limitations, and high costs) (Bonnel, Lee-Gosselin, Madre, & Zmud, 2009). However in recent years due to advancement in technology, it is possible to collect data by several improved data collection methods (for instance smart phone-based, GPS-based, and web-based surveys) (Bourbonnais & Morency, 2013; Cottrill et al., 2013; Stopher, Prasad, Wargelin, & Minser, 2013). But, these methods are still evolving and require more research to understand their benefits, problems and challenges (Zmud, Lee-Gosselin, Munizaga, & Carrasco, 2013). In addition, the use of technology by travellers on a daily basis generates a massive amount of data (big data) at low cost (Colak, Alexander, Alvim, Mehndiretta, & González, 2015). These huge longitudinal data offer opportunity to develop more dynamic models and research is emerging towards using such data in travel demand modelling (Pozdnoukhov, 2015; Toole et al., 2015). However, more research is required on how to adapt the existing travel demand modelling frameworks to utilize such data sources (Toole et al., 2015).

This research also discusses the potential reasons behind the large differences found in some cases at different levels, as discussed below:

- The large variations most likely indicate real differences in behaviours exist between individuals of the Montreal Island and the Toronto Area. These two cities are distinctly different in terms of socio-demographic, cultural, economic development, employment pattern, commuting pattern and so on, as indicated in the literature (Balakrishnan, Maxim, & Jurdi, 2005; Heisz, 2006; Rose, 1999; Shearmur, 2006). To further understanding, this research also compares some key socio-demographic characteristics between the Montreal Island (using 2003 O-D survey) and the Toronto Area (using 1996 TTS) and observes significant differences in their characteristics (for instance age groups, employment status, possession of driver's license, possession of transit pass, car ownership, and presence of children) between these two cities. These socio-demographic and other differences may lead to the differences in activity-travel behaviours between these cities. The scale difference between the Toronto Area (7,125 square km) and the Montreal Island (500 square km) may also lead to the large differences observed in the validation results. In addition, differences in built environments in these two cities could also be another factor behind the large differences found in the validation results, as activity participation is highly correlated with the built environment (Ewing & Cervero, 2010; Gehrke & Welch, 2016; Merlin, 2014). There are many built environment factors (for instance land use, population density, accessibility, pedestrian-friendly environment, transit friendliness and so on) which have impacts on activity participation; however we did not examined the associated literature as it was not the purpose of this research.

- This research uses the model parameters from Toronto settings from an earlier year (i.e. 1996) to simulate activity schedules of the individuals for the year of 2003, while transferring the TASHA model to the context of the Montreal Island. We assume that the model parameters (observed activity attributes distributions and location choice parameters) remain constant over time. However, this research further investigates and demonstrates that the observed distributions of activity attributes (activity frequency, start time and duration) are changing over time in the GMA. In addition, it discusses the possible reasons for these changing trends in terms of demographic, socio-economic, land use, technology, and policy changes in the GMA. It is also useful to discuss that there has been a significant increase in information and communication technologies (ICT)-use (for instance use of smartphone and other communication media, mobile computing, access to internet and use, and wireless networks) during the period from 1996 to 2003. The substantial increase in ICT-use may lead to changes in individuals' activity-travel decisions (for instance changes in location choice, timing and duration of activities) and also result in new patterns of activity and travel in space and time (Kwan, 2002; Kwan, Dijst, & Schwanen, 2007). However to date, there has been a limited impact of the ICT-use on activity-travel decisions; it is expected that the impact would be stronger in the near future (Rasouli & Timmermans, 2014).

- We also assume that the activity attributes distributions in the Toronto Area and the Montreal Island are the same and therefore uses the activity attributes distributions from the 1996 TTS of the GTA while applying the TASHA model in the context of the Montreal Island. However, all 262 distributions might not be the same in both cities. This research further investigates this hypothesis and compares the similar distributions between the Montreal Island (distributions from the 2003 O-D survey) and the Toronto Area (distributions from the 1996 TTS) for three activity attributes (frequency, start time and duration). The comparison mainly focuses on the distributions for shopping and other activities between two cities, as large differences are mostly found for some activity attributes of these activities. The analysis indicates that activity frequency distributions of these activities are mostly similar in both cities, while start time and duration distributions are mostly different. The finding of this further investigation clearly reveals the behavioural differences in activity-travel patterns between both cities.
- In TASHA, the population segmentation in each activity type to develop activity attributes distributions is done by cross-classification of a set of variables such as activity type, person, household and schedule attributes. The segmentation was based on intuition and testing. Thus, there could be further improvement in population segmentation procedure to develop activity attributes distributions, which may improve the model performance.
- The rule-based model, TASHA applies a single set of strategies of activity scheduling for all individuals in a household. However, it is quite plausible that different people may have different strategies for scheduling their activities. Therefore, there could be further improvement in the TASHA modelling framework using different rules of scheduling for different population segments.

The TASHA application results and the analysis of the potential reasons behind the large differences found in some cases at different levels direct us towards further investigations and/or improvements of some elements of the activity-based modelling framework (for instance, the evolution of activity generation attributes over time and development of more systematic procedure for population segmentation), as discussed below.

8.2 Activity generation behaviours and their trends

Activity-based models, in general, consist of two major components (activity generation and activity scheduling). However, in several activity-based modelling frameworks, activity generation models get relatively less attention than scheduling models. In the activity-based model TASHA, activity generation component also receives less attention than activity scheduling component.

TASHA employs Monte-Carlo simulation to generate activity patterns based on the observed distributions of activity attributes (frequency of different types of activity, start time and duration) from the 1996 TTS of the GTA and the distributions of activity generation attributes are assumed to remain constant over time (thus, temporally transferable). We keep the same assumption and thus use these observed distributions of activity attributes while transferring the TASHA model in the context of the Montreal Island. However, this research further examines the hypothesis of temporal stability of activity generation attributes over time. It empirically investigates the changes in the distributions of activity generation attributes (frequency, start time and duration) over a 10-year period using the 1998, 2003 and 2008 O-D travel surveys of the GMA. The trend analysis at three time points in a 10-year period reveals that distributions of activity attributes for work, school, shopping, and other activities are significantly changing over time. Activity behaviours of individuals and households are becoming complex for several reasons such as demographic and socio-economic changes, growing congestion, introducing new transport facilities and services, technology changes and innovative policy instruments (for instance TDM, ITS technology and, HOV lanes) (Shiftan & Ben-Akiva, 2011). Thus, rather than assuming the stability in activity attributes distributions over time, we suggest preparing activity attributes for the application of an activity-based model, TASHA, such that they reflect temporal changes in travel behaviour of the GMA (for instance by developing and integrating an improved activity generation model, sensitive to the changes in activity behaviours).

This research also focuses on how the activity attributes vary with individuals' socio-demographic characteristics, which explores some interesting trends of the GMA. This research also conducts a comparative analysis of key socio-demographic characteristics and land-uses in the GMA over a 10-year period, which also reveals some important changes in socio-demographic characteristics and land uses over a 10-year period.

8.3 Population segmentation to develop activity generation attributes

Activity-based models are often estimated for the entire population or pre-defined segments. In TASHA, population segmentation to develop activity attributes distributions was also done by cross-classification of a set of variables such as activity type, person, household and schedule attributes; this process was based on intuition and testing. We assume that if we systematically segment the population to develop activity attributes distributions, it may improve the TASHA model performance.

This research develops a methodology combining two methods (multiple sequence alignment method (SAM) and multinomial logit model (MNL)) to segment the population of the Island of Montreal using the 2003 O-D travel survey. The analysis carries on in two steps. First, it applies the multiple SAM to segment the population based on the similarities in individuals' daily activity patterns (represented by three activity attributes, frequency, start time and duration). Second, it estimates the multinomial logit models for the distinct segments of individuals based on their socio-demographic characteristics.

Eleven distinct segments (*Work 1, Work 2, Work 3, School 1, School 2, Shopping 1, Shopping 2, Other 1, Other 2, Home* and *Outlier*) are obtained from the cluster analysis using the multiple SAM based on the similarities in individuals' daily activity patterns. Multinomial logit models are estimated for these segments of individuals (excluding *Outlier* which includes a variety of uncommon activity sequences) based on socio-demographic characteristics. The population segmentation model with two basic demographic variables (age groups with 5-year interval and gender) performs better than the other estimated models and provides useful insights into the determinants of belonging to different segments. The population segmentation model is also validated using a holdout sample and the validation results indicate that the model provides quite reasonable prediction.

The systematic procedure of population segmentation shown in this research contributes in several aspects. It provides more behaviourally driven segments, which are obtained based on different dimensions of activity patterns (frequency, start time, and duration). This procedure considers both activity patterns and socio-demographic characteristics of the individuals to systematically segment the population. In addition, the population segmentation model is quite simple regarding data requirements (variables, age and gender), and estimation, thus it would be easier to use this model for prediction purposes. This model can also be used for other purposes (for instance to identify a relevant group of interest to simulate their activity schedules and/or conduct detail investigation of their travel behaviours).

CHAPTER 9 CONCLUSIONS AND RECOMMENDATIONS

This research aims to contribute in the enhancement of the current modelling approach of travel demand of the Greater Montreal Area (GMA) by applying the activity-based model TASHA in this region. It is also an effort to validate and enhance the activity-based travel demand modelling framework by demonstrating a validation procedure and proposing some improvements. The ultimate goal of this research is to increase the practical application of activity-based travel demand models in the real world contexts. The following sub-sections include research contributions, research limitations, some perspectives for future research, and final remarks.

9.1 Research contributions

The results obtained from this research contribute in several aspects, as discussed below:

- Little research assesses the spatial transferability, as a validation test, of activity-based models. This research empirically assesses the spatial transferability of the activity-based model, TASHA and contributes in this under-researched area. It provides useful insights in transferring an activity-based model (i.e. TASHA) to another context (for instance Montreal). Other activity-based travel demand models can be validated using such test. If the models are transferable, those could be implemented in other contexts; it would save time, cost and expertise required to develop a model for those contexts.
- Validations of activity-based models are mostly done at the macro-level (aggregation of the entire population). A few activity-based models focus on the validations at different levels (such as aggregation by population segments). This research demonstrates a validation procedure of an activity-based model (i.e. TASHA) at three different levels of aggregation (macro-level (aggregation of the entire population), meso-level (aggregation by population segments by age group and gender, and by home location), and micro-level (individuals)). This validation procedure can also be adapted for other activity-based travel demand models.
- Observed distributions of activity attributes (activity frequency, start time and duration) from the 1996 TTS used as inputs in the activity generation component of the TASHA model are assumed to remain constant over time. This research examines this hypothesis by examining

the changes in the distributions of activity generation attributes over time in the GMA. It reveals that the activity attributes distributions for work, school, shopping, and other activities are significantly changing over time. It also investigates the possible reasons for the changes in the activity attributes distributions over time and reveals some useful insights.

- In the current version of TASHA, population segmentation to develop activity attributes distributions was based on intuition and testing, it was not based on any systematic segmentation process. This research develops a methodology to segment the population based on similarity in individuals' activity patterns (represented by three activity attributes, activity frequency, start time, and duration) and socio-demographic characteristics to develop activity attribute distributions for an activity-based model. The simple population segmentation model can be integrated in the TASHA modelling or other activity-based modelling frameworks to segment the population to develop activity attributes distributions. In addition, it can be applied to identify a relevant group of interest to simulate their activity schedules and/or conduct detail investigation of their travel behaviours.
- The MTQ in the GMA still uses the trip-based modelling tools, which are a combination of aggregate and disaggregate approaches. However, trip-based approach is criticized for several reasons and activity-based approach is recognized as a more powerful framework to better understand the behavioural basis for individual decisions in participation in activities in space and time. But, there has been no effort of developing or applying such modelling approach in this region. This research explores the opportunity of implementing an activity-based model in Montreal by applying the activity-based model, TASHA in this region.

9.2 Research limitations

Research limitations are discussed below:

- TASHA has been transferred to the context of the Montreal Island using the observed distributions of activity attributes (such as frequency, start time, and duration of different types of activities) and location choice model parameters from the 1996 TTS (earlier year compared to 2003) of the GTA. We assume that the activity attributes distributions from the 1996 TTS of the GTA and location choice parameters estimated using the 1996 TTS are the same for the Montreal Island. Similar to Toronto, this research also assumes that these observed activity attributes distributions remain constant over time. However, this research further investigates and clearly reveals the behavioural differences in activity-travel patterns between the Montreal Island and the Toronto Area. Therefore, the TASHA model must be applied in this region using model parameters and activity attributes distributions (frequency, start time and duration) from the local dataset. However, this research also demonstrates that the observed distributions of activity attributes (activity frequency, start time and duration) are changing over time in the GMA. The attributes must be prepared for the activity-based modelling framework (i.e. TASHA) such that they reflect temporal changes in travel behaviour resulting from other changes in the GMA.

- TASHA has been applied in the context of the Montreal Island (500 square km), which is 14 times smaller than the Toronto Area (7,125 square km).
- In the TASHA modelling framework, feedback (travel time) goes from trip assignment component to activity scheduling and mode choice components (Roorda et al., 2008). However, this research only focuses on the application of the first three components i.e. activity generation, activity location choice, and activity scheduling of the TASHA modelling framework without integrating or coupling with trip assignment model. However, if simulated travel times on the network need a change in departure time or more extreme activity rescheduling decisions; this approach might result in inconsistencies (Rasouli & Timmermans, 2014).
- Policy sensitivity is one of the key motivations in the development of activity-based models. However, this research does not assess the policy sensitivity of the TASHA model.
- This research examines the trends of activity attributes distributions over time and which activities and socio-demographic cohorts are experiencing the important changes. It reveals several interesting trends however cannot provide the reasons for some of the observed phenomena, which requires further research.
- Multinomial logit models are estimated to develop the population segmentation model in this research. However, we have observed similarities between some segments; it is possible that the multinomial logit model may not be the best method due to its independence from irrelevant alternatives (IIA) property. Nested logit models might be more appropriate.

- The validation of the population segmentation model is done using the 2003 O-D survey (the same year on which the model has been developed).

9.3 Directions for future research

Some perspectives for future research are discussed below:

- TASHA application in the Montreal context using local parameters and integrating the population segmentation model

We assume that the model performance of TASHA to simulate individuals' activity schedules would be better in Montreal if we re-estimate model parameters (location choice parameters) and use activity attributes distributions (frequency, start time and duration) from the local dataset (i.e. 2003 O-D survey). This research already prepares the activity attributes distributions using the 2003 O-D survey to compare the Montreal and Toronto distributions. Future research could calibrate the location choice parameters using the Montreal dataset. It could then apply the TASHA model in the context of the Montreal Island using the local parameters and activity attributes distributions and further validates the model performance.

We also assume that activity attributes distributions developed based on population segments, obtained by a systematic process may improve the TASHA model performance. This research develops a simple population segmentation model; future research could integrate this model within the TASHA modelling framework to segment the population to develop local activity attributes distributions. It could then repeat the application process in Montreal using these activity attributes distributions and further evaluate the model performance. However, population segmentation model could also be validated using a future year dataset (for instance, 2008 O-D survey of Montreal). In addition, the population segmentation model could further be improved by estimating the nested logit models, as we observe similarities between some segments (for instance same dominant activity for the segments *Work 1*, *Work 2* and *Work 3*).

This research applies the TASHA model in the context of the Montreal Island; future research could also extend the model application to the Greater Montreal Area (GMA).

Future research could also integrate trip assignment model with the activity-based travel demand model (i.e. TASHA) and use feedbacks from the network model, which could lead to more consistent outputs.

- Developing an activity generation model

This research indicates that the observed distributions of activity attributes (activity frequency, start time and duration) are changing over time in the GMA for several reasons. Rather than using the observed activity attributes distributions in the generation component of TASHA, future research could develop an improved activity generation model, sensitive to the changes in individuals' activity behaviours of the GMA.

- Policy sensitivity of the TASHA model

Policy sensitivity towards emerging policies was one of the key motivations in development of activity-based models. In order to increase application of such models in practice, it is also important to assess the policy sensitivity of such models (such as TASHA). Therefore, future research could examine some policy scenarios with the adapted TASHA with parameters calibrated using the local datasets from the Montreal Island. Some policy scenarios can be tested, including:

- Demographic scenarios (such as impact of ageing population),
- Land-use scenarios (such as introducing new shopping mall, concentrated development in a centre, mixed development), and
- Demand-oriented scenarios (such as alternative working hours (for instance flexible working hours), HOV lanes (for instance ridesharing), parking restrictions, and congestion pricing in AM period).

- Further investigation of several observed trends

This research reveals that the activity attributes distributions of different activities are changing over time and highlights several interesting trends, which require further investigation to better understand some phenomena. For example, this research indicates that the work activity pattern in a typical weekday is changing over time. Work activity frequency "1" has increased and frequency "2+" decreased over time; workers, (especially more men than women) less often return home or do other activities during lunch hour. Future research could examine differences in work activity attributes for each of the five days of the week and their changes over time.

9.4 Final remarks

Activity-based models, which are recognized as a more powerful framework in capturing individuals' activity behaviours are the next generation travel demand models. It is recommended for Montreal to make a shift from trip-based to activity-based approach of travel demand modelling. The activity-based model (i.e. TASHA) application in this region provides some promising results along with some limitations. It is felt that with some improvements (for instance using local parameters and activity attributes distributions, and integrating newly developed population segmentation model), the TASHA model would perform better in simulating individuals' activity schedules in Montreal.

REFERENCES

Agence Métropolitaine de Transport [AMT]. (2015). Portrait de la mobilité. Retrieved January 14, 2015, from https://www.amt.qc.ca/fr/a-propos/portrait-mobilite

Arentze, T., Hofman, F., Van Mourik, H., & Timmermans, H. (2002). Spatial transferability of the Albatross model system: Empirical evidence from two case studies. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 1-7.

Arentze, T., & Timmermans, H. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological, 38*(7), 613-633. doi: 10.1016/j.trb.2002.10.001

Arentze, T., & Timmermans, H. (2008). *ALBATROSS: Overview of the model, application and experiences*. Paper presented at the Innovations in Travel Modeling 2008 Conference, Portland, Oregon, USA.

Arentze, T., & Timmermans, H. (2009). A need-based model of multi-day, multi-person activity generation. *Transportation Research Part B: Methodological, 43*(2), 251-265. doi: 10.1016/j.trb.2008.05.007

Arminger, G. (1994). Specification and estimation of mean structures: Regression models. In G. Arminger, C. C. Clogg & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* NY, USA: Plenum Press.

Auld, J., & Mohammadian, A. K. (2012). Activity planning processes in the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Research Part A: Policy and Practice, 46*(8), 1386-1403. doi: http://dx.doi.org/10.1016/j.tra.2012.05.017

Auld, J., & Mohammadian, A. K. (2009). Framework for the development of the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Letters: The International Journal of Transportation Research, 1*(3), 243-253.

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

Badoe, D. A., & Miller, E. J. (1998). An automated segmentation procedure for studying variations in mode choice behavior. *Journal of Advanced Transportation, 32*, 190-215.

Balakrishnan, T. R., Maxim, P., & Jurdi, R. (2005). Social class versus cultural identity as factors in the residential segregation of ethnic groups in Toronto, Montreal and Vancouver for 2001. *Canadian Studies in Population, 32*(2), 203-227.

Beinborn, E. A. (1995). A transportation modeling primer. Retrieved January 10, 2010, from http://www4.uwm.edu/cuts/utp/models.pdf

Ben-Akiva, M. E., & Bowman, J. (1998). Activity-based travel demand model systems. In P. Marcotte & S. Nguyen (Eds.), *Equilibrium and Advanced Transportation Modeling* (pp. 27-46). Boston/ Dordrecht/London: Kluwer Academic Publishers.

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, Massachusetts, USA: The MIT Press.

Berkhin, P. (2006). A survey of clustering data mining techniques. In J. Kogan, C. Nicholas & M. Teboulle (Eds.), *Grouping Multidimensional Data: Recent Advances in Clustering* (pp. 25-71). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bhat, C. R. (1995). A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research Part B: Methodological, 29*(6), 471-483. doi: http://dx.doi.org/10.1016/0191-2615(95)00015-6

Bhat, C. R. (1998a). *Activity-based travel demand modeling.* Paper presented at the First Symposium on Integrated Land Use-Transport Models, Portland, Oregon, USA.

Bhat, C. R. (1998b). A model of post home-arrival activity participation behavior. *Transportation Research Part B: Methodological, 32*(6), 387-400.

Bhat, C. R. (1999). *A comprehensive and operational analysis framework for generating the daily activity travel profiles of workers.* Paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC, USA.

Bhat, C. R., Guo, J. Y., Srinivasan, S., & Sivakumar, A. (2004). Comprehensive econometric microsimulator for daily activity-travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1894*, 57-66. doi: 10.3141/1894-07

Bhat, C. R., & Koppelman, F. S. (1993). A conceptual framework of individual activity program generation. *Transportation Research Part A: Policy and Practice, 27*(6), 433-446.

Bhat, C. R., & Koppelman, F. S. (1999a). Activity-based modeling of travel demand. In R. Hall (Ed.), *Handbook of Transportation Research*. New York: Kluwer Academic Publisher.

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

Bhat, C. R., & Lawton, K. T. (2000). Passenger travel demand forecasting *Millennium Paper*. Washington D.C., USA: Transportation Research Board, National Research Council.

Bhat, C. R., Paleti, R., Pendyala, R. M., & Goulias, K. G. (2013). SimAGENT activity-based travel demand analysis *SimAGENT Core Models* (Vol. 4): GeoTrans Laboratory, University of California, Santa Barbara.

Bierlaire, M. (2015). BisonBiogeme 2.4: estimating a first model. Retrieved November 4, 2015, from http://biogeme.epfl.ch

Bonnel, P. (2004). *Prévoir la demande de transport*. Paris, France: Presses Ponts et Chaussées.

Bonnel, P., Lee-Gosselin, M., Madre, J.-L., & Zmud, J. (2009). Keeping up with a changing world: Challenges in the design of transport survey methods *Transport Survey Methods: Keeping Up with a Changing World* (pp. 3-13): Emerald.

Bourbonnais, P.-L., & Morency, C. (2013). Web-Based Travel Survey: A Demo. In J. Zmud, M. Lee-Gosselin & J. A. Carrasco (Eds.), *Transport Survey Methods: Best practice for decision making* (pp. 207-223). United Kingdom Emerald Group Publishing Limited, Bingley.

Bowman, J. L. (1998). *The day activity schedule approach to travel demand analysis.* (Doctor of Philosophy), Massachusetts Institute of Technology, Cambridge, Massachusetts.

Bowman, J. L., & Ben-Akiva, M. E. (2000). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice, 35*(1), 1-28.

Bowman, J. L., Bradley, M., Shiftan, Y., Lawton, T. K., & Ben-Akiva, M. E. (1999). Demonstration of an activity based model system for Portland. *World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research, 3*.

Bowman, J. L., Bradley, M. A., Castiglione, J., & Yoder, S. L. (2014). *Making advanced travel forecasting models affordable through model transferability.* Paper presented at the 93rd Transportation Research Board Annual Meeting, Washington D.C., USA.

Bradley, M. A., & Bowman, J. L. (2006). *A summary of design features of activity-based microsimulation models for U.S. MPOs.* Paper presented at the TRB Conference on Innovations in Travel Demand Modeling, Austin, Texas.

Bradley, M. A., & Vovsha, P. (2005). A model for joint choice of daily activity pattern types of household members. *Transportation, 32*(5), 545-571.

Cascetta, E. (2009). *Transportation system analysis: Models and applications* (Vol. Second): Springer.

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

Chapleau, R. (1986). *Transit network analysis and evaluation with a totally disaggregate approach.* Paper presented at the Research for Tomorrow's Transport Requirements. Proceedings of the Fourth World Conference on Transport Research.

Chapleau, R., Allard, B., Trépanier, M., & Morency, C. (2001). Les logiciels d'enquête transport comme instruments incontournables de la planification analytique. *Recherche - Transports - Sécurité, 70*, 59-77. doi: http://dx.doi.org/10.1016/S0761-8980(01)90227-4

Childress, S., Sabina, E., Kurth, D., Rossi, T., & Malm, J. (2010). *DRCOG activity-based model calibration/ validation.* Paper presented at the Innovations in Travel Modeling Conference, Transportation Research Board, Tempe, Arizona.

Colak, S., Alexander, L. P., Alvim, B. G., Mehndiretta, S. R., & González, M. C. (2015). Analyzing cell phone location data for urban travel: Current methods, limitations, and opportunities. *Transportation Research Record: Journal of the Transportation Research Board*.

Cottrill, C. D., Pereira, F. C., Zhao, F., Dias, I. F., Lim, H. B., Ben-Akiva, M. E., & Zegras, P. C. (2013). Future mobility survey: Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board, 2354*, 59-67.

Data Management Group. (2014). Transportation Tomorrow Survey (TTS) 2011: Design and conduct of the survey: University of Toronto, Toronto, Canada.

Davidson, W., Vovsha, P., Freedman, J., & Donnelly, R. (2010, September 29 – October 1, 2010). *CT-RAMP family of activity-based models.* Paper presented at the Australasian Transport Research Forum, Canberra, Australia.

Dibbelt, J., Pajor, T., Strasser, B., & Wagner, D. (2013). Intriguingly simple and fast transit routing. In V. Bonifaci, C. Demetrescu & A. Marchetti-Spaccamela (Eds.), *Experimental Algorithms: 12th International Symposium, SEA 2013, Rome, Italy, June 5-7, 2013. Proceedings* (pp. 43-54). Berlin, Heidelberg: Springer Berlin Heidelberg.

Doherty, S. T., Nemeth, E., Roorda, M. J., & Miller, E. J. (2004). Design and assessment of the Toronto area computerized household activity scheduling survey. *Transportation Research Record: Journal of the Transportation Research Board, 1894*, 140-149.

Domencich, T., & McFadden, D. L. (1975). *Urban Travel Demand - A Behavioural Analysis*. Amsterdam: North-Holland Publishing Co.

Eberhard, L. K. (2002). *A 24-h household-level activity based travel demand model for the GTA.* (Masters Thesis), University of Toronto, Toronrto, Canada.

Ettema, D., Bergers, A., & Timmermans, H. (1993). Simulation model of activity scheduling behaviour. *Transportation Research Record: Journal of the Transportation Research Board, 1413*, 1-11.

Ettema, D., Borgers, A., & Timmermans, H. (1996). SMASH (Simulation Model of Activity Scheduling Heuristics): Some simulations. *Transportation Research Record: Journal of the Transportation Research Board, 1551*, 88-94.

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

Florian, M., & Morosan, C. D. (2014). On uniqueness and proportionality in multi-class equilibrium assignment. *Transportation Research Part B: Methodological, 70*, 173-185. doi: http://dx.doi.org/10.1016/j.trb.2014.06.011

Fried, M., Havens, J., & Thall, M. (1977). Travel behavior - A synthesized theory *Final Report*. Washington: NCHRP, Transportation Research Board.

Gangrade, S., Kasturirangan, K., & Pendyala, R. M. (2000). Coast-to-coast comparison of time use and activity patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1718*, 34-42.

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

Gehrke, S. R., & Welch, T. F. (2016). The built environment determinants of activity participation and walking near the workplace. *Transportation*, 1-16. doi: 10.1007/s11116-016-9687-5

Golledge, R. G., Kwan, M.-P., & Gärling, T. (1994). Computational process modeling of household travel decisions using a geographic information system. *Papers in Regional Science, 73*(2), 99-117.

Golob, T. F. (2001). Joint models of attitudes and behavior in evaluation of the San Diego I-15 congestion pricing project. *Transportation Research Part A: Policy and Practice, 35*(6), 495-514. doi: http://dx.doi.org/10.1016/S0965-8564(00)00004-5

Grist, D., & Demetsky, M. J. (1981). *Energy sensitivity of transportation planning techniques*. Virginia, United States: Virginia Highway and Transportation Research Council, US Department of Transportation, Federal Highway Administration.

Habib, K. M. N. (2007). *Modelling activity generation processes.* (Doctor of Philosophy Dissertation), University of Toronto, Toronto, Canada.

Habib, K. M. N., & Miller, E. J. (2008). Modeling daily activity program generation considering within-day and day-to-day dynamics in activity-travel behaviour. *Transportation, 35*(4), 467-484. doi: 10.1007/s11116-008-9166-8

Habib, K. M. N., & Miller, E. J. (2009). Modelling activity generation: A utility based model for activity agenda formation. *Transportmetrica, 5*(1), 3-23. doi: 10.1080/18128600802591384

Hagerstrand, T. (1970). What about people in regional science? . *Papers in Regional Science, 24*(1), 7-21.

Hanson, S. (1982). The determinants of daily travel-activity patterns: Relative location and socio-demographic factors. *Urban Geography, 3*(3), 179-202. doi: 10.2747/0272-3638.3.3.179

Heisz, A. (2006). Canada's global cities: Socio-economic conditions in Montréal, Toronto and Vancouver *Trends and Conditions in Census Metropolitan Areas*. Ottawa: Statistics Canada.

Ishaq, R., Bekhor, S., & Shiftan, Y. (2012a). A flexible model structure approach for discrete choice models. *Transportation, 40*(3), 609-624. doi: 10.1007/s11116-012-9431-8

Ishaq, R., Bekhor, S., & Shiftan, Y. (2012b). Flexible model structure for discrete-choice models. *Proceedings of the Institution of Civil Engineers: Transport, 165*(1), 39-47.

Joh, C.-H., Arentze, T., Hofman, F., & Timmermans, H. (2002). Activity pattern similarity: a multidimensional sequence alignment method. *Transportation Research Part B: Methodological, 36*(5), 385-403. doi: http://dx.doi.org/10.1016/S0191-2615(01)00009-1

Joh, C.-H., Arentze, T., & Timmermans, H. (2001). Pattern recognition in complex activity-travel patterns: A comparison of Euclidean distance, signal processing theoretical, and multidimensional sequence alignment methods. *Transportation Research Record: Journal of the Transportation Research Board, 1752*, 16-22.

Jones, P. M., Dix, M. C., Clarke, M. I., & Heggie, I. G. (1983). *Understanding travel behaviour*. Aldershot, UK: Gower.

Jones, P. M., Koppelman, F., & Orfeuil, J.-P. (1990). Activity analysis: State of the art and future directions. In P. Jones (Ed.), *Developments in Dynamic and Activity Based Approaches to Travel Analysis* (pp. 34-55). Avebury: Oxford Studies in Transport.

Jonnalagadda, N., Freedman, J., Davidson, W. A., & Hunt, J. D. (2001). Development of microsimulation activity-based model for San Francisco: Destination and mode choice models. *Transportation Research Record: Journal of the Transportation Research Board, 1777*, 25-35.

Karasmaa, N. (2001). *The spatial transferability of the Helsinki Metropolitan Area mode choice models.* Paper presented at the 5th workshop of the Nordic Research Network on Modeling Transport, Land-use, and the Environment, Nynashamn, Sweden.

Karasmaa, N. (2007). Evaluation of transfer methods for spatial travel demand models. *Transportation Research Part A: Policy and Practice, 41*(5), 411-427. doi: http://dx.doi.org/10.1016/j.tra.2006.09.009

Kitamura, R. (1996). *Applications of models of activity behavior for activity-based demand forecasting*. Paper presented at the Activity-Based Travel Forecasting Conference, New Orleans, Louisiana.

Kitamura, R., & Fujii, S. (1998). Two computational process models of activity-travel behavior. In T. Garling, T. Laitila & K. Westin (Eds.), *Theoretical Foundations of Travel Choice Modeling* (pp. 251-279). Bingley, UK: Emerald Group Publishing.

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

Kitamura, R., Pas, E. I., Lula, C. V., Lawton, T. K., & Benson, P. E. (1996). The sequenced activity mobility simulator (SAMS): An integrated approach to modeling transportation, land use and air quality. *Transportation, 23*(3), 267-291. doi: 10.1007/BF00165705

Koppelman, F. S., & Bhat, C. R. (2006). A self instructing course in mode choice modeling: Multinomial and nested logit models. Washington D.C., USA: US Department of Transport, Federal Transit Administration.

Koppelman, F. S., & Wilmot, C. G. (1982). Transferability analysis of disaggregate choice models. *Transportation Research Record: Journal of the Transportation Research Board, 895*, 18-24.

Kuppam, A. R., & Pendyala, R. M. A structural equations analysis of commuters' activity and travel patterns. *Transportation, 28*(1), 33-54. doi: 10.1023/a:1005253813277

Kurth, D. L., Childress, S., Sabina, E. E., & Rossi, T. (2006). *Proposed validation and sensitivity testing of Denver region activity-based models.* Paper presented at the Innovations in Travel Demand Modeling Conference, Transportation Research Board, Austin, TX.

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

Kwan, M.-P. (2002). Time, information technologies and the geographies of everyday life. *Urban Geography, 23*, 471-482.

Kwan, M.-P., Dijst, M., & Schwanen, T. (2007). The interaction between ICT and human activity-travel behavior. *Transportation Research Part A: Policy and Practice, 41*(2), 121-124. doi: http://dx.doi.org/10.1016/j.tra.2006.02.002

Le Vine, S., Sivakumar, A., Roorda, M. J., & Polak, J. W. (2010). *Toward the implementation of an activity-based travel demand model for London: Development of the household activity scheduling module*. Paper presented at the 42nd conference of the Universities' transport study group, Plymouth, United Kingdom.

Lieberman, W., Schumacher, D., Hoffman, A., & Wornum, C. (2001). Creating a new century of transit opportunity: Strategic planning for transit. *Transportation Research Record, 1747*, 60-67.

Mandel, B. N. (1998). *Interdependency of Airport Choice and Travel Demand*. Paper presented at the Symposium "Taking Stock of Air Liberalisation" at the ICAO Montreal, Canada

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in Economics*. New York, USA: Academic Press.

McNally, M. (2007a). The activity-based approach. In D. A. Hensher & K. J. Button (Eds.), *Handbook of transport modelling* (Vol. 2nd edition, pp. 53-70). Oxford, UK: Pergamon.

McNally, M. (2007b). The four-step model. In D. A. Hensher & K. J. Button (Eds.), *Handbook of transport modelling* (Vol. 2nd edition, pp. 35-52). Oxford, UK: Pergamon.

Merlin, L. A. (2014). Can the built environment influence nonwork activity participation? An analysis with national data. *Transportation, 42*(2), 369-387. doi: 10.1007/s11116-014-9554-1

Miller, E. J., & Roorda, M. J. (2003). Prototype model of household activity-travel scheduling. *Transportation Research Record: Journal of the Transportation Research Board, 1831*, 114-121. doi: 10.3141/1831-13

Miller, E. J., Vaughan, J., King, D., & Austin, M. (2015). *Implementation of a "next generation" activity-based travel demand model: The Toronto case*. Paper presented at the Conference of the Transportation Association of Canada (TAC), Charlottetown, PEI, Canada.

Ministère des transports du Québec [MTQ]. (2014). Prévisions de la demande en transport des personnes: Scénario tendanciel Montréal 2008-2031. Montréal, Québec, Canada

Mohammadian, A. K., Auld, J., & Yagi, S. (2009). Recent progress on activity-based microsimulation models of travel demand and future prospects. In B. W. Sloboda (Ed.), *Transportation Statistics* (pp. 151-171): J. Ross Publishing, Inc.

Morency, C., & Trépanier, M. (2013). Enrichissement du processus de modélisation de la demande de transport à l'aide des modèles d'activités. Montreal, QC: Polytechnique Montreal.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48*(3), 443-453. doi: http://dx.doi.org/10.1016/0022-2836(70)90057-4

Ortúzar, J. d. D., & Willumsen, L. G. (2011). *Modelling transport* (Vol. 4th Edition). New York, USA: Wiley.

Páez, A. (2006). Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. *Journal of Transport Geography, 14*(3), 167-176. doi: http://dx.doi.org/10.1016/j.jtrangeo.2005.11.002

Pas, E. I. (1984). The effect of selected sociodemographic characteristics on daily travel-activity behavior. *Environment*

and Planning A, 16, 571-581.

Pendyala, R. M., Kitamura, R., Kikuchi, A., Yamamoto, T., & Fujii, S. (2005). Florida Activity Mobility Simulator: Overview and preliminary validation results. *Transportation Research Record: Journal of the Transportation Research Board, 1921*, 123-130. doi: 10.3141/1921-14

Pimparé, M., & Thiffault, J. (2000). *Une enquête Origine-Destination qui fait du chemin: Tout sur son itinéraire.* Paper presented at the Assemblée annuelle de ta SSC, Recueil de la section des Méthodes d'enquête, MTQ, , Montreal, Canada.

Pinjari, A. R., & Bhat, C. R. (2011). Activity-based travel demand analysis. In A. de Palma, R. Lindsey, E. Quinet & R. Vickerman (Eds.), *Handbook of Transport Economics* (pp. 213-248 (Ch. 210)): Edward Elgar Publishing.

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

Pozdnoukhov, A. (2015). *Activity-based travel demand modeling with cellular data*. Paper presented at the Workshop on the Role of Big Data in Transportation Planning, Sacramento, USA.

Primerano, F., Taylor, M., Pitaksringkarn, L., & Tisato, P. (2008). Defining and understanding trip chaining behaviour. *Transportation, 35*(1), 55-72. doi: 10.1007/s11116-007-9134-8

Rashidi, T. H., Auld, J., & Mohammadian, A. (2013). Effectiveness of Bayesian updating attributes in data transferability applications. *Transportation Research Record: Journal of the Transportation Research Board, 2344*, 1-9.

Rasouli, S., & Timmermans, H. (2014). Activity-based models of travel demand: promises, progress and prospects. *International Journal of Urban Sciences, 18*(1), 31-60.

Recker, W. W. (1995). The household activity pattern problem: General formulation and solution. *Transportation Research Part B: Methodological, 29*(1), 61-77. doi: 10.1016/0191-2615(94)00023-S

Recker, W. W., McNally, M. G., & Root, G. S. (1986a). A model of complex travel behavior: Part I--Theoretical development. *Transportation Research Part A: General, 20*(4), 307-318. doi: 10.1016/0191-2607(86)90089-0

Recker, W. W., McNally, M. G., & Root, G. S. (1986b). A model of complex travel behavior: Part II--An operational model. *Transportation Research Part A: General, 20*(4), 319-330. doi: 10.1016/0191-2607(86)90090-7

Roorda, M. J. (2005). *Activity-based modelling of household travel.* (Doctor of Philosophy), University of Toronto, Toronto, Canada.

Roorda, M. J., Doherty, S. T., & Miller, E. J. (2002). *Operationalizing household activity scheduling models: Addressing assumptions and the use of new sources of behavioural data*. Paper presented at the International Colloquium on the Behavioural Foundations of Integrated Land-Use and Transportation Models: Assumptions and New Conceptual Frameworks, Quebec City, Canada.

Roorda, M. J., & Miller, E. J. (2006). Past Presidents' Award for merit in transportation engineering: Assessing transportation policy using an activity-based microsimulation model of travel demand. *ITE Journal (Institute of Transportation Engineers), 76*, 16-21.

Roorda, M. J., Miller, E. J., & Habib, K. M. N. (2008). Validation of TASHA: A 24-h activity scheduling microsimulation model. *Transportation Research Part A: Policy and Practice, 42*, 360-375. doi: 10.1016/j.tra.2007.10.004

Rose, D. (1999). Urban hierarchies and the changing characteristics of 'urban professionals' in Toronto and Montreal. *Canadian Journal of Regional Science, XXII-1,2*, 133-142.

Ryley, T. (2006). Use of non-motorised modes and life stage in Edinburgh. *Journal of Transport Geography, 14*(5), 367-375. doi: http://dx.doi.org/10.1016/j.jtrangeo.2005.10.001

Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees *Molecular Biology and Evolution, 4*(4), 406-425.

Salvini, P. A., & Miller, E. J. (2005). ILUTE: An operational prototype of a comprehensive microsimulation model of urban systems. *Networks and Spatial Economics, 5*, 217-234.

Saneinejad, S., & Roorda, M. J. (2009). Application of sequence alignment methods in clustering and analysis of routine weekly activity schedules. *Transportation Letters, 1*(3), 197-211.

Schlich, R. (2001). *Analysing intrapersonal variability of travel behavior using the sequence alignment method.* Paper presented at the Proceedings of the AET European Transport Conference, Cambridge, UK.

Schlich, R. (2003). *Homogenous groups of travellers*. Paper presented at the 10th International Conference on Travel Research, Lucerne.

Shearmur, R. (2006). Travel from home: An economic geography of commuting distances in montreal. *Urban Geography, 27*(4), 330-359.

Shiftan, Y., & Ben-Akiva, M. E. (2011). A practical policy-sensitive, activity-based, travel-demand model. *The Annals of Regional Science, 47*(3), 517-541.

Shiftan, Y., Ben-Akiva, M. E., Proussaloglou, K., Jong, G. d., Popuri, Y., Kasturirangan, K., & Bekhor, S. (2003). *Activity-based modeling as a tool for better understanding travel behaviour*. Paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Shiftan, Y., Outwater, M. L., & Zhou, Y. (2008). Transit market research using structural equation modeling and attitudinal market segmentation. *Transport Policy, 15*(3), 186-195.

Sikder, S., Pinjari, A., Srinivasan, S., & Nowrouzian, R. (2013). Spatial transferability of travel forecasting models: A review and synthesis. *International Journal of Advances in Engineering Sciences and Applied Mathematics, 5*(2-3), 104-128. doi: 10.1007/s12572-013-0090-6

Smith, T., & Waterman, M. (1981). Identification of common molecular subsequences. *Journal of Molecular Biology, 147*, 195-197.

Srinivasan, S., & Bhat, C. R. (2005). Modeling household interactions in daily in-home and out-of-home maintenance activity participation. *Transportation, 32*(5), 523-544.

Statistics Canada. (2011). Population and dwelling counts, for Canada and census divisions. *Canada 2011 Census.* Retrieved July 30, 2013, from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=701

Stopher, P. R., Prasad, C., Wargelin, L., & Minser, J. (2013). Conducting a GPS-only household travel survey In J. Zmud, M. Lee-Gosselin & J. A. Carrasco (Eds.), *Transport Survey*

Methods: Best Practice for Decision Making (pp. 91-113). United Kingdom: Emerald Group Publishing Limited, Bingley.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research, 24*, 4876-4882.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research, 22*, 4673-4680.

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies, 58, Part B*, 162-177. doi: http://dx.doi.org/10.1016/j.trc.2015.04.022

Train, K. E. (2003). *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press.

Tremblay, P. (2007). Modélisation des transports à l'échelle régionale : Survol des approches utilisées à Transports-Québec. *42e Congrès de l'AQTR.* Retrieved February 2, 2011, from http://www.aqtr.qc.ca/documents/Congres2007/Conferences2007/42Congres/pdfs/Tremblay_Pierre.pdf

Trépanier, M., Chapleau, R., & Morency, C. (2008). Tools and methods for a transportation household survey. *Journal of the Urban and Regional Information Systems Association (URISA), 20*(1), 35-43.

Tye, W. B., Sherman, L., Kinnucan, M., Nelson, D., & Tardiff, T. J. (1982). Application of disaggregate travel demand models *National Cooperative Highway Research Report 253*. Washington D.C. , USA.: Transportation Research Board, National Research Board, National Research Council.

Vovsha, P., Bradley, M. A., & Bowman, J. L. (2004). *Activity-based travel forecasting models in the United States: Progress since 1995 and prospects for the future* Paper presented at the EIRASS Conference on Progress in Activity-Based Analysis, Vaeshartelt Castle, Maastricht, The Netherlands.

Vovsha, P., Petersen, E., & Donnelly, R. (2004). Impact of intra-household interactions on individual daily activity travel patterns. *Transportation Research Record: Journal of the Transportation Research Board, 1898*, 87-97.

Wilson, C. (1998). Analysis of travel behavior using sequence alignment methods. *Transportation Research Record: Journal of the Transportation Research Board, 1645*, 52-59.

Wilson, C. (1999). Activity patterns of Canadian women: An application of ClustalG sequence alignment software. *Transportation Research Record: Journal of the Transportation Research Board, 1777*, 55-67.

Wilson, C., Harvey, A., & Thompson, J. (2005). *ClustalG: software for analysis of activities and sequential events*. Research Development Initiatives Program, Social Science and Humanities Research Council of Canada, Halifax.

Xu, R., & Wunsch II, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks, 16*, 645-678.

Yagi, S., & Mohammadian, A. K. (2007a). *Microsimulation and policy applications of an activity-based modeling system of travel demand.* Paper presented at the 86th Transportation Research Board Annual Meeting, Washington DC, United States.

Yagi, S., & Mohammadian, A. K. (2007b). *Validation and policy applications of an activity-based microsimulation model of travel demand.* Paper presented at the 11th World Conference on Transportation Research, Berkeley, California.

Yagi, S., & Mohammadian, A. K. (2008a). Joint models of home-based tour mode and destination choices: Applications to a developing country. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 29-40.

Yagi, S., & Mohammadian, A. K. (2008b). Modeling daily activity-travel tour patterns incorporating activity scheduling decision rules. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 123-131.

Yagi, S., & Mohammadian, A. K. (2010). An activity-based microsimulation model of travel demand in the Jakarta Metropolitan Area. *Journal of Choice Modelling, 3*(1), 32-57. doi: http://dx.doi.org/10.1016/S1755-5345(13)70028-9

Yasmin, F., Morency, C., & Roorda, M. J. (2012). *An empirical analysis of change in activity attribute distributions over time*. Paper presented at the 13th International Conference on Travel Behaviour Research organized by the International Association for Travel Behavior Research (IATBR), Toronto, Canada.

Yasmin, F., Morency, C., & Roorda, M. J. (2013). *Spatial transferability of activity-based models: Moving TASHA from Toronto to Montreal*. Paper presented at the 13th World Conference on Transport Research, Rio de Janeiro, Brazil.

Yasmin, F., Morency, C., & Roorda, M. J. (2015a). Assessment of spatial transferability of an activity-based model, TASHA. *Transportation Research Part A: Policy and Practice, 78*, 200-213. doi: 10.1016/j.tra.2015.05.008

Yasmin, F., Morency, C., & Roorda, M. J. (2015b). Trend analysis of activity generation attributes over time. *Transportation*, 1-21. doi: 10.1007/s11116-015-9624-z

Yasmin, F., Morency, C., & Roorda, M. J. (submitted). Macro-, meso-, and micro-level validation of an activity-based travel demand model. *Transportmetrica A: Transport Science*.

Zhang, Y., & Mohammadian, A. (2008). Bayesian updating of transferred household travel data. *Transportation Research Record: Journal of the Transportation Research Board, 2049*, 111-118.

Zmud, J., Lee-Gosselin, M., Munizaga, M. A., & Carrasco, J. A. (2013). Transport surveys: Considerations for decision makers and decision making. In J. Zmud, M. Lee-Gosselin & J. Carrasco (Eds.), *Transport Survey Methods: Best practice for decision making* (pp. 3-19). United Kingdom Emerald Group Publishing Limited, Bingley.

APPENDIX A – PROBABILITY ESTIMATION OF THE POPULATION SEGMENTATION MODELS

Probability estimation of the final population segmentation model, *Model 1*

Model 1 : **Deterministic component of the utility of different segments (variables - age groups and gender)**

$V1_{Work\ 1}$ = - 2.83 - 6.44 * age5-9 -0.39 * age10-14 + 2.05 * age15-19 + 3.16 * age20-24 + 3.60 * age25-29 + 3.71 * age30-34 + 3.85 * age35-39 + 3.70 * age40-44 + 3.94 * age45-49 + 3.57 * age50-54 + 2.98 * age55-59 + 2.44 * age60-64 + 0.44 * men .... (A.1)

$V1_{Work\ 2}$ = - 5.34 - 4.09 * age5-9 - 3.85 * age10-14 + 3.09 * age15-19 + 2.98 * age20-24 + 3.22 * age25-29 + 3.02 * age30-34 + 3.01 * age35-39 + 2.80 * age40-44 + 3.17 * age45-49 + 2.68 * age50-54 + 2.12 * age55-59 + 2.03 * age60-64 + 0.64 * men .... (A.2)

$V1_{Work\ 3}$ = - 6.52 - 2.76 * age5-9 - 2.55 * age10-14 - 3.37 * age15-19 + 1.66 * age20-24 + 2.68 * age25-29 + 2.47 * age30-34 + 3.37 * age35-39 + 3.18 * age40-44 + 1.76 * age45-49 + 1.58 * age50-54 + 2.66 * age55-59 + 2.72 * age60-64 + 0.25 * men .... (A.3)

$V1_{School\ 1}$ = - 6.51 + 9.14 * age5-9 + 9.37 * age10-14 + 8.26 * age15-19 + 6.77 * age20-24 + 5.77 * age25-29 + 5.17 * age30-34 + 4.80 * age35-39 + 3.51 * age40-44 + 3.55 * age45-49 + 2.67 * age50-54 - 4.86 * age55-59 + 1.62 * age60-64 + 0.23 * men ... (A.4)

$V1_{School\ 2}$ = - 10.00 + 8.27 * age5-9 + 8.50 * age10-14 + 9.40 * age15-19 + 8.87 * age20-24 + 7.77 * age25-29 + 7.33 * age30-34 + 6.63 * age35-39 + 5.74 * age40-44 - 3.55 * age45-49 + 5.06 * age50-54 - 4.14 * age55-59 + 5.10 * age60-64 + 0.26 * men .... (A.5)

$V1_{Shopping\ 1}$ = -1.18 - 1.28 * age5-9 - 7.57 * age10-14 - 1.22 * age15-19 -1.04 * age20-24 - 0.44 * age25-29 - 0.40 * age30-34 - 0.05 * age35-39 - 0.39 * age40-44 0.12 * age45-49 - 0.09 * age50-54 + 0.12 * age55-59 - 0.01 * age60-64 - 0.14 * men .... (A.6)

$V1_{Shopping\ 2}$ = -2.87 - 6.15 * age5-9 - 5.92 * age10-14 - 6.78 * age15-19 - 7.70 * age20-24 - 0.55 * age25-29 0.04 * age30-34 0.17 * age35-39 0.12 * age40-44 0.30 * age45-49 - 0.17 * age50-54 - 0.36 * age55-59 - 0.12 * age60-64 - 0.03 * men .... (A.7)

$V1_{Other\ 1}$ = - 1.04 -0.06 * age5-9 - 0.24 * age10-14 - 0.05 * age15-19 + 0.06 * age20-24 - 0.14 * age25-29 + 0.57 * age30-34 + 0.55 * age35-39 + 0.32 * age40-44 + 0.38 * age45-49 + 0.16 * age50-54 + 0.27 * age55-59 + 0.44 * age60-64 + 0.16 * men .... (A.8)

$V1_{Other\ 2}$ = - 3.06 - 5.96 * age5-9 - 5.73 * age10-14 - 0.77 * age15-19 - 1.00 * age20-24 - 1.05 * age25-29 + 0.23 * age30-34 - 0.89 * age35-39 - 1.77 * age40_44 - 0.20 * age45-49 - 0.67 * age50-54 + 0.01 * age55-59 - 0.34 * age60-64 - 0.03 * men .... (A.9)

$V1_{Home}$ = 0 .... (A.10)

Model 1: Probability estimation of each segment

$$P1(Work\ 1) = \frac{e^{(V1_{Work\ 1})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

.... (A.11)

$$P1(Work\ 2) = \frac{e^{(V1_{Work\ 2})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

.... (A.12)

$$P1(Work\ 3) = \frac{e^{(V1_{Work\ 3})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

.... (A.13)

$$P1(School\ 1) = \frac{e^{(V1_{\text{School 1}})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

…………………………………………………………… (A.14)

$$P1(School\ 2) = \frac{e^{(V1_{\text{School 2}})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

…………………………………………………………… (A.15)

$$P1(Shopping\ 1) = \frac{e^{(V1_{\text{Shopping 1}})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

…………………………………………………………… (A.16)

$$P1(Shopping\ 2) = \frac{e^{(V1_{\text{Shopping 2}})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

…………………………………………………………… (A.17)

$$P1(Other\ 1) = \frac{e^{(V1_{\text{Other 1}})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

…………………………………………………………… (A.18)

$$P1(Other\ 2) = \frac{e^{(V1_{\text{Other 2}})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

…………………………………………………………… (A.19)

$$P1(Home) = \frac{e^{(V1_{\text{Home}})}}{e^{(V1_{Work\ 1})} + e^{(V1_{Work\ 2})} + e^{(V1_{Work\ 3})} + e^{(V1_{School\ 1})} + e^{(V1_{School\ 2})} + e^{(V1_{Shopping\ 1})} + e^{(V1_{Shopping\ 2})} + e^{(V1_{Other\ 1})} + e^{(V1_{Other\ 2})} + e^{(V1_{Home})}}$$

…………………………………………………………… (A.20)

where,

$P1(Work\ 1)$ = probability of belonging to the segment *Work 1* (*Model 1*)

$P1(Work\ 2)$ = probability of belonging to the segment *Work 2* (*Model 1*)

$P1(Work\ 3)$ = probability of belonging to the segment *Work 3* (*Model 1*)

$P1(School\ 1)$ = probability of belonging to the segment *School 1* (*Model 1*)

$P1(School\ 2)$ = probability of belonging to the segment *School 2* (*Model 1*)

$P1(Shopping\ 1)$ = probability of belonging to the segment *Shopping 1*(*Model 1*)

$P1(Shopping\ 2)$ = probability of belonging to the segment *Shopping 2* (*Model 1*)

$P1(Other\ 1)$ = probability of belonging to the segment *Other 1* (*Model 1*)

$P1(Other\ 2)$ = probability of belonging to the segment *Other 2* (*Model 1*)

$P1(Home)$ = probability of belonging to the segment *Home* (*Model 1*)

$V1_{Work\ 1}$ = deterministic component of the utility of segment *Work 1* (*Model 1*)

$V1_{Work\ 2}$ = deterministic component of the utility of segment *Work 2* (*Model 1*)

$V1_{Work\ 3}$ = deterministic component of the utility of segment *Work 3* (*Model 1*)

$V1_{School\ 1}$ = deterministic component of the utility of segment *School 1* (*Model 1*)

$V1_{School\ 2}$ = deterministic component of the utility of segment *School 2* (*Model 1*)

$V1_{Shopping\ 1}$ = deterministic component of the utility of segment *Shopping 1* (*Model 1*)

$V1_{Shopping\ 2}$ = deterministic component of the utility of segment *Shopping 2* (*Model 1*)

$V1_{Other\ 1}$ = deterministic component of the utility of segment *Other 1* (*Model 1*)

$V1_{Other\ 2}$ = deterministic component of the utility of segment *Other 2* (*Model 1*)

$V1_{Home}$ = deterministic component of the utility of segment *Home* (*Model 1*)

Detail definitions of the explanatory variables of *Model 1* can be found in Table 7.5 (Chapter 7).

Probability estimation of the population segmentation model, *Model 2*

Model 2 **: Deterministic component of the utility of different segments (variables - age groups by gender)**

$V2_{Work\ 1}$ = - 3.29 - 9.32 * m5-9 - 8.89 * m10-14 + 2.86 * m15-19 + 3.85 * m20-24 + 4.23 * m25-29 + 4.86 * m30-34 + 4.62 * m35-39 + 4.50 * m40-44 + 4.95 * m45-49 + 4.46 * m50-54 + 3.95 * m55-59 + 3.29 * m60-64 + 1.25 * m65more -9.10 * w5-9 + 0.89 * w10-14 + 2.60 * w15-19 + 3.79 * w20-24 + 4.36 * w25-29 + 3.93 * w30-34 + 4.42 * w35-39 + 4.25 * w40-44 + 4.33 * w45-49 + 4.03 * w50-54 + 3.38 * w55-59 + 2.95 * w60-64 .... (A.21)

$V2_{Work\ 2}$ = -5.93 -6.64 * m5-9 - 6.20 * m10-14 + 2.80 * m15-19 + 3.97 * m20-24 + 4.48 * m25-29 + 4.37 * m30-34 + 4.26 * m35-39 + 4.02 * m40-44 + 4.65 * m45-49 + 2.71 * m50-54 + 3.65 * m55-59 + 3.45 * m60-64 + 1.59 * m65more - 6.40 * w5-9 - 6.40 * w10-14 + 4.55 * w15-19 + 3.78 * w20-24 + 3.52 * w25-29 + 3.55 * w30-34 + 3.50 * w35-39 + 3.35 * w40-44 + 3.47 * w45-49 + 3.89 * w50-54 + 2.26 * w55-59 + 2.33 .... (A.22)

$V2_{Work\ 3}$ = -5.93 - 6.78 * m5-9 -6.33 * m10-14 -7.29 * m15-19 - 8.21 * m20-24 - 8.39 * m25-29 + 2.29 * m30-34 + 3.45 * m35-39 + 3.43 * m40-44 - 7.75 * m45-49 - 8.08 * m50-54 + 2.04 * m55-59 + 2.06 * m60-64 -9.61 * m65more - 6.54 * w5-9 - 6.53 * w10-14 - 7.34 * w15-19 + 1.84 * w20-24 + 3.00 * w25-29 + 1.76 * w30-34 + 1.89 * w35-39 - 8.35 * w40-44 + 1.68 * w45-49 + 1.59 * w50-54 + 2.26 * w55-59 + 2.33 * w60-64 .... (A.23)

$V2_{School1}$ = - 5.93 + 8.64 * m5-9 + 9.03 * m10-14 + 7.88 * m15-19 + 6.22 * m20-24 + 5.35 * m25-29 + 5.13 * m30-34 + 4.26 * m35-39 + 2.51 * m40-44 + 3.04 * m45-49 - 7.95 * m50-54 - 7.90 * m55-59 -7.87 * m60-64 - 9.44 * m65more + 8.74 * w5-9 + 8.78 * w10-10 + 7.71 * w15-19 + 6.35 * w20-24 + 5.24 * w25-29 + 4.32 * w30-34 + 4.37 * w35-39 + 3.35 * w40-44 + 3.07 * w45-49 + 2.69 * w50-54 - 8.38 * w55-59 + 1.64 * w60-64 .... (A.24)

$$V2_{School2} = -12.90 + 11.70 * m5\text{-}9 + 10.70 * m10\text{-}14 + 12.50 * m15\text{-}19 + 11.60 * m20\text{-}24 + 11.00 * m25\text{-}29 + 10.90 * m30\text{-}34 + 9.74 * m35\text{-}39 + 8.81 * m40\text{-}44 -3.27 * m45\text{-}49 + 9.01 * m50\text{-}54 - 3.58 * m55\text{-}59 + 9.05 * m60\text{-}64 - 5.69 * m65more + 10.50 * w5\text{-}9 + 11.90 * w10\text{-}14 + 12.40 * w15\text{-}19 + 12.10 * w20\text{-}24 + 10.50 * w25\text{-}29 + 9.84 * w30\text{-}34 + 9.57 * w35\text{-}39 + 8.73 * w40\text{-}44 - 3.98 * w45\text{-}49 - 4.09 * w50\text{-}54 - 4.14 * w55\text{-}59 - 4.04 * w60\text{-}64 \qquad \text{(A.25)}$$

$$V2_{Shopping1} = - 1.31 - 0.63 * m5\text{-}9 - 10.90 * m10\text{-}14 - 1.13 * m15\text{-}19 - 1.63 * m20\text{-}24 - 0.71 * m25\text{-}29 - 0.94 * m30\text{-}34 -0.07 * m35\text{-}39 - 1.01 * m40\text{-}44 + 0.29 * m45\text{-}49 + 0.04 * m50\text{-}54 - 0.10 * m55\text{-}59 + 0.08 * m60\text{-}64 + 0.19 * m65more - 11.10 * w5\text{-}9 - 11.00 * w10\text{-}14 -1.17 * w15\text{-}19 - 0.59 * w20\text{-}24 - 0.07 * w25\text{-}29 - 0.09 * w30\text{-}34 + 0.10 * w35\text{-}39 + 0.07 * w40\text{-}44 + 0.15 * w45\text{-}49 - 0.04 * w50\text{-}54 + 0.34 * w55\text{-}59 + 0.07 * w60\text{-}64 \qquad \text{(A.26)}$$

$$V2_{Shopping2} = - 3.04 - 9.59 * m5\text{-}9 - 9.08 * m10\text{-}14 - 10.00 * m15\text{-}19 - 10.90 * m20\text{-}24 - 11.10 * m25\text{-}29 - 0.60 * m30\text{-}34 + 0.78 * m35\text{-}39 + 0.02 * m40\text{-}44 + 0.15 * m45\text{-}49 + 0.73 * m50\text{-}54 - 10.80 * m55\text{-}59 - 0.14 * m60\text{-}64 + 0.37 * m65more - 9.28 * w5\text{-}9 - 9.33 * w10\text{-}14 -10.00 * w15\text{-}19 - 11.00 * w20\text{-}24 + 0.40 * w25\text{-}29 + 0.47 * w30\text{-}34 - 0.31 * w35\text{-}39 + 0.46 * w40\text{-}44 + 0.58 * w45\text{-}49 - 1.31 * w50\text{-}54 + 0.28 * w55\text{-}59 + 0.13 * w60\text{-}64 \qquad \text{(A.27)}$$

$$V2_{Other\ 1} = - 1.15 + 0.12 * m5\text{-}9 + 0.34 * m10\text{-}14 + 0.21 * m15\text{-}19 + 0 * m20\text{-}24 + 0.06 * m25\text{-}29 + 0.84 * m30\text{-}34 + 0.57 * m35\text{-}39 + 0.47 * m40\text{-}44 + 0.56 * m45\text{-}49 + 0.33 * m50\text{-}54 + 0.81 * m55\text{-}59 + 0.74 * m60\text{-}64 + 0.41 * m65more + 0.14 * w5\text{-}9 - 0.56 * w10\text{-}14 + 0.05 * w15\text{-}19 + 0.42 * w20\text{-}24 + 0.01 * w25\text{-}29 + 0.069 * w30\text{-}34 + 0.84 * w35\text{-}39 + 0.51 * w40\text{-}44 + 0.54 * w45\text{-}49 + 0.33 * w50\text{-}54 + 0.11 * w55\text{-}59 + 0.52 * w60\text{-}64 \qquad \text{(A.28)}$$

***V2**$_{Other2}$* = - 3.15 - 9.49 * m5-9 - 8.99 * m10-14 + 0.02 * m15-19 - 10.80 * m20-24 - 1.07 * m25-29 - 0.48 * m30-34 - 10.70 * m35-39 - 10.90 * m40-44 - 0.43 * m45-49 + 0.34 * m50-54 + 0.36 * m55-59 - 0.02 * m60-64 + 0.20 * m65more - 9.18 * w5-10 - 9.24 * w10-14 - 9.93 * w15-19 - 0.25 * w20-24 - 0.87 * w25-29 + 0.59 * w30-34 - 0.20 * w35-39 - 1.04 * w40-44 + 0.00 * w45-49 - 11.20 * w50-54 - 0.12 * w55-59 - 0.44 * w60-64 .... (A.29)

***V2**$_{Home}$* = 0 .... (A.30)

Model 2: **Probability estimation of each segment**

$$P2(Work\ 1) = \frac{e^{(V2_{Work\ 1})}}{e^{(V2_{Work\ 1})} + e^{(V2_{Work\ 2})} + e^{(V2_{Work\ 3})} + e^{(V2_{School\ 1})} + e^{(V2_{School\ 2})} + e^{(V2_{Shopping\ 1})} + e^{(V2_{Shopping\ 2})} + e^{(V2_{Other\ 1})} + e^{(V2_{Other\ 2})} + e^{(V2_{Home})}}$$

.... (A.31)

$$P2(Work\ 2) = \frac{e^{(V2_{Work\ 2})}}{e^{(V2_{Work\ 1})} + e^{(V2_{Work\ 2})} + e^{(V2_{Work\ 3})} + e^{(V2_{School\ 1})} + e^{(V2_{School\ 2})} + e^{(V2_{Shopping\ 1})} + e^{(V2_{Shopping\ 2})} + e^{(V2_{Other\ 1})} + e^{(V2_{Other\ 2})} + e^{(V2_{Home})}}$$

.... (A.32)

$$P2(Work\ 3) = \frac{e^{(V2_{Work\ 3})}}{e^{(V2_{Work\ 1})} + e^{(V2_{Work\ 2})} + e^{(V2_{Work\ 3})} + e^{(V2_{School\ 1})} + e^{(V2_{School\ 2})} + e^{(V2_{Shopping\ 1})} + e^{(V2_{Shopping\ 2})} + e^{(V2_{Other\ 1})} + e^{(V2_{Other\ 2})} + e^{(V2_{Home})}}$$

.... (A.33)

$$P2(School\ 1) = \frac{e^{(V2_{\mathrm{School}\ 1})}}{e^{(V2_{Work\ 1})} + e^{(V2_{Work\ 2})} + e^{(V2_{Work\ 3})} + e^{(V2_{School\ 1})} + e^{(V2_{School\ 2})} + e^{(V2_{Shopping\ 1})} + e^{(V2_{Shopping\ 2})} + e^{(V2_{Other\ 1})} + e^{(V2_{Other\ 2})} + e^{(V2_{Home})}}$$

.... (A.34)

$$P2(School\ 2) = \frac{e^{(V2_{\text{School 2}})}}{e^{(V2_{Work\,1})} + e^{(V2_{Work\,2})} + e^{(V2_{Work\,3})} + e^{(V2_{School\,1})} + e^{(V2_{School\,2})} + e^{(V2_{Shopping\,1})} + e^{(V2_{Shopping\,2})} + e^{(V2_{Other\,1})} + e^{(V2_{Other\,2})} + e^{(V2_{Home})}}$$

.... (A.35)

$$P2(Shopping\ 1) = \frac{e^{(V2_{\text{Shopping 1}})}}{e^{(V2_{Work\,1})} + e^{(V2_{Work\,2})} + e^{(V2_{Work\,3})} + e^{(V2_{School\,1})} + e^{(V2_{School\,2})} + e^{(V2_{Shopping\,1})} + e^{(V2_{Shopping\,2})} + e^{(V2_{Other\,1})} + e^{(V2_{Other\,2})} + e^{(V2_{Home})}}$$

.... (A.36)

$$P2(Shopping\ 2) = \frac{e^{(V2_{\text{Shopping 2}})}}{e^{(V2_{Work\,1})} + e^{(V2_{Work\,2})} + e^{(V2_{Work\,3})} + e^{(V2_{School\,1})} + e^{(V2_{School\,2})} + e^{(V2_{Shopping\,1})} + e^{(V2_{Shopping\,2})} + e^{(V2_{Other\,1})} + e^{(V2_{Other\,2})} + e^{(V2_{Home})}}$$

.... (A.37)

$$P2(Other\ 1) = \frac{e^{(V2_{\text{Other 1}})}}{e^{(V2_{Work\,1})} + e^{(V2_{Work\,2})} + e^{(V2_{Work\,3})} + e^{(V2_{School\,1})} + e^{(V2_{School\,2})} + e^{(V2_{Shopping\,1})} + e^{(V2_{Shopping\,2})} + e^{(V2_{Other\,1})} + e^{(V2_{Other\,2})} + e^{(V2_{Home})}}$$

.... (A.38)

$$P2(Other\ 2) = \frac{e^{(V2_{\text{Other 2}})}}{e^{(V2_{Work\,1})} + e^{(V2_{Work\,2})} + e^{(V2_{Work\,3})} + e^{(V2_{School\,1})} + e^{(V2_{School\,2})} + e^{(V2_{Shopping\,1})} + e^{(V2_{Shopping\,2})} + e^{(V2_{Other\,1})} + e^{(V2_{Other\,2})} + e^{(V2_{Home})}}$$

.... (A.39)

$$P2(Home) = \frac{e^{(V2_{\text{Home}})}}{e^{(V2_{Work\,1})} + e^{(V2_{Work\,2})} + e^{(V2_{Work\,3})} + e^{(V2_{School\,1})} + e^{(V2_{School\,2})} + e^{(V2_{Shopping\,1})} + e^{(V2_{Shopping\,2})} + e^{(V2_{Other\,1})} + e^{(V2_{Other\,2})} + e^{(V2_{Home})}}$$

.... (A.40)

where,

$P2(Work\ 1)$ = probability of belonging to the segment *Work 1* (*Model 2*)

$P2(Work\ 2)$ = probability of belonging to the segment *Work 2* (*Model 2*)

$P2(Work\ 3)$ = probability of belonging to the segment *Work 3* (*Model 2*)

$P2(School\ 1)$ = probability of belonging to the segment *School 1* (*Model 2*)

$P2(School\ 2)$ = probability of belonging to the segment *School 2* (*Model 2*)

$P2(Shopping\ 1)$ = probability of belonging to the segment *Shopping 1*(*Model 2*)

$P2(Shopping\ 2)$ = probability of belonging to the segment *Shopping 2* (*Model 2*)

$P2(Other\ 1)$ = probability of belonging to the segment *Other 1* (*Model 2*)

$P2(Other\ 2)$ = probability of belonging to the segment *Other 2* (*Model 2*)

$P2(Home)$ = probability of belonging to the segment *Home* (*Model 2*)

$V2_{Work\ 1}$ = deterministic component of the utility of segment *Work 1* (*Model 2*)

$V2_{Work\ 2}$ = deterministic component of the utility of segment *Work 2* (*Model 2*)

$V2_{Work\ 3}$ = deterministic component of the utility of segment *Work 3* (*Model 2*)

$V2_{School\ 1}$ = deterministic component of the utility of segment *School 1* (*Model 2*)

$V2_{School\ 2}$ = deterministic component of the utility of segment *School 2* (*Model 2*)

$V2_{Shopping\ 1}$ = deterministic component of the utility of segment *Shopping 1* (*Model 2*)

$V2_{Shopping\ 2}$ = deterministic component of the utility of segment *Shopping 2* (*Model 2*)

$V2_{Other\ 1}$ = deterministic component of the utility of segment *Other 1* (*Model 2*)

$V2_{Other\ 2}$ = deterministic component of the utility of segment *Other 2* (*Model 2*)

$V2_{Home}$ = deterministic component of the utility of segment *Home* (*Model 2*)

Detail definitions of the explanatory variables of *Model 2* can also be found in Table 7.5 (Chapter 7).