A PRISM- and GAP-BASED MODEL of SHOPPING DESTINATION CHOICE

by

Joshua Wang

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science

Department of Civil Engineering
University of Toronto

Canadä

A Prism- and Gap-Based Model of Shopping Destination Choice

Joshua Wang

Master of Applied Science

Department of Civil Engineering
University of Toronto

2011

Abstract

This thesis presents a prism- and gap-based approach for modelling shopping destination choice in the Travel/Activity Scheduler for Household Agents (TASHA). The gap-location choice model improves upon TASHA's existing destination choice model in 3 key ways: 1) Shifting from a zone-based to a disaggregate location choice model, 2) Categorizing shopping trips into meaningful types, and 3) Accounting for scheduling constraints in choice set generation and location choice. The model replicates gap and location choices reasonably well at an aggregate level and shows that a simple yet robust model can be developed with minimal changes to TASHA's existing location choice model. The gap-based approach to destination choice is envisioned as a small but significant step towards a more comprehensive location choice model in a dynamic activity scheduling environment.

Acknowledgments

I am very grateful to everyone who has been a part of my enjoyable journey of completing my Master degree. I have learned a great deal about transportation and life, and the privilege to work with great mentors and colleagues.

First, I would like to thank Professor Miller, an incredible supervisor and mentor, guiding me through both my undergraduate and Master theses. I am continually amazed by how much I have yet to learn. Thanks, Eric, for continually stretching my mind and intellectual curiosity.

I would also like to thank Professors Roorda and Habib. Both of you have been mentors to me along my journey in life and academia. Your experiences and enthusiasm are truly inspiring.

I am grateful for the funding I received from NSERC's Industrial Postgraduate Scholarship in partnership with IBI Group. I have learned a lot through my internship with IBI Group through industry mentors in Bruce Mori and Jesse Coleman.

I am indebted to Nikola Kramaric, who was very helpful and patient in teaching me the ropes of TASHA and C#. I am also grateful to Rebecca Hamer from ALOGIT for answering the plethora of software questions I had. And I am thankful to Professor Tony Hernandez from Ryerson's Centre for the Study of Commercial Activity for generously providing commercial floor space data for the GTA.

I am also indebted to my colleagues in our Transport Group, in particular: Bilal Farooq, Rinaldo Calvacante, Chris Bachmann, Bryce Sharman, and Yuan Tian. Each of you have helped me in various ways and inspired my research efforts.

Finally, I would like to thank my family and friends who have been supportive throughout my entire academic and life journeys.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Appendix Tables and Figures

List of Appendix Tables and Figures, cont'd

1   Introduction

Travel demand models are important tools used by transportation planners to predict future growth in transport demand, forecast traveller responses to various policies, and assess socioeconomic and environmental impact of the transportation system. Insights from travel demand model results are used by decision makers to answer key questions to complex transportation issues. Some of these questions include:

- *How effective are different demand management strategies in reducing congestion?*

- *Will there be sufficient demand to warrant investment in new infrastructure?*

- *What is the environmental impact of traffic pollution?*

As urban regions continue to expand, modellers will need to adapt travel demand models to address the variety of questions posed by regional transportation authorities. Conventional trip-based models of travel demand, used in the standard four-stage modelling approach, are limited in representing traveller behaviour and testing transportation policies.

Transportation is a derived demand, arising from the need to participate in activities, such as work, school, and shopping activities. Travel patterns, such as congestion and time-of-day variations of traffic conditions, arise from different individuals participating in activities over space and time. To capture the complex behaviour of travellers, travel demand needs to be modelled at the activity level.

Activity-based approaches to travel demand modelling have proven to be better than conventional approaches in transportation policy analysis and forecasting traveller behaviour responses (Habib, 2011). Activity-based models have been implemented in different parts of the world: South Florida (FAMOS) (Pendyala et al., 2005), Portland (Bowman et al., 1998), Texas (CEMDAP) (Bhat et al., 2004), Netherlands (ALBATROSS) (Arentze and Timmermans, 2000), and Toronto (TASHA) (Roorda et al., 2008).

While the strength of the activity-based approach is clearly recognized and used in many regions in the United States (e.g. San Francisco, Columbus, Atlanta, and Denver), many Canadian regions still use trip-based or simple tour-based methods. In Canada, researchers have developed a fully operational activity-based model for Toronto, known as the Travel/Activity Scheduler for

Household Agents (TASHA). While activity-based models have not been adopted in the Toronto region, there are several improvements that can be made to TASHA to better represent traveller behaviour. One of the most significant improvements that can be made is to improve the behavioural realism of TASHA's location choice model, since location choice is a fundamental decision that influences other travel choices, such as choices of mode and route.

The main purpose of this thesis is to develop a robust shopping location choice model for the TASHA. TASHA is a fully operational activity-based microsimulation model based on the Transportation Tomorrow Survey (TTS) (Data Management Group, 1996). As with most activity-based models, TASHA consists of an activity generation component and an activity scheduling component. While TASHA can exploit data from activity diaries, TASHA simulates the activity generation and scheduling process using the same trip-based data used for conventional four-stage models and is robust enough to be used as an alternative to conventional models currently used in the Greater Toronto Area (GTA) (Roorda et al., 2008).

The motivation behind improving TASHA's shopping location choice model is twofold:

1. Location choice is a foundational choice in the activity scheduling process. Many choices, such as mode and route choices, derive from location choice. The current implementation of TASHA's location choice model predicts destination choices at the zone level, whereas subsequent choices depend on the individual, scheduling constraints, and household interactions. To align with the overall microsimulation framework of TASHA, a natural modification to the location choice model is to change from an aggregate model to a disaggregate model.

2. While shopping comprises only a small percentage of total trips, shopping trips are growing more rapidly than any other shopping type (Buliung et al., 2007). With the growing number of shopping trips and the development of new urban centres, it is important that shopping activities are properly modelled in addition to other trip purposes, such as Work and School. Furthermore, unlike work and school trips, shopping trips exhibit a great deal of heterogeneity.

This thesis aims to develop a model to answer the question: *How do individuals choose shopping locations?* This behavioural model can be used to predict shopping location choices at an

individual level. While this research was inspired by the need to improve TASHA's shopping location choice model, the main objective is to explore location choice behaviour and examine some simplifying assumptions made in choosing locations in the activity-based modelling framework.

One of the key assumptions in TASHA is the independence of location choice from other choices in the activity generation and scheduling process. In this thesis, we propose that shopping location is chosen in the context of an individual's schedule. In particular, we postulate that an individual first chooses a time window (or gap); the choice of shopping location depends on the gap chosen. This notion arises from our understanding that individuals exhibit different shopping location choice behaviours depending on the type of shopping, individual scheduling constraints, time of day, and the day of week. Chapter 3 will give additional details to our proposed prism- and gap-based shopping location choice model.

The remainder of the thesis is organized as follows: Chapter 2 presents a review of TASHA, shopping location choice models, and choice set formation models; Chapter 3 presents TASHA's current location choice model and the proposed approach, along with a description of the data; Chapter 4 presents the specification of the model and parameter estimates; Chapter 5 presents simulates how well the model predicts shopping location choices; Chapter 6 presents opportunities for future research; and Chapter 7 concludes with the main contributions of this research.

2 Literature Review

2.1 Approaches to Travel Demand Modelling

Two main approaches exist in Travel Demand Modelling:

1. Four-Stage Model (Urban Transportation Modelling System), and

2. Activity-based Models.

A large body of literature discusses these models in great depth, including Meyer and Miller (2001) and Timmermans (2005). For this review, these approaches will be discussed in relation to the Travel/Activity Scheduler for Household Agents (TASHA).

2.1.1 Four-Stage Model

TASHA was designed to improve the existing four-stage travel demand modelling system used in the Toronto area. The four-stage model, or the Urban Transportation Modelling System (UTMS), was developed in the mid-20th century and is still the state-of-practice in many urban regions, including the Greater Toronto Area (GTA) (Meyer and Miller, 2001). The UTMS models different components of travel demand, including trip generation, trip distribution, mode choice, and route choice. A key difference between TASHA and the UTMS is that the UTMS models travel demand at an aggregate level of zone-to-zone flows, whereas TASHA models activity and travel choices at a disaggregate level of trips by individual persons. Individual choices modelled in TASHA can be aggregated to give similar output to the conventional four-stage models (Bowman and Ben-Akiva, 1996).

Furthermore, four-stage models predict trip outcomes, while disaggregate models such as TASHA predict travel and activity choices, thereby capturing the underlying behaviour leading to trip outcomes. These behavioural models are more sensitive to transportation policies and are more transferable to different spatial and temporal contexts than the conventional four-stage models.

2.1.2   Activity-based Models

Activity-based models take into account the behavioural component of travel demand, which is a derived demand arising from basic needs or desires that lead to individual participation in activities over space and time (Ettema and Timmermans, 1997). These models are more sensitive to policies and travel behaviour responses. There are two main approaches to activity-based models of travel demand:

1. An Econometric Approach

2. A Rule-based Scheduling Process Approach

2.1.2.1   Econometric Approach

The econometric approach uses random utility maximization (RUM) models to capture the choices resulting in schedule formation. These models usually involve a series of sequential discrete choice models to predict activity travel and activity patterns. Ben-Akiva and Lerman (2001) provide a good overview of random utility and discrete choice theories.

The Comprehensive Econometric Microsimulator for Daily Activity travel Patterns (CEMDAP) is one of the few operational activity-based models relying mostly on an econometric approach to model activity scheduling (Bhat et al., 2004). CEMDAP's activity scheduling component consists of a number of sub-models, which consider a fixed sequence of activity scheduling, defined by the modeller. This approach efficiently captures various components of travel behaviour in the econometric model components.

One shortfall of CEMDAP, and other econometric models, is that these models do not provide a comprehensive framework to handle the dynamics of the scheduling process. Rather, these econometric models generally seek to replicate activity and travel patterns without considering the schedule formation process itself. Habib and Miller (2008, 2009) and Habib (2011) propose a unified econometric framework to address some of these issues; however this model is not yet operational. A few operational activity-based models addressing the scheduling formation process are discussed in the next section.

2.1.2.2   Rule-based Scheduling Process Approach

The rule-based scheduling process approach uses a variety of computational techniques to model the activity scheduling process. As many travel behaviour researchers have long recognized, to capture the underlying behaviour in schedule formation, researchers need to model the process behind activity schedule formation leading to travel and activity patterns (Pas, 1985; Axhausen and Garling, 1992; Bhat and Lawton, 2000). Most fully operational process models use one of two approaches:

1. Computational process models, such as ALBATROSS (Arentze and Timmermans, 2000), or

2. Simple empirical rules, as in TASHA (Roorda et al., 2008).

ALBATROSS uses data mining techniques and probabilistic decision trees to capture the sequence of decisions leading to activity schedule formation. The core of ALBATROSS is driven a scheduling engine, which is a system of models connected by a comprehensive set of conditional IF-THEN statements. The activity schedule is built by assuming fixed activities, such as work, and simulating schedule formation through probabilistic decision trees. The fixed activities and conditional heuristics derive from travel diary data, meaning that this model will need similar resources for implementation in other contexts.

TASHA is an activity-based model, which uses some simple empirical rules to represent the scheduling process. These rules are informed by the CHASE data, which is a weeklong travel and activity dataset containing individual and household scheduling information. Unlike ALBATROSS, TASHA uses simpler rules and can be implemented without activity diary data.

More recently, Auld and Mohammadian (2009) proposed an activity-based modelling framework, which is a hybrid of behavioural schedule process rules and econometrics. This model builds on the TASHA framework and makes significant improvements in activity planning and scheduling dynamics. Habib (2011) suggests a more unified econometric modelling framework to address the fundamental issues related to scheduling dynamics.

Nevertheless, TASHA represents a significant advance in activity-based modelling. The next section will provide a general overview of the TASHA framework.

2.2 TASHA

TASHA is an agent-based microsimulation model that captures the schedule formation process and complex interactions such as vehicle allocation, household allocation, and tour formation. Figure 2.1 provides a high level overview of TASHA's conceptual framework.
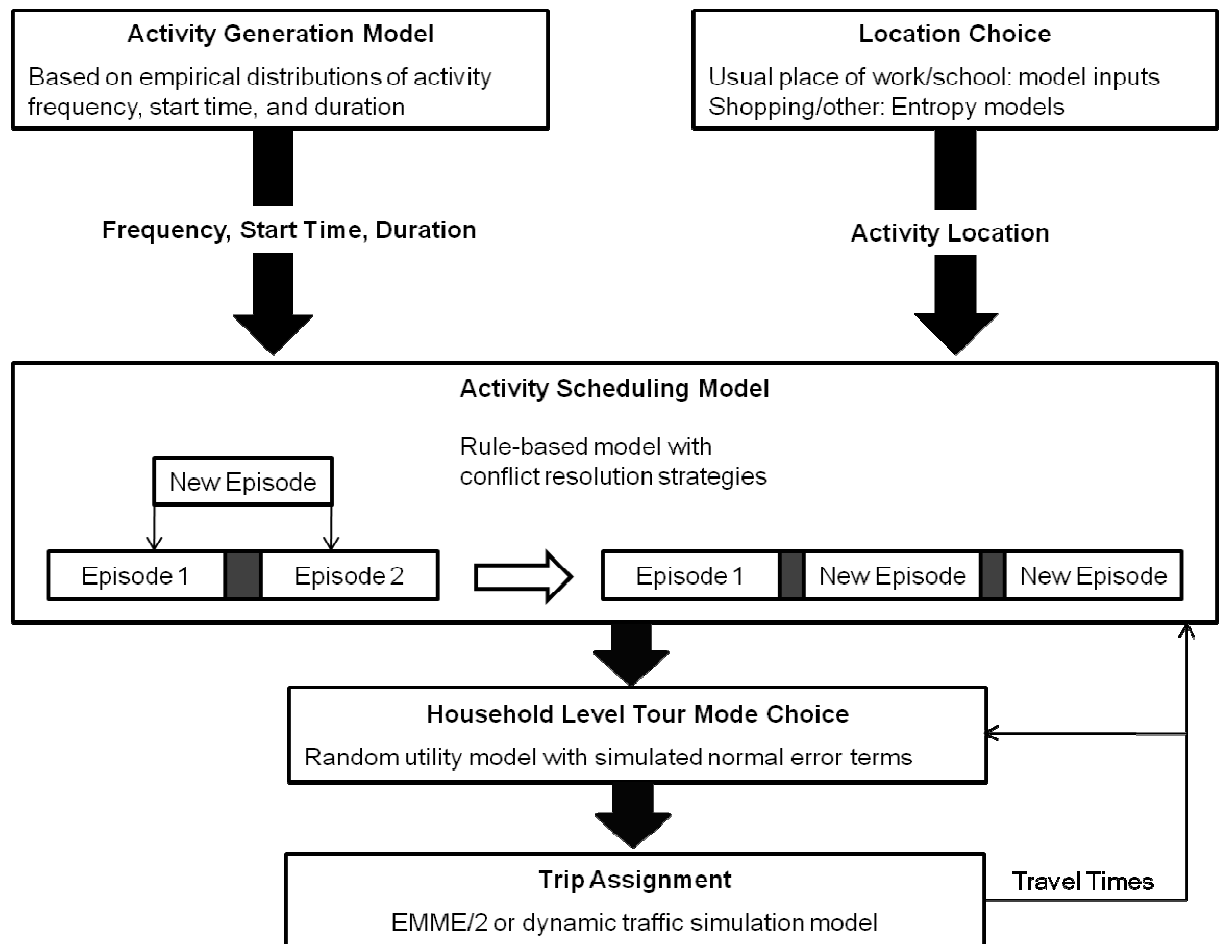
Figure 2.1 Conceptual Framework of TASHA

One of the key concepts in TASHA is that of a project. A **project** is defined as a set of logically connected activities with a common goal or purpose (Axhausen, 1998). Miller (2005a, 2005b) extends this concept to argue that all personal and household activities can be logically formed in a smaller set of projects, such as a Work or School project. TASHA uses this project-based approach in the schedule formation process.

Each project generates a set of potential activity episodes. An **episode** is an instance of a particular activity with the following attributes: activity type, duration, location, start time, end

time, and a set of participants. A conceptual example of a project can be "cleaning up the house". A set of potential activity episodes may include: cleaning the dishes, doing the laundry, washing the floor, and rearranging the living space. Each of these episodes would have associated attributes, including start time, duration, and a set of participants.

For this thesis, the focus is on the Shopping project, which generally only has one episode (i.e. shopping). In this case, the project concept can be disregarded without loss of generality. This organizing principle is more useful in more complicated projects, such as the Work project, and in projects that involve other individuals.

As shown in Figure 2.1, the TASHA framework consists of two key components: activity generation and activity scheduling. Activities are generated from empirical distributions of activity type, frequency, and duration. These activities are grouped into agendas and inserted into an individual's **provisional schedule**, defined as the set of episodes scheduled for execution at any time t. In the current version of TASHA, empirical rules derived from the CHASE dataset are used to define the scheduling priorities of different projects (Roorda et al., 2008).

An important point to notice in Figure 2.1 is that activity location is also modelled as an attribute of the activity rather than an outcome of the scheduling process. While this assumption holds true for fixed and routine activities, for many other activities, location is chosen as part of the schedule formation process. TASHA's location choice model is discussed further in Section 3.1.

As activities are inserted into an individual's provisional schedule by project priority, there is no consideration of the existing schedule and time budget, leading to conflicts. In TASHA, simple heuristics are developed that shift, split, shorten, or delete activities to resolve conflicts. Recent work in other models, such as ADAPTS, have incorporated a conflict resolution model in the scheduling process (Auld and Mohammadian, 2009). Nevertheless, the underlying issue of time budget considerations in the schedule formation process remains unsolved.

Once activities from all projects have been scheduled, travel mode is chosen using a household tour-based model that incorporates joint mode choice for joint activities, vehicle allocation, and carpooling. A more detailed description of the mode choice model can be found in Roorda et al. (2006).

Finally, TASHA's modelling system interacts with a traffic assignment model to forecast demand for the transportation infrastructure. TASHA is designed to interact with conventional traffic assignment models, such as EMME/2 as shown by Roorda et al. (2008), and dynamic traffic assignment models, such as MATSIM as shown by Hao et al. (2011). In either case, the system is run iteratively, and includes feedback of travel times into the activity scheduling and mode choice models.

2.3   Shopping Location Choice Models

Location (or destination) choice is an important and often neglected component of travel demand models. Within the activity-based framework, the scheduling and travel decisions are influenced by the choice of activity location. Unlike Work or School activities, which usually have fixed locations, Shopping activity locations are generally more flexible in space and time.

Researchers have been investigating destination choice models for non-work activities for many years. Ansah (1977) used generalized choice sets to develop functional classifications of shopping destinations. In this formulation, shopping destinations were represented and classified at the store level and exploit retail store data, which is not always available. Furthermore, most operational travel demand models including TASHA capture activity location choices at the zone level.

Fotheringham (1983) developed the competing destinations model to better represent spatial choice behaviour. This model accounts for systematic similarities and differences between destinations and has been extended to various applications, including the destination choice model in the recently developed ADAPTS activity-based model (Auld, 2011).

Destination choice models have also been extended to include more advanced model formulations, including a model accounting for error correlation in workplace location (Bolduc et al, 1996), a unified econometric framework for non-work locations incorporating spatial cognition, heterogeneity in preference behaviour, and spatial interaction (Sivakumar and Bhat, 2003), and the development of mixed generalized extreme value models for residential location choice (Sener et al., 2009).

The location choice model developed in this thesis does not account for many of the shopping behaviour variability that these more advanced models incorporate. Rather, the proposed model offers a simple approach to capture the heterogeneity in shopping activities while considering time and space constraints within an individual's activity schedule.

2.4 Choice Set Formation Models

While significant advances have been made in location choice models, many of these models do not adequately consider (if at all) choice set formation. This is not an issue for problems involving smaller choice sets, where the actual choice set can be approximated by the universal set. Examples include mode choice problems and location choice problems involving the choice between several shopping centres.

However, choice set formation is a greater issue in destination choice models, where the set of locations are much larger than the actual zones considered. For example, the universal choice set used in this thesis is the 1548 traffic zones in the GTA. It is behaviourally unrealistic for an individual to consider all 1548 locations, let alone compare each location with respect to all other locations.

Several models have been developed to address choice set formation for destination choice (Thill, 1992; Pagliara and Timmermans, 2009). Some approaches to choice set formation include:

1. Problem decomposition into choice-set generation and choice generation sub-problems (Burnett and Hanson, 1979; 1982), with an econometric formulation of these two processes by Manski (1977),

2. Models that account for limited information processing ability, such as the satisficing model (Simon, 1955) and the elimination-by-aspects model (Tversky, 1972),

3. Models that incorporate learning to incorporate the influence of information gathering (Meyer, 1980),

4. Behavioural theory of random constraints to explain the probabilistic nature of choice sets (Swait and Ben-Akiva, 1987),

5. Latent choice sets models to predict the unobserved set of considered locations (Ben-Akiva and Boccara, 1995), and

6. Time-space prisms, conceptualized by Hagerstränd (1970), to constrain the choice set to the set of feasible alternatives based on scheduling constraints, as in the PCATS and FAMOS models (Kitamura and Fujii, 1998; Pendyala et al., 2005).

Many of these choice set formation models require additional data of alternatives available and / or considered by the individual, or impute assumptions on choice set behaviour. In this thesis, choice set is handled by incorporating scheduling constraints defined by time-space prisms and availability constraints defined by land use data. This approach is purely deterministic and does not make any assumptions in the choice set formation process, while significantly reducing the choice set size (see Section 4.4).

2.5   Research Gaps and Contributions

An overview was provided above of TASHA in the context of some operational activity-based travel demand models. TASHA is an agent-based microsimulation model that uses simple rules to model the schedule formation process for individuals and households. Finally, a review of location choice and choice set formation models were provided in the context of shopping destination choice.

The aim of the model developed in this thesis is to improve the behavioural realism of TASHA's location choice model for shopping trips. The contributions of this research can be summarized as follows:

- Destination choice sets are constrained using time-space prisms, as in ADAPTS and PCATS (Auld and Mohammadian, 2009; and Kitamura and Fijii, 1998). See Section 3.2.3 for details.

- The Shopping activity type is further disaggregated to capture the heterogeneity inherent in shopping trips. See Section 3.3.2 for details.

- A gap-location choice model is developed using a nested logit model to model the location choice dependency on the chosen gap or time window. This modelling relaxes

the assumption of location choice being independent of the schedule formation process. To our knowledge, the gap-location choice model presented in this thesis is an innovative approach to modelling shopping destination choice. See Chapter 3 for details.

3 A Prism- and Gap-based Approach to Shopping Destination Choice

This chapter introduces a gap-based approach to modelling shopping location choice based on Hagerstrand's time-space prism. Section 3.1 introduces the current implementation of the location choice model in TASHA. The subsequent sections present the motivation for the proposed gap-location choice model, and a description of the modelling framework and data used to test this model.

3.1 TASHA's Current Location Choice Model

The current location choice model in TASHA uses an entropy formulation as documented in Eberhard (2002). In this formulation, the probability that a person living in zone i chooses location j is defined by Equation 3.1:

$$P_{j|i} = \frac{\exp\left(\sum_k \delta_{jk}[\alpha_k + \beta_k \log(E_j) + \varphi_k \log(P_j) + \gamma_k d_{ij}]\right)}{\sum_{j\prime} \exp\left(\sum_k \delta_{j\prime k}[\alpha_k + \beta_k \log(E_{j\prime}) + \varphi_k \log(P_{j\prime}) + \gamma_k d_{ij\prime}]\right)} \qquad [3.1]$$

where
δ_{jk} = 1 if zone j belongs to 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
$\alpha_k, \beta_k, \varphi_k, \gamma_k$ = parameters to be estimated
k = 1, if the zone is the city core
   = 2, if shopping mall floor space > 100,000 sq. ft.
   = 3, otherwise

The location choice model shown in Equation 3.1 is equivalent to a zone-based logit model and does not consider the location prior or posterior to the shopping episode, using a distance variable based on an individual's home location. This assumption is valid for simple tours where the home location serves as the anchor point(s) in the trip chain. However, for more complicated tours or trip chains, which do not involve home-based trips, the location choice model needs to account for the possibility of different anchor point locations.

Additionally, the current location choice model does not consider choice set formation; the choice set for the shopping location choice model is the universal set of all zones in the study boundary. While selecting from the universal set is a reasonable assumption for mode choice models, this assumption is behaviourally unsound for destination choice models, where there can often be over 1000 alternatives. Individuals rarely consider all alternatives nor do they have the computational capacity to evaluate each alternative with respect to all other locations. The current model also includes locations that do not have retail stores and locations that are not accessible by the individual due to scheduling constraints. The proposed model aims to address these issues in an effort to improve the behavioural realism of the shopping location choice model.

3.2   Proposed Gap-Location Choice Model

The proposed model simulates location and gap choices at the individual level, while accounting for time budget constraints. The location choice model retains the same basic model structure of the multinomial logit model with two key differences:

1. The revised model predicts location choice at an individual level, instead of the zone-based model currently implemented in TASHA. This approach is better suited for the microsimulation framework in TASHA and better captures the heterogeneity inherent in shopping activities.

2. A nested logit model is used to model gap and location choices. Rather than choosing location exogenously to the scheduling process, this model considers scheduling constraints in location choice by choosing an available time window (gap) from an individual's provisional schedule. The chosen shopping location depends on the gap chosen.

3.2.1   Motivation for Gap Choice

The main premise for the proposed gap-location choice model is that location choice is not an independent process. Location can be chosen jointly with activity type, transportation mode, and/or start time, among other activity attributes. In this thesis, we propose that location choice is dependent on gap choice: *Where you shop depends on when you choose to shop*.

Further, we propose that an individual chooses from a set of discrete time windows (gaps) rather than choosing a specific start time. The decision is more likely *"I will shop in the evening"* instead of *"I will shop at 7:36pm"*. Eventually a specific start time will be chosen; however, when choosing a shopping location, an individual's start time is usually more vaguely defined and subject to change. From our review of the literature, the gap-location choice model is a novel approach to modelling shopping destination choice.

3.2.2   Revisiting the Scheduling Process

Shopping location choice is currently modelled exogenously to the activity scheduling process. For example, in Figure 2.1, the location of the "New Episode" (e.g. shopping episode) is determined without considering prior or posterior episodes in the provisional schedule (e.g. Episodes 1 and 2). Once the location has been determined by the zone-based logit model, this location remains fixed during the activity scheduling process.

The proposed model considers location endogenously to the scheduling process. Referring to Figure 2.1 as an example, the location of the "New Episode" would be determined when the episode is inserted into the provisional schedule. The choice of location would not only depend on the attractiveness of shopping zones, but would also depend on the current provisional schedule, which would consist of other higher priority activities that were inserted before the new shopping episode.

3.2.3   Choice Set Generation

In the gap-location choice model, an individual's choice set (**consideration set**) is defined by the intersection between the **awareness set** and the **feasible set** (i.e. Consideration Set = Feasible Set ∩ Awareness Set). For a location to be considered, the location needs to be known to the individual (in the awareness set) and the location must be accessible given individual scheduling constraints (in the feasible set). The time-space prism conceptualized by Hägerstrand (1970) is used to define the set of feasible locations.

As shown in Figure 3.1, an individual's choice of shopping location must fall within the feasible region, which is the projection of the time-space prism on the spatial plane. An individual's time space path must lie within the time-space prism, which is defined by an individual's speed and the anchor points of the chosen gap. These anchor points are defined by the prior and posterior

episodes (e.g. Work and Home). Additionally, the location needs to be close enough to allow for sufficient time to shop. In the proposed gap-location choice model, the duration of the shopping activity is assumed to be known.
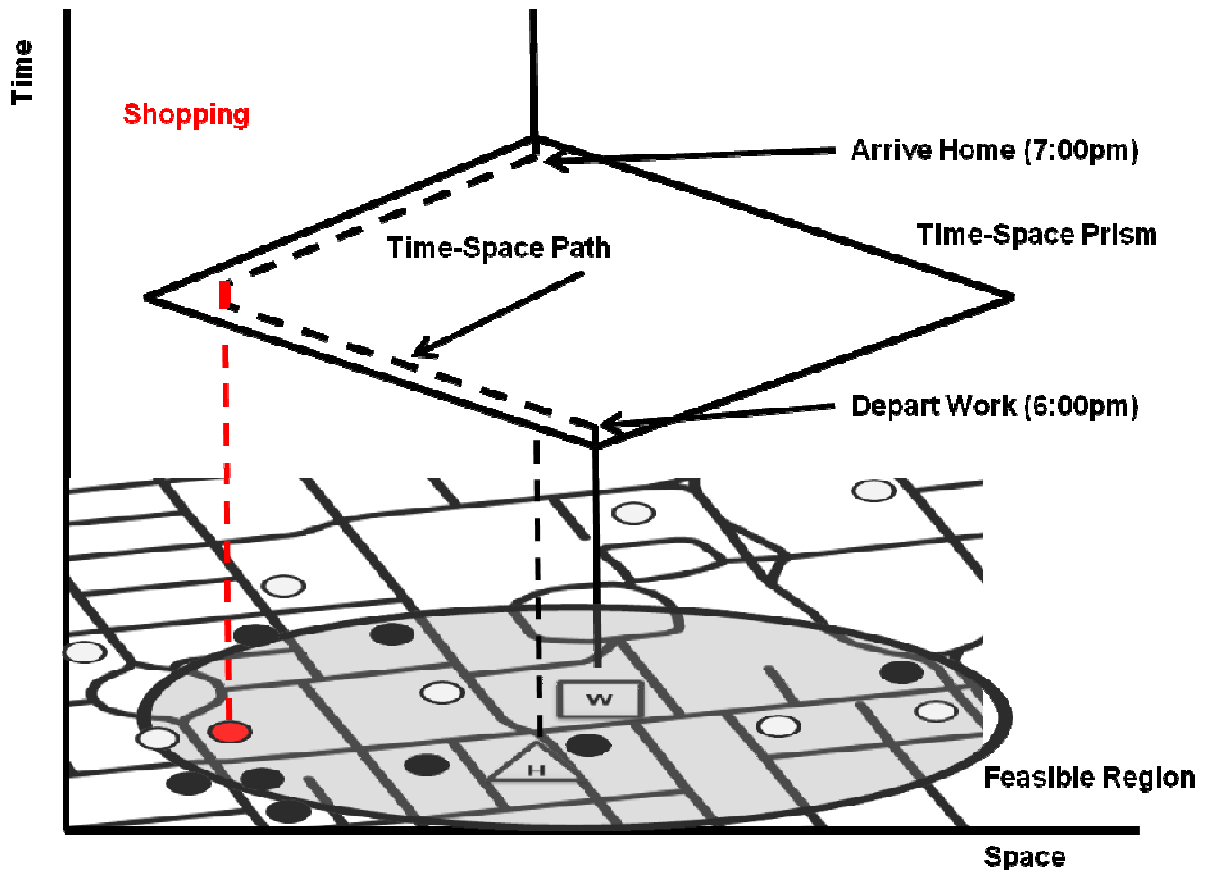
Figure 3.1 Time-Space Prism Constraint to Define the Feasible Set of Shopping Locations

For simplicity, the awareness set model used in this thesis assumes full information is known to the individual. In other words, given a type of shopping activity, an individual knows all the zones that have retail stores of this shopping type. In Figure 3.1, the black circles represent the locations in an individual's awareness set, while all circles that lie within the ellipse (i.e. feasible region) represent an individual's feasible set. Therefore, the consideration set consists of all of the black circles within the feasible region. The location in red represents the chosen shopping location.

The constraints on choice set used in this thesis are entirely deterministic. This is a strong assumption since anchor points (or vertices of the time-space prism) are spatially and temporally flexible. Stochastic frontier models can be used to simulate the unobservable temporal

boundaries of time-space prisms (Kitamura et al., 2000; Pendyala et al., 2002; Yamamoto et al., 2004). Furthermore, as mentioned in the literature review, an individual's choice set can be modelled using latent choice sets to represent an individual's unobserved awareness of available alternatives (Ben-Akiva and Boccara, 1995). Additional data collection would be required to develop a probabilistic awareness set model. Both stochastic frontier models and latent choice sets are areas for further research.

3.3   Exploiting Travel Activity Data

TASHA is currently calibrated using data from the 2001 Transportation Tomorrow Survey (TTS) (Data Management Group, 2001). TASHA uses TTS data to create a synthetic population and simulates activity participation and travel decisions for individuals and households. However, when the proposed gap-location choice model was still under development, TASHA was not modular – the location choice model could not be run independently from the rest of TASHA.

As the modelling capabilities of TASHA were being developed, we designed a standalone program to test our gap-location choice model. However, since we did not have access to the synthetic population and the scheduling model, we needed to generate provisional schedules for each individual to test our location choice model, which depends on the gaps in the provisional schedule. Conventional travel surveys (e.g. TTS) do not contain activity scheduling information. Therefore, to generate provisional schedules, we needed data from a detailed activity diary.

3.3.1   Using CHASE Data

The gap-location choice model was developed using data based on the 2002-2003 Computerized Household Activity Scheduling Elicitor (CHASE) survey conducted in Toronto. This dataset was used because it captures an individual's scheduling decisions over the course of an entire week. The CHASE dataset contains activity diary data for 416 individuals from 262 households, with 361individuals from 244 households engaged in at least one shopping episode throughout the weeklong survey. A detailed description of the survey methodology and the data are available in Doherty et al. (2004) and Doherty and Miller (2000).

The 2001TTS zone system (Data Management Group, 2001) is used to define location alternatives. All geocoded data are mapped to the 2001 TTS zone system using Geographic Information Systems (GIS) software. The study area chosen was the Greater Toronto Area (GTA) with 1548 zones because it provided sufficient coverage for most of the 1194 observed shopping locations in the CHASE data. Figure 3.2 shows the TTS zone system, GTA region, and the observed shopping locations in CHASE. It is noted that not all shopping locations fall within the GTA region (<1%). These 11 shopping episodes are ignored in the shopping location choice model since the observed location is not in the universal set of possible locations.
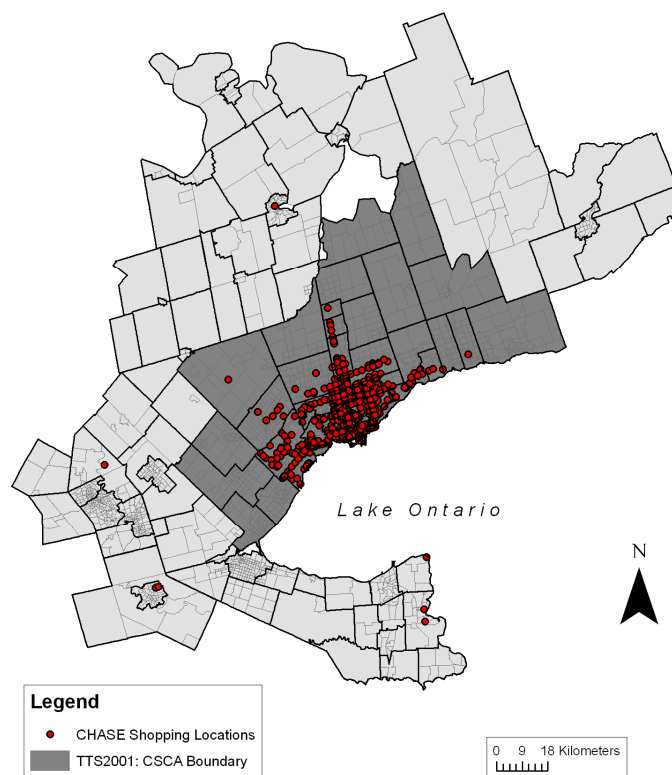
Figure 3.2 Observed Shopping Locations in the GTA

Another reason that the GTA boundary was chosen for the study area is data availability. Data on commercial floor space by TTS zone was generously provided by Tony Hernandez from Ryerson's Centre for the Study of Commercial Activity (CSCA). Data from the CSCA is available only for the GTA region.

3.3.2   Shopping Activity Types

The CHASE dataset defines eight generic activity types within the Shopping activity category: Convenience Store, Minor Groceries, Major Groceries, Houseware, Clothing / Personal, Drug Store, Internet, and Other. Internet shopping episodes, accounting for 2 of the 1196 shopping episodes, were ignored in this analysis since TASHA only models out-of-home activities. The remaining seven generic activity types were analyzed to see if any generic types could be grouped based on similarities in trip characteristics, including trip distance and travel time.

3.3.2.1   Travel Time vs Activity Duration

As shown in Kitamura et al. (1998), time of day, activity duration, and home location have a significant influence on destination choice for non-home-based trips. Hence, a preliminary examination of the generic shopping types began with examining the relationship between travel time and activity duration. As shown in Figures A1.1 to A1.7 in Appendix A, there is a slight indication that individuals are willing to travel farther for activities with longer durations. Also, as shown in Table A1, the average travel time is significantly less than the average activity duration for all shopping types, reinforcing the disutility of travel time. However, these results could not be used to group the existing generic shopping types as it would require new activity types to be defined based on the relationship between travel time and activity duration.

3.3.2.2   Mode Share and the Assumption of Auto Travel Times

As shown in Figure A2 in Appendix A, Auto dominates as the main mode of transportation for shopping trips across all shopping types. Since mode choice is not considered in the gap-location choice model, the model specification for the location choice model presented in this thesis uses travel times for the Auto mode only. Based on Figure A2, this is a reasonable assumption for Transit trips. However, this assumption may not be valid for other types of modes, such as Walk, as shown in Table A2 and reinforced in Table A5. Conceptually, one could include walk times by calculating the distance travelled and using an assumed walk speed. However, many of the shopping trips accessed by walking are short distance trips and include intrazonal trips. These intrazonal trips would have walk travel times of 0 min (using conventional zone-based assignment methods to compute travel times), which is unrealistic. Thus, given the lack of a pedestrian network and an integrated mode and location choice model, auto travel times were used for all shopping trips for consistency.

3.3.2.3   Activity Planning

As shown in Figure A3 in Appendix A, many of the shopping trips are spontaneous or planned on the same day, with relatively few routine shopping trips. While the activity planning horizon influences shopping destination choice, the location choice model developed in this thesis does not explicitly account for the planning horizon, unlike more recently developed models (Auld and Mohammadian, 2009). For further research, the activity planning horizon can be investigated as an attribute for shopping activity type classification. However, the goal of the classification approach used in this thesis is to use the existing classification provided in the CHASE dataset and group similar generic activity types.

3.3.2.4   Final Activity Type Classification

For the final activity type classification, Kolmogorov-Smirnov tests were used to compare the similarities and differences in the distributions of trip characteristics. Given the small dataset of 416 individuals and 262 households, these tests were used to guide the activity type classification. Table 3.1 shows the results of these statistical tests comparing the travel times between different generic activity types. The table lists the ratio between the D-statistic and the threshold statistic using a significance level of $\alpha = 0.1$. Ratios greater than 1 show that the generic activity types have different travel time distributions at the 0.1 significance level, while ratios less than 1 suggest that these types share similar travel time characteristics.

Table 3.1 Kolmogorov-Smirnov Tests of Travel Times by Shopping Activity Type

D(Travel Time) : D($\alpha = 0.1$)	40	41	42	43	44	45	47
40: Convenience Store		**1.0297**	**1.6371**	**1.2242**	**1.6424**	0.4038	**1.2236**
41: Minor Groceries			**1.1666**	0.4552	**1.1195**	**1.0704**	0.6299
42: Major Groceries				0.4271	0.2781	**1.7354**	0.8392
43: Houseware					0.4377	**1.2634**	0.2318
44: Clothing / Personal						**1.7171**	0.8190
45: Drug Store							**1.2772**
47: Other							

The final activity type classifications used in this model are shown in Table 3.2. As shown in Table 3.1, Convenience Store trips share similar travel time characteristics as Drug Store trips and are significantly different from the other generic activity types. Also, while Minor and Major Groceries have different travel time distributions at $\alpha = 0.1$, the difference was fairly small (ratio

= 1.2). Kolmogorov-Smirnov tests were also performed for travel distance distributions for short distance trips (<1km) (see Table A4) and reinforced this classification for Convenience trips, yet showed significant differences between Minor and Major Grocery trips. The final groupings were chosen to reflect different types of shopping activities (e.g. Convenience trips are categorized under Type 1). It is noted that there are some overlaps in the generic activity classifications (e.g. Minor Groceries can include Convenience Store trips, Drug Store trips, and even Major Grocery trips) and the proposed classification is ultimately imputed.

This shopping activity classification is supplemented by Enhanced Point of Interest (EPOI) data provided by DMTI Spatial (2003). The EPOI data contains geographic locations of retail stores in Canada, disaggregated by 64 types. These types were mapped to the 2001 TTS zone system and to the shopping types to provide the number of stores by shopping type per zone (see Figures A4.1 to A4.6 in Appendix A). Table 3.2 shows the summary of the activity type mapping between the shopping types, the CHASE generic activity types, and the EPOI data. Table A6 provides a more detailed mapping of the shopping activity types.

Table 3.2 Shopping Types Mapped to CHASE and Retail Location Datasets

Shopping Type	CHASE Generic Activity Type	Number of Trips	EPOI (SIC Codes) Number of Types
1	40: Convenience Store	40	4
	45: Drug Store	46	
2	41: Major Groceries	284	9
	42: Minor Groceries	234	
3	44: Clothing / Personal	190	12
4	43: Houseware	89	14
5	47: Other	262	19

The EPOI data provide two key pieces of information: 1) Available zones for shopping by type, and 2) Number of stores by shopping type and zone. Shopping type availability is used to constrain the choice set, while the number of stores is included in the location choice model as an attraction variable.

3.4   Using CHASE to Simulate Location Choice Behaviour

The current operational model of TASHA is not modular. Hence, to test the proposed location choice model, a standalone model was developed in C# using the same class structure as TASHA. This model is developed using the CHASE dataset, instead of the much larger TTS dataset, to take advantage of the detailed activity and scheduling information in CHASE.

Three main tables are used from CHASE: 1) Person table, 2) Household table, and 3) Schedule table. Data from the Person and Household tables were used to synthesize the sample population. The Schedule table was used to generate the provisional schedule and the list of shopping episodes per individual. The details of the input files derived from the CHASE tables are shown in Appendix B.

3.4.1   Generating the Provisional Schedule

The **provisional schedule** is an individual's list of scheduled activities at a particular instance in time (Miller, 2005a). In this application, the instance in time is right before the shopping episode is scheduled. CHASE has ten general activity types: Basic Needs, Work / School, Household Obligations, Drop-off / Pick-up, Shopping, Services, Recreation / Entertainment, Social, Just for Kids, and Other. The following activity types were assumed to take priority over Shopping: Work / School, Household Obligations, Drop-off / Pick-up, and Services (see Figure 3.3). This is similar to the project priority used in TASHA (Roorda et al., 2008).

These higher priority activities are inserted into an individual's provisional schedule and are blocked periods. Open periods (or gaps) are time windows where the shopping episode can be scheduled (see Figure 3.3). The gaps in an individual's weekly schedule form the basis for the gap choice set and the temporal boundaries for the time-space prism (see Figure 3.4)

It is important to note that simple and sensible heuristics are used to generate the provisional schedule (or activity skeleton schedule), rather than explicitly modelling the schedule formation process. Habib and Miller (2006) have developed hazard-based models to represent the daily skeleton schedule formation for workers. While activity skeleton formation is an important behavioural process, modelling the schedule formation process is left for further research. The simple heuristics used in this implementation attempt to replicate the schedule formation process in TASHA and show the significance of small changes to the existing location choice model.
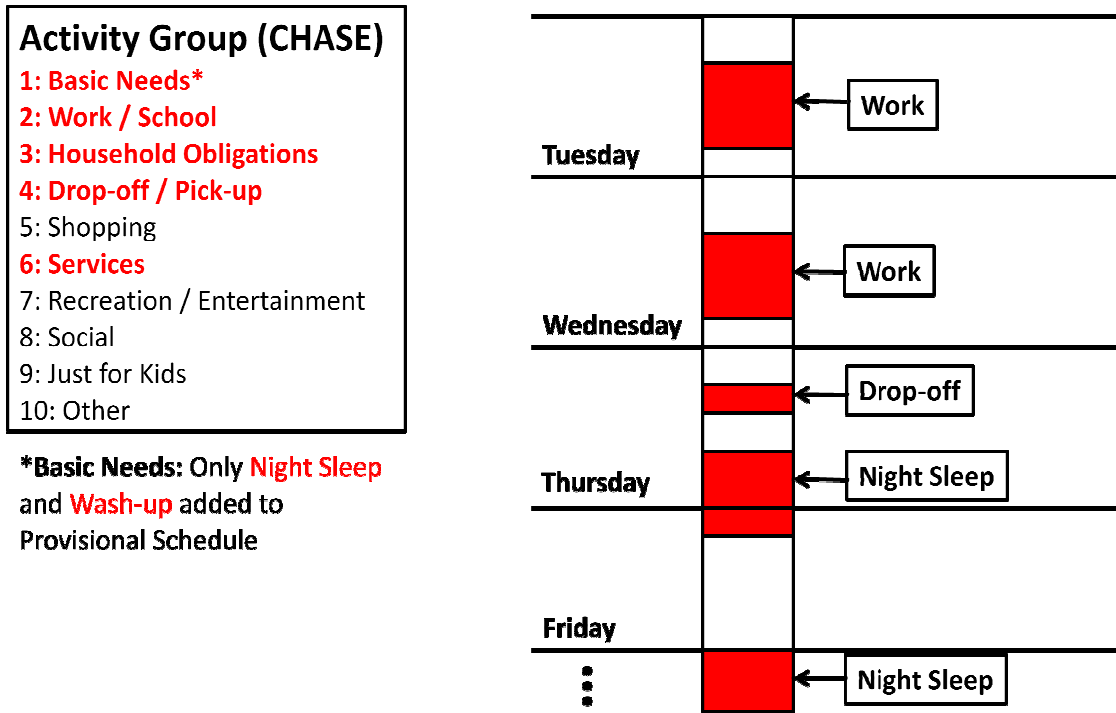
Figure 3.3 Adding Activities to the Provisional Schedule
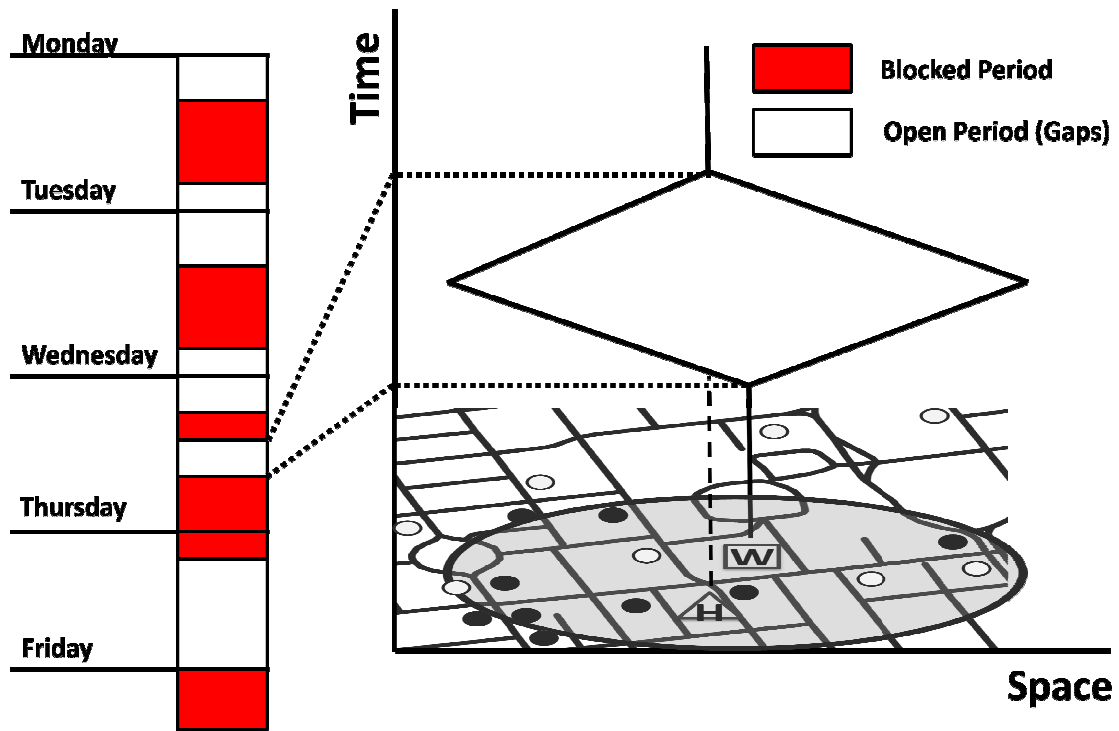
Figure 3.4 Open Periods in Provisional Schedule Generate Feasible Set of Locations

Finally, only out-of-home activities are scheduled in an individual's provisional schedule. TASHA only models out-of-home activities, assuming an individual is at home otherwise. The only in-home activities scheduled are Night Sleep and Wash-up, which ensure that an activity is not scheduled while an individual is sleeping. In the future, this can be replaced by retail store availability constraints on shopping time by shopping type (retail store hour information at the zone level was unavailable).

Of the 34,279 episodes in the Schedule table, 17,502 episodes were scheduled, including out-of-home activities and the in-home activities mentioned above. There is a total of 1194 shopping episodes (and observations) between 416 individuals, with each shopping episode scheduled being assumed to be an independent event. This is a strong assumption since individuals have been observed to group some shopping activities and to shop with others. Joint shopping activities can be handled when this model is integrated into the modelling framework. Trip chaining effects arising from shopping activities grouped with other activities are not handled in the proposed model. However, in this model, trip chaining behaviour is accounted for in the location choice model by using the travel time to and from the prior and posterior episodes in the provisional schedule.

3.4.2   Modelling Framework

The software is designed to allow for model estimation and simulation to test the proposed gap-location choice model. The modelling framework uses an object-oriented approach, adopting many of the classes used in TASHA. Class diagrams of the key classes are shown in Figures C1.1 to C1.6 in Appendix C, with additions to the TASHA classes shown in red. To ensure consistency with the current implementation of TASHA, this program was developed in C# and was designed to emulate TASHA's scheduling process using travel diary data.

3.4.2.1   ALOGIT for Model Estimation

The gap and location choice models were estimated using ALOGIT 4.2C (ALOGIT, 2007), which can estimate the multinomial logit (MNL) and nested logit (NL) models used in this thesis. ALOGIT allows for both sequential and simultaneous estimation of the nested logit model. However, the models presented in this thesis use a sequential approach.

While it is desirable to exploit ALOGIT's ability to simultaneously estimate nested logit models with unlimited choice set sizes, we found that the constraints imposed on the gap-location choice model were not sufficient to allow for model estimation. With a maximum choice set size of 63,468 choices, arising from a maximum of 41 gaps and up to 1548 locations within each gap, we found ourselves returning to the issue of unrealistic choice set sizes. It is behaviourally unrealistic for individuals to choose from a set of over 60,000 alternatives, let alone meaningfully distinguish one alternative from another. Further research can be done to model the choice sets of gaps and locations.

3.4.2.2   Location Choice Set Generation

The consideration set (C) is the union between the feasible set (F) and the awareness set (A). The feasible set of locations is defined by Hägerstrand's time-space prisms. The temporal boundaries of the time-space prisms are defined by the gaps that arise from the generation of the provisional schedule. Auto travel times (to and from the shopping location) and shopping episode duration are used to define the time-space prisms. In the current model, the awareness set is assumed to be the set of all available shopping zones as defined by the EPOI data (i.e. assumption of perfect information of retail store availability at the zone level). Section 4.4 shows the significant reduction in choice set size resulting from these feasibility and availability constraints.

The choice set (consideration set) is defined by the set of all zones that are in the prism's feasible region and have at least one store of the same shopping type as the shopping episode to be scheduled. The computer program generates this choice set by selecting from a set of zones (see Figure B1, Appendix B, for the attributes of the zone file) that satisfy the individual's scheduling constraint.

 Figure 3.5 shows the modules and data sources used to generate the choice sets of locations (from the Choice Set Generator) and gaps (from the Provisional Schedule). Figure C2 in Appendix C shows the program flow chart, with each component described in pseudocode in Figures C3.5 to C3.8.
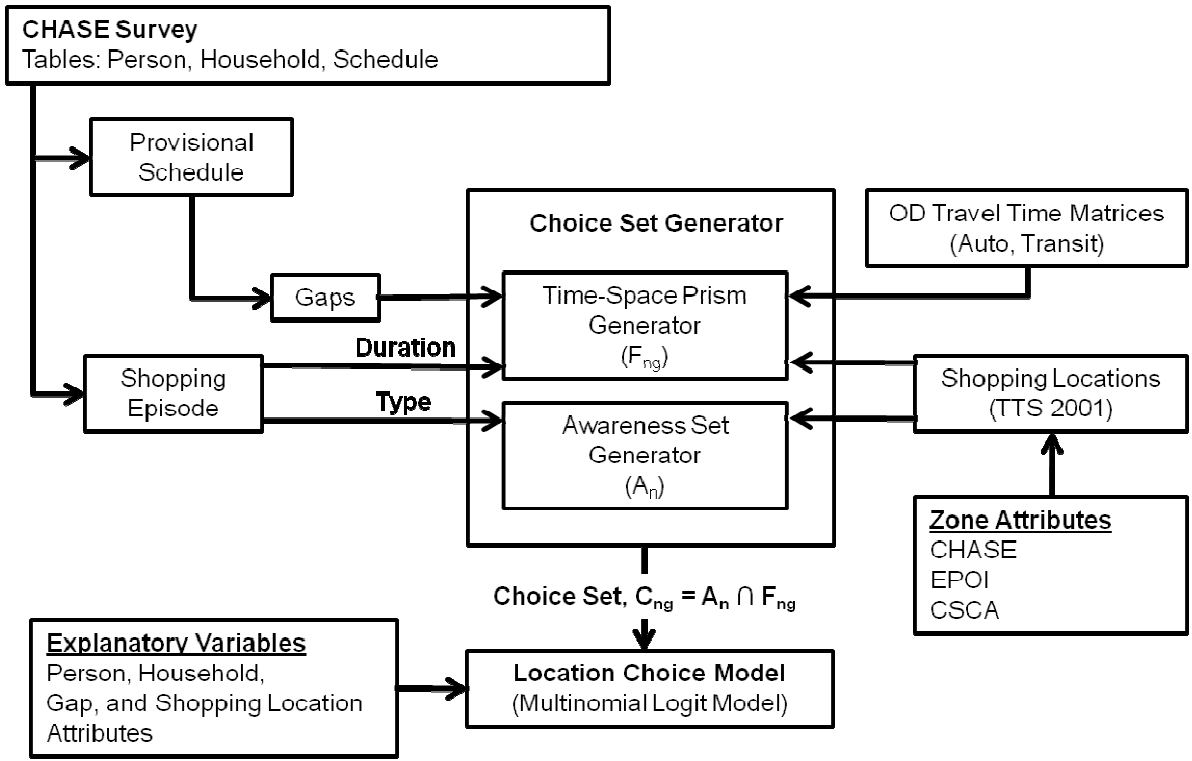
Figure 3.5 Proposed Modelling Framework for Shopping Location Choice

4 Gap-Location Choice Model for Shopping Trips

The gap-location choice model is a unique approach to modelling shopping location choice decisions. In this modelling framework, we propose that individuals choose locations in the context of a time period choice (e.g. "*I will do grocery shopping Thursday night and will go to the No Frills since I will be in the area*"). While individuals exhibit different travel and activity behaviour, this approach makes a step forward in improving the behavioural realism of location choice decisions for shopping activities.

Furthermore, we hypothesize that gap choice is more realistic than a specific start time choice (e.g. "*I will shop in the evening*", rather than "*I will shop at 7:26pm*"). It is more likely than individuals choose from a discrete set of time periods, rather than along a continuous time interval. Ultimately, an activity start time will need to be generated when the episode is scheduled. However, in the context of location choice decisions, the activity "start time" is usually chosen from a set of time windows or gaps.

We began by testing a nested choice structure of gap and location, with gap choice in the upper level and location choice in the lower level. This assumes gap is chosen first, with location choice conditioned upon the gap chosen. This assumption is examined in Section 4.2.

To test this theory, we use the nested logit (NL) formulation to model the location choice conditioned by the gap choice. The NL model is based on the multinomial logit (MNL) models, which originate from econometric theory used to predict a consumer's choice from a set of discrete alternatives. The MNL and NL models are explained in greater depth by Ben-Akiva and Lerman (1985).

One of the drawbacks behind MNL models is the assumption of the error terms in the utility functions of all alternatives. The MNL model assumes that these errors are independently and identically distributed (IID) with a Type I Extreme Value (Gumbel) distribution. The NL model relaxes the independence assumption by allowing correlation between alternatives within a nest, with each nest being IID. In the context of the gap-location choice problem, this implies that locations within a nest share some unobservable attribute(s) related to the gap. It is important to note that this assumption does not always hold (e.g. when the gaps exhibit similar characteristics, such as weekday evening gaps with similar time window lengths).

In mathematical terms, an individual's choice of a gap and location for a shopping activity can be represented by U_{gl}, the utility of location l in gap g (Note: the subscript specific to the individual is implicit and omitted for simplicity). From random utility theory, the actual utility perceived by the individual (U) is the sum of the systematic utility (V) and a random component (ε), which accounts for random unobserved variables influencing utility. Furthermore, as shown in Equation 4.1, the systematic utility of alternative l in nest g (V_{gl}) can be decomposed as the utility common to all locations in the gap (V_g), the utility specific to each location within the gap (V_l), and the utility specific to a location within a particular gap (V_{gl}).

$$U_{gl} = V_l + V_g + V_{gl} + \varepsilon_l + \varepsilon_g + \varepsilon_{gl} = V_l + V_g + V_{gl} + \varepsilon_g + \varepsilon_{gl} = V_g + V_{l/g} + \varepsilon_g + \varepsilon_{l/g} \qquad [4.1]$$

The nested logit structure used in this thesis (gaps in the upper level and locations in the lower level) assumes that the errors arising from the utility attributed location alternatives alone are negligible (i.e. $\varepsilon_l = 0$). Unobserved random components attributed to location choices are captured in the error term of the total utility (ε_{gl}, which reduces to $\varepsilon_{l/g}$). The error term for the gap utility (ε_g) captures the unobserved attributes shared by alternatives within each gap nest. The full derivation of the nested logit model is explained in depth in Ben-Akiva and Lerman (1985). The key results of this derivation are the systematic utility specifications of the upper (gap, V_g) and lower (location, $V_{l/g}$) levels, and their associated choice probabilities.

The general specifications for systematic utilities of the gaps (V_g) and locations ($V_{l/g}$) are as follows (Sections 4.1 and 4.3 provide the exact specifications used for the location and gap location utilities respectively):

$$V_g = \alpha_1 EpGapDurRatio_g + \alpha_2 TimePeriodDayDummys_g + \varphi_g LogSum \qquad [4.2]$$

> α_1 is the vector of parameters associated with the *EpGapDurRatio* based on income and gender.
>
> *EpGapRatio$_g$* is the ratio between the shopping episode duration to the gap duration. The episode duration includes auto travel times to and from the shopping location.
>
> α_2 is the vector of parameters associated with *TimePeriodDayDummys*, based on whether the gap is on a weekday or weekend, and whether it falls within one of four time periods (Morning – 6am to 12pm, Afternoon – 12pm to 5pm, Evening – 5pm to 9pm, Late – 9pm to 6am).

TimePeriodDayDummys$_g$ are dummy variables that indicate whether a portion of the gap falls within a certain time period.

φ_g is the parameter for the logsum variable or inclusive value

LogSum is the inclusive value, which is the expected maximum utility derived from the location alternatives in nest *g*. In this model, a simplifying assumption is made by assuming the inclusive value parameter is same across all nests

$$V_{l/g} = \beta_1 AutoCost_{gl} + \beta_2 AutoTime_{gl} + \beta_3 Log(NumRetailStores_l) + \beta_4 Log(RetailArea_l) \qquad [4.3]$$

β_1 is the vector of parameters associated with *AutoCost$_{gl}$*

AutoCost$_{gl}$ is the auto cost (\$) associated with travelling to location *l* from the prior location and from location *l* to the posterior location. The prior and posterior locations are defined by the start and end locations of the time window.

β_2 is the vector of parameters associated with *AutoTime$_{gl}$*

AutoTime$_{gl}$ is the auto travel time (minutes) associated with travelling to location *l* from the prior location and from location *l* to the posterior location. The prior and posterior locations are defined by the start and end locations of the time window.

β_3 is the vector of parameters associated with *Log(NumRetailStores$_l$)*

Log(NumRetailStores$_l$) is the natural logarithm of the number of retail stores by shopping type defined by EPOI data

β_4 is the vector of parameters associated with *Log(RetailArea$_l$)*

Log(RetailArea$_l$) is the natural logarithm of the commercial area (sq ft) from CSCA data

Notice that the auto travel time variable (*AutoTime$_{gl}$*) considers trip chaining behaviour by using the prior and posterior episodes as anchor points. Another variable used in explaining trip chaining is the deflected or incremental travel time, which is the total travel time minus the travel time without the shopping episode. In the final choice probabilities, this total travel time gives the same results as the incremental travel time, since the travel time from the prior to posterior episodes is constant across all alternatives within a gap.

Equations 4.4 to 4.6 define the choice probabilities and are used by ALOGIT for model parameter estimation (see Tables D1 and D2 in Appendix D for ALOGIT code), and by the C# program for simulation of gap and location choice decisions (see Chapter 5). The probability that an individual chooses location *l* given the choice of gap *g* is

$$P_{gl} = P_g \cdot P_{l|g} \tag{4.4}$$

With $P_{l|g} = \dfrac{\lambda_g e^{V_{l|g}}}{\sum_{l' \in L_g} e^{\lambda_g V_{l'|g}}}$ [4.5]

$$P_g = \frac{e^{\beta V_g}}{\sum_{g' \in G} e^{\beta V_{g'}}} = \frac{e^{\beta X_g + \varphi_g \log\left(\sum_{l' \in L_g} e^{V_{l'|g}}\right)}}{\sum_{g' \in G} e^{\beta X_{g'} + \varphi_g \log\left(\sum_{l' \in L_{g'}} e^{V_{l'|g'}}\right)}} \tag{4.6}$$

Where β is the scale parameter for the upper level and normalized to 1, and

λ_g is the scale parameter associated with gap g, and assumed to be the same across all nests.

4.1 Location Choice Model

As previously discussed in Section 3.4.2.1, the nested logit model developed in this thesis is estimated sequentially. The first step in estimating the gap-location choice model is to develop a model specification and estimate model parameters for the lower level model.

We began by testing the travel time variable and found that the travel time for the entire trip chain (to and from the shopping location) was a significant variable. Auto cost for the entire trip chain was added to improve the model specification.

The number of retail stores and commercial area were then included in the model. These variables are size variables; hence the natural logarithm of each of these variables entered the final model specification. This logarithmic relationship is sensible as one can imagine a zone being very attractive with a certain amount of commercial space and retail stores; however, the marginal benefit decreases for an individual interested in choosing a shopping location.

The final model specification shown in Table 4.1 also accounts for individual attributes, such as income and gender. A more detailed model was developed for the actual simulation, as shown in Table D1 in Appendix D. The models in Table 4.1 and in Table D1 give similar behavioural results. The specification presented in Table 4.1 is shown here as it is simpler without making too many strong assumptions.

The key difference between the two models is that the model presented in Table D1 assumes both auto cost and auto travel time vary by both gender and income. The model presented in Table 4.1 assumes that gender does not significantly impact the influence of auto cost and income does not significantly impact the influence of auto travel time. Rather, the similarities and differences of gender perceptions of travel time and the influence of income on the impact of auto cost were investigated.

Table 4.1 Final Location Choice Model

	All Types		Type 1		Type 2	
Parameter	Estimate	T-Ratio	Estimate	T-Ratio	Estimate	T-Ratio
Auto Cost						
Low Income (<$30,000)	-0.09840	-2.8	-0.06497	-0.4	-0.1535	-2.4
Med Income ($30,000 to $60,000)	-0.05826	-1.8	-0.06771	-0.5	-0.06123	-1.2
High Income (≥$60,000)	-0.007534	-0.2	- (0.2913)	- (1.2)	- (0.08433)	- (1.6)
Auto Travel Time						
Male	-0.1824	-19.4	-0.4027	-7.2	-0.2152	-14.7
Female	-0.1800	-23.5	-0.3893	-8.2	-0.2311	-17.5
Log(number of retail stores)	0.7590	21.5	- (-0.1864)	- (-1.3)	0.8723	14.8
Log(retail floor space)	0.1179	12.8	0.04761	1.7	0.1008	8.2
Log-Likelihood	-2661.9360		-286.9652		-1199.2636	
Rho-Squared w.r.t. Zero	0.5378		0.4923		0.5619	

Table 4.1 Final Location Choice Model, cont'd

	Type 3		Type 4		Type 5	
Parameter	Estimate	T-Ratio	Estimate	T-Ratio	Estimate	T-Ratio
Auto Cost						
Low Income (<$30,000)	-0.1718	-2.4	-0.2474	-1.6	0.03575	0.5
Med Income ($30,000 to $60,000)	-0.09991	-1.5	-0.3843	-2.3	0.06912	1.1
High Income (≥$60,000)	-0.2651	-2.9	-0.2139	-1.4	-0.02273	-0.2
Auto Travel Time						
Male	-0.09564	-4.9	-0.1034	-3.6	-0.1557	-8.2
Female	-0.1139	-7.6	-0.07943	-3.0	-0.1476	-10.6
Log(number of retail stores)	1.020	12.9	0.7125	5.5	0.6113	9.6
Log(retail floor space)	0.1680	6.8	0.2476	4.2	0.1469	6.5
Log-Likelihood	-429.9943		-123.1604		-488.6452	
Rho-Squared w.r.t. Zero	0.5303		0.7055		0.5650	

As Table 4.1 shows, given the small dataset of 1194 observations, the model predicts location choices reasonably well. Most variables are statistically significant at the 90% confidence interval for the two-tailed t-statistic (T-Ratio, $t = 1.64$). Given the small dataset, some parameters that did not meet the statistical criterion were retained in the model as they provided behavioural insight into the location choice decision process. The model parameters are also reasonable with negative values for auto travel time and auto cost, and positive values for the zone-specific accessibility variables (i.e. number of retail stores and retail floor space).

The goodness-of-fit test used is the adjusted Rho-squared value taken with respect to zero, since the model does not have any alternative specific constants. These adjusted Rho-squared values are very strong compared to typical destination choice models, indicating the value of modelling at the individual level and of using schedule-based time-space prisms to define choice sets.

The model parameter estimates shown in Table 4.1 reveal the following insights into shopping location choice behaviour:

- Convenience and Grocery shoppers exhibit greater sensitivity to the total travel time than individuals making shopping trips of other types. This suggests that trips made to zones that incur minimal additional travel time, while shopping locations of other types (including Personal, Houseware, and Others) are chosen more for the specific shopping purpose rather than for trip convenience.

- Individuals are also more generally sensitive to travel time than travel costs. Personal and Houseware shopping trips are an exception to this, with larger values of travel cost parameters. The difference likely lies in the purpose-specific nature of the shopping trips. Convenience and Grocery trips are likely to occur more regularly as they satisfy more fundamental human needs such as health (e.g. drug store) and the need for food (e.g. grocery store). Personal and Houseware trips can be viewed as 'lower priority' trips that satisfy other "needs" such as comfort and luxury. Further investigation on shopping type priority will be required to validate this theory.

- The number of retail stores is positive for all models with the exception of the Convenience shopping location model. This variable was omitted for the Convenience trip model since a negative parameter was not realistic for the number of retail stores

variable. This suggests that the number of convenience stores does not have a significant influence on the location choice of convenience stores trips. Rather, the influence of travel time dominates, indicating the importance of scheduling constraints on convenience store location choice.

- Commercial floor space (used as a proxy variable for retail floor space) is statistically significant for all models. Personal and Houseware trips are especially sensitive to retail floor space. This makes sense since many stores that sell clothes or house items have large stores or are part of large department stores. On the other hand, convenience store trips are not strongly influenced by floor space, as suggested by the small parameter for floor space.

- In this model, both males and females (from the CHASE dataset) show very similar degrees of sensitivity to travel time, with the gender-specific parameters not being significant from one another at the 99% confidence level. This suggests that the distinction between males and females are less pronounced compared with the early 20th century, where gender roles were more defined. Gender may not be sufficient or adequate to explain the heterogeneity in the value of travel time, which may now depend on other variables such as individual income, marital status, and employment type. These influences can be investigated in further research.

- On the other hand, individuals with different income levels showed varying sensitivity to auto travel cost. In particular, individuals with lower income showed higher sensitivity to travel costs for Convenience and Grocery trips, while Personal and Houseware trips showed less pronounced (and even reversed) differences. The auto cost parameters for the Convenience and Grocery location choice models show that auto cost is not significant for those with higher incomes (as shown by the positive values of high income auto cost parameters). On the other hand, both Low and High income individuals exhibit higher degrees of sensitivity to auto cost for Personal and Houseware trips. This suggests that there are other unobserved factors influencing the perception of auto cost for these two shopping types apart from auto cost relative to individual income.

Stepping away from these behavioural insights, the key point is that these parameter estimates show that a very robust and behavioural location choice model can be developed by disaggregating shopping trips by type and by constraining choice sets. This in itself is a significant improvement to TASHA's current location choice model, described in Section 3.1. The next two sections will describe the insights gained by considering time window choice directly in the proposed gap-location choice modelling framework.

4.2   Gap Choice Considerations

The next step in the sequential estimation of the gap-location choice model is to estimate the upper level gap choice model. The model development began with testing different variables that would characterize the gap, including the gap length and the shopping episode duration relative to the gap length. The final gap model used the ratio between the shopping episode duration, including the auto travel times to and from the shopping location, and the gap duration. The start and end locations and times of these gaps are defined by the prior and posterior episodes in an individual's provisional schedule.

The final model included dummy variables for time of day and day of week. Due to the relatively small sample, weekday and weekend dummies were used for day of week dummies. Time of day was divided into four periods: Morning (6am to 12 pm), Afternoon (12pm to 5pm), Evening (5pm to 9pm), and Late (9pm to 6am). The dummy variables used in the final model accounted for day of week and time of day interactions.

Finally, the inclusive value was added to the gap choice model to account for the utilities of the available locations in each gap. It is important to note that the number of locations in each nest varies depending on the gap length, since locations are constrained by time-space prisms. This difference in choice set size biases the results, since the gap duration influences the size of the inclusive value, which should only be influenced by the utility of the location alternatives within each gap and not the utility of attributed to the gap itself. Improving the choice set formation of the gap and location choice sets is an important area of future research.

Table D2 in Appendix D shows the parameter estimates for the gap choice model, with the inclusive value included for all types. This is shown to test the proposed gap-location choice model structure (i.e. gaps in the upper level, and locations in the lower level). Table 4.2

summarizes the inclusive value parameter estimates for each model (by type and gender). For the gap-location model structure, the inclusive value parameter needs to be between 0 and 1. The inclusive value parameter is also the ratio between the scale parameters for the lower (location) and the upper (gap) levels. Since the scale parameters need to be greater than 0, the inclusive value also needs to be greater than 0. Further, the inclusive value parameter needs to be less than or equal to one to ensure consistency with random utility theory (e.g. if the inclusive value parameter was greater than one, this would imply an inverted model with location choice in the upper level and gap choice in the lower level).

Table 4.2 Inclusive Value Estimates for Upper Level Gap Model

Shopping Type	IV Value	T-Ratio	Num Obs	Model Structure
All Types				
Male	0.6714	1.6	471	Nested: Gap-Location
Female	-0.06495	-0.5	670	Independent (IV = 0)
Type 1				
Male	2.576	0.3	37	Independent (IV = 0)
Female	-0.4958	-1.3	46	Independent (IV = 0)
Type 2				
Male	-0.8319	-0.7	225	Independent (IV = 0)
Female	0.6861	1.0	287	Nested: Gap-Location
Type 3				
Male	1.575	2.5	60	Nested: Location-Gap – Re-estimate!
Female	-0.2385	-0.7	136	Independent (IV = 0)
Type 4				
Male	1.522	0.8	51	Nested: Location-Gap – Re-estimate!
Female	1.227	1.4	37	Nested: Location-Gap – Re-estimate!
Type 5				
Male	-0.5093	-0.6	98	Independent (IV = 0)
Female	-0.1495	-0.5	164	Independent (IV = 0)

Therefore, based on the value of the inclusive value parameter (IV), there are 3 possible model structures:

1. Independent Models (IV $\leq$ 0): Estimate the gap and location choices separately, removing the inclusive value from the gap choice model specification.

2. Nested Model, Gap-Location (0 < IV $\leq$ 1): Keep the current model. Note that in the case where IV = 1, this would suggest gap and location are chosen simultaneously and as independent gap and location choice bundles.

3. Inverted Nested Model, Location- (IV > 1): Test inverted model by using the gap model to calculate inclusive values for the location choice mode and re-estimate the location choice model.

Table 4.2 shows the model structure suggested by the IV for each model. The two-tailed *t*-statistic ($t = 1.645$, 90% confidence level) was used to help guide the significance of the inclusive value parameter.

IV estimates for Types 3 (Personal) and 4 (Houseware) suggest that an inverted model would be more appropriate. This suggests location is chosen before gap for these types of trips. This is not unreasonable since one could imagine having a particular store associated with a shopping type (e.g. *"I need to pick up some house items at the Home Depot [location decided]. Now when is a good time to do this?"*).

However, when the inverted model was tested, inclusive values were not estimable from the data. A larger dataset and an improved choice set model would help with additional testing. Further, in the location-gap approach, location choices are constrained only by availability and not by the time-space prism constraints. Shopping location would be chosen exogenously to the scheduling of the shopping episode, as is currently done in TASHA.

While additional research would be required to test the theory of gap-location choice, it is noted that only the "All Types – Male" and "Type 2 (Grocery) – Female" had significant inclusive value parameter estimates. While this might suggests that the gap-location choice framework applies to males and not females when predicting shopping location, it is important to look at the IV parameters for males for other shopping types. From examining the IV parameters of males across the 5 shopping types, one can see that the "sensible" IV estimate for the "All Types – Male" model is due to the aggregation of negative and >1 IV estimates.

On the other hand, the nested gap-location choice model seems to apply to females doing groceries (i.e. the choice of time period take precedence over the choice of grocery shopping location). However, even this model is only somewhat significant, since the T-Ratio is only 1.0 despite the relatively larger number of observations. In the final model validation, only female grocery trips adopt the nested gap-location choice model. The other shopping location models predict gap choice independently of the available shopping locations in each gap.

While this may seem discouraging, it is also encouraging since it suggests that the independent treatment of location choice in TASHA is not entirely unreasonable. The prism- and gap-based approach to shopping location choice would help constrain choice set and improve behavioural realism by modelling at the individual level. However, to validate the proposed nested gap-location choice structure, this model should be tested in the current TASHA implementation with the much larger TTS dataset.

4.3   Final Model Parameter Estimates

Table 4.3 presents the final gap model estimates based on the model structures. For the final model, the gap model was re-estimated without the IV parameter for all models that did not have the nested gap-location model structure, as shown in Table 4.2. The final gap choice model parameters reveal the following insights into gap choice behaviour for shopping trips:

- Episode-to-gap duration ratio parameters were significant and negative for all shopping types. The negative ratio value is sensible and implies that gaps are more attractive if they have a relatively larger time window length than the total time spent on the shopping activity (including shopping duration and travel time). A ratio parameter with a larger negative value implies an increased sensitivity to scheduling constraints, favouring gaps with longer time windows.

- Based on the ratio parameter estimates, individuals with lower income exhibit a higher sensitivity to scheduling constraints. This can be attributed to more fixed work hours or commitments. Interestingly, male convenience shoppers are very sensitive to scheduling constraints for Convenience trips, preferring longer relative gap lengths for their Convenience trips.

- Another interesting observation is the increasing sensitivity experienced by higher income males for Houseware trips. This trend suggests that males with higher income would choose gaps that allowed for greater flexibility for the shopping activity. Males with lower incomes may have a pre-planned list of house items to be bought (need-based spending), while males with higher incomes would presumably have more discretionary income to spend on items as they encounter them (need-based + discretionary spending).

- Notice, for larger negative values of this ratio, the model alone is not sufficient to distinguish between an individual operating on tight scheduling constraints, and an individual choosing larger gap lengths to accommodate the greater variability associated with the shopping activity. The previous explanations were assumptions based on higher income individuals having greater flexibility in work schedule and larger amounts of discretionary income. An area to be examined in further research is the influence of income on shopping duration variability by shopping type.

- Based on the time period dummies, individuals generally favour shopping in the Evening and Late periods for weekday trips and the Morning and Afternoon periods for weekend trips. This is reasonable since individuals usually have work commitments during the weekday. In the weekend, stores generally close earlier and individuals may have personal non-shopping commitments in the later parts of the day.

- Time period dummies are not very significant for the gap choice of Convenience trips, which depend more on individual scheduling constraints or on the variance of the shopping activity duration.

Overall, the gap choice model provides insight into gap choice behaviour for shopping trips, despite the relatively low Rho-squared goodness-of-fit values. To improve upon this model in the future, a larger dataset and an improved gap choice set generation model would be required.

The theory of time window choice should be further investigated and considered for implementation in an activity-based model, such as TASHA. When considering location choice, individuals usually consider from a set of discrete time windows, rather than choosing from a continuous time domain. Assigning a specific activity start time better reflects the execution of the activity episode rather than an activity start time choice jointly made with location choice. Discrete time window (gap) choice should be considered in activity-based models seeking to model schedule formation more dynamically.

Table 4.3 Final Gap Choice Model

Parameter	Male		Female	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
All Types				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-2.007	-3.0	-2.078	-5.1
Medium Income ($30,000 to $60,000)	-1.668	-4.3	-1.446	-4.5
High Income (≥$60,000)	-1.842	-4.2	-1.476	-3.0
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-0.2964	-0.5	-0.1870	-0.4
Eve (5pm-9pm)	2.134	3.2	1.370	2.3
Lat (9pm-6am)	1.287	6.5	1.228	7.2
Weekend				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	1.232	2.4	0.7143	1.7
Eve (5pm-9pm)	-1.010	-2.2	-1.048	-2.5
Lat (9pm-6am)	-0.5236	-1.3	-0.8021	-2.6
IV	0.6714	1.6	-	-
Log-Likelihood	-1101.5916		-1662.8219	
Rho-Squared	0.0924		0.0595	
Num. Observations	471		670	
Parameter	**Male**		**Female**	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 1				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-8.329	-0.7	-3.757	-1.8
Medium Income ($30,000 to $60,000)	-3.448	-2.2	-1.223	-1.1
High Income (≥$60,000)	-9.655	-1.4	-1.088	-0.5
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-	-
Eve (5pm-9pm)	-	-	-	-
Lat (9pm-6am)	-	-	-	-
Weekend				
Mor (6am-12pm)	-	-	1.770	2.5
Aft (12pm-5pm)	2.034	3.0	-	-
Eve (5pm-9pm)	-	-	-	-
Lat (9pm-6am)	-	-	-	-
IV	-	-	-	-
Log-Likelihood	-87.4168		-123.5878	
Rho-Squared	0.1340		0.0663	
Num. Observations	37		46	

Table 4.3 Final Gap Choice Model, cont'd

Parameter	Male		Female	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 2				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-2.574	-2.5	-3.886	-4.3
Medium Income ($30,000 to $60,000)	-1.183	-2.4	-1.986	-3.6
High Income (≥$60,000)	-1.397	-2.6	-2.195	-3.0
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	0.2671	0.4	-0.5267	-0.7
Eve (5pm-9pm)	0.5272	0.9	1.100	1.2
Lat (9pm-6am)	-	-	0.9313	3.7
Weekend				
Mor (6am-12pm)	1.280	4.6	-	-
Aft (12pm-5pm)	0.7854	1.4	1.434	2.3
Eve (5pm-9pm)	-	-	-0.8287	-1.2
Lat (9pm-6am)	-	-	-	-
IV	-	-	0.6861	1.0
Log-Likelihood	-540.0725		-710.5559	
Rho-Squared	0.0854		0.0809	
Num. Observations	225		287	
Parameter	**Male**		**Female**	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 3				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-1.610	-1.3	-1.701	-2.1
Medium Income ($30,000 to $60,000)	-1.159	-1.0	-1.705	-2.3
High Income (≥$60,000)	-0.6341	-0.6	-0.8295	-0.9
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-	-
Eve (5pm-9pm)	-	-	2.079	1.7
Lat (9pm-6am)	-	-	-	-
Weekend				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	0.9886	1.4	0.5034	0.9
Eve (5pm-9pm)	-0.9182	-1.0	-2.131	-2.0
Lat (9pm-6am)	-0.8679	-1.1	-0.9229	-1.6
IV	-	-	-	-
Log-Likelihood	-138.6488		-316.0355	
Rho-Squared	0.0291		0.0352	
Num. Observations	60		136	

Table 4.3 Final Gap Choice Model, cont'd

Parameter	Male		Female	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 4				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-1.366	-0.7	-2.287	-1.2
Medium Income ($30,000 to $60,000)	-2.684	-1.8	-1.869	-1.2
High Income (≥$60,000)	-5.633	-2.8	-1.602	-0.4
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-2.776	-1.5
Eve (5pm-9pm)	-	-	2.389	1.2
Lat (9pm-6am)	-	-	-1.038	-0.7
Weekend				
Mor (6am-12pm)	-	-	3.190	2.0
Aft (12pm-5pm)	2.183	1.8	2.117	1.7
Eve (5pm-9pm)	-1.598	-1.3	-	-
Lat (9pm-6am)	-0.6953	-0.5	-	-
IV	-	-	-	-
Log-Likelihood	-113.0698		-83.2299	
Rho-Squared	0.1302		0.1434	
Num. Observations	51		37	
Parameter	**Male**		**Female**	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 5				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-3.521	-1.7	-0.9608	-1.5
Medium Income ($30,000 to $60,000)	-3.865	-3.4	-1.048	-1.8
High Income (≥$60,000)	-5.365	-3.2	-0.6641	-0.6
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-0.8257	-0.9
Eve (5pm-9pm)	-	-	0.6039	0.4
Lat (9pm-6am)	-	-	1.445	4.1
Weekend				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	1.508	2.3	1.211	1.6
Eve (5pm-9pm)	-1.316	-1.5	-0.8242	-1.0
Lat (9pm-6am)	0.4042	0.4	-0.3973	-0.6
IV	-	-	-	-
Log-Likelihood	-220.5617		-418.9921	
Rho-Squared	0.1159		0.0430	
Num. Observations	98		164	

4.4 Impact of Choice Set Constraints

TASHA's existing location choice model uses the universal set as the considered choice set (e.g. 1548 zones in the GTA). The choice set constraints used in this application generate more realistic choice sets, without making any behavioural assumptions on the choice set formation process. To illustrate the reduction in choice set from these constraints, time-space prisms were generated for each gap in an individual's provisional schedule for an entire week. Figure 4.1 shows the reduction in choice set from just the time-space prism constraints (i.e. average of 1193 zones per nest). Figure 4.2 shows the significant improvement from adding the activity type constraint, which removes a zone from the choice set if it does not have a retail store (i.e. average of 90 zones per nest).

While the choice set sizes are still unrealistic for some gaps, these figures show the reduction in choice set with time-space prism and retail availability constraints. Further research will be required to investigate the choice set formation process of how gaps are considered and which locations enter the choice set in each considered gap. Additional data will help in gaining insights into an individual's awareness of gap and location alternatives.
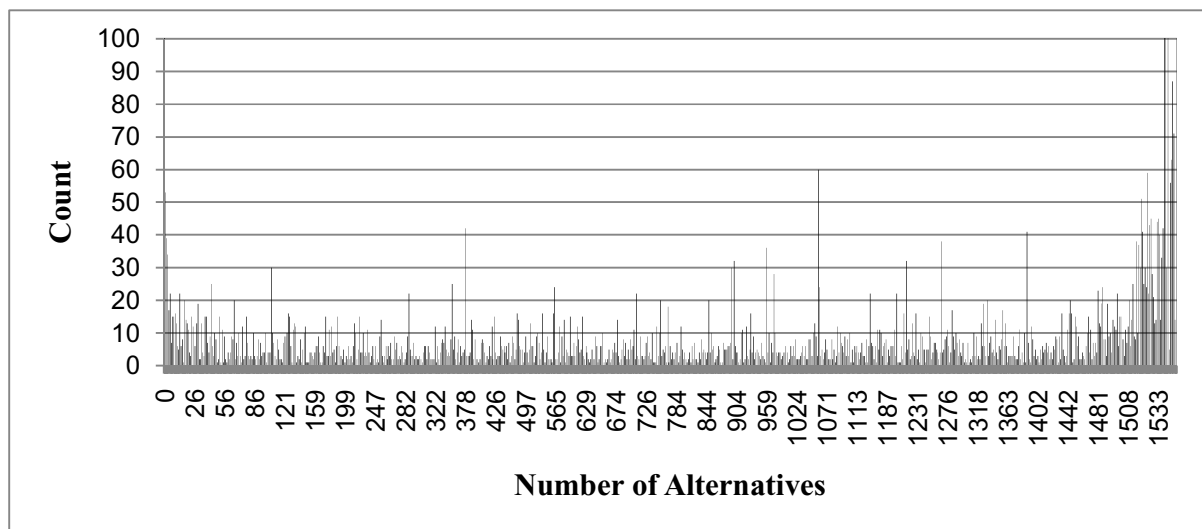
Figure 4.1 Time-Space Prism Constraints
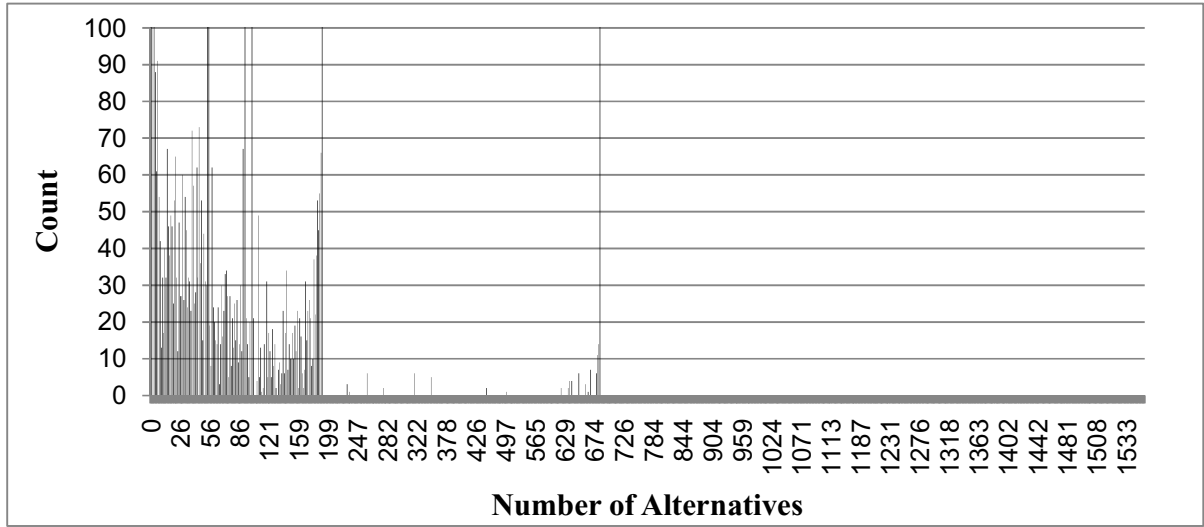
Figure 4.2 Time-Space Prism and Activity Type Constraints

5 Model Validation

As shown in the previous chapter, the gap-location choice models have fairly reasonable Rho-squared values. While this "goodness-of-fit" measure shows that the model reasonably replicates the data, what do these values actually mean? While modellers may be interested in these statistical measures, practitioners and policy makers are more likely to be interested on how well the model predicts gap and location choices. Given the data available, a natural step would be to attempt to replicate the gap and location choice patterns observed in the CHASE dataset.

5.1 Method

To validate the model against observed gaps and locations choice patterns in the CHASE data, the gap and location choices were simulated using Equations 4.5 and 4.6 and taking random draws from a uniform distribution. The pseudocode for the simulation component is shown in Figure C3.8 in Appendix C. As shown in Figure C2, the same program is used for choice set generation for both model estimation and simulation.

To determine a reasonable number of iterations for model simulation, the distributions of distances from the chosen zone were generated for each shopping type using 1, 10, 100, and 1000 iterations (see Table E1 and Figures E1.1 to E1.5 in Appendix E). The shapes of the distributions were visually inspected and observed to converge at 100 iterations. Hence the number of iterations chosen was 100.

5.2 Percent Right for Gap and Location Choices

Table 5.1 summarizes the percent right of gap and location choices for each shopping type.

Table 5.1 Percent Right of Simulated Gap and Location Choices

Iterations: 100					
Shopping Type	**1**	**2**	**3**	**4**	**5**
Num Total	6883	53400	20400	8900	26700
Num Loc Right	296	1362	350	206	414
% Loc Right	4.30%	2.55%	1.72%	2.31%	1.55%
Num Gap Right	822	5376	2189	1418	2619
% Gap Right	11.94%	10.07%	10.73%	15.93%	9.81%
Avg Dist (km) from Pred to Chosen	7.22	8.01	10.23	9.92	9.46
Max Dist (km) from Pred to Chosen	83.08	104.00	93.52	89.30	94.15

As Table 5.1 shows, the proportions of gap and location choices predicted correctly are quite low. This result shows one of the weaknesses of destination choices models (and the proposed gap choice model). Models of this type have a large number of alternatives and are likely to share some unobserved attributes. For example, location choice is modelled at the zone level, whereas individuals choose a store or retail centre.

Also, gaps were generated using simplistic rules to build a provisional schedule. This allows for gaps to be concentrated on certain days and for meaningless gaps to be generated. Moving forward, two key questions to answer are: 1) *What defines a gap?* and 2) *What defines a shopping location?* The problem of choice set characterization is nontrivial.

5.3   Gap Choice Predictions

Despite the low percent right values for predicting the exact gaps chosen, Tables 5.2 and 5.3 show the model predicts gap choice reasonably well at an aggregate level. Additional data can be found in Tables E2.1 to E2.4 in Appendix E.

Table 5.2 Gap Choice Predictions by Time Period

Chosen	Time Period			
Shopping Type	Mor	Aft	Eve	Lat
Type 1	8%	6%	5%	7%
Type 2	43%	48%	53%	44%
Type 3	19%	17%	18%	18%
Type 4	9%	10%	8%	7%
Type 5	22%	19%	16%	24%
Total	26%	12%	8%	55%
Predicted	**Time Period**			
Shopping Type	Mor	Aft	Eve	Lat
Type 1	11%	11%	7%	8%
Type 2	46%	43%	44%	45%
Type 3	15%	18%	29%	17%
Type 4	10%	14%	9%	7%
Type 5	18%	14%	11%	24%
Total	22%	11%	9%	58%
Difference	**Time Period**			
Shopping Type	Mor	Aft	Eve	Lat
Type 1	3%	4%	2%	0%
Type 2	3%	-5%	-9%	1%
Type 3	-4%	1%	11%	-1%
Type 4	2%	4%	0%	-1%
Type 5	-4%	-5%	-5%	0%
Total	-4%	0%	1%	3%

Table 5.3 Gap Choice Predictions by Day of Week

	Chosen		Predicted		Difference	
Shopping Type	Weekday	Weekend	Weekday	Weekend	Weekday	Weekend
Type 1	9%	5%	5%	12%	-3%	7%
Type 2	45%	44%	44%	48%	-1%	4%
Type 3	17%	18%	21%	9%	4%	-10%
Type 4	7%	8%	5%	14%	-3%	6%
Type 5	22%	25%	25%	17%	3%	-7%
Total	61%	39%	70%	30%	8%	-8%

An interesting point to note is that the gap choice model exhibits the creates degree of variability on Saturday (see Table E2.4). This reinforces the importance of developing a strong weekend model to account for the heterogeneity in weekend shopping and travel behaviour.

5.4   Location Choice Predictions

Similarly, for location choice predictions, despite the low percent right values, the model predicts the general spatial patterns reasonably well as shown in Figure 5.1. Similar figures are shown in Figures E3.1 to E3.5 in Appendix E for the spatial distribution of predicted locations by shopping type.

Furthermore, average distance from the chosen shopping locations ranged from 7 km to 10 km. When examining the distributions of these average distances, a relatively strong peak is observed around 3km to 6km (refer to Figures E1.1 to E1.5). These are fairly reasonable results given the large choice set and small dataset.

Furthermore, a significant amount of trips were observed to be 25 km or more away from the chosen location. This is likely due to the inclusion of all zones in the GTA and the random variability inherent in the model. It is also possible that individuals are choosing between shopping malls or power retail centres (Buliung et al., 2007), instead of the choosing between retail stores, which is simulated at the zone level.

Finally, the attractiveness of each zone could be better characterized by including other variables, such as commercial area by shopping type, and price of goods by zone and shopping type. The data for these variables were not available during model development.
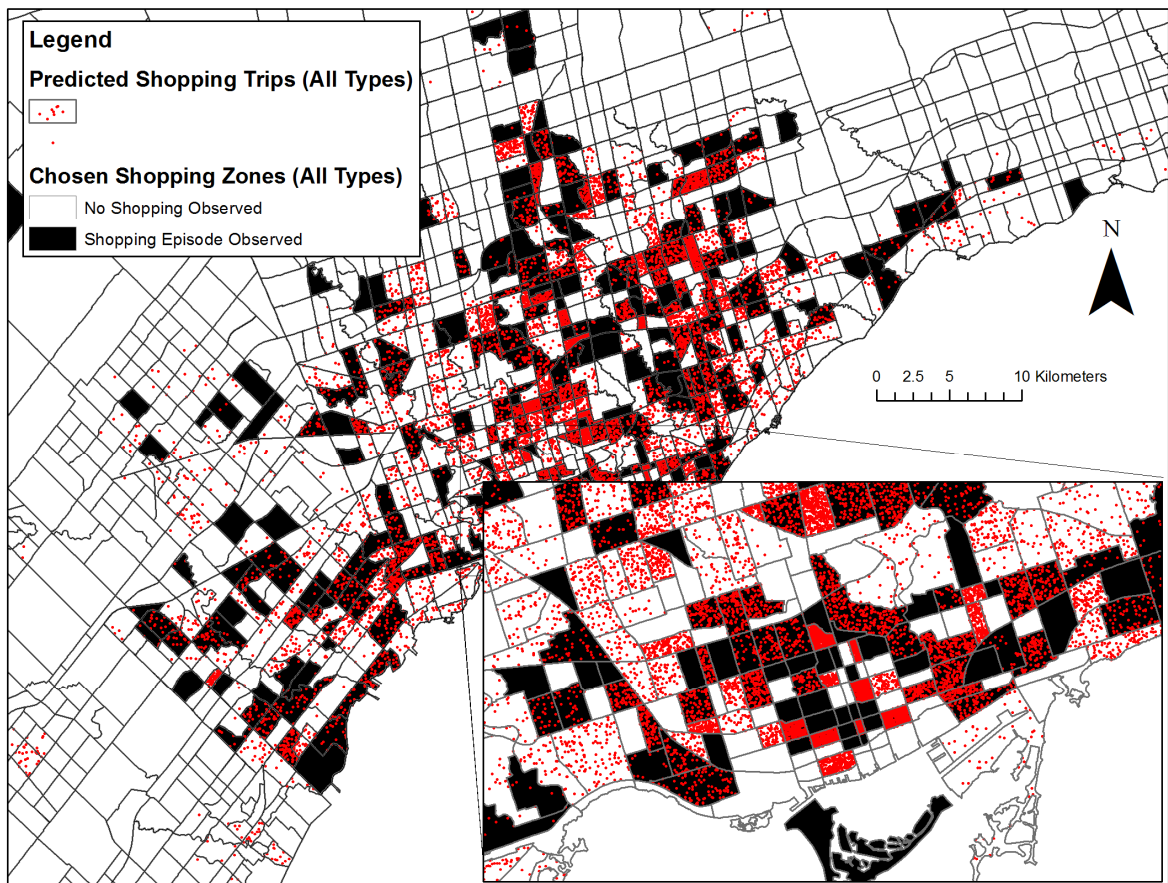
Figure 5.1 Spatial Distribution of Predicted and Chosen Shopping Locations

5.5 Travel Time Distributions

As a final goodness-of-fit measure, the travel time distributions were examined and compared using *t*-tests to compare the means and visual inspection, which provide an approximate measure of the similarity between the two distributions. The travel time distributions presented below are for the "All Types" model (Figures 5.2 to 5.4). Travel time distributions by shopping type are shown in Figures E2.1 to E2.3 in Appendix E.

Figure 5.2 shows the impact of using zone-based travel times by comparing the observed travel times to the shopping locations recorded in the CHASE dataset with the travel times to the shopping locations generated using zone-to-zone travel times. Figure 5.2 shows that zone-based travel times have smaller frequencies for travel times between 0-10 min, and significantly higher travel times for 10-15 min travel times. This anomaly can be explained by intrazonal trips being

recorded as 0 min and the threshold travel time for interzonal shopping trips being around 10 min. This accounts for the significantly higher travel time observed in the CHASE dataset.
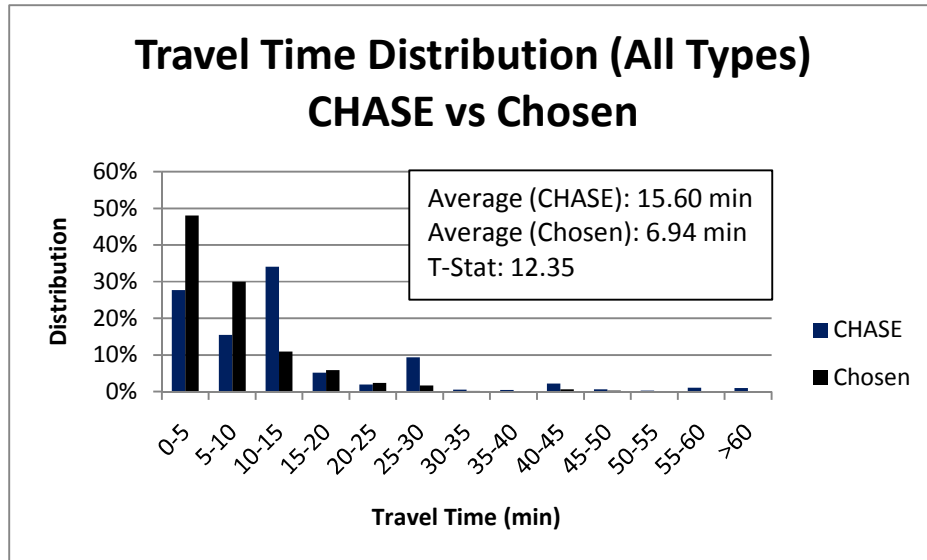
Figure 5.2 Auto Travel Time Distribution to Shopping: CHASE vs Chosen

Figures 5.3 and 5.4 compare the auto travel times between the predicted and chosen locations. Figure 5.3 only considers the travel time to the shopping location, while Figure 5.4 considers the total travel time to and from the shopping location.
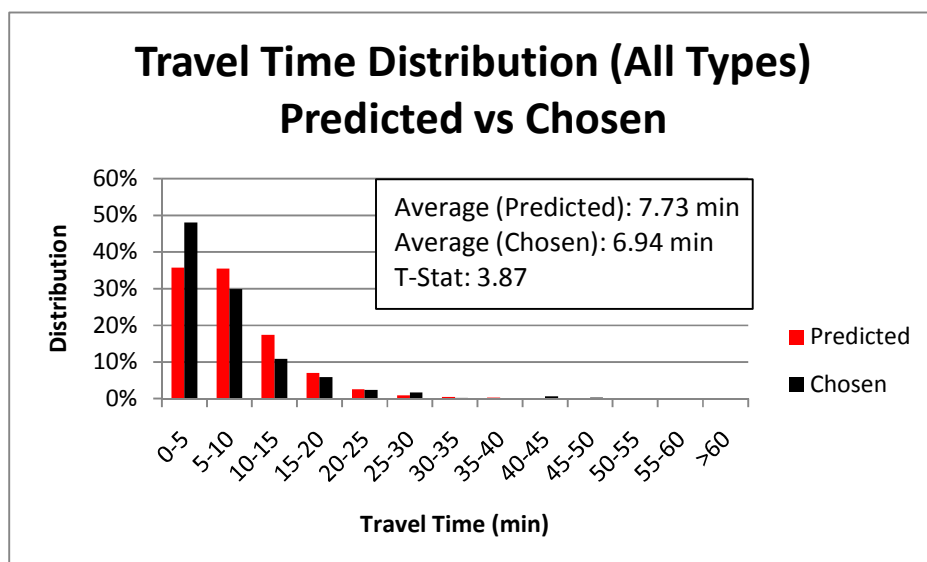
Figure 5.3 Auto Travel Time Distribution to Shopping: Predicted vs Chosen

As Figure 5.3 shows, the model under-predicts short distance trips (0-5 min) and slightly over predicts medium distance trips (5-15 min). As Figure 5.4 shows, this same trend is observed for the travel time distributions that include trip chaining effects (i.e. under-predicting in the 0-10 min range, and over-predicting in the 15-30 min range). Despite some biases from using zone-based travel times, the similar travel time distributions show that the model predicts travel times at an aggregate level reasonably well.
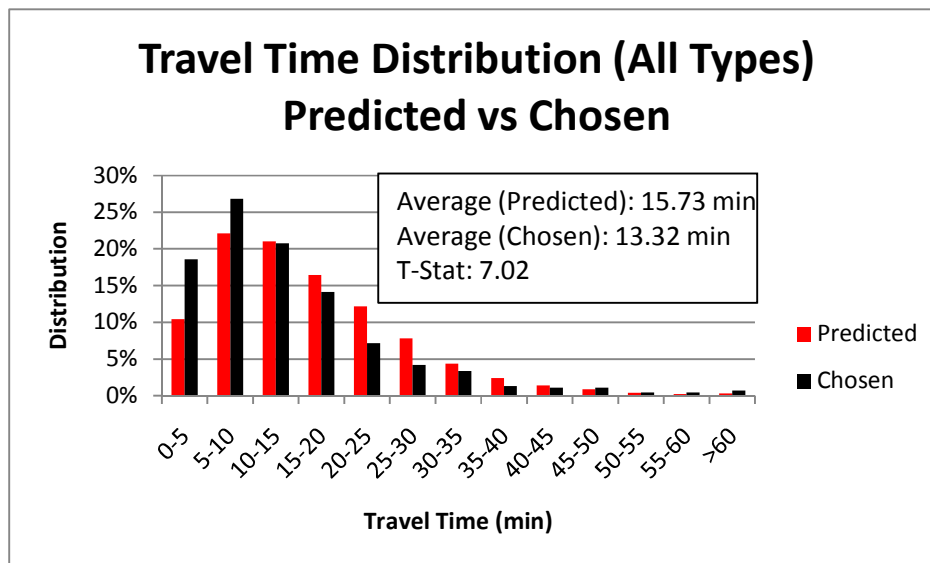
Figure 5.4 Auto Time Distribution to/from Shopping: Predicted vs Chosen

6   Future Directions

The gap-location choice model presented in this thesis is a novel approach to modelling shopping destination choice in an activity-based modelling framework. This model has the potential to add to the behavioural realism in activity schedule formation models. As mentioned throughout this thesis, there are many assumptions and limitations in the gap-location choice model presented in this thesis. The areas of future work can be categorized as follows:

1. Improving choice set formation

2. Moving towards dynamic scheduling

3. Integrating with TASHA

6.1   Improvements to Choice Set Formation

Choice set formation remains unaddressed in this thesis. Rather, the choice sets were constrained using time-space prisms and the availability of shopping opportunities. While this improves the realism of the choice set by reducing the choice set size, the process of choice set formation is an important area of future research.

First, gap choice set formation is based on simple rules used to generate a provisional schedule. Improvements to the gap choice set formation would require a model to generate the provisional schedule. This requires an understanding of activity priority and how this translates to precedence in activity scheduling. Generic store hours by shopping type would help constrain the gap choice set.

Second, location choice set formation is based on deterministic time-space and availability constraints. Full awareness was assumed in this model. To improve upon this, stated preference data could be collected to gauge an individual's awareness of shopping opportunities for different shopping types.

Finally, the characterization of shopping alternatives remains a challenge. How should locations be modelled? At the zone level? At the store level? At the shopping mall or power retail centre level? What are gaps? What constitutes "peak period", "Morning", "Late", and other time periods?

6.2   Integration with TASHA

The gap-location choice model developed in this thesis was designed to be integrated with the C# version of TASHA. TASHA should be modified to test just the location choice model and also the gap-location choice model. To do this, there needs to be a provisional schedule available to TASHA prior to the location choice decision.

Currently, TASHA schedules inserts shopping activities into the schedule with a predetermined location and start time, using simple heuristics for conflict resolution. Rather, TASHA should account for an individual's provisional schedule prior to assigning a location and start time. This would allow TASHA to account for the dynamics of the schedule formation process and provide a better opportunity to test the theory of choosing from a discrete set of time windows.

Finally, TASHA uses the TTS data, which has significantly more observations than the CHASE dataset. Using a larger number of observations would improve the model validity and provide stronger model estimates. While TTS only has one "Shopping" type, additional research can examine ways of disaggregating the shopping type or using the "All Types" model developed in this thesis.

6.3   Use of Time-Space Prisms

In the process of integrating with TASHA, improvements can be made with the C# code developed for this thesis. One improvement to highlight is how time-space prisms are generated and stored. In particular, time-space prisms are generated each time the software is run. While storing zone-to-zone travel times in a cache file help with program efficiency, these prisms should also be stored as well.

Furthermore, it may not be necessary to generate time-space prisms for every gap. As Figure 4.1 in Section 4.4 showed, time-space prisms only helped limit choice sets for certain gaps. Gaps that are too large do not need time-space prism constraints. Also, some shopping activities do not need time-space prisms and may take higher priority than previously scheduled activities, which have higher precedence. For example, if finding the finding the right set of tools (i.e. Shopping) takes longer than expected, and happens to be more important (higher priority) than a previously scheduled (higher precedence) School activity (e.g. finishing this thesis), the School activity would be shortened to accommodate the unexpected change in shopping duration.

The time-space prisms in this thesis assumed that the temporal boundaries were fixed by previously scheduled activities and the prism shape defined by auto travel times. Flexibility in temporal vertices can be modelled using stochastic frontier models, as in the operational activity-based model, FAMOS (Pendyala et al., 2005). Further, auto times used to define the time-space prisms can be defined by mode; however, this would require integration with the mode choice model. Integrating location choice with TASHA's tour-based mode choice model is a very nontrivial problem and is left as a problem for the future generation.

6.4   Shopping Activity Type

The shopping type classifications here were based on the CHASE dataset. However, this classification has many shortfalls since it is based on the user's entry into the travel and activity diary. For example, an individual going to the Shoppers' Drug Mart may be making a "Convenience" trip or "Minor Grocery" trip. While the classification improves the model's explanatory power, there remains a lot of unexplained heterogeneity in shopping behaviour.

Activity attributes can be used to help better classify shopping trips. For example, activity duration and planning horizon can be considered to define types (e.g. Convenience – Short, Spontaneous – Long, Spontaneous – Short, Routine – Short, etc.).

Relationships between activity attributes can also be examined. For example, Figures A1.1 to A1.7 show travel times with respect to shopping episode durations for different shopping types. This relationship can be used as a basis to classify activities based on how far individuals are willing to travel for different activity durations.

Using trip attributes for shopping activity type classification would be extremely valuable for disaggregating the Shopping trip type used in the TTS. The challenge for future research would be to develop meaningful classifications that can be validated with detailed descriptions found in activity diary data (i.e. CHASE).

7   Conclusions and Contributions

A prism- and gap-based approach was developed to model shopping location choice decisions. In the current implementation of TASHA, location choice is treated as an independent process, chosen exogenously to the scheduling process. To improve the behavioural realism of schedule process formation, location choice needs to be considered in the context of scheduling decisions. The gap-location choice model adds to the behavioural realism TASHA's existing location choice in three key ways:

First, *destination choices are modelled at an individual level*. Location choice is a fundamental decision which influences the entire scheduling process. Subsequent decisions, such as mode and route choices, depend on the choice of shopping location. To ensure consistency with the agent-based modelling approach used in TASHA and other activity-based models, it is necessary to also model location choices at the individual level.

Second, *shopping trips are disaggregated into 5 shopping types*. Shopping location choice behaviour varies by the type of shopping trip. The preliminary shopping types used in this model emphasize the importance of categorizing shopping trip behaviour to capture some of the heterogeneity in shopping location choice.

Finally, *a nested logit model is developed to model gap and location choices*. To our knowledge, predicting location choice based on gap choice is a novel approach to modelling shopping location choice. This prism- and gap-based approach is a step forward towards a more comprehensive model of shopping location choice behaviour that would integrate with a dynamic scheduling model.

The gap-location choice model developed in this thesis is a simple yet robust model developed through a few small changes that add to the behavioural realism of shopping location choice. The software application has been designed using the same architecture as the current C# version of TASHA to facilitate seamless integration into the existing model. The relatively simple and intuitive changes proposed in this thesis will improve TASHA's behavioural modelling capability by addressing one of its fundamental weaknesses: modelling shopping destination choice behaviour.

References

ALOGIT (2007). *ALOGIT 4.2 User Manual*. ALOGIT Software & Analysis Ltd.

Ansah, J.A. (1977). Destination choice set definition and travel behaviour modelling. *Transportation Research*, **11**, 127-140.

Arentze, T. and H. Timmermans (2000). *ALBATROSS: A Learning-Based Transportation Oriented Simulation System*. EIRASS, Eindhoven.

Auld, J.A. (2002). *ADAPTS: Agent-Based Dynamic Activity Planning and Travel Scheduling – Model Development and Data Collection*. PhD Thesis, University of Illinois, Chicago.

Auld, J.A. and A. Mohammadian (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. (1998). Can we ever obtain the data we would like to have? In: *Theoretical Foundations of Travel Choice Modelling* (T. Garling, T. Laitila and K. Westin, eds), pp. 305-323. Pergamon, Elsevier Science, Oxford, UK.

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

Ben-Akiva, M.E. and B. Boccara (1995). Discrete Choice Models with Latent Choice Sets. *International Journal of Research in Marketing*, **12**, 9-24.

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

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

Bhat, C.R. and K.T. Lawton (2000). *Passenger Travel Demand Forecasting*. Transportation Research Board, National Research Council (USA). Millennium Paper.

Bolduc, D., B. Fortin, M-A. Fournier (1996). The Effect of Incentive Policies on Practice Location of Doctors: A Multinomial Probit Analysis. *Journal of Labor Economics*, **14**(4), 703-732.

Bowman, J.L., M.E. Ben-Akiva (1996). Activity-Based Travel Forecasting. Proceedings of the *Activity-Based Travel Forecasting Conference*. Travel Model Improvement Program, US Department of Transportation, June.

Bowman, J.L., M. Bradley, Y. Shiftan, T.K. Lawton, and M.E. Ben-Akiva (1998). Demonstration of an activity based model system for Portland. In: *8ᵗʰ World Conference on Transport Research*, Antwerp, July.

Buliung, R.N., T. Hernandez, and J. Mitchell (2007). Exploring the relationship between big-box retailing and consumer travel demand in the Greater Toronto Area. *Canadian Journal of Transportation*. **1**, 122-41.

Burnett, K.P. and S. Hanson (1979). Rationale for an alternative mathematical approach to movement as complex human behavior. *Transportation Research Record: Journal of the Transportation Research Board*. **723**, 11-24.

Burnett, K.P. and S. Hanson (1982). The analysis of travel as an example of complex human behaviour in spatially constrained situations: definitions and measurement issues. *Transportation Research A*, **16**, 87-102.

Data Management Group (2001). *Transportation Tomorrow Suvey: Design and Conduct of the Survey*. Joint Program in Transportation, University of Toronto, Toronto.

DMTI Spatial (2003). *Enhanced Point of Interest Version 2.2 User Manual*. DMTI Spatial Inc., Markham, Ontario.

Doherty, S.T and E.J. Miller (2000). A computerized household activity scheduling survey. *Transportation*. **21**(1), 75-97.

Doherty, S.T., E. Nemeth, M.J. Roorda, and E.J. Miller (2004). Design and assessment of 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*, Master's Thesis, University of Toronto, Toronto.

Ettema, D.F. and H. Timmermans (1997). *Activity-based approaches to travel analysis*. Pergamon/Elsevier, Amsterdam.

Fotheringham, A.S. (1983). A new set of spatial interaction models: the theory of competing destinations. *Environment and Planning A*, **15**, 15-36.

Habib, K.M.N. (2011). A random utility maximization (RUM) based dynamic activity scheduling model: Application in weekend activity scheduling, *Transportation*. **38**(1), 123-151.

Habib, K.M.N. and E.J. Miller (2006). Modeling workers' skeleton travel-activity schedules. *Transportation Research Record: Journal of the Transportation Research Board*. **1985**, 88-97.

Habib, K.M.N. and E.J. Miller (2008). Modelling daily activity program generation considering within day and day-to-day dynamics in activity travel behavior. *Transportation*, **35**, 467-484.

Habib, K.M.N. and E.J. Miller (2009). Modelling activity generation: a utility based model of activity-agenda formation. *Transportmetrica*, **5**(1), 3-23.

Hägerstrand, T. (1970). What about people in regional science? *Papers of the Regional Science Association*, **24**, 7-21.

Hao, J.Y., M. Hatzopoulou, M., and E.J. Miller (2010). Integrating an Activity-Based Travel Demand Model with Dynamic Traffic Assignment and Emissions Models: An Implementation in the Greater Toronto Area. *Transportation Research Record: Journal of the Transportation Research Board*. **2176**, 1-13.

Kitamura, R., C. Chen and R. Narayanan (1998). The effects of time of day, activity duration and home location on travelers' destination choice behaviour. An Implementation in the

Greater Toronto Area. *Transportation Research Record: Journal of the Transportation Research Board*. **1645**, 76-81.

Kitamura, R. and S. Fujii (1998). Two computational process models of activity-travel choice. In: *Theoretical Foundations of Travel Choice Modeling* (T. Garling, T. Latila and K. Westin, eds.), pp. 251-279, Pergamon, Oxford.

Kitamura, R., T. Yamamoto, K. Kishizawa and R.M. Pendyala (2000). Stochastic frontier models of prism vertices. *Transportation Research Record: Journal of the Transportation Research Board*. **1718**, 18-26.

Manski, C.F. (1997). The Structure of Random Utility Models. *Theory and Decision*, **8**, 229-254.

Meyer, R. (1980). Theory of destination choice-set formation under informational constraints. *Transportation Research Record: Journal of the Transportation Research Board*. **750**, 6-12.

Meyer, M. and E.J. Miller (2001). *Urban Transportation Planning: A Decision Oriented Approach. Second Edition*, McGraw Hill, New York.

Miller, E.J. (2005a). Project-based activity scheduling for person agents. Paper presented at the *16th International Symposium on Transportation and Traffic Theory*, College Park, Maryland, July.

Miller, E.J. (2005b). Propositions for modelling household decision making. In: *Behavioural Foundations of Integrated Land-use and Transportation Models: Assumptions and New Conceptual Frameworks* (M. Lee-Gosselin, and S.T. Doherty, eds.), Elsevier, New York.

Miller, E.J. and M.J. Roorda (2003). A prototype model of 24-h household activity scheduling for the Toronto Area. *Transportation Research Record: Journal of the Transportation Research Board*. **1831**, 114-121.

Pagliara, F. and H.J.P. Timmermans (2009). Choice Set Generation in Spatial Contexts: A Review. *Transportation Letters. International Journal of Transportation Research*, **1**(3), 181-196.

Pas, E. (1985). State of the Art and Research Opportunities in Travel Demand: Another Perspective. *Transportation Research Part A*, **19**, 460-464.

Pendyala, R.M., R. Kitamura, A. Kikuchi, T. Yamamoto, and S. Fujii (2005). FAMOS: The Florida activity mobility simulator. *Transportation Research Record: Journal of the Transportation Research Board*. **1921**, 123-130.

Pendyala, R.M., T. Yamamoto and R. Kitamura (2002). On the formation of time-space prisms to model constraints on personal activity-travel engagement. *Transportation*, **29**(1), 73-94.

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

Roorda, M.J., E.J. Miller, and N. Kruchten (2006). Incorporating within-household interactions into a mode choice model using a genetic algorithm for parameter estimation. *Transportation Research Record: Journal of the Transportation Research Board*. **1895**, 171-179.

Salvini, P.A. and E.J. Miller (2005). ILUTE: an operational prototype of a comprehensive microsimulation model of urban systems. *Network and Spatial Economics*. **5**, 217-234.

Sener, I.N., R.M. Pendyala and C.R. Bhat (2009). Accommodating Spatial Correlation Across Choice Alternatives in Discrete Choice Models: Application to Modelling Residental Location Choice Behavior. Paper presented at the 88th Annual Meeting of the Transportation Research Board (DVD).

Simon, H.A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, **69**(1), 99-118.

Swait, J.D. and M.E. Ben-Akiva (1987). Incorporating random constraints in discrete models of choice set generation. *Transportation Research B*, **21**, 91-102.

Sivakumar, A. and C.R. Bhat (2003). Comprehensive, Unified Framework for Analyzing Spatial Location Choice. *Transportation Research Record: Journal of the Transportation Research Board*. **2003**, 103-111.

Thill, J-C. (1992). Choice set formation for destination choice modeling. *Progress in Human Geography*, **16**, 361-382.

Timmermans, H.J.P. (2005). *Progress in activity-based analysis*. Oxford: Elsevier.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, **79**(4), 281-299.

Yamamoto, T., R. Kitamura and R.M. Pendyala (2004). Comparative analysis of time-space prism vertices for out-of-home activity engagement on working days and non-working days. *Environment and Planning B: Planning and Design*, **31**(2), 235-250.

Appendices

Appendix A: Activity Type Analysis

Figure A1.1 Distributions: *Travel Time vs Duration* **(Generic Type = 40)**

Convenience Store (Generic Type = 40)
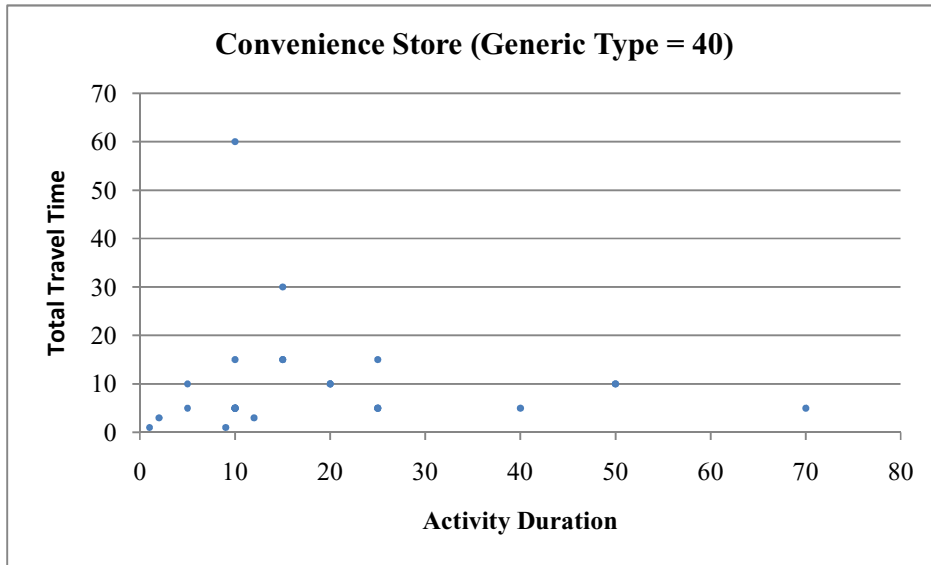
Figure A1.2 Distributions: *Travel Time vs Duration* **(Generic Type = 41)**

Minor Groceries, <10 items (Generic Type = 41)
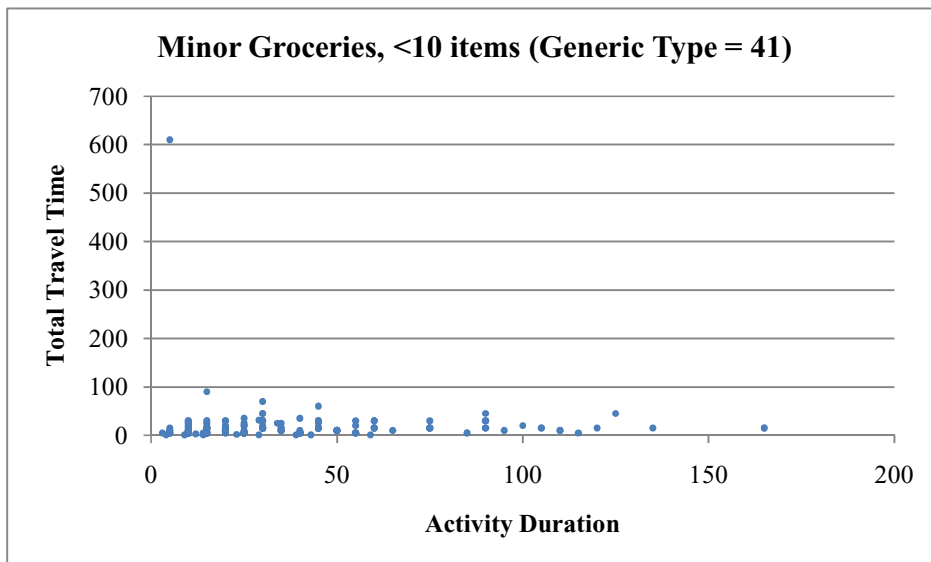
Figure A1.3 Distributions: *Travel Time vs Duration* **(Generic Type = 42)**
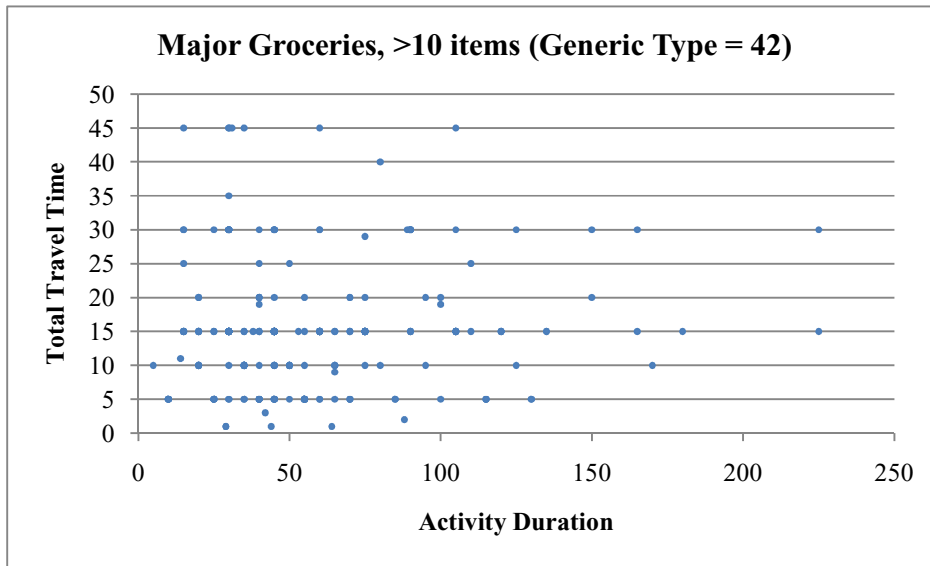
Figure A1.4 Distributions: *Travel Time vs Duration* **(Generic Type = 43)**
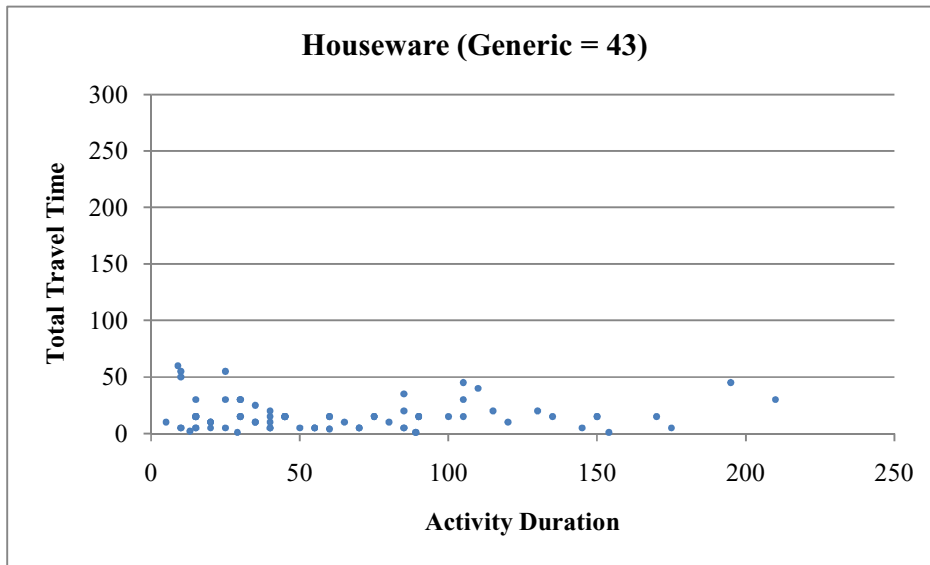
Figure A1.5 Distributions: *Travel Time vs Duration* **(Generic Type = 44)**

Clothing / Personal (Generic = 44)
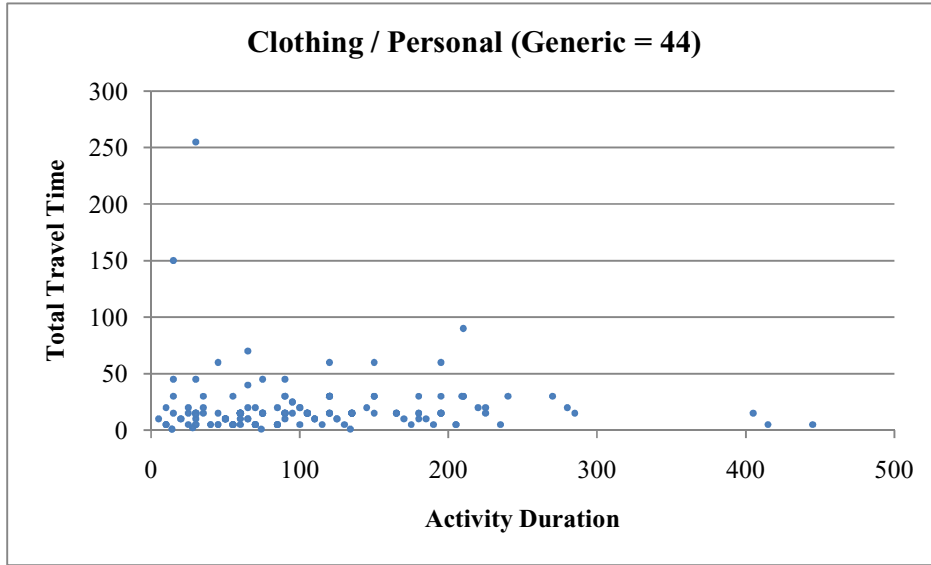
Figure A1.6 Distributions: *Travel Time vs Duration* **(Generic Type = 45)**

Drug Store (Generic = 45)
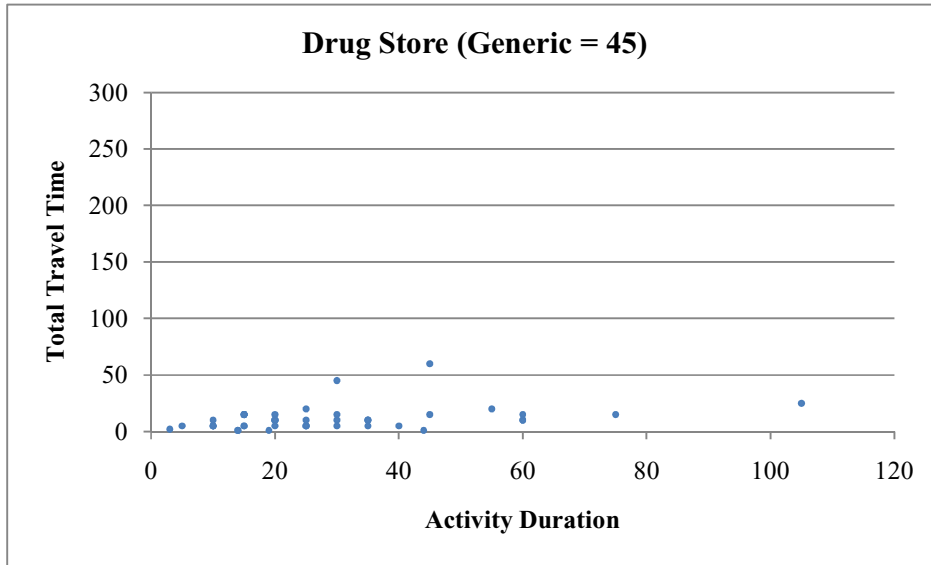
Figure A1.7 Distributions: *Travel Time vs Duration* **(Generic Type = 47)**
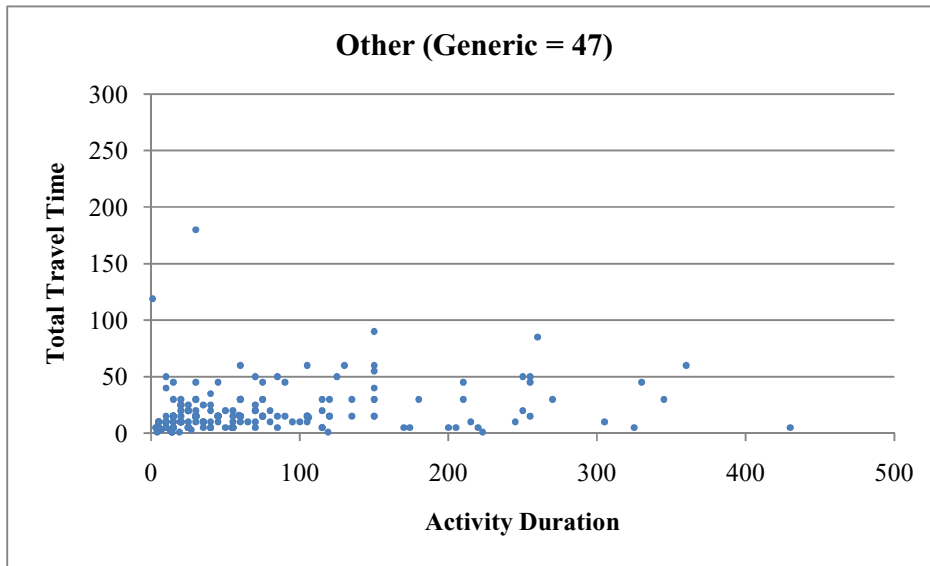
Table A1 Travel Time and Activity Duration Summary

Shopping Type	Total Travel Time					Activity Duration				
	Count	Average	Variance	Min	Max	Count	Average	Variance	Min	Max
Convenience Store	40	11.9	132.5	1	60	40	19.8	272.6	1	75
Minor Groceries	234	16.9	1647.9	1	610	240	35.3	789.4	3	165
Major Groceries	284	15.0	82.3	1	45	295	58.3	1332.5	5	225
Houseware	89	15.9	165.4	1	60	91	61.6	2191.3	5	210
Clothing / Personal	190	18.5	555.3	1	255	207	100.7	5232.1	5	445
Drug Store	46	10.8	115.3	1	60	49	27.5	390.8	3	105
Internet Shopping	0	#DIV/0!	#DIV/0!	0	0	2	100.0	2450.0	65	135
Other	262	17.4	341.1	1	180	273	65.2	5286.7	1	430
Undefined	0	#DIV/0!	#DIV/0!	0	0	2	107.5	312.5	95	120

Table A2 Detailed Mode Share for Shopping Trips by Type

Mode		Shopping Type / Generic_Type	Convenience Store (40)	Minor Groceries (41)	Major Groceries (42)	Houseware (43)	Clothing / Personal (44)	Drug Store (45)	Other (47)	Total	Group
Mode		Generic_Type	40	41	42	43	44	45	47	Total	
Car	1	10	13	131	189	51	101	24	152	661	
Van	1	11	2	24	46	18	33	7	31	161	
Truck	1	12		1	7		1	1	1	11	910
<name>'s	1	14			2	1		1	6	10	
SUV	1	15	1	15	15	9	12	2	11	65	
Other auto	1	29			1		1			2	
GO Train	3	31					1			1	
TTC	3	35	1	13	2		9		8	33	39
Mississau	3	39			1		3			4	
Other pub	3	49			1					1	
Walk	9	91	22	41	17	10	27	11	44	172	
Taxi	9	93		1	1		2		2	6	196
Bike	9	94	1	6	2				4	13	
Other wal	9	99		2					3	5	
		Total	40	234	284	89	190	46	262	1145	

Figure A2 Mode Share Summary for Shopping Trips by Type
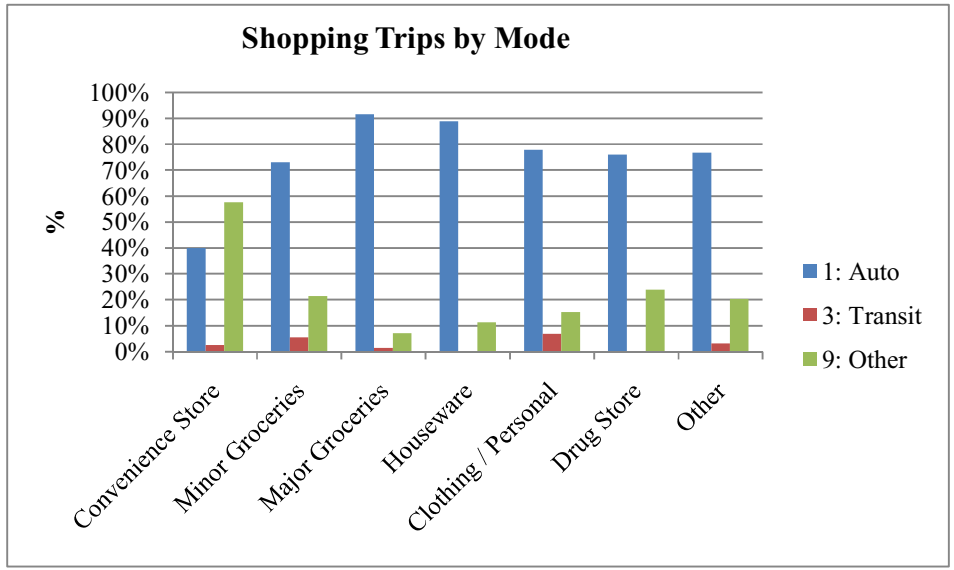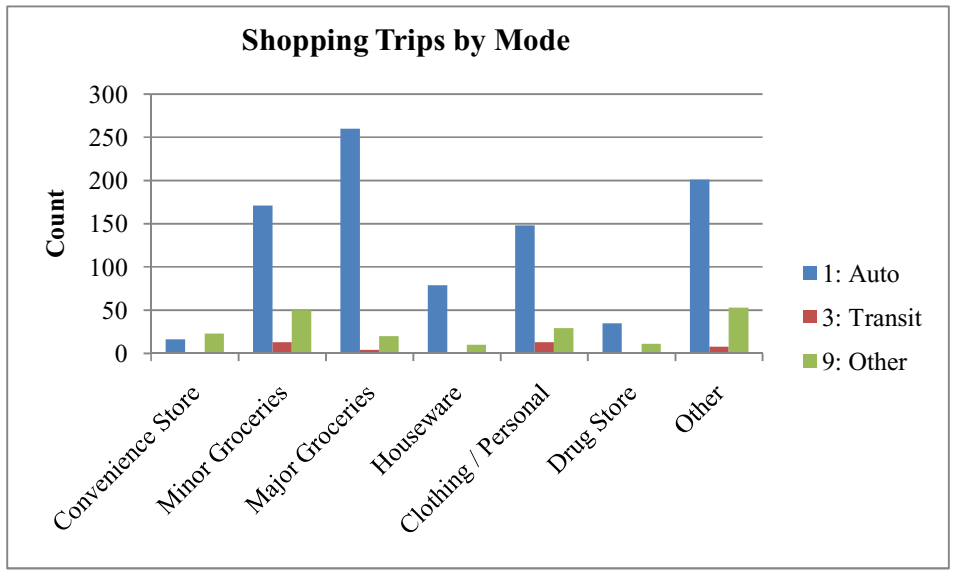
Figure A3 Activity Planning by Shopping Type
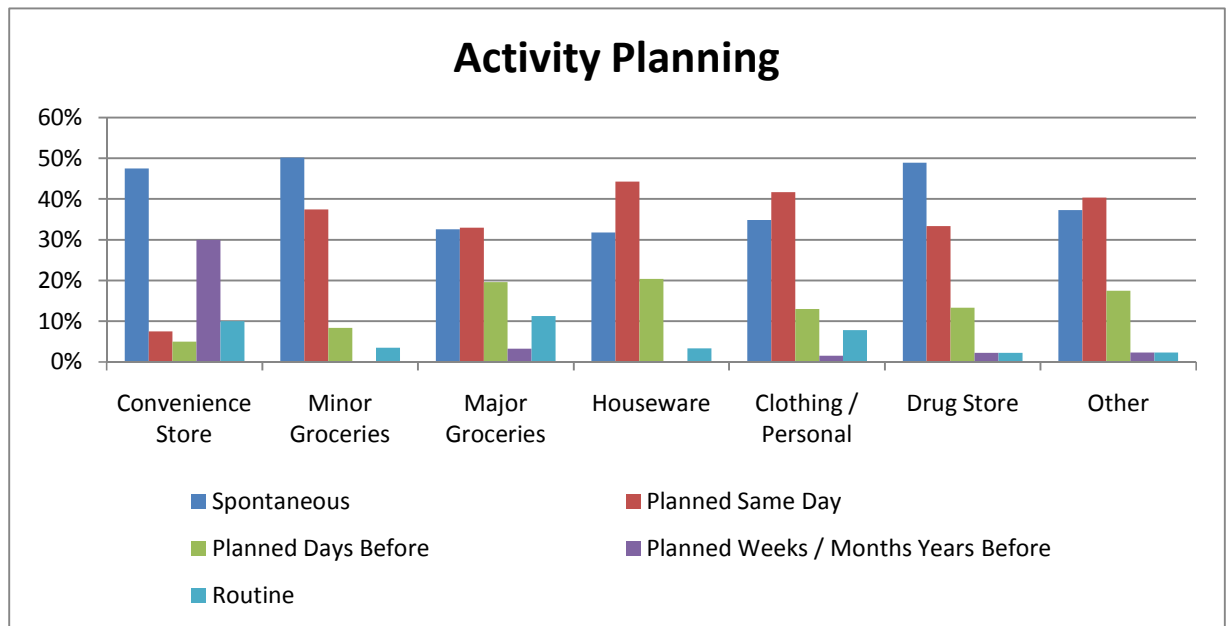
Table A3 Kolmogorov-Smirnov Tests: *Travel Time Distributions*

D_TT	40	41	42	43	44	45	47
40		0.2150	0.3373	0.2843	0.3489	0.1065	0.2534
41			0.1258	0.0692	0.1339	0.2107	0.0692
42				0.0633	0.0319	0.3365	0.0877
43					0.0687	0.2799	0.0347
44						0.3446	0.0955
45							0.2491
47							

D_alpha	40	41	42	43	44	45	47
40		0.2088	0.2060	0.2322	0.2124	0.2638	0.2071
41			0.1078	0.1520	0.1196	0.1968	0.1099
42				0.1482	0.1147	0.1939	0.1045
43					0.1570	0.2215	0.1497
44						0.2007	0.1166
45							0.1950
47							

D_TT : D_alpha	40	41	42	43	44	45	47
40		1.0297	1.6371	1.2242	1.6424	0.4038	1.2236
41			1.1666	0.4552	1.1195	1.0704	0.6299
42				0.4271	0.2781	1.7354	0.8392
43					0.4377	1.2634	0.2318
44						1.7171	0.8190
45							1.2772
47							

Table A4 Kolmogorov-Smirnov Tests: *Travel Distance Distributions (<1km)*

D_0-1	40	41	42	43	44	45	47
40		0.3368	0.5329	0.4449	0.3829	0.3299	0.3833
41			0.2388	0.2054	0.1629	0.2708	0.0917
42				0.1667	0.1943	0.3526	0.1878
43					0.1851	0.2738	0.2012
44						0.2159	0.1137
45							0.2250
47							

D_alpha	40	41	42	43	44	45	47
40		0.2384	0.2368	0.2863	0.2556	0.2964	0.2427
41			0.1410	0.2139	0.1708	0.2273	0.1508
42				0.2121	0.1685	0.2256	0.1481
43					0.2330	0.2771	0.2187
44						0.2453	0.1768
45							0.2318
47							

D_0-1 : D_alpha	40	41	42	43	44	45	47
40		1.4126	2.2508	1.5541	1.4978	1.1130	1.5791
41			1.6938	0.9600	0.9537	1.1912	0.6081
42				0.7860	1.1533	1.5631	1.2678
43					0.7945	0.9881	0.9199
44						0.8801	0.6432
45							0.9705
47							

Table A5 Kolmogorov-Smirnov Tests: *Shopping Mode*

Distance

n	D_TT	1	3	9
40	1		0.1015	0.3946
233	3			0.4027
284	9			

		684	222	239
	D_alpha	40	41	42
684	40		0.0942	0.0917
222	41			0.1137
239	42			

D_TT : D_alpha	40	41	42
40		1.0771	4.3045
41			3.5412
42			

Travel Time

n	D_TT	1	3	9
40	1		0.0454	0.1204
233	3			0.5676
284	9			

		684	222	239
	D_alpha	40	41	42
684	40		0.0942	0.0917
222	41			0.1558
239	42			

D_TT : D_alpha	40	41	42
40		0.4818	1.3134
41			3.6431
42			

Speed

n	D_TT	1	3	9
40	1		0.0833	0.3842
233	3			0.3805
284	9			

		684	222	239
	D_alpha	40	41	42
684	40		0.0942	0.0917
222	41			0.1137
239	42			

D_TT : D_alpha	40	41	42
40		0.8839	4.1911
41			3.3460
42			

Table A6 Activity Type Mapping

Shopping Activity Type	CHASE Generic Activity Types	SIC Codes (See DMTI manual for description of SIC codes[1])
1	40: Convenience Store 45: Drug Store	5311, 5331, 5912, 5921
2	41: Minor Groceries 42: Major Groceries	5311, 5411, 5421, 5431, 5441, 5451, 5461, 5499, 5921
3	44: Clothing / Personal	5311, 5611, 5621, 5631, 5641, 5651, 5661, 5699, 5731, 5734, 5735, 5736
4	43: Houseware	5211, 5231, 5251, 5261, 5271, 5311, 5531, 5712, 5713, 5714, 5719, 5722, 5731, 5734
5	47: Other	5311, 5399, 5932, 5941, 5942, 5943, 5944, 5945, 5946, 5947, 5948, 5949, 5961, 5962, 5992, 5993, 5994, 5995, 5999

---

[1] DMTI Spatial (2003). *Enhanced Point of Interest Version 2.2 User Manual*. DMTI Spatial Inc., Markham, Ontario.

Figure A4.1 Spatial Distribution of the Number of Retail Stores (All Types)
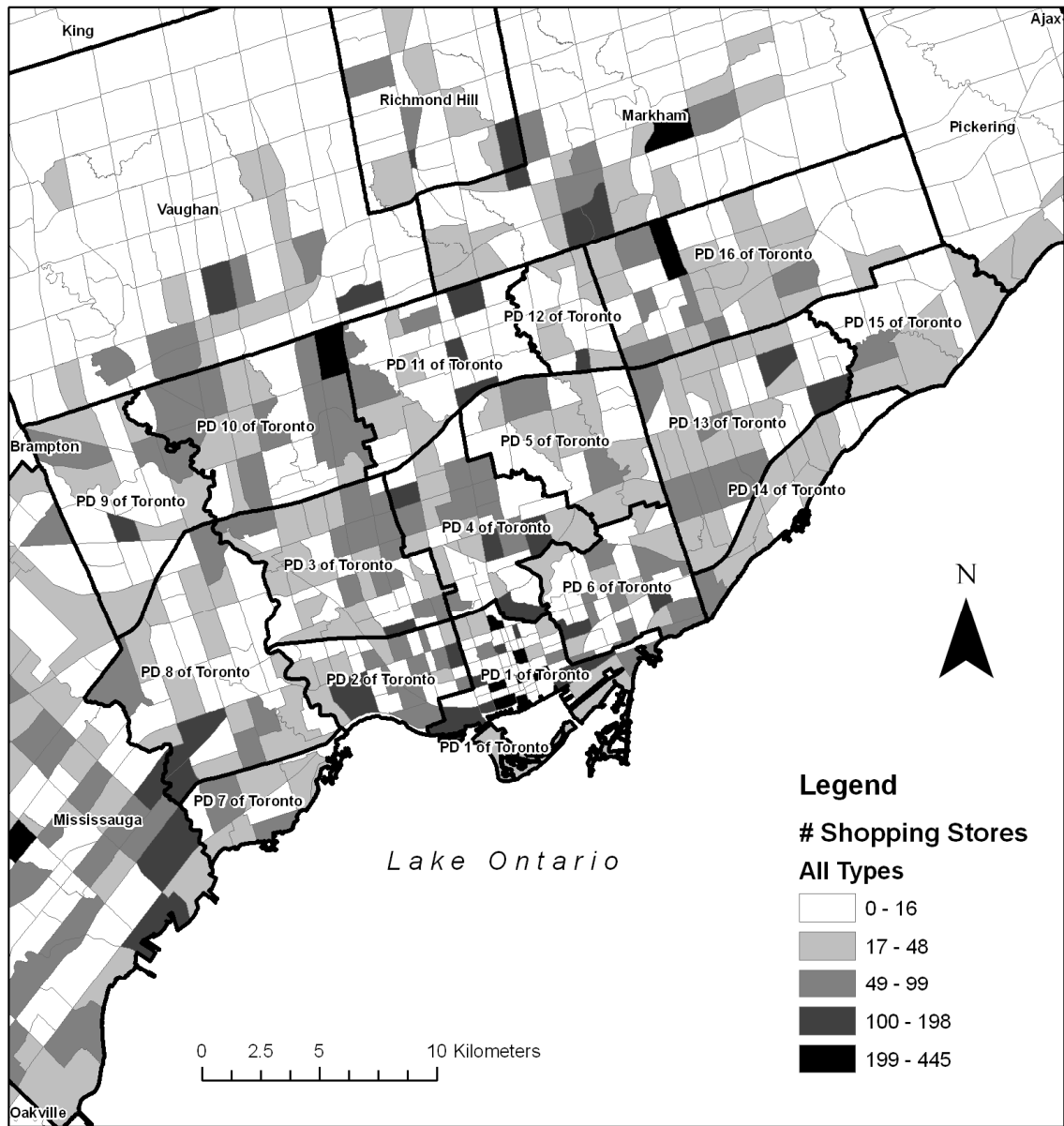
Figure A4.2 Spatial Distribution of the Number of Retail Stores (Type 1)
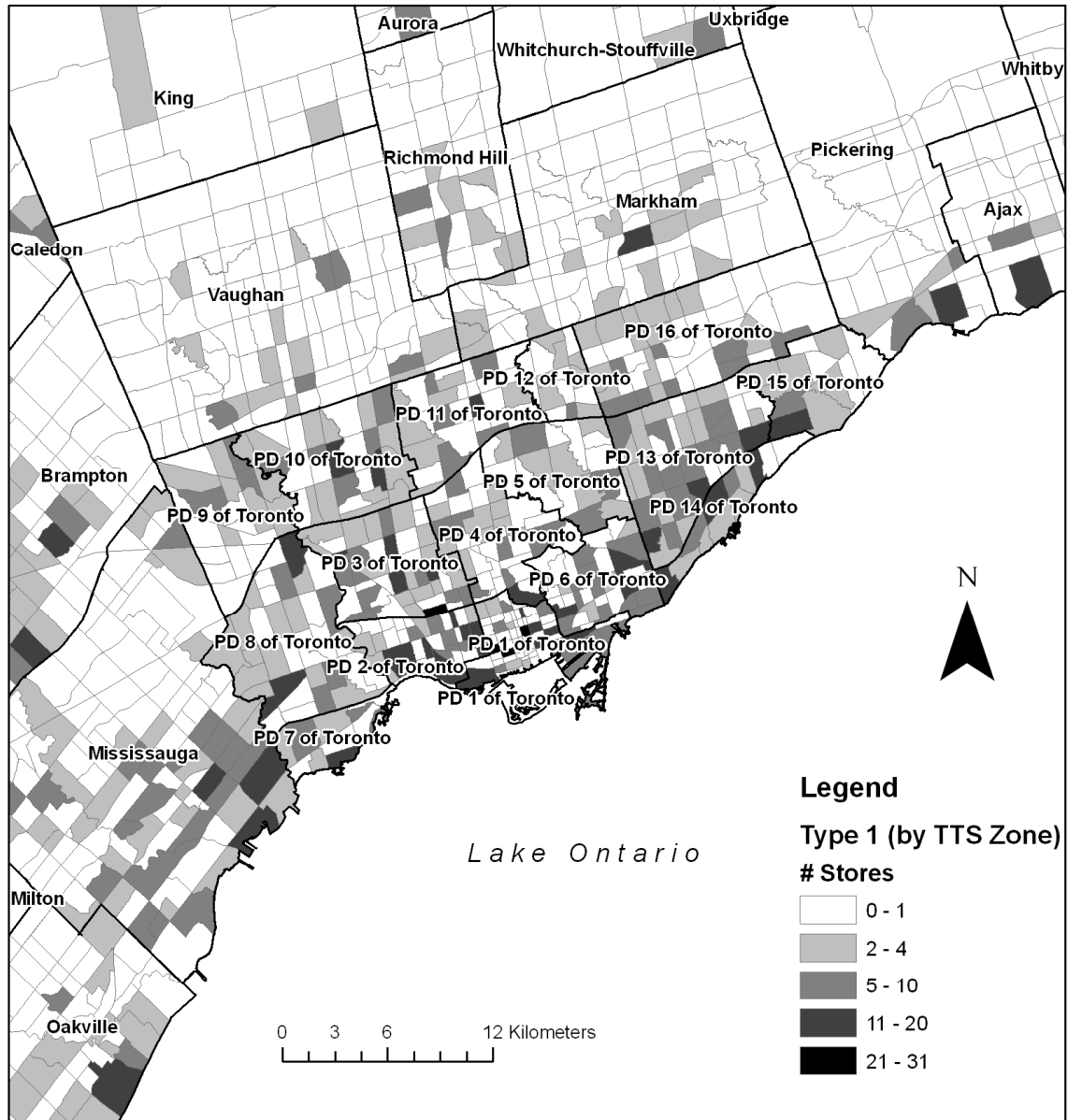
Figure A4.3 Spatial Distribution of the Number of Retail Stores (Type 2)
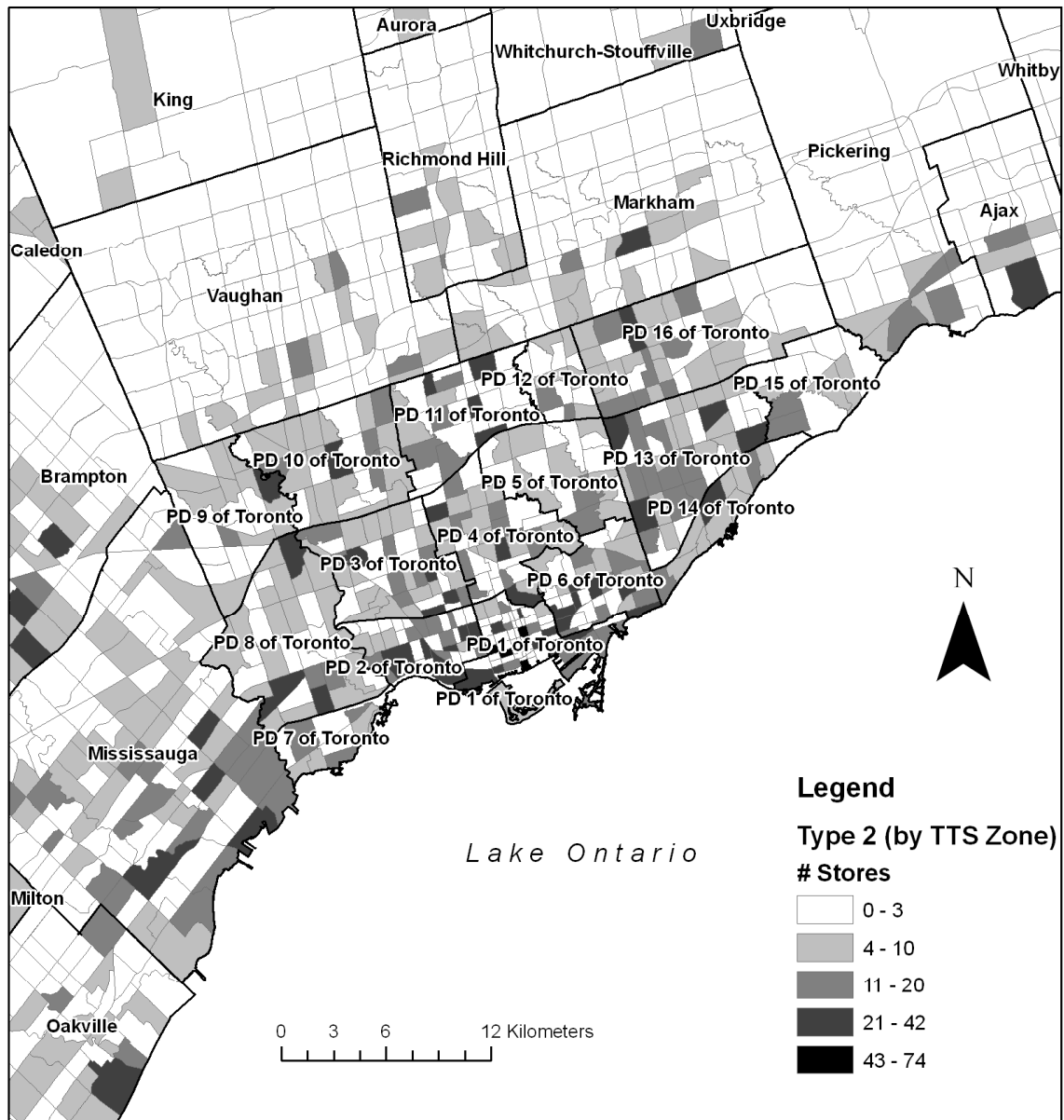
Figure A4.4 Spatial Distribution of the Number of Retail Stores (Type 3)
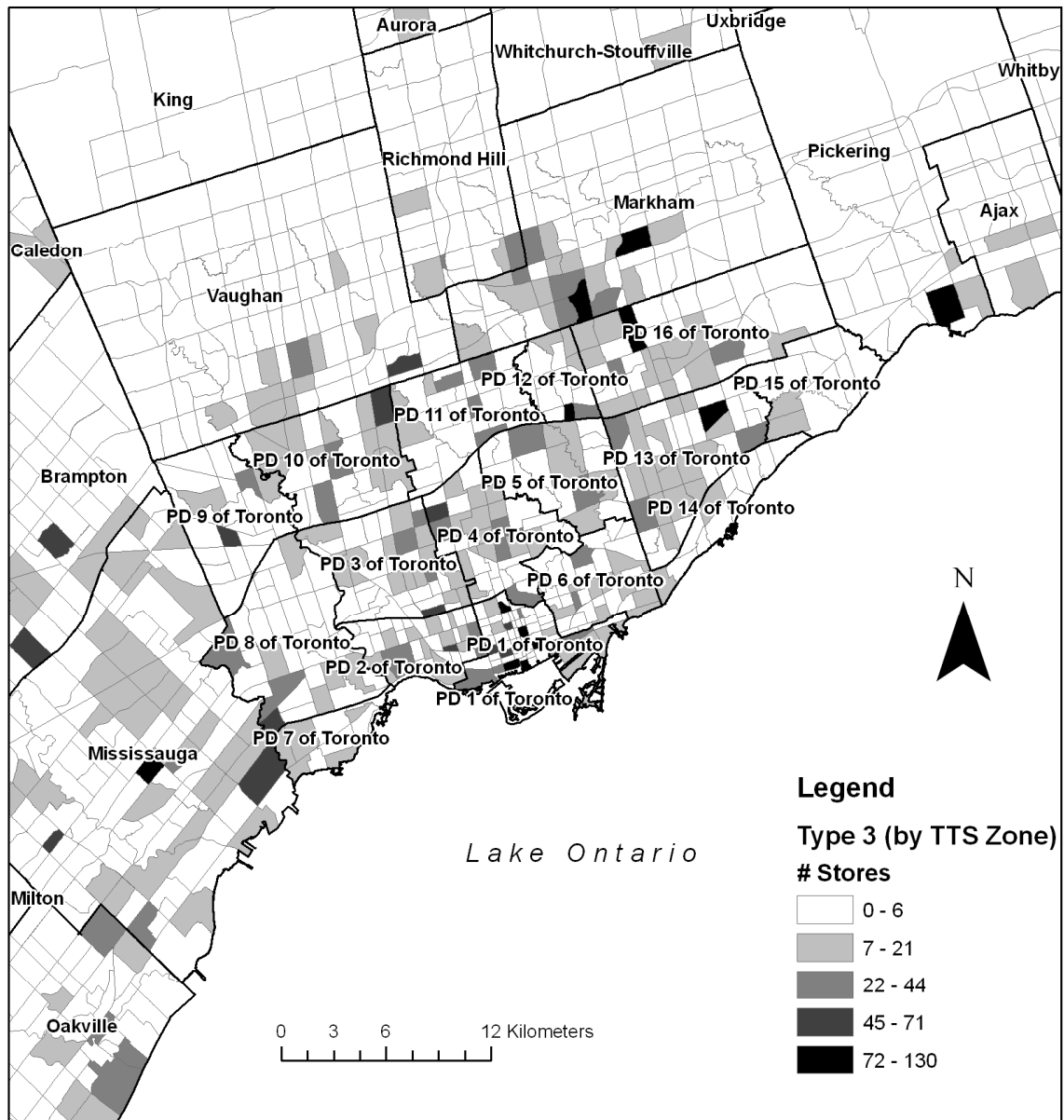
Figure A4.5 Spatial Distribution of the Number of Retail Stores (Type 4)
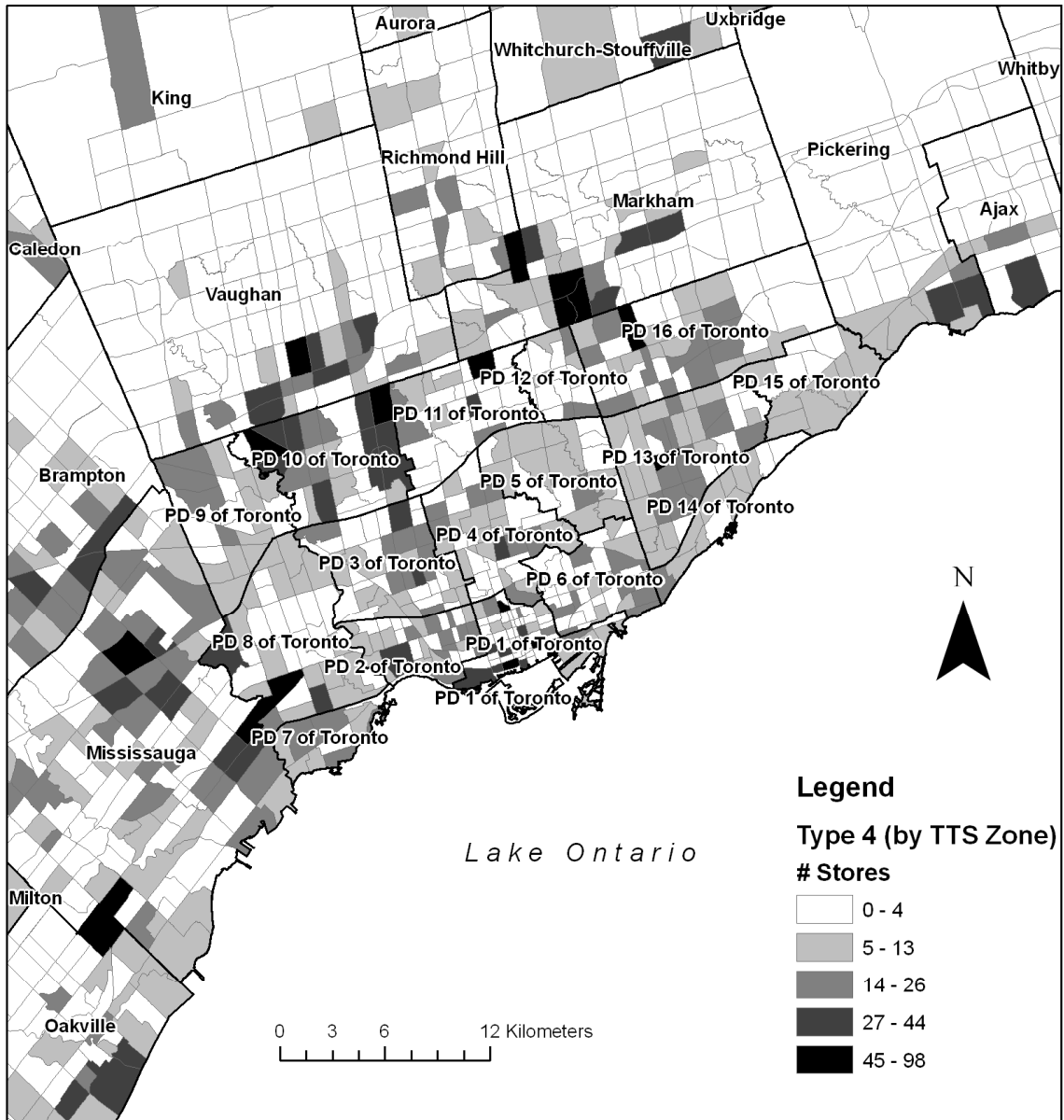
Figure A4.6 Spatial Distribution of the Number of Retail Stores (Type 5)

Legend

Type 5 (by TTS Zone)
Stores

- 0 - 7
- 8 - 24
- 25 - 51
- 52 - 102
- 103 - 244
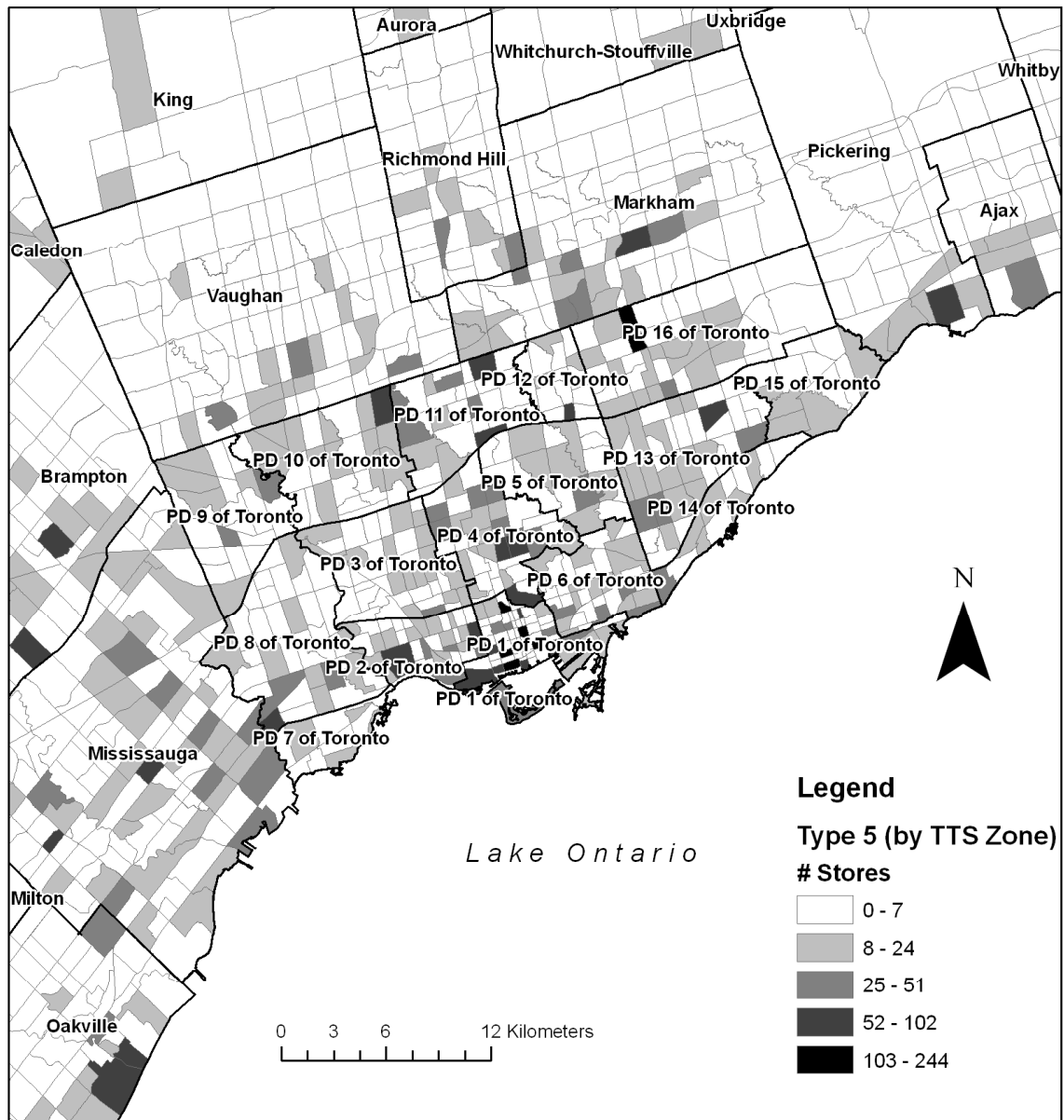
Appendix B: Input Files from CHASE

Figure B1 GTA Zone File ("Toronto2001Zone-GTA.csv")

Variable No.	Variable Name	Description (TASHA Variable)
1	TTSZone	ZoneNumber
2	PD	PlanningDistrict
3	Pop	Population
4	Work-G	WorkGeneral
5	Work-M	WorkManufacturing
6	Work-P	WorkProfessional
7	Work-R	WorkRetail
8	Work-Unk	WorkUnknown
9	Emp	Employment
10	Emp-G	EmploymentGeneral
11	Emp-M	EmploymentManufacturing
12	Emp-P	EmploymentProfessional
13	Emp-R	EmploymentRetail
14	Emp-Unk	EmploymentUnknown
15	X	X
16	Y	Y
17	Internal_Dist	InternalDistance
18	RetailActLevel	RetailActivityLevel
19	OtherActLevel	OtherActivityLevel
20	WorkActLevel	WorkActivityLevel
21	Internal_Dens	InternalDensity
22	ParkingCost	ParkingCost
23	ArtirialPercentage	ArtirialPercentage
New Variables		
24-29	TypeN_Dummy	A dummy variable indicating whether a TTS Zone has a store of Type N, N = {ALL, 1, 2, 3, 4, 5}, based on EPOI data
30-35	Count_TTS_N	Number of stores of Type N in the TTS Zone, N = {ALL, 1, 2, 3, 4, 5}, based on EPOI data
36-41	Count_PD_N	Number of stores of Type N in the Planning District (PD), N = {ALL, 1, 2, 3, 4, 5,}, based on EPOI data
42	TotalCommSqFt	Total Commercial Square Footage from CSCA

Figure B2 Person File ("CHASE_Persons.csv")

Variable No.	Variable Name	Description
1	PersonID	Unique Person identifier based on household and member IDs
2	HH_ID	Household ID
3	TTSZone	TTS 2001 Zone Number
4	PD	TTS 2001 Planning District
5	Region	TTS 2001 Region
6	H_Size	Household Size
7	H_kids	Number of kids in household
8	H_teens	Number of teens in household
9	H_Adlts	Number of adults in household
10	H_Autos	Number of vehicles of household
11	H_Dwel	Type of dwelling
12	H_OwnRnt	Household ownership type
13	H_Type	Household Type
14	H_ResDur	Years Living in Residence
15	H_CitDur	Years Living in City
16	H_RegDur	Years Living in Region
17	Status	Individual's status
18	I_Role	Role of individual in household
19	I_Gender	Gender (1: Male, 2: Female)
20	I_Type	Person Type
21	I_Autos	Number of vehicles of individual
22	I_Age	Age
23	I_Income	Annual Income ($)
24	I_License	Driver's Licence
25	I_EmpSta	Employment Status
26	I_StuSta	Student
27	I_Education	Education Level
28	I_TranPs	Transit pass ownership
29	I_Cell	Cell phone ownership

Figure B3 Schedule File ("CHASE_Schedule-OOH.csv")*

Variable No.	Variable Name	Description
1	PersonID	Unique Person identifier based on household and member IDs
2	HH_ID	Household ID
3	ScheduleID	Schedule ID
4	member_oi	Household member ID
5	CellCo	Type of coordination involved
6	Day_	Activity Day (0:First Day, …, 6:Last Day)
7	Weekday	Weekday (1:Sunday, …, 7:Saturday)
8	StartH	Start Time (hhmm, 24 hour)
9	EndH	End Time
10	Duration	Activity Duration
11	TTT	Total Travel Time
12	Activity	Activity Type
13	ActivityGeneric	Generic Activity Type
14	Shopping	Shopping Dummy (1: Episode is a Shopping Activity, 0: Otherwise)
15	ShoppingType	Shopping Type (1:Convenience, 2:Groceries, 3:Personal, 4:Houseware, 5:Other)
16	Mode	Mode used
17	Discussed	Activity involved other individuals
18	ChildrenTotal	Activity involved children
19	Location	Location
20	PD	TTS 2001 Planning District based on geocoded location
21	GTA01	TTS 2001 Zone based on geocoded location

*OOH = out-of-home, in-home activities excluded except for night-sleep / washing-up

Appendix C: Modelling Framework

Figure C1.1 Class Diagrams: *AutoData*

AutoData
-CacheTemplate : string -Data : OD Cache -ODC : string
+AutoData (template : string) +ConvertToHours (minutes : float) : float +Generate (sourceDirectory : string, targetFile : string) +GetCost (origin : IZone, destination : IZone, time : float) : float +GetTimePeriod (time : float) : TravelTimePeriod +GetTravelTime (origin : IZone, destination : IZone, time : float) : float +Load (filename : string) +Release () +Reload ()

Figure C1.2 Class Diagrams: *Episode*

Episode
+chosenGap : TimeWindow +GapChoiceSet : List<TimeWindow> +ChosenAlt : int = -999 +ChosenAlt_Gap : int = -999 +ChosenAlt_Zone : int = -999 +IV2 +IV2sim +shopCDF_MNL : double [1548] +gapCDF : double[41]
+Episode (personid : int, hh_id : int, scheduleid : int, member_oi : int, cellco : int, days : int, weekday : int, starth : int, endh : int, duration : int, ttt : int, activity : int, activitygeneric : int, shopping : int, shoppingtype : int, mode : int, discussed : int, childrentotal : int, location : int, pd : int, gta01 : int) : Episode

Figure C1.3 Class Diagrams: *Person*

Person
+EpisodeMap : SortedList<TashaTime, Episode> +ShoppingDummy : bool -ShoppingEpisodes : List<Episode> -GapList : List<TimeWindow>
+Person (personid : int, hh_id : int, ttszone : int, pd : int, region : int, h_size : int, h_kids : int, h_teens : int, h_adlts : int, h_autos : int, h_dwel : int, h_ownrnt : int, h_type : int, h_resdur : int, h_citdur : int, h_regdur : int, status : int, role : int, gender : int, i_type : int, i_autos : int, i_age : int, i_income : int, i_license : int, i_empsta : int, i_stusta : int, i_education : int, i_tranps : int, i_cell : int) : Person

Figure C1.4 Class Diagrams: *TashaTime*

TashaTime
-days : byte -hours : byte -minutes : byte -seconds : byte -weekday : byte -scheduleID : int
+TashaTime (time1 : float, time2 : float, scheduleID : int, days : int, weekday : int) : TashaTime +ToFloat (time : TashaTime) : float +ToFloat2 (time : TashaTime) : float +ToFloat3 (time : TashaTime) : float ++ (t1 : TashaTime, t2: TashaTime) : TashaTime +- (t1 : TashaTime, t2: TashaTime) : TashaTime +>= (t1 : TashaTime, t2: TashaTime) : TashaTime +> (t1 : TashaTime, t2: TashaTime) : TashaTime +<= (t1 : TashaTime, t2: TashaTime) : TashaTime +< (t1 : TashaTime, t2: TashaTime) : TashaTime +== (t1 : TashaTime, t2: TashaTime) : TashaTime +!= (t1 : TashaTime, t2: TashaTime) : TashaTime

Figure C1.5 Class Diagrams: *TimeWindow*

TimeWindow
+StartTime : TashaTime
+EndTime : TashaTime
+Duration : TashaTime
+DurationF : float
+StartLocation : int
+EndLocation : int
+chosengap : bool = false
+IV : double = 0
+morning : bool = false
+afternoon : bool = false
+evening : bool = false
+late : bool = false
+FeasibleZone : List<int>
+locCDF : double[1548]
+TimeWindow (start : TashaTime, end : TashaTime) : TimeWindow

Figure C1.6 Class Diagrams: *Zone*

Zone
-ZoneNumber : short -PlanningDistrict : int -Population : int -WorkGeneral : int -WorkManufacturing : int -WorkProfessional : int -WorkRetail : int -WorkUnknown : int -Employment : int -ManufacturingEmployment : int -ProfessionalEmployment : int -RetailEmployment : int -UnknownEmployment : int -GeneralEmployment : int -X : int -Y : int -InternalDistance : int -RetailActivityLevel : int -OtherActivityLevel : int -WorkActivityLevel : int -ParkingCost : int -IntrazonalDensity : int -ArterialRoadRatio : int -TypeALL_Dummy : short -Type1_Dummy : short -Type2_Dummy : short -Type3_Dummy : short -Type4_Dummy : short -Type5_Dummy : short -Count_TTS_ALL : short -Count_TTS_1 : short -Count_TTS_2 : short -Count_TTS_3 : short -Count_TTS_4 : short -Count_TTS_5 : short -Count_PD_ALL : short -Count_PD_1 : short -Count_PD_2 : short -Count_PD_3 : short -Count_PD_4 : short -Count_PD_5 : short -TotalCommSqFt : short
+GetDistance (x : IZone) : float +GetDistance (x : IZone) : float +GetStraightDistance (x : IZone) : float +Zone (ZoneID : int, Data : float[]) : Zone -ConvertToZone (ZoneID : int, data : float[]) : Zone

Figure C2 Program Flow Chart
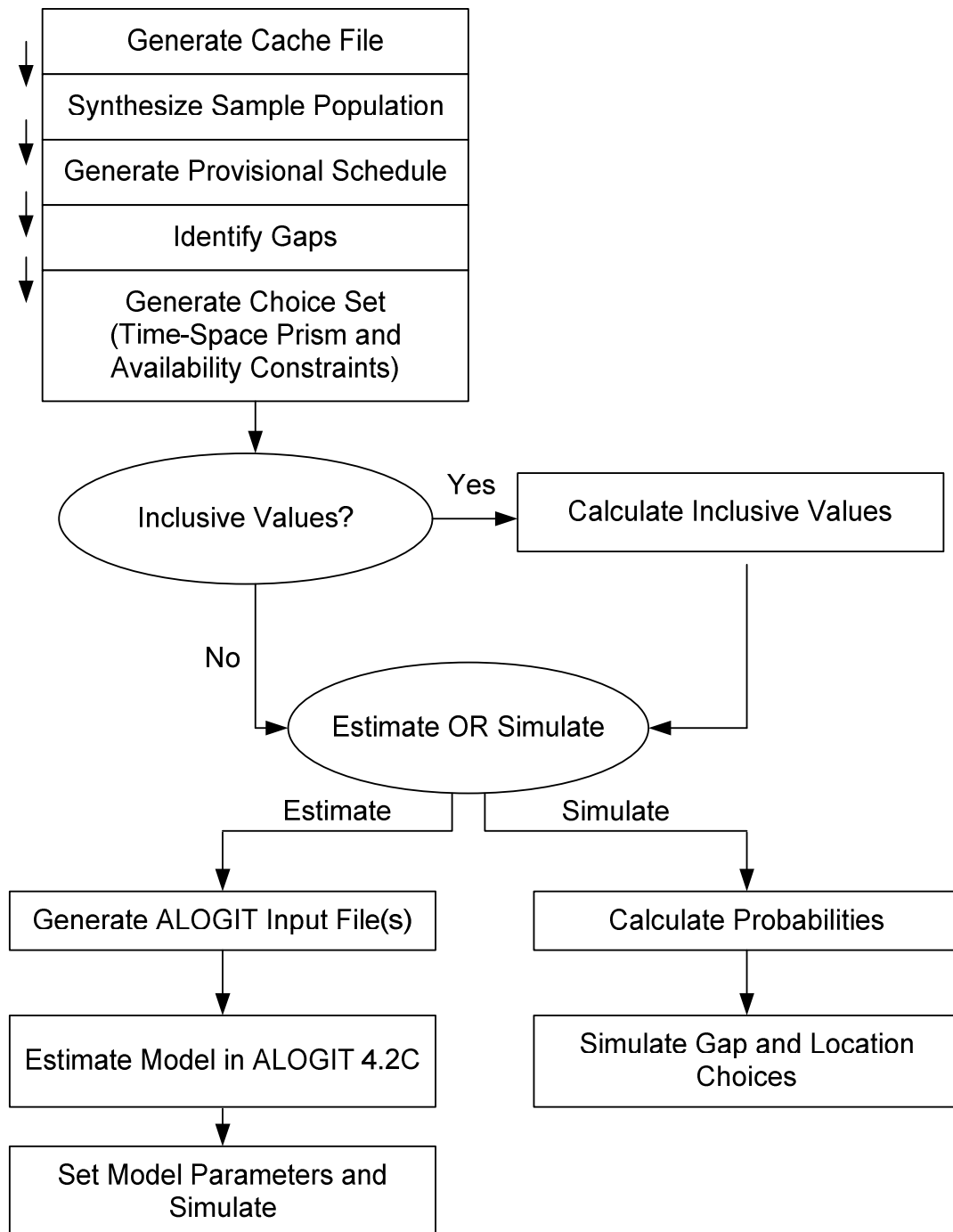
Figure C3.1 Pseudocode: *Generate Cache File*

```
// === Cache File Generator ===
// This method uses the cache file generator in TASHA

ZoneRetriever.Object = zoneRetriever
zoneRetriever.LoadCacheFile(inputfile = "Data/Toronto2001Zones-GTA.csv",
                            cachefile = "CacheData/Zones2001-GTA.zfc")
AutoData.Object = autodata
autodata.GenerateODdata(file = "CacheData/Auto.odc")
```

Figure C3.2 Pseudocode: *Synthesize Sample Population*

```
// === Generate Population ===

// Person Attributes
ReadInput.PersonFile(inputfile = "CHASE_Persons.csv")
NEW SortedList = people<int, Person>
FOREACH entry IN inputfile
      people.Add(PersonID, CSV_ConvertToPerson())
NEXT entry
```

Figure C3.3 Pseudocode: *Generate Provisional Schedule*

```
// === Generate Provisional Schedule ===

// Convert schedule file to list of Episode objects
ReadInput.ScheduleFile(input = "CHASE_Schedule-OOH.csv")
NEW List = all_episodes<Episode>
FOREACH entry IN inputfile
      all_episodes.Add(CSV_ConvertToEpisode())
NEXT entry

// Add Episode objects to individual provisional schedule (EpisodeMap)
NEW SortedList = EpisodeMap<TashaTime, Episode>
FOREACH e IN all_episodes
      IF (e == Shopping)
            people[e.PersonID].ShoppingDummy = true
            people[e.PersonID].ListOfShoppingEpisodes.Add(e)
      ELSE IF (e.Activity == 1 OR 2 OR 3 OR 4 OR 6)
            TashaTime.Object = start_time()
            people[e.PersonID].EpisodeMap.Add(start_time, e)
      END IF
NEXT e
```

Figure C3.4 Pseudocode: *Identify Gaps*

```
// === Identify Gaps ===

FOREACH p IN people
      // Define anchor points in weekly schedule
      TashaTime.Object = t_last(0,0,0,0,0)
      TashaTime.Object = t1(StartTimeOfFirstEpisodeOfEpisodeMap)
      int loc1 = StartLocationOfFirstEpisode
      int loc2
      FOREACH e in p.EpisodeMap
            IF (firstEpisode)
                  skip
            ELSE
                  TashaTime.Object = t2(EpisodeStartTime)
                  TashaTime.Object = t_last1(EpisodeEndTime)
                  t_last = t_last1
                  loc2 = e.GTA01
                  TimeWindow.Object = gap(t1, t2)
                  gap.StartLocation = loc1
                  gap.EndLocation = loc2
                  p.GapList.Add(gap)
                  t1 = t_last1
                  loc1 = loc2
                  loc2 = p.TTSZone
                  IF (secondLastEpisode)
                        TashaTime.Object = t_end(2359,2350,0,6,7)
                        TimeWindow.Object = lastGap(t_last,t_end)
                        p.GapList.Add(lastGap)
                  END IF
            END IF
      NEXT e
NEXT p
```

Figure C3.5 Pseudocode: *Generate Choice Set*

```
// === Generate Choice Set ===
// Gaps: GapChoiceSet, List of TimeWindow objects (gaps) in choice set
// Zones: FeasibleZones, List of zone numbers of zones in consideration set
// Consideration set = Awareness Set AND Feasible Set
// Awareness Set = all zones with shopping (of type n) available
// Feasible Set = all zones within time-space prism

int nZ = 0 // Tracks chosen zone FOR ALOGIT input file
int nG = 0 // Tracks chosen gap FOR ALOGIT input file
FOREACH p IN people
      FOREACH sh_ep in p.ShoppingEpisodes
            nG = 0
            nZ = zonelist.IndexOf(sh_ep.GTA01) + 1
            sh_ep.ChosenAlt_Zone = nZ
            TashaTime.Object = start_time(ShoppingEpisodeStartTime)
            FOREACH g in p.GapList
                  g.FeasibleZones.Clear()
                  float aTime1 = 0
                  float aTime2 = 0
                  IF (g.DurationF > sh_ep.Duration)
                        FOREACH z IN zonelist
                              atime1 = autodata.GetTravelTime(
                                    zoneRetriever.GetZone(g.StartLocation),
                                    zoneRetriever.GetZone(z),
                                    start_time.ToFloat3())
                              atime2 = autodata.GetTravelTime(
                                    zoneRetriever.GetZone(z),
                                    zoneRetriever.GetZone(g.EndLocation),
                                    start_time.ToFloat3())
                              // Generate Choice Set
                              // Time-Space Prism Constraint (Feasibility)
                              IF (aTime1+aTime2+sh_ep.Duration <= g.DurationF)
                                    // Shopping Type Constraint (Awareness)
                                    int zoneShopType = z.ShoppingType
                                    IF (zoneShopType == sh_ep.ShoppingType)
                                        g.FeasibleZones.Add(z)
                                        inGapChoiceSet = true
                                    END IF
                              ELSE IF (z ==  sh_ep.GTA01)
                                    // Include chosen zone in choice set
                                    g.FeasibleZones.Add(z)
                              END IF
                        NEXT z
                        IF (inGapChoiceSet)
                              sh_ep.GapChoiceSet.Add(g)
                              nG++
                              IF (start_time >= g.StartTime AND
                                          start_time <=g.EndTime)
                                    sh_ep.chosenGapDuration = g.DurationF
                                    sh_ep.chosenGap = g // TimeWindow object
                                    sh_ep.ChosenAlt_Gap = nG // Gap Index
                              END IF
                        END IF
                  END IF
            NEXT g
```

```
                     // Index of chosen location FOR simultaneous estimation
                IF (sh_ep.ChosenAlt_Gap > 0 AND sh_ep.ChosenAlt_Zone > 0)
                        sh_ep.ChosenAlt = (sh_ep.ChosenAlt_Gap - 1)*1548 +
                                          sh_ep.ChosenAlt_Zone
                ELSE
                        sh_ep.ChosenAlt = -999
                END IF
        NEXT sh_ep

NEXT p
```

Figure C3.6 Pseudocode: *Generate Inclusive Values*

```
// === Generate Inclusive Values - Lower Level: Locations ===
FOREACH p IN people
      FOREACH sh_ep IN p.ShoppingEpisodes
            FOREACH g in sh_ep.GapChoiceSet
                  FOREACH z in g.FeasibleZones
                        InitializeVariables()
                        InitializeParameters()
                        SetParameters()
                        float aCost1 = autdata.GetCost(
                                    zoneRetriever.GetZone(start_loc),
                                    zoneRetriever.GetZone(z),
                                    start_time.ToFloat3())
                        float aCost2 = autdata.GetCost(
                                    zoneRetriever.GetZone(z),
                                    zoneRetriever.GetZone(end_loc),
                                    start_time.ToFloat3())
                        float aTime1 = autdata.GetTravelTime(
                                    zoneRetriever.GetZone(start_loc),
                                    zoneRetriever.GetZone(z),
                                    start_time.ToFloat3())
                        float aTime2 = autdata.GetTravelTime(
                                    zoneRetriever.GetZone(z),
                                    zoneRetriever.GetZone(end_loc),
                                    start_time.ToFloat3())
                        acost = aCost1 + aCost2
                        atime = aTime1 + aTime2
                        // Set LnstoreT
                        IF (z.ShoppingTypeCount ==  0)
                            LnstoreT = 0
                        ELSE
                            LnstoreT = Math.Log(z.ShoppingType.Count)
                        END IF
                        // Set LCSCAarea
                        IF (z.TotalCommSqFt ==  0)
                            LCSCAarea = 0
                        ELSE
                            LCSCAarea = Math.Log(z.TotalCommSqFt)
                        END IF
                        // Calculate Utilities
                        tempUtility = p_acost * acost +
                                    p_atime * atime +
                                    p_LnstoreT * LnstoreT +
                                    p_LCSCAarea * LCSCAarea
                        sumUtility = tempUtility + sumUtility
```

```
                    NEXT z
                    IF (sumUtility == 0 OR doubleIsNaN(Math.Log(sumUtility))
                            g.IV = 0
                    ELSE
                            g.IV = Math.Log(sumUtility)
                    END IF
            NEXT g
        NEXT sh_ep
NEXT p

// === Generate Inclusive Values - Lower Level: Gaps (Inverted Model) ===
FOREACH p IN people
        FOREACH sh_ep IN p.ShoppingEpisodes
                sumUtil = 0
                FOREACH g in sh_ep.GapChoiceSet
                    InitializeVariables()
                    InitializeParameters()
                    EpGap_Numerator = sh_ep.Duration +
                                        autoTravelTimeToFromChosenZone()
                    EpGap_Denominator = g.DurationF
                    float EpGapRatio = EpGap_Numerator / EpGap_Denominator
                    SetDayTimeDummys()
                    SetParameters()
                    SetVariables()

                    expUtil = Math.Exp(p_EpGapTTGapRatio * EpGapTTGapRatio +
                                    p_DayTimeWKD_MOR * DayTimeWKD_MOR +
                                    p_DayTimeWKD_AFT * DayTimeWKD_AFT +
                                    p_DayTimeWKD_EVE * DayTimeWKD_EVE +
                                    p_DayTimeWKD_LAT * DayTimeWKD_LAT +
                                    p_DayTimeWDY_MOR * DayTimeWDY_MOR +
                                    p_DayTimeWDY_AFT * DayTimeWDY_AFT +
                                    p_DayTimeWDY_EVE * DayTimeWDY_EVE +
                                    p_DayTimeWDY_LAT * DayTimeWDY_LAT)
                    sumUtil = sumUtil + expUtil
                NEXT g
                sh_ep.IV2 = Math.Log(sumUtil)
        NEXT sh_ep
NEXT p
```

Figure C3.7 Pseudocode: *Generate ALOGIT Input Files*

```
// === Generate ALOGIT Input File - Location Choice ===
StreamWriter alogit_MNL (outputfile = "shopMNL.dat")
int nObs = 0
int nShEp
bool[] avail = new bool[1548]
FOREACH p IN people
      nShEp = 0
      FOREACH sh_ep IN p.ShoppingEpisodes
            TashaTime.Object = start_time(EpisodeStartTime)
            nObs++
            alogit_MNL.Write("{0},{1}{2},{3}", nObs, p.PersonID, nShEp,
                                        zonelist.IndexOf(sh_ep.GTA01)+1)
            nShEp++
            // === Set Availability ===
            IF (sh_ep.chosenGapDuration == 0)
                  FOR (int i = 0; i < 1548; i++)
                        avail[i] = false;
                        alogit_MNL.Write(",1")
                  LOOP

            ELSE
                  FOREACH z IN zonelist
                        IF (sh_ep.chosenGap.FeasibleZones.Contains(z))
                              alogit_MNL.Write(",1")
                        ELSE
                              alogit_MNL.Write(",0")
                        END IF
                  NEXT z
            END IF

            // === Episode and Individual Attributes ===
            // Shopping Type
            alogit_MNL.Write(",{0}", sh_ep.ShoppingType)
            // Household Size
            alogit_MNL.Write(",{0}", p.H_Size)
            // Gender
            alogit_MNL.Write(",{0}", p.I_Gender)
            // License
            alogit_MNL.Write(",{0}", p.I_License)
            // Number of Cars
            alogit_MNL.Write(",{0}", p.I_Autos)
            // log(Income)
            IF (p.I_Income > 0)
                  alogit_MNL.Write(",{0}", Math.Log(p.Value.I_Income))
            ELSE
                  alogit_MNL.Write(",0")
            END IF
            // Income Level
            IF (p.I_Income >= 60000)
                  alogit_MNL.Write(",3")
            ELSE IF (p.I_Income >= 30000)
                  alogit_MNL.Write(",2")
            ELSE
                  alogit_MNL.Write(",1")
            END IF
```

```
// === Zone Attributes ===
int start_loc
int end_loc;
IF (sh_ep.chosenGapDuration != 0)
      start_loc = sh_ep.chosenGap.StartLocation
      end_loc = sh_ep.chosenGap.EndLocation
ELSE
      start_loc = p.TTSZone
      end_loc = p.TTSZone
END IF
// aTime1
FOREACH z IN zonelist
      float aTime1 = autodata.GetTravelTime(
            zoneRetriever.GetZone(start_loc),
            zoneRetriever.GetZone(z), start_time.ToFloat3())
      alogit_MNL.Write(",{0}", aTime1)
NEXT z
// aTime1 + aTime2
FOREACH z IN zonelist
      float aTime1 = autodata.GetTravelTime(
            zoneRetriever.GetZone(start_loc),
            zoneRetriever.GetZone(z), start_time.ToFloat3())
      float aTime2 = autodata.GetTravelTime(
            zoneRetriever.GetZone(z),
            zoneRetriever.GetZone(end_loc),
            start_time.ToFloat3())
      alogit_MNL.Write(",{0}", aTime1 + aTime2);
NEXT z
// aTime1 + aTime2 - aTime3
FOREACH z IN zonelist
      float aTime1 = autodata.GetTravelTime(
            zoneRetriever.GetZone(start_loc),
            zoneRetriever.GetZone(z), start_time.ToFloat3())
      float aTime2 = autodata.GetTravelTime(
            zoneRetriever.GetZone(z),
            zoneRetriever.GetZone(end_loc),
            start_time.ToFloat3())
      float aTime3 = autodata.GetTravelTime(
            zoneRetriever.GetZone(start_loc),
            zoneRetriever.GetZone(end_loc),
            start_time.ToFloat3())
      float aT_Inc = aTime1 + aTime2 - aTime3;
      IF (aT_Inc < 0)
            alogit_MNL.Write(",0");
      ELSE
            alogit_MNL.Write(",{0}", aT_Inc);
      END IF
NEXT z
// aCost
FOREACH z IN zonelist
      float aCost1 = autodata.GetCost(
            zoneRetriever.GetZone(start_loc),
            zoneRetriever.GetZone(z), start_time.ToFloat3())
      float aCost2 = autodata.GetCost(
            zoneRetriever.GetZone(z),
            zoneRetriever.GetZone(end_loc),
            start_time.ToFloat3())
      alogit_MNL.Write(",{0}", aCost1 + aCost2);
```

```
NEXT z
// aCost/Income
FOREACH z IN zonelist
      float aCost1 = autodata.GetCost(
            zoneRetriever.GetZone(start_loc),
            zoneRetriever.GetZone(z), start_time.ToFloat3())
      float aCost2 = autodata.GetCost(
            zoneRetriever.GetZone(z),
            zoneRetriever.GetZone(end_loc),
            start_time.ToFloat3())
      IF (p.Value.I_Income > 0)
            alogit_MNL.Write(",{0}",
                              (aCost1 + aCost2) / p.Income)
      ELSE
            alogit_MNL.Write(",0")
      END IF
NEXT z
// Num Stores
FOREACH z IN zonelist
      float temp
      IF (zoneRetriever.GetZone(z).Count_TTS_ALL == 0)
            temp = (float)Math.E
      ELSE
            temp = zoneRetriever.GetZone(z).Count_TTS_ALL
      END IF
      alogit_MNL.Write(",{0}", Math.Log(temp))
NEXT z
// Num Stores by Type
FOREACH z IN zonelist
      float temp = 0;
      IF (shop_type == 1)
            temp = zoneRetriever.GetZone(z).Count_TTS_1
      ELSE IF (shop_type == 2)
            temp = zoneRetriever.GetZone(z).Count_TTS_2
      ELSE IF (shop_type == 3)
            temp = zoneRetriever.GetZone(z).Count_TTS_3
      ELSE IF (shop_type == 4)
            temp = zoneRetriever.GetZone(z).Count_TTS_4
      ELSE IF (shop_type == 5)
            temp = zoneRetriever.GetZone(z).Count_TTS_5
      END IF
      IF (temp == 0)
            alogit_MNL.Write(",0")
      ELSE
            alogit_MNL.Write(",{0}", Math.Log(temp))
      END IF
NEXT z
// CSCA - Retail Area
FOREACH z IN zonelist
      IF (zoneRetriever.GetZone(z).TotalCommSqFt == 0)
            alogit_MNL.Write(",0"); }
      ELSE
            alogit_MNL.Write(",{0}",
                  Math.Log(zoneRetriever.GetZone(z).TotalCommSqFt))
      END IF
NEXT z
// Population
FOREACH z IN zonelist
```

```
                    IF (zoneRetriever.GetZone(z).Population == 0)
                          alogit_MNL.Write(",0"); }
                    ELSE
                          alogit_MNL.Write(",{0}",
                                Math.Log(zoneRetriever.GetZone(z).Population))
                    END IF
              NEXT z
              // Employment
              FOREACH z IN zonelist
                    IF (zoneRetriever.GetZone(z).Employment == 0)
                          alogit_MNL.Write(",0")
                    ELSE
                          alogit_MNL.Write(",{0}",
                                Math.Log(zoneRetriever.GetZone(z).Employment)); }
              NEXT z

              // === Inclusive Values ===
              alogit_MNL.Write(",{0}", sh_ep.IV2)
              alogit_MNL.Write("\n")
              alogit_MNL.Flush()
        NEXT sh_ep
NEXT p

// === Generate ALOGIT Input File – Gap Choice ===
StreamWriter alogit_NLSeq (outputfile = "shopNLSeq.dat")
int nShEp = 0
int maxNumGaps = 41
int nGap = 0
FOREACH p IN people
        FOREACH sh_ep IN p.ShoppingEpisodes
              nShEp++
              alogit_NLSeq.Write("{0},{1},{2}", nShEp, p.Value.PersonID,
                                                sh_ep.ShoppingType);

              // == Chosen gap ==
              alogit_NLSeq.Write(",{0}", sh_ep.ChosenAlt_Gap);

              // === Availability ===
              nGap = 0
              FOREACH g IN sh_ep.GapChoiceSet
                    nGap++
                    alogit_NLSeq.Write(",1")
              NEXT g
              FOR (int i = nGap; i < maxNumGaps; i++)
                    alogit_NLSeq.Write(",0")
              LOOP

              // == Individual Attributes
              // Gender
              alogit_NLSeq.Write(",{0}", p.Value.I_Gender);
              // Income
              IF (p.Value.I_Income >= 60000)
                    alogit_NLSeq.Write(",3")
              ELSE IF (p.Value.I_Income >= 30000)
                    alogit_NLSeq.Write(",2")
              ELSE
                    alogit_NLSeq.Write(",1")
              END IF
```

```
// === (Episode+TravelTime):Gap Duration Ratio ===
nGap = 0
float attime = 0
float atime1 = 0
float atime2 = 0
TashaTime.Object = start_time(EpisodeStartTime)
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    atime1 = autodata.GetTravelTime(
        zoneRetriever.GetZone(g.StartLocation),
        zoneRetriever.GetZone(sh_ep.GTA01), start_time.ToFloat3())
    atime2 = autodata.GetTravelTime(
        zoneRetriever.GetZone(sh_ep.GTA01),
        zoneRetriever.GetZone(g.EndLocation),
        start_time.ToFloat3())
    attime = atime1 + atime2
    IF (sh_ep.Duration > 0)
         alogit_NLSeq.Write(",{0}",
            (sh_ep.Duration+attime) / g.DurationF )
    ELSE
        alogit_NLSeq.Write(",0")
    END IF
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP


// === Day of Week Dummies ===
// -- Sunday
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF (sh_ep.Weekday == 1) { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP
// -- Monday
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF (sh_ep.Weekday == 2) { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP
// -- Tuesday
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF (sh_ep.Weekday == 3) { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP
```

```
// -- Wednesday
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF (sh_ep.Weekday == 4) { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP
// -- Thursday
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF (sh_ep.Weekday == 5) { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP
// -- Friday
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF (sh_ep.Weekday == 6) { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP
// -- Saturday
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF (sh_ep.Weekday == 7) { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
{
    alogit_NLSeq.Write(",0")
LOOP
// -- Weekend Dummy
nGap = 0
FOREACH g IN sh_ep.GapChoiceSet
    nGap++
    IF ((sh_ep.Weekday == 1) || (sh_ep.Weekday == 7))
    { alogit_NLSeq.Write(",1") }
    ELSE { alogit_NLSeq.Write(",0") }
NEXT g
FOR (int i = nGap i < maxNumGaps i++)
    alogit_NLSeq.Write(",0")
LOOP

// === Time of Day Dummies ===
#region
FOREACH g IN sh_ep.GapChoiceSet
{
    float gapStart = g.StartTime.ToFloat() * 100
```

```
        float gapDuration = g.DurationF
        IF ((gapStart >= 600) && (gapStart < 1200))
        {
            g.morning = true
            // Duration criteria (g.DurationF)
            IF (gapDuration >= 360) { g.afternoon = true }
            IF (gapDuration >= 660) { g.evening = true }
            IF (gapDuration >= 900) { g.late = true }
        }
        ELSE IF ((gapStart >= 1200) && (gapStart < 1700))
        {
            g.afternoon = true
            // Duration criteria
            IF (gapDuration >= 300) { g.evening = true }
            IF (gapDuration >= 540) { g.late = true }
            IF (gapDuration >= 1080) { g.morning = true }
        }
        ELSE IF ((gapStart >= 1700) && (gapStart < 2100))
        {
            g.evening = true
            // Duration criteria
            IF (gapDuration >= 240) { g.late = true }
            IF (gapDuration >= 780) { g.morning = true }
            IF (gapDuration >= 1140) { g.afternoon = true }
        }
        ELSE IF ((gapStart >= 2100) || (gapStart < 600))
        {
            g.late = true
            // Duration criteria
            IF (gapDuration >= 540) { g.morning = true }
            IF (gapDuration >= 900) { g.afternoon = true }
            IF (gapDuration >= 1200) { g.evening = true }
        }
    NEXT g

    // -- Morning
    nGap = 0
    FOREACH g IN sh_ep.GapChoiceSet
        nGap++
        IF (g.morning) { alogit_NLSeq.Write(",1") }
        ELSE { alogit_NLSeq.Write(",0") }
    NEXT g
    FOR (int i = nGap i < maxNumGaps i++)
        alogit_NLSeq.Write(",0")
    LOOP
    // -- Afternoon
    nGap = 0
    FOREACH g IN sh_ep.GapChoiceSet
        nGap++
        IF (g.afternoon) { alogit_NLSeq.Write(",1") }
        ELSE { alogit_NLSeq.Write(",0") }
    NEXT g
    FOR (int i = nGap i < maxNumGaps i++)
        alogit_NLSeq.Write(",0")
    LOOP
    // -- Evening
    nGap = 0
    FOREACH g IN sh_ep.GapChoiceSet
```

```
        nGap++
        IF (g.evening) { alogit_NLSeq.Write(",1") }
        ELSE { alogit_NLSeq.Write(",0") }
    NEXT g
    FOR (int i = nGap i < maxNumGaps i++)
        alogit_NLSeq.Write(",0")
    LOOP
    // -- Late
    nGap = 0
    FOREACH g IN sh_ep.GapChoiceSet
        nGap++
        IF (g.late) { alogit_NLSeq.Write(",1") }
        ELSE { alogit_NLSeq.Write(",0") }
    NEXT g
    FOR (int i = nGap i < maxNumGaps i++)
        alogit_NLSeq.Write(",0")
    LOOP
    #endregion

    // Inclusive Value (FOR each gap)
    #region
    // -- This value should be calculated FOR each gap in the
    GapChoiceSet
    nGap = 0
    FOREACH g IN sh_ep.GapChoiceSet
        nGap++
        alogit_NLSeq.Write(",{0}",g.IV)
    NEXT g
    FOR (int i = nGap i < maxNumGaps i++)
        alogit_NLSeq.Write(",0")
    LOOP
    #endregion

    alogit_NLSeq.Write("\n")
    alogit_NLSeq.Flush()

    NEXT sh_ep
NEXT p
```

Figure C3.8 Pseudocode: *Simulate Gap and Location Choices*

```
// Gap Choice Model
double[] gapUtilities = new double[41]
double sumGapUtil = 0
double[] gapPDF = new double[41]
double[] gapUtilitiesall = new double[41]

// Location Choice Model
int[] choiceSet = new int[1548]
double[] tempUtilities = new double[1548]
double[] Utilities = new double[1548]
double sumUtilities = 0
double[] tempPDF = new double[1548]

// === Calculate probabilities and utilities ===
FOREACH p IN people
      FOREACH sh_ep IN p.ShoppingEpisodes
            // Gap Choice Model
            Initialize.GapVariables()
            Initialize.GapParameters()
            // -- Calculate Gap Utilities
            int i = 0
            FOREACH g IN sh_ep.GapChoiceSet
                  Initialize.GapVariables()
                  Initialize.GapParameters()
                  // Ratio includes travel time in numerator
                  Calculate.EpGapDuration()
                  Set.DayTimeDummys()
                  Set.VariablesParameters()
                  // Calculate Utility
                  expUtil = Math.Exp(p_EpGapTTGapRatio * EpGapTTGapRatio +
                                     p_DayTimeWKD_MOR * DayTimeWKD_MOR +
                                     p_DayTimeWKD_AFT * DayTimeWKD_AFT +
                                     p_DayTimeWKD_EVE * DayTimeWKD_EVE +
                                     p_DayTimeWKD_LAT * DayTimeWKD_LAT +
                                     p_DayTimeWDY_MOR * DayTimeWDY_MOR +
                                     p_DayTimeWDY_AFT * DayTimeWDY_AFT +
                                     p_DayTimeWDY_EVE * DayTimeWDY_EVE +
                                     p_DayTimeWDY_LAT * DayTimeWDY_LAT)
                  sumGapUtil = sumGapUtil + expUtil
                  gapUtilities[i] = expUtil
                  i++
            NEXT g
            // -- Calculate Gap Choice Probabilities
            // -- -- PDF
            FOR (int k = 0; k < 41; k++)
                  IF (sumGapUtil != 0)
                  { gapPDF[k] = gapUtilities[k] / gapUtilities.Sum()}
                  ELSE { gapPDF[k] = 0; }
            LOOP
            // -- -- CDF
            sh_ep.gapCDF[0] = gapPDF[0];
            sh_ep.gapCDFall[0] = gapPDFall[0];
            FOR (int k = 1; k < 41; k++)
                  sh_ep.gapCDF[k] = sh_ep.gapCDF[k - 1] + gapPDF[k];
            LOOP
```

```
                    // Location Choice Model (gap-dependent)
                    Set.LocationParameters()
                    IF (zonelist.Contains(sh_ep.GTA01)
                          FOREACH g IN sh_ep.GapChoiceSet
                                i = 0
                                Set.GapStartEndLocations()
                                TashaTime.Object start_time(g.StartTime)


                                // Calculate location utilities and probabilities
                                // -- Calculate Utilities
                                i = 0;
                                FOREACH z IN zonelist
                                        Calculate.AutoTravelTimes()
                                        Calculate.AutoTravelCosts()
                                        Calculate.LogShoppingStoreCountType()
                                        Calculate.LogRetailArea()
                                        // Calculate Utilities
                                        tempUtilities[i] = Math.Exp(p_acost * acost
                                                          + p_atime * atime
                                                          + p_LnstoreT * LnstoreT
                                                          + p_LCSCAarea * LCSCAarea)
                                        Utilities[i] = tempUtilities[i] * choiceSet[i]
                                        i++;
                                NEXT z
                                sumUtilities = tempUtilities.Sum();

                                // -- Calculate CDF
                                IF (sumUtilities <= 0)
                                        SumUtilErrorCount++;
                                ELSE IF (sumUtilities > 0)
                                        double CDF_prior = 0;
                                        FOR (i = 0; i < 1548; i++)
                                              tempPDF[i] = tempUtilities[i] /
                                                             sumUtilities;
                                              g.locCDF[i] = tempPDF[i] + CDF_prior;
                                              CDF_prior = g.locCDF[i];
                                        LOOP
                                END IF
                          NEXT g
                    END IF
              NEXT sh_ep
        NEXT p

// === Simulate Gap and Location Choices ===
int numIterations = 1 // Tested 1, 10, 100, 1000
Random.Object = randObj(seed = 1)
float rand_num
TimeWindow.Object = predGap
Initialize.DistributionArrays()


// Travel Time Distributions – CHASE, Chosen; Chosen Zone Distribution
FOREACH p IN people
      FOREACH sh_ep IN p.ShoppingEpisodes
            CHASE_traveltime.WriteLine("{0},{1}", sh_ep.ShoppingType,
                                      sh_ep.TTT)
            Chosen_zone.WriteLine("{0},{1}", sh_ep.ShoppingType,
                                      sh_ep.GTA01)
```

```
           IF (sh_ep.chosenGap != null)
                 TashaTime.Object = start_time(EpisodeStartTime)
                 float atime1 = autodata.GetTravelTime(
                     zoneRetriever.GetZone(sh_ep.chosenGap.StartLocation),
                     zoneRetriever.GetZone(sh_ep.GTA01),
                     start_time.ToFloat3())
                 float atime2 = autodata.GetTravelTime(
                     zoneRetriever.GetZone(sh_ep.GTA01),
                     zoneRetriever.GetZone(sh_ep.chosenGap.EndLocation),
                     start_time.ToFloat3())
                 Chosen_traveltime.WriteLine = ("{0},{1},{2}",
                                                sh_ep.ShoppingType,
                                                aTime1, aTime1 + aTime2)
           END IF
       NEXT sh_ep
NEXT p

// Simulate for n iterations
FOR (int iter = 0; iter < numIterations; iter++)
       int nobs = 0
       FOREACH p IN people
             FOREACH sh_ep IN p.ShoppingEpisodes
                   nobs++
                   IF (zonelist.Contains(sh_ep.GTA01)
                         // Predict Gap
                         rand_num = (float)randObj.NextDouble()
                         predGapIndex = -1
                         FOR (int i = 0; i < 41; i++)
                               IF (rand_num <= sh_ep.gapCDF[i]
                                     // Index for GapChoiceSet
                                     predGapIndex = i
                                     i = 41
                               END IF
                         LOOP
                         IF (predGapIndex = -1)
                               Error(NoGapPredicted)
                         END IF
                         predGap = sh_ep.GapChoiceSet[predGapIndex]
                         // Initialize Time Periods
                         bool mor = false
                         bool aft = false
                         bool eve = false
                         bool lat = false
                         Set.TimePeriodDummy(predGap) // Accounts for overlap
                         // Array for distribution of PredictedGap
                         // Time Period Distribution
                         IF (mor)
                               timeOfDayList[sh_ep.ShoppingType - 1, 0]++
                         END IF
                         IF (aft)
                               timeOfDayList[sh_ep.ShoppingType - 1, 1]++
                         END IF
                         IF (eve)
                               timeOfDayList[sh_ep.ShoppingType - 1, 2]++
                         END IF
                         IF (lat)
                               timeOfDayList[sh_ep.ShoppingType - 1, 3]++
                         END IF
```

```
// Weekday distribution
dayOfWeekList[sh_ep.ShoppingType - 1,
predGap.StartTime.Weekday - 1]++

// Predict Zone based on Predicted Gap
numTotal++
rand_num = (float)randObj.NextDouble()
predZone = -1
FOR (int i = 0; i < 1548; i++)
     IF (rand_num <= predGap.locCDF[i])
           predZone = zonelist[i]
           zoneOutPredicted.WriteLine("{0},{1}",
               sh_ep.ShoppingType, predZone)
           float atime1 = autodata.GetTravelTime(
               zoneRetriever.GetZone(
               predGap.StartLocation),
               zoneRetriever.GetZone(predZone),
               predGap.StartTime.ToFloat3())
           float atime2 = autodata.GetTravelTime(
               zoneRetriever.GetZone(predZone),
               zoneRetriever.GetZone(
               predGap.EndLocation),
               predGap.StartTime.ToFloat3()
               + (sh_ep.Duration/60))
           // Predicted Travel Time Distributions
           ttOutPredicted.WriteLine(("{0},{1},{2}",
               sh_ep.ShoppingType, atime1,
               atime1 + atime2)
           i = 1548
     END IF
LOOP
IF (predGap == sh_ep.chosenGap)
     numGapRight[sh_ep.ShoppingType]++
END IF
IF (predZone == -1)
     Error(NoZonePredicted)
ELSE
     // lineDist: distance from predicted zone
     //          to chosen zone
     IF (predZone == sh_ep.GTA01)
           numRight[sh_ep.ShoppingType]++
           lineDist = 0
     ELSE
           lineDist = zoneRetriever.GetZone(
               sh_ep.GTA01).GetStraightLineDistance(
               zoneRetriever.GetZone(predZone))
     END IF
     // Array for distance distributions
     IF (sh_ep.ShoppingType == 1)
           lineDistList1.Add(lineDist)
     ELSE IF (sh_ep.ShoppingType == 2)
           lineDistList2.Add(lineDist)
     ELSE IF (sh_ep.ShoppingType == 3)
           lineDistList3.Add(lineDist)
     ELSE IF (sh_ep.ShoppingType == 4)
           lineDistList4.Add(lineDist)
     ELSE IF (sh_ep.ShoppingType == 5)
           lineDistList5.Add(lineDist)
```

```
                          END IF
                     END IF
               ELSE
                     numError2++ // Chosen zone not in GTA region
               END IF
          NEXT sh_ep
     NEXT p
LOOP
Calculate.PercentRight_Gap()
Calculate.PercentRight_Location()
Calculate.AverageDistance()
Output.DistanceDistributions()
Output.TimeOfDayDistributions()
Output.DayOfWeekDistributions()
```

Appendix D: Gap-Location Choice Model

Figure D1 ALOGIT Code: Location Choice Model

```
------------------------------------------------------------
------------------------------------------------------------
- MNL Model of Location Choice to generate IV parameter FOR Upper Level
- Also used to estimate inverted model (Location-Gap Choice)
------------------------------------------------------------
$title Shopping Location Choice
$subtitle (@shopMNL.alo)
$print transFORms
$gen.stats utilities all
$estimate

-************** C O E F F I C I E N T S **********************
------ location coefficients
-- auto travel time by income and gender
$coeff taincrM1
+       taincrM2
+       taincrM3
+       taincrF1
+       taincrF2
+       taincrF3
-- auto travel cost by income and gender
+       tacostM1
+       tacostM2
+       tacostM3
+       tacostF1
+       tacostF2
+       tacostF3
-- land use variables (num stores by type and retail area)
+       LnstoreT=0
+       LCSCAarea=0
-- inclusive value (to test location-gap model, inverted gap-location model)
+       pIV

-************** D A T A   I N P U T **************************
$keep nzones = 1548
$array AVAILT(nzones) ATIME1(nzones) ATTOT(nzones) ATINCR(nzones)
$array ACOST(nzones) acstINC(nzones)
$array NSTORES(nzones) NSTOREST(nzones) AREA(nzones) POP(nzones) EMP(nzones)
file (name=shopMNLv6.dat, handle=shopMNL6)
OBS_NUM P_SHEP_ID ALT AVAILT
SH_TYPE H_SIZE I_GENDER I_LIC I_AUTO L_INC INC_LEV
ATIME1 ATTOT ATINCR
ACOST acstINC
NSTORES NSTOREST AREA POP EMP
IV IVall

-************** A V A I L A B I L I T Y **********************
------ location availability

Avail(1) = ifeq(AVAILT(1),1)
Avail(2) = ifeq(AVAILT(2),1)
...
Avail(1548) = ifeq(AVAILT(1548),1)
```

```
------ C H O I C E
choice = ALT
id = OBS_NUM

------ E X C L U S I O N
- Exclude an observation if the chosen location is not in the prism
exclude(1) = ifeq(ALT,0)
- Include only observations of certain shopping types (i.e. the line below
would only include observations of shopping type = 1)
exclude(2) = ifne(SH_TYPE,1)
- Include only observations of a certain gender (i.e. the line below would
include only male observations, gender = 1)
exclude(3) = ifne(I_GENDER,1)

-************* V A R I A B L E   T R A N S F O R M A T I O N S ***

------ declare arrays
-- auto travel time
$array ATINCRM1(nzones) ATINCRM2(nzones) ATINCRM3(nzones)
$array ATINCRF1(nzones) ATINCRF2(nzones) ATINCRF3(nzones)
-- auto travel cost
$array ACOSTM1(nzones) ACOSTM2(nzones) ACOSTM3(nzones)
$array ACOSTF1(nzones) ACOSTF2(nzones) ACOSTF3(nzones)

do dest = 1, nzones
    ifeq(I_GENDER, 1) then
      ifeq(INC_LEV, 1) then
            ATINCRM1(dest) = ATINCR(dest)
            ATINCRM2(dest) = 0
            ATINCRM3(dest) = 0
            ATINCRF1(dest) = 0
            ATINCRF2(dest) = 0
            ATINCRF3(dest) = 0
            ACOSTM1(dest) = ACOST(dest)
            ACOSTM2(dest) = 0
            ACOSTM3(dest) = 0
            ACOSTF1(dest) = 0
            ACOSTF2(dest) = 0
            ACOSTF3(dest) = 0
      end
      ifeq(INC_LEV, 2) then
            ATINCRM1(dest) = 0
            ATINCRM2(dest) = ATINCR(dest)
            ATINCRM3(dest) = 0
            ATINCRF1(dest) = 0
            ATINCRF2(dest) = 0
            ATINCRF3(dest) = 0
            ACOSTM1(dest) = 0
            ACOSTM2(dest) = ACOST(dest)
            ACOSTM3(dest) = 0
            ACOSTF1(dest) = 0
            ACOSTF2(dest) = 0
            ACOSTF3(dest) = 0
      end
      ifeq(INC_LEV, 3) then
            ATINCRM1(dest) = 0
            ATINCRM2(dest) = 0
            ATINCRM3(dest) = ATINCR(dest)
```

```
            ATINCRF1(dest) = 0
            ATINCRF2(dest) = 0
            ATINCRF3(dest) = 0
            ACOSTM1(dest) = 0
            ACOSTM2(dest) = 0
            ACOSTM3(dest) = ACOST(dest)
            ACOSTF1(dest) = 0
            ACOSTF2(dest) = 0
            ACOSTF3(dest) = 0
       end
    end
    ifeq(I_GENDER, 2) then
      ifeq(INC_LEV, 1) then
            ATINCRM1(dest) = 0
            ATINCRM2(dest) = 0
            ATINCRM3(dest) = 0
            ATINCRF1(dest) = ATINCR(dest)
            ATINCRF2(dest) = 0
            ATINCRF3(dest) = 0
            ACOSTM1(dest) = 0
            ACOSTM2(dest) = 0
            ACOSTM3(dest) = 0
            ACOSTF1(dest) = ACOST(dest)
            ACOSTF2(dest) = 0
            ACOSTF3(dest) = 0
      end
      ifeq(INC_LEV, 2) then
            ATINCRM1(dest) = 0
            ATINCRM2(dest) = 0
            ATINCRM3(dest) = 0
            ATINCRF1(dest) = 0
            ATINCRF2(dest) = ATINCR(dest)
            ATINCRF3(dest) = 0
            ACOSTM1(dest) = 0
            ACOSTM2(dest) = 0
            ACOSTM3(dest) = 0
            ACOSTF1(dest) = 0
            ACOSTF2(dest) = ACOST(dest)
            ACOSTF3(dest) = 0
      end
      ifeq(INC_LEV, 3) then
            ATINCRM1(dest) = 0
            ATINCRM2(dest) = 0
            ATINCRM3(dest) = 0
            ATINCRF1(dest) = 0
            ATINCRF2(dest) = 0
            ATINCRF3(dest) = ATINCR(dest)
            ACOSTM1(dest) = 0
            ACOSTM2(dest) = 0
            ACOSTM3(dest) = 0
            ACOSTF1(dest) = 0
            ACOSTF2(dest) = 0
            ACOSTF3(dest) = ACOST(dest)
      end
    end
end
```

```
-********** U T I L I T Y   F U N C T I O N S **************************

------ utilities
----------- Note: Had estimation issues when trying to index utilities and
alternative availability.

------ Location Utilities
U(1) = LnstoreT*NSTOREST(1) +LCSCAarea*AREA(1)
 +tacostM1*ACOSTM1(1) +tacostM2*ACOSTM2(1)
 +tacostM3*ACOSTM3(1) +tacostF1*ACOSTF1(1)
 +tacostF2*ACOSTF2(1) +tacostF3*ACOSTF3(1)
 +taincrM1*ATINCRM1(1) +taincrM2*ATINCRM2(1)
 +taincrM3*ATINCRM3(1) +taincrF1*ATINCRF1(1)
 +taincrF2*ATINCRF2(1) +taincrF3*ATINCRF3(1)
 +pIV*IV +pIVall*IVall
U(2) = LnstoreT*NSTOREST(2) +LCSCAarea*AREA(2)
 +tacostM1*ACOSTM1(2) +tacostM2*ACOSTM2(2)
 +tacostM3*ACOSTM3(2) +tacostF1*ACOSTF1(2)
 +tacostF2*ACOSTF2(2) +tacostF3*ACOSTF3(2)
 +taincrM1*ATINCRM1(2) +taincrM2*ATINCRM2(2)
 +taincrM3*ATINCRM3(2) +taincrF1*ATINCRF1(2)
 +taincrF2*ATINCRF2(2) +taincrF3*ATINCRF3(2)
 +pIV*IV +pIVall*IVall
...
U(1548) = LnstoreT*NSTOREST(1548) +LCSCAarea*AREA(1548)
 +tacostM1*ACOSTM1(1548) +tacostM2*ACOSTM2(1548)
 +tacostM3*ACOSTM3(1548) +tacostF1*ACOSTF1(1548)
 +tacostF2*ACOSTF2(1548) +tacostF3*ACOSTF3(1548)
 +taincrM1*ATINCRM1(1548) +taincrM2*ATINCRM2(1548)
 +taincrM3*ATINCRM3(1548) +taincrF1*ATINCRF1(1548)
 +taincrF2*ATINCRF2(1548) +taincrF3*ATINCRF3(1548)
 +pIV*IV +pIVall*IVall
```

Figure D2 ALOGIT Code: Gap Choice Model

```
------------------------------------------------------------
------------------------------------------------------------
- Gap Choice Model: Upper Level in Nested Logit Sequential Estimation
- Also used to estimate Lower Level to test Inverted Model
------------------------------------------------------------
$title Shopping Location Choice
$subtitle (@shopNLSeq.alo)
$print transFORms
$gen.stats utilities all
$estimate

-************ C O E F F I C I E N T S *********************

------ location coefficients
------ episode duration (+travel time) to gap duration ratio
$coeff pEpGap=0
+      pEpGap1
+      pEpGap2
+      pEpGap3
------ weekday
+      pMor=0
+      pAft
+      pEve
+      pLat
------ weekend
+      pWkd1=0
+      pWkd2
+      pWkd3
+      pWkd4
------ inclusive value
+      pIV=0

-************ D A T A   I N P U T ***************************

$keep ngaps = 41
$array AVAILT(ngaps) EpGap(ngaps)
$array WKD(ngaps)
$array SAT(ngaps) SUN(ngaps)
$array MON(ngaps) TUE(ngaps) WED(ngaps) THU(ngaps) FRI(ngaps)
$array MOR1(ngaps) AFT1(ngaps) EVE1(ngaps) LAT1(ngaps)
$array IV(ngaps)
file (name=shopNLSeq.dat, handle=shopNLS)

OBS_NUM P_ID SH_TYPE
ALT AVAILT I_GENDER I_INCOME EpGap
SUN MON TUE WED THU FRI SAT WKD
MOR1 AFT1 EVE1 LAT1
IV

-************ A V A I L A B I L I T Y *********************

------ location availability

Avail(1) = ifeq(AVAILT(1),1)
Avail(2) = ifeq(AVAILT(2),1)
```

```
...
Avail(41) = ifeq(AVAILT(41),1)

------ C H O I C E
choice = ALT
id = OBS_NUM

------ E X C L U S I O N
- Exclude an observation if the chosen location is not in the prism
exclude(1) = ifeq(ALT,0)
- Exclude by gender (i.e. line below excludes all observations except female,
gender = 2)
exclude(2) = ifne(I_GENDER, 2)
- Exclude by income (i.e. line below is commented out, and if included in the
code, would only include individuals with high income, income = 3)
-exclude(2) = ifne(I_INCOME, 3)
- Exclude by shopping type (i.e. line below only includes shopping type 5 in
the model estimation)
exclude(3) = ifne(SH_TYPE,5)

-************** V A R I A B L E   T R A N S F O R M A T I O N S ***
$array EpGap1(ngaps) EpGap2(ngaps) EpGap3(ngaps)
$array MOR(ngaps) AFT(ngaps) EVE(ngaps) LAT(ngaps)
$array WKD1(ngaps) WKD2(ngaps) WKD3(ngaps) WKD4(ngaps)
$array WDY(ngaps)

------ EpGap by Income
ifeq(I_INCOME,1) then
      do gap = 1, ngaps
            EpGap1(gap) = EpGap(gap)
            EpGap2(gap) = 0
            EpGap3(gap) = 0
      end
end
ifeq(I_INCOME,2) then
      do gap = 1, ngaps
            EpGap1(gap) = 0
            EpGap2(gap) = EpGap(gap)
            EpGap3(gap) = 0
      end
end
ifeq(I_INCOME,3) then
      do gap = 1, ngaps
            EpGap1(gap) = 0
            EpGap2(gap) = 0
            EpGap3(gap) = EpGap(gap)
      end
end

------ DayOfWeek and TimeOfDay interaction
-- WKD1(gap): Morning-Weekend
do gap = 1, ngaps
      WKD1(gap) = Mor1(gap)*WKD(gap)
end
-- WKD2(gap): Afternoon-Weekend
do gap = 1, ngaps
      WKD2(gap) = Aft1(gap)*WKD(gap)
end
```

```
-- WKD3(gap): Evening-Weekend
do gap = 1, ngaps
     WKD3(gap) = Eve1(gap)*WKD(gap)
end
-- WKD4(gap): Late-Weekend
do gap = 1, ngaps
     WKD4(gap) = Lat1(gap)*WKD(gap)
end

-- Define Weekday
do gap = 1, ngaps
     ifeq(WKD1(gap),1) then
          WDY(gap) = 1
     else
          WDY(gap) = 0
     end
end

-- Mor(gap): Weekday-Morning
do gap = 1, ngaps
     Mor(gap) = Mor1(gap)*WDY(gap)
end
-- Aft(gap): Weekday-Afternoon
do gap = 1, ngaps
     Aft(gap) = Aft1(gap)*WDY(gap)
end
-- Eve(gap): Weekday-Evening
do gap = 1, ngaps
     Eve(gap) = Eve1(gap)*WDY(gap)
end
-- Lat(gap): Weekday-Late
do gap = 1, ngaps
     Lat(gap) = Lat1(gap)*WDY(gap)
end

-********** U T I L I T Y   F U N C T I O N S **************************

------- Gap Utilities
U(1) = pEpGap*EpGap(1)
 +pEpGap1*EpGap1(1) +pEpGap2*EpGap2(1) +pEpGap3*EpGap3(1)
 +pWkd1*WKD1(1) +pWkd2*WKD2(1)
 +pWkd3*WKD3(1) +pWkd4*WKD4(1) +pMor*Mor(1)
 +pAft*Aft(1) +pEve*Eve(1) +pLat*Lat(1)
 +pIV*IV(1)
U(2) = pEpGap*EpGap(2)
 +pEpGap1*EpGap1(2) +pEpGap2*EpGap2(2) +pEpGap3*EpGap3(2)
 +pWkd1*WKD1(2) +pWkd2*WKD2(2)
 +pWkd3*WKD3(2) +pWkd4*WKD4(2) +pMor*Mor(2)
 +pAft*Aft(2) +pEve*Eve(2) +pLat*Lat(2)
 +pIV*IV(2)
...
U(41) = pEpGap*EpGap(41)
 +pEpGap1*EpGap1(41) +pEpGap*EpGap2(41) +pEpGap3*EpGap3(41)
 +pWkd1*WKD1(41) +pWkd2*WKD2(41)
 +pWkd3*WKD3(41) +pWkd4*WKD4(41) +pMor*Mor(41)
 +pAft*Aft(41) +pEve*Eve(41) +pLat*Lat(41)
 +pIV*IV(41)
```

Table D1 Location Choice Model used for Simulation

Parameter	All Types		Type 1		Type 2	
	Estimate	T-Ratio	Estimate	T-Ratio	Estimate	T-Ratio
Auto Cost						
Male						
Low Income (<$30,000)	-	-	-	-	-	-
Med Income ($30,000 to $60,000)	-	-	-	-	-	-
High Income (≥$60,000)	-	-	-	-	-	-
Female						
Low Income (<$30,000)	-0.09635	-2.2	-	-	-0.2232	-2.6
Med Income ($30,000 to $60,000)	-0.09380	-2.0	-0.1171	-0.6	-0.1105	-1.5
High Income (≥$60,000)	-	-	-	-	-	-
Auto Travel Time						
Male						
Low Income (<$30,000)	-0.2478	-9.2	-1.152	-1.5	-0.2525	-6.6
Med Income ($30,000 to $60,000)	-0.2283	-15.9	-0.3902	-6.6	-0.2673	-12.1
High Income (≥$60,000)	-0.1491	-14.8	-0.3692	-3.6	-0.1624	-9.9
Female						
Low Income (<$30,000)	-0.1833	-14.3	-0.4124	-6.5	-0.2441	-10.3
Med Income ($30,000 to $60,000)	-0.1595	-13.0	-0.3258	-5.2	-0.1839	-8.9
High Income (≥$60,000)	-0.2594	-11.1	-0.6087	-2.7	-0.2623	-8.3
Log(number of retail stores)	0.7681	21.7	-	-	0.8817	14.9
Log(retail floor space)	0.1205	13.0	0.04994	1.8	0.1044	8.4
Log-Likelihood	-2661.9369		-290.7554		-1211.1822	
Rho-Squared w.r.t. Zero	0.5378		0.4856		0.5576	

Table D1 Location Choice Model used for Simulation, cont'd

Parameter	Type 3		Type 4		Type 5	
	Estimate	T-Ratio	Estimate	T-Ratio	Estimate	T-Ratio
Auto Cost						
Male						
Low Income (<$30,000)	-	-	-	-	-	-
Med Income ($30,000 to $60,000)	-	-	-0.3743	-1.3	-0.2857	-1.3
High Income ($\geq$$60,000)	-0.2470	-1.9	-0.1667	-0.8	-	-
Female						
Low Income (<$30,000)	-0.1684	-1.9	-0.2033	-1.2	-	-
Med Income ($30,000 to $60,000)	-0.1070	-1.1	-0.4503	-1.6	-	-
High Income ($\geq$$60,000)	-0.4584	-1.7	-	-	-	-
Auto Travel Time						
Male						
Low Income (<$30,000)	-0.2467	-4.8	-	-	-0.1569	-3.6
Med Income ($30,000 to $60,000)	-0.1346	-3.9	-0.07491	-1.4	-0.1070	-2.9
High Income ($\geq$$60,000)	-0.05319	-1.8	-0.1041	-2.4	-0.1469	-6.6
Female						
Low Income (<$30,000)	-0.08488	-3.6	-0.06774	-1.7	-0.1616	-8.9
Med Income ($30,000 to $60,000)	-0.1339	-5.2	-0.06927	-1.3	-0.1259	-8.7
High Income ($\geq$$60,000)	-0.1758	-2.6	-	-	-0.2261	-3.8
Log(number of retail stores)	1.041	13.0	0.8006	6.3	0.6234	9.7
Log(retail floor space)	0.1649	6.7	0.2786	4.9	0.1534	6.6
Log-Likelihood	-419.8102		-143.8703		-484.8085	
Rho-Squared w.r.t. Zero	0.5414		0.6560		0.5684	

Table D2 Gap Choice Model Estimated with Inclusive Values

Parameter	Male		Female	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
All Types				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-2.007	-3.0	-2.078	-5.1
Medium Income ($30,000 to $60,000)	-1.668	-4.3	-1.446	-4.5
High Income (≥$60,000)	-1.842	-4.2	-1.476	-3.0
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-0.2964	-0.5	-0.1915	-0.4
Eve (5pm-9pm)	2.134	3.2	1.371	2.3
Lat (9pm-6am)	1.287	6.5	1.227	7.2
Weekend				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	1.232	2.4	0.7194	1.7
Eve (5pm-9pm)	-1.010	-2.2	-1.050	-2.5
Lat (9pm-6am)	-0.5236	-1.3	-0.8016	-2.6
IV	0.6714	1.6	-0.06495	-0.5
Log-Likelihood	-1101.5916		-1662.7207	
Rho-Squared	0.0924		0.0595	
Num. Observations	471		670	
Parameter	**Male**		**Female**	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 1				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-8.329	-0.7	-4.203	-1.9
Medium Income ($30,000 to $60,000)	-3.372	-2.2	-1.267	-1.2
High Income (≥$60,000)	-9.503	-1.4	-2.439	-1.0
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-	-
Eve (5pm-9pm)	-	-	-	-
Lat (9pm-6am)	-	-	-	-
Weekend				
Mor (6am-12pm)	-	-	1.731	2.5
Aft (12pm-5pm)	2.019	2.9	-	-
Eve (5pm-9pm)	-	-	-	-
Lat (9pm-6am)	-	-	-	-
IV	2.576	0.3	-0.4958	-1.3
Log-Likelihood	-87.3058		-122.5120	
Rho-Squared	0.1351		0.0744	
Num. Observations	37		46	

Table D2 Gap Choice Model Estimated with Inclusive Values, cont'd

Parameter	Male		Female	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 2				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-2.547	-2.5	-3.886	-4.3
Medium Income ($30,000 to $60,000)	-1.186	-2.4	-1.986	-3.6
High Income (≥$60,000)	-1.421	-2.6	-2.195	-3.0
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	0.2681	0.4	-0.5267	-0.7
Eve (5pm-9pm)	0.5286	0.9	1.100	1.2
Lat (9pm-6am)	-	-	0.9313	3.7
Weekend				
Mor (6am-12pm)	1.742	3.4	-	-
Aft (12pm-5pm)	2.532	6.7	1.434	2.3
Eve (5pm-9pm)	-	-	-0.8287	-1.2
Lat (9pm-6am)	-	-	-	-
IV	-0.8319	-0.7	0.6861	1.0
Log-Likelihood	-539.8232		-710.5559	
Rho-Squared	0.0858		0.0809	
Num. Observations	225		287	
Parameter	**Male**		**Female**	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 3				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-1.009	-0.8	-1.684	-2.1
Medium Income ($30,000 to $60,000)	-1.465	-1.1	-1.740	-2.3
High Income (≥$60,000)	-2.043	-1.7	-0.8685	-0.9
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-	-
Eve (5pm-9pm)	-	-	2.076	1.7
Lat (9pm-6am)	-	-	-	-
Weekend				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	0.9627	1.4	0.5051	1.0
Eve (5pm-9pm)	-0.8252	-0.8	-2.137	-2.0
Lat (9pm-6am)	-0.8852	-1.1	-0.9319	-1.6
IV	1.575	2.5	-0.2385	-0.7
Log-Likelihood	-135.7255		-315.8100	
Rho-Squared	0.1416		0.0359	
Num. Observations	60		136	

Table D2 Gap Choice Model Estimated with Inclusive Values, cont'd

Parameter	Male		Female	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 4				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-0.4420	-0.2	-2.903	-1.5
Medium Income ($30,000 to $60,000)	-2.715	-1.8	-2.166	-1.4
High Income (≥$60,000)	-5.652	-2.9	-1.053	-0.2
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-2.823	-1.5
Eve (5pm-9pm)	-	-	2.303	1.2
Lat (9pm-6am)	-	-	-1.173	-0.8
Weekend				
Mor (6am-12pm)	-	-	3.284	2.0
Aft (12pm-5pm)	2.227	1.8	2.117	1.7
Eve (5pm-9pm)	-1.572	-1.2	-	-
Lat (9pm-6am)	-0.6120	-0.4	-	-
IV	1.522	0.8	1.227	1.4
Log-Likelihood	-112.6431		-82.4077	
Rho-Squared	0.1335		0.1518	
Num. Observations	51		37	
Parameter	**Male**		**Female**	
	Estimate	**T-Ratio**	**Estimate**	**T-Ratio**
Type 5				
Episode Duration (+Auto Travel Time) to Gap Duration				
Low Income (<$30,000)	-3.511	-1.7	-0.9937	-1.6
Medium Income ($30,000 to $60,000)	-3.883	-3.4	-1.036	-1.8
High Income (≥$60,000)	-5.439	-3.3	-0.6943	-0.6
Weekday				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	-	-	-0.8488	-0.9
Eve (5pm-9pm)	-	-	0.6179	0.4
Lat (9pm-6am)	-	-	1.450	4.1
Weekend				
Mor (6am-12pm)	-	-	-	-
Aft (12pm-5pm)	1.496	2.3	1.244	1.6
Eve (5pm-9pm)	-1.321	-1.5	-0.8383	-1.0
Lat (9pm-6am)	0.3918	0.4	-0.3890	-0.6
IV	-0.5093	-0.6	-0.1495	-0.5
Log-Likelihood	-220.4235		-418.8653	
Rho-Squared	0.1165		0.0433	
Num. Observations	98		164	

Appendix E: Simulation Output

Table E1 Goodness of Fit Tables

Iterations: 1					
Shopping Type	1	2	3	4	5
Num Total	67	534	204	89	267
Num Loc Right	5	15	2	1	13
% Loc Right	7.46%	2.81%	0.98%	1.12%	4.87%
Num Gap Right	7	56	17	8	33
% Gap Right	10.45%	10.49%	8.33%	8.99%	12.36%
Avg Dist (km) from Pred to Chosen	5.82	8.59	10.17	10.73	9.75
Max Dist (km) from Pred to Chosen	32.24	81.45	78.00	64.28	65.04
Iterations: 10					
Shopping Type	1	2	3	4	5
Num Total	685	5340	2040	890	2670
Num Loc Right	29	143	37	19	43
% Loc Right	4.23%	2.68%	1.81%	2.13%	1.61%
Num Gap Right	82	523	199	139	262
% Gap Right	11.97%	9.79%	9.75%	15.62%	9.81%
Avg Dist (km) from Pred to Chosen	7.05	7.95	10.35	10.26	9.55
Max Dist (km) from Pred to Chosen	73.93	88.73	78.00	79.00	70.31
Iterations: 100					
Shopping Type	1	2	3	4	5
Num Total	6883	53400	20400	8900	26700
Num Loc Right	296	1362	350	206	414
% Loc Right	4.30%	2.55%	1.72%	2.31%	1.55%
Num Gap Right	822	5376	2189	1418	2619
% Gap Right	11.94%	10.07%	10.73%	15.93%	9.81%
Avg Dist (km) from Pred to Chosen	7.22	8.01	10.23	9.92	9.46
Max Dist (km) from Pred to Chosen	83.08	104.00	93.52	89.30	94.15
Iterations: 1000					
Shopping Type	1	2	3	4	5
Num Total	68862	534000	204000	89000	267000
Num Loc Right	2850	13191	3705	2216	4096
% Loc Right	4.14%	2.47%	1.82%	2.49%	1.53%
Num Gap Right	8013	53631	22493	14155	26644
% Gap Right	11.64%	10.04%	11.03%	15.90%	9.98%
Avg Dist (km) from Pred to Chosen	7.35	7.99	10.29	9.94	9.57
Max Dist (km) from Pred to Chosen	96.58	104.00	104.45	103.28	100.27

Figure E1.1 Distributions of Distance from Predicted to Chosen Zones (Type 1)
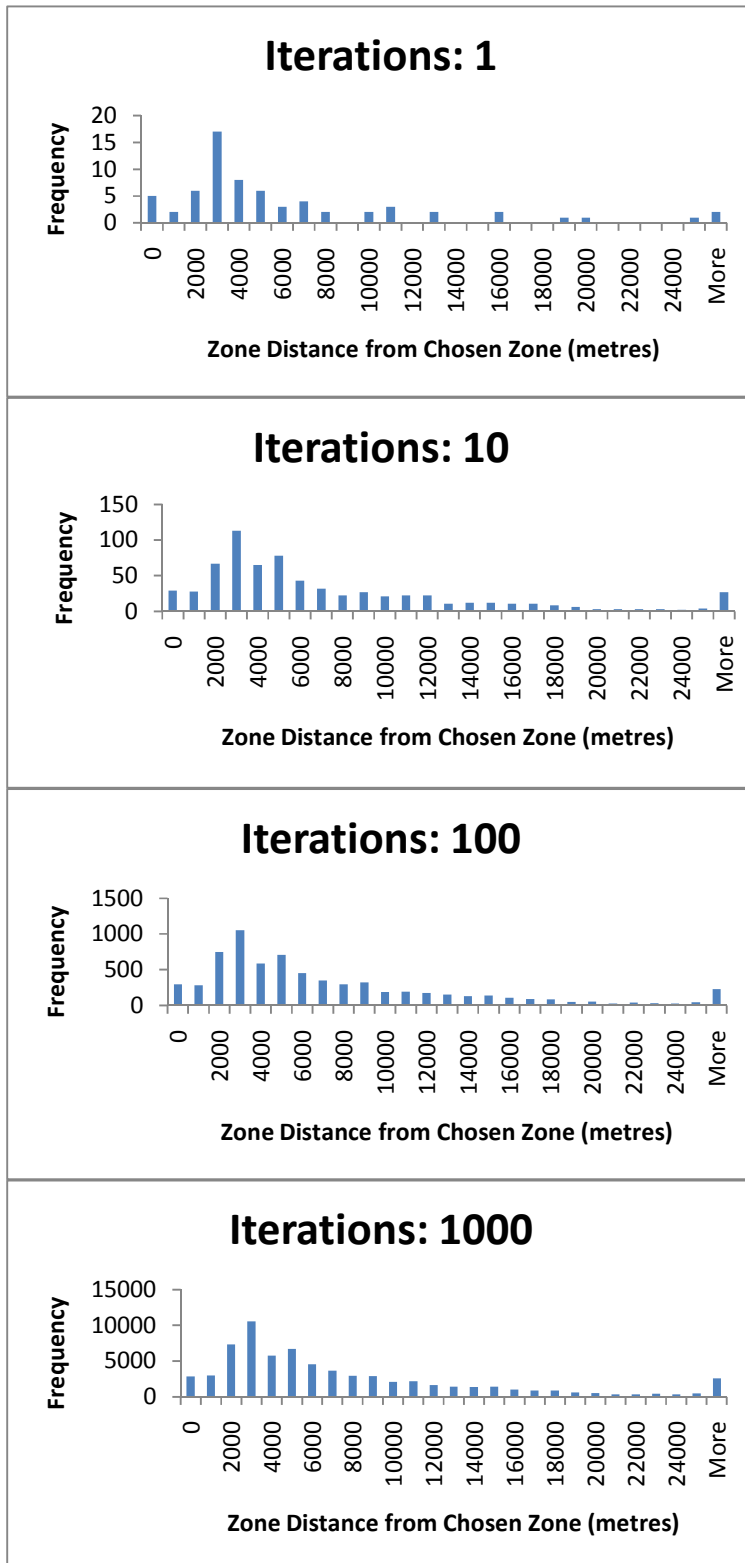
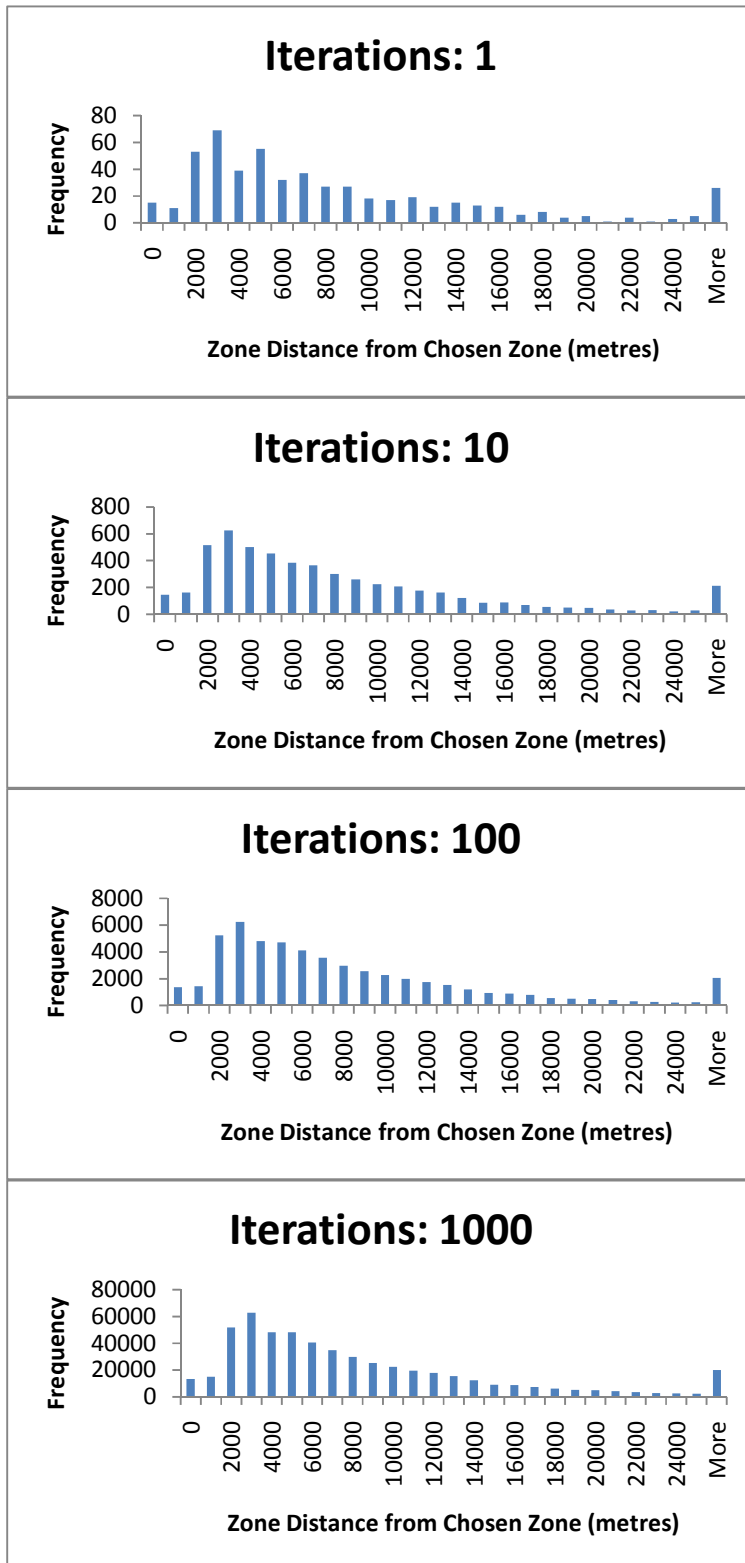**Figure E1.2 Distributions of Distance from Predicted to Chosen Zones (Type 2)**

Figure E1.3 Distributions of Distance from Predicted to Chosen Zones (Type 3)

Figure E1.4 Distributions of Distance from Predicted to Chosen Zones (Type 4)
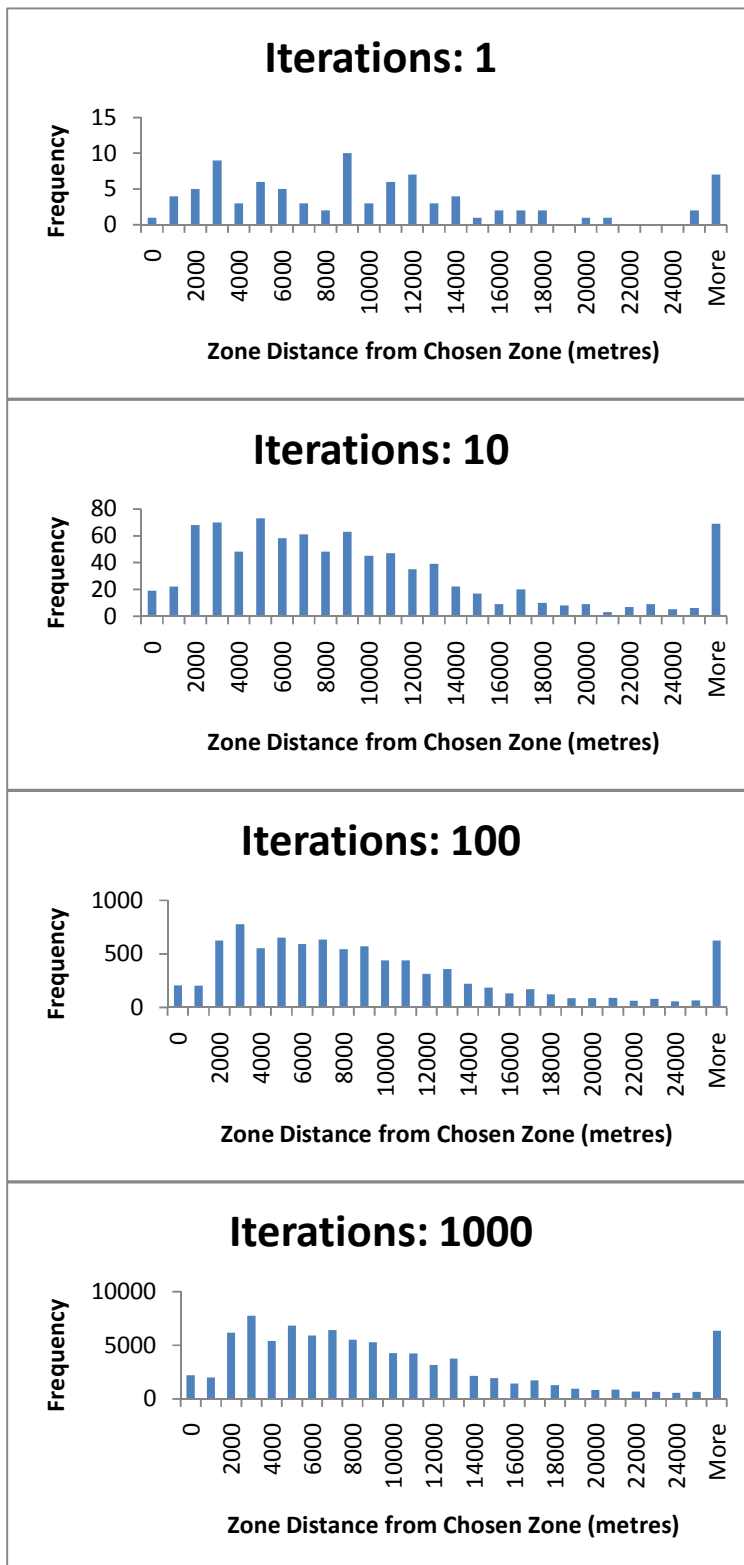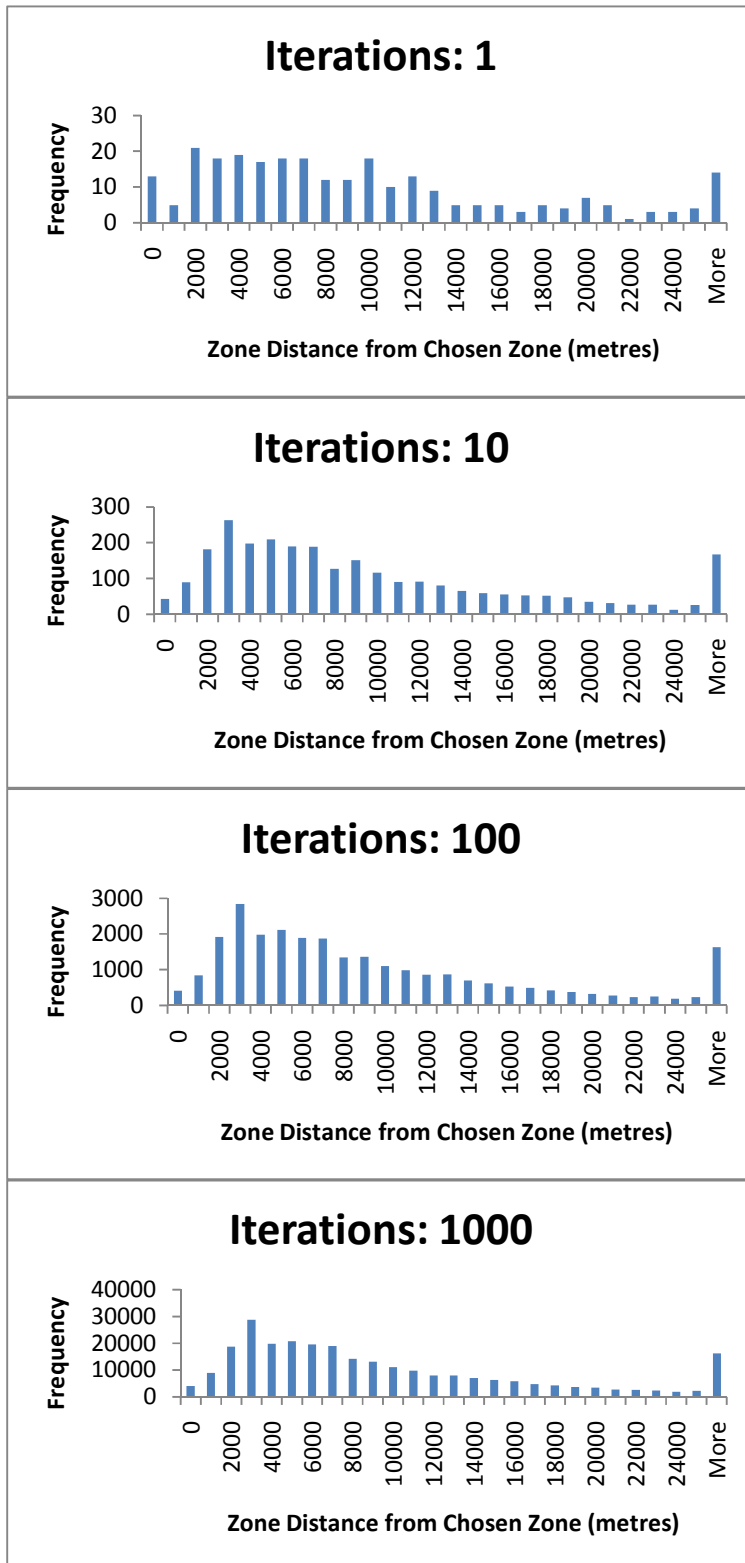
Figure E1.5 Distributions of Distance from Predicted to Chosen Zones (Type 5)

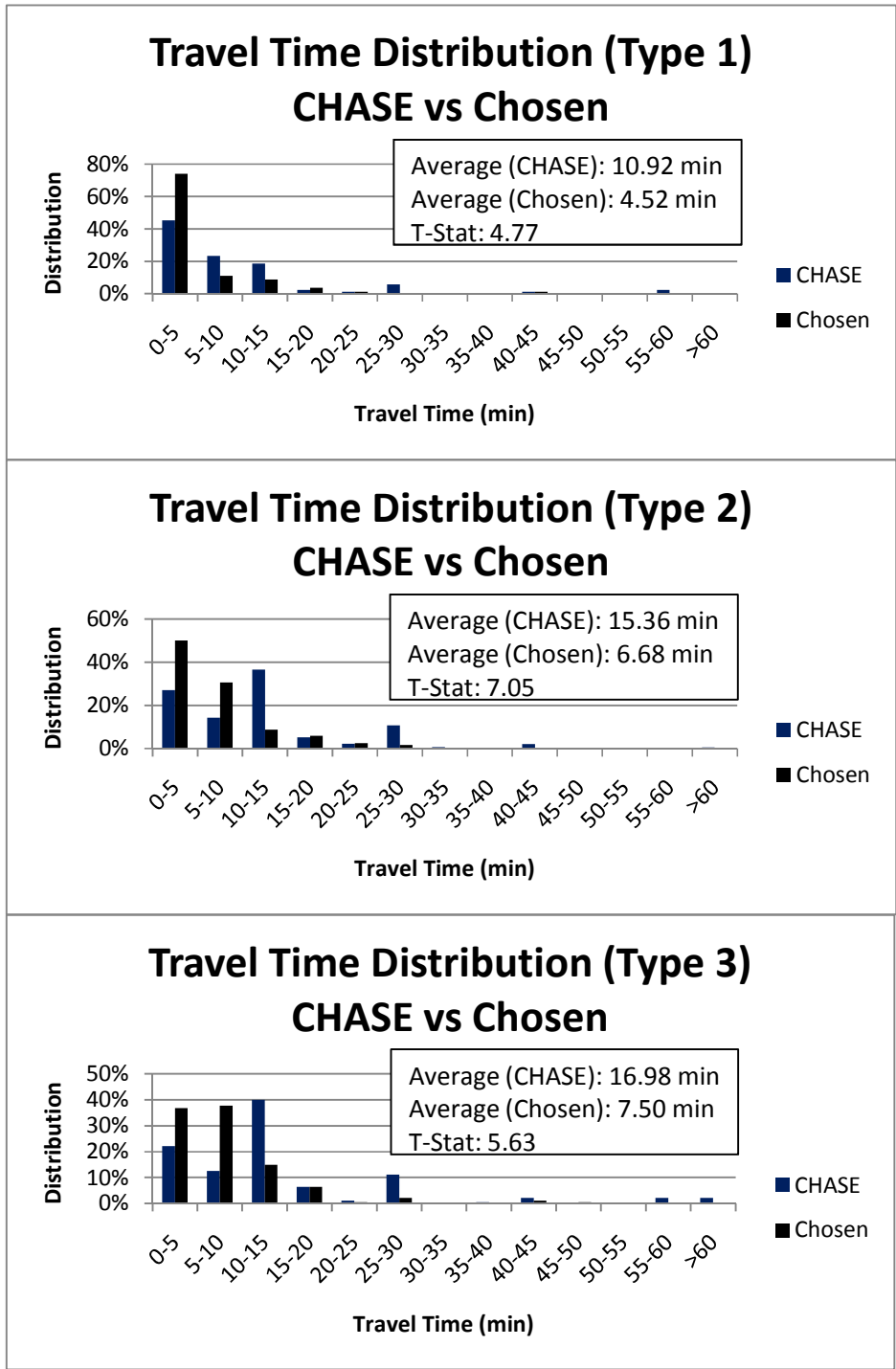**Figure E2.1 Auto Travel Time Distribution to Shopping: CHASE vs Chosen**

Figure E2.1 Auto Travel Time Distribution to Shopping: CHASE vs Chosen, cont'd

Travel Time Distribution (Type 4) CHASE vs Chosen

Average (CHASE): 15.06 min
Average (Chosen): 8.21 min
T-Stat: 4.68

Travel Time Distribution (Type 5) CHASE vs Chosen

Average (CHASE): 16.70 min
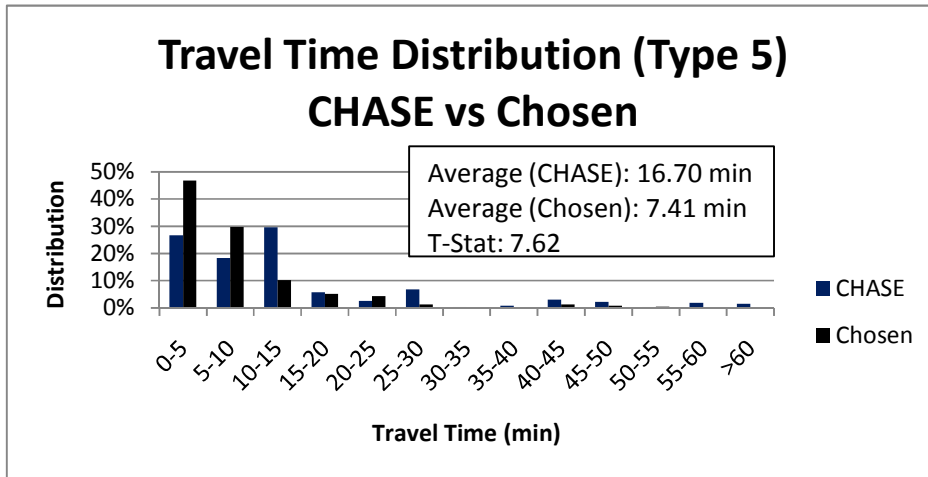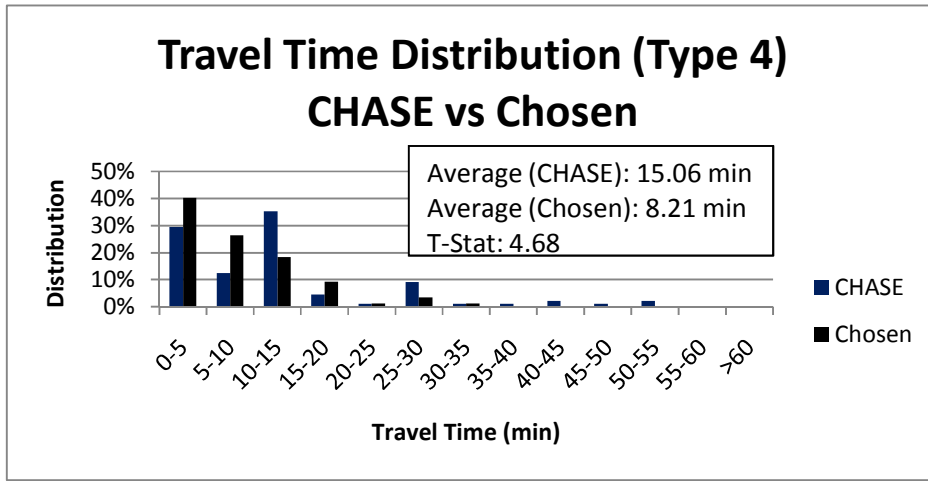Average (Chosen): 7.41 min
T-Stat: 7.62

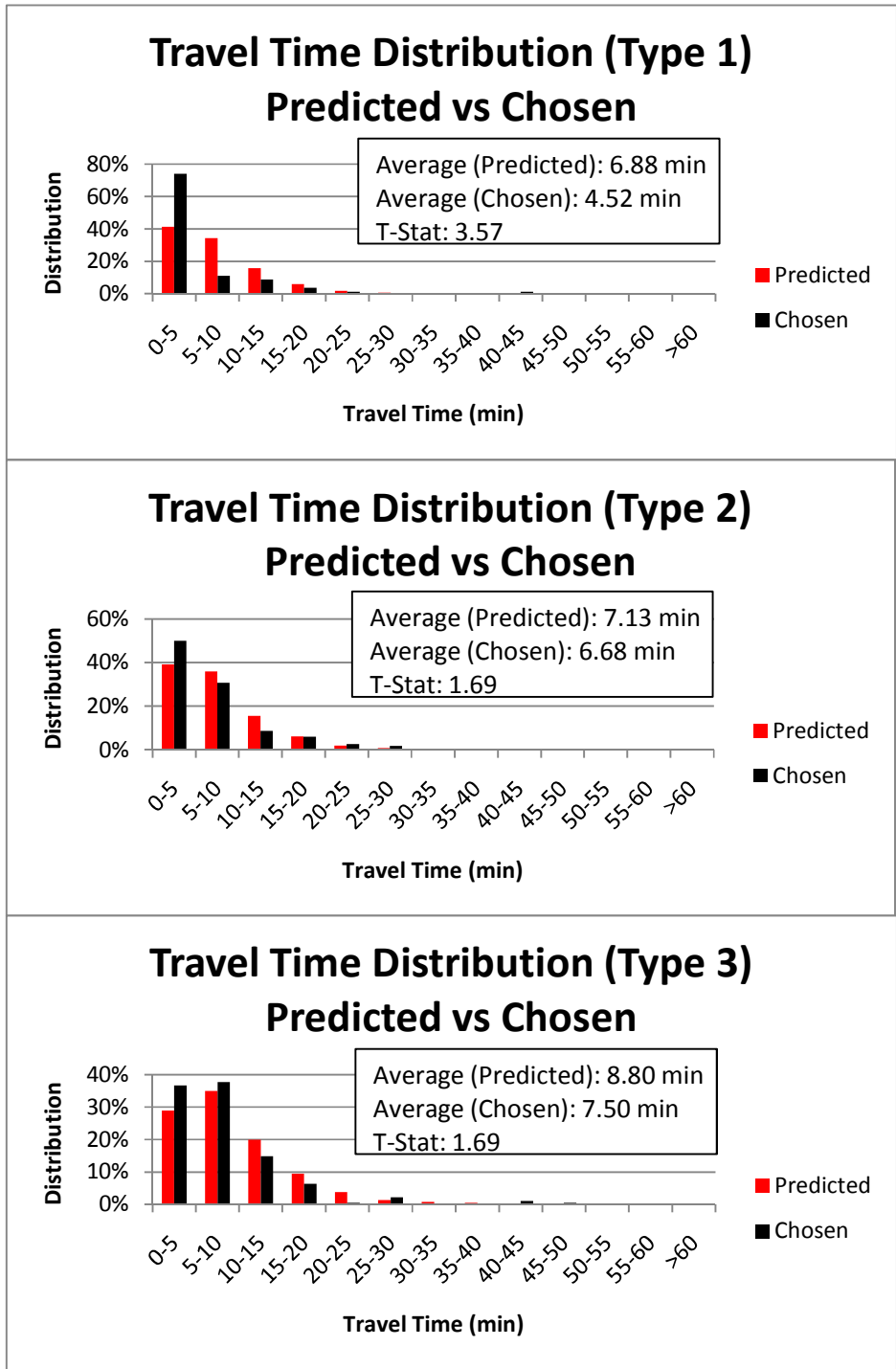**Figure E2.2 Auto Travel Time Distribution to Shopping: Predicted vs Chosen**

Figure E2.2 Auto Travel Time Distribution to Shopping: Predicted vs Chosen, cont'd

Travel Time Distribution (Type 4)
Predicted vs Chosen

Average (Predicted): 8.65 min
Average (Chosen): 8.21 min
T-Stat: 0.62

Travel Time Distribution (Type 5)
Predicted vs Chosen

Average (Predicted): 8.00 min
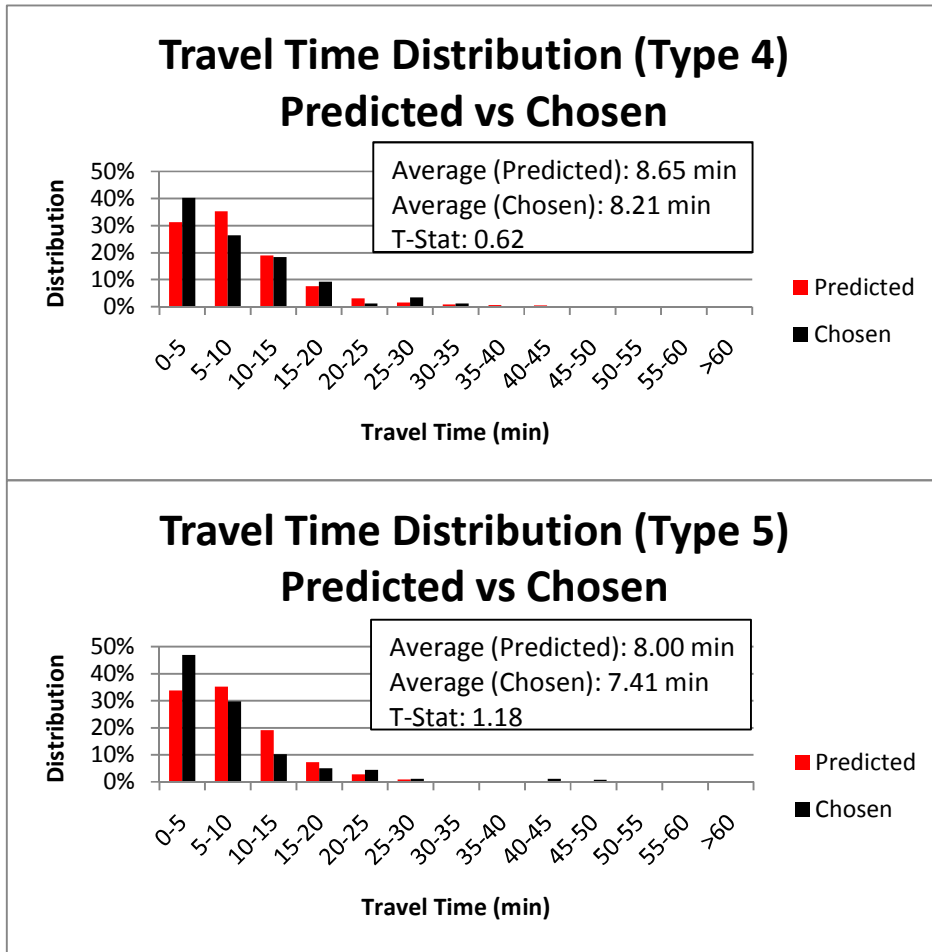Average (Chosen): 7.41 min
T-Stat: 1.18

Figure E2.3 Auto Travel Time Distributions to/from Shopping: Predicted vs Chosen
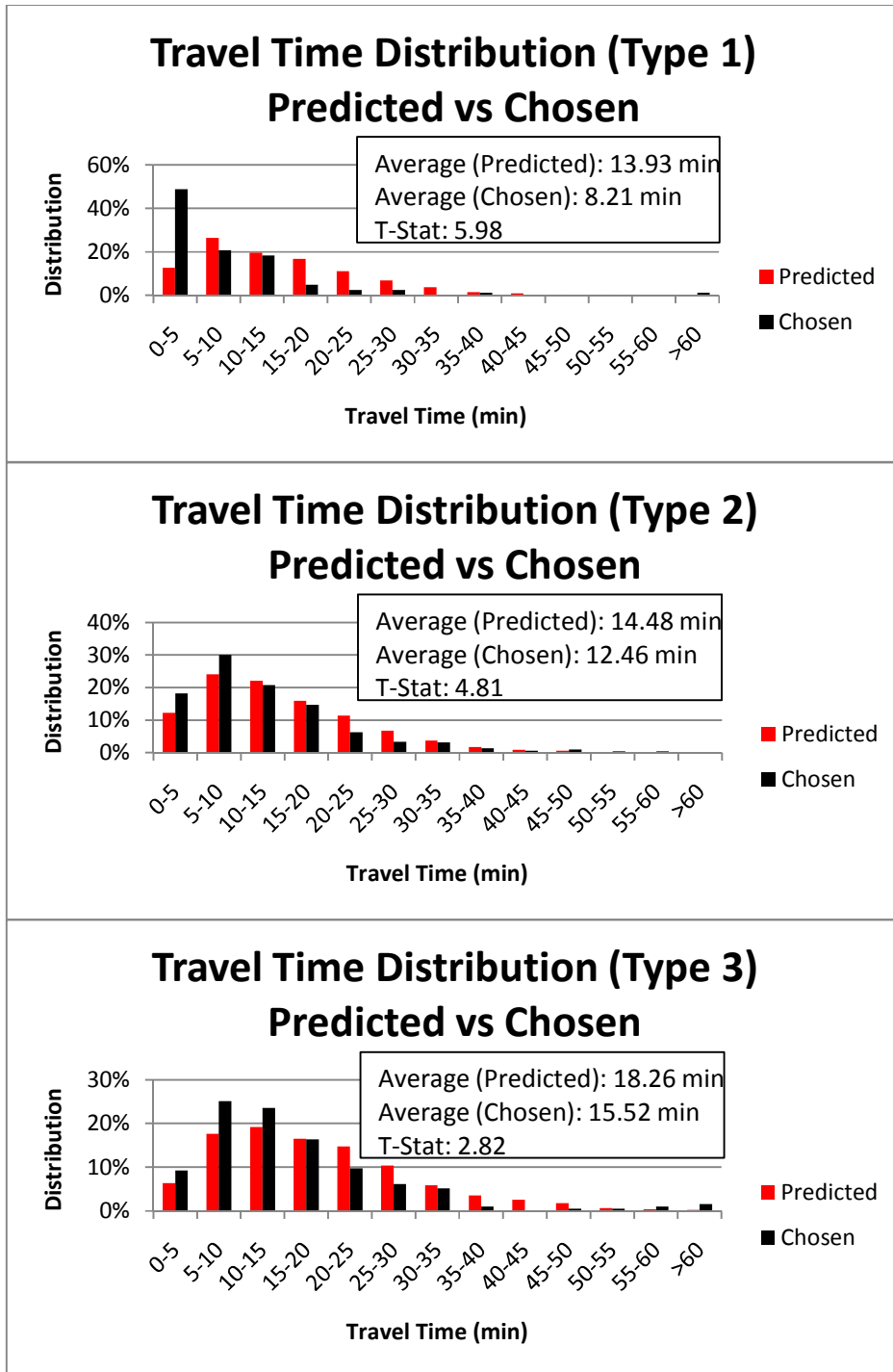
Figure E2.3 Auto Travel Time Distributions to/from Shopping: Predicted vs Chosen, cont'd
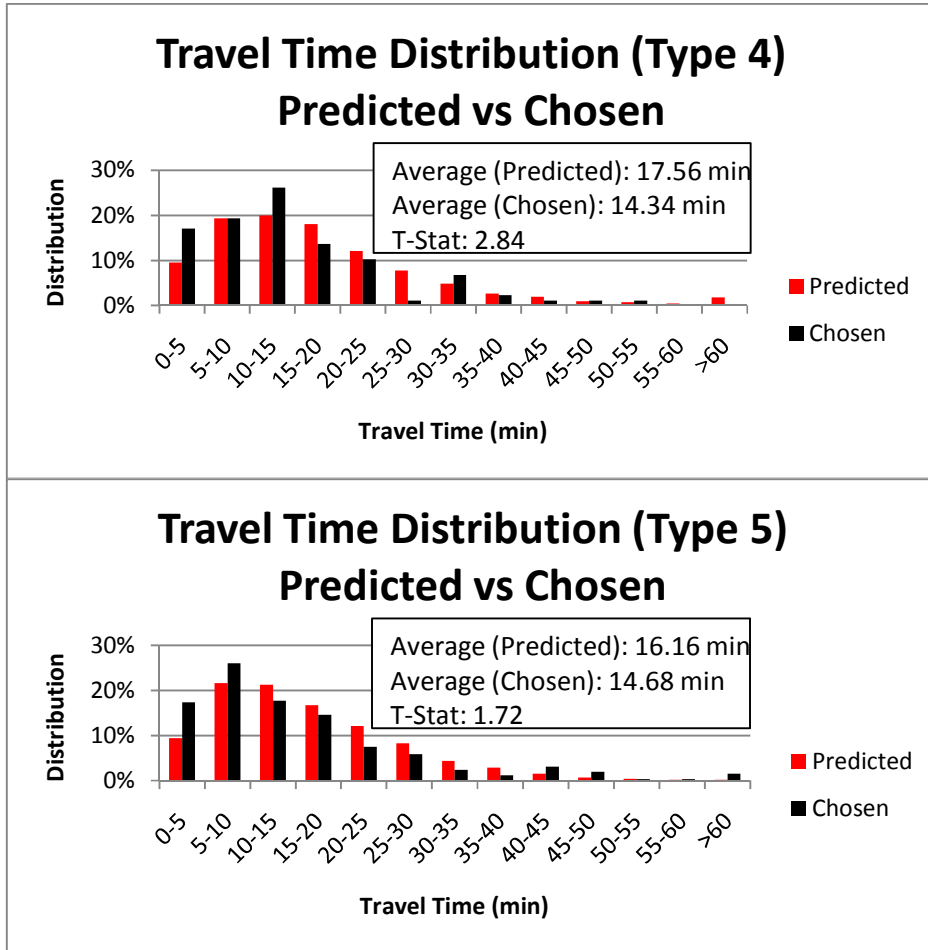
Table E2.1 Predicted Gap Choices (Counts) – by Time Period

Chosen	Time Period				
Shopping Type	Mor	Aft	Eve	Lat	Total
Type 1	40	14	8	74	136
Type 2	207	104	78	451	840
Type 3	90	36	26	181	333
Type 4	43	21	12	74	150
Type 5	106	42	23	245	416
Total	486	217	147	1025	1875
Predicted	**Time Period**				
Shopping Type	Mor	Aft	Eve	Lat	Total
Type 1	4629	2259	1223	8043	16154
Type 2	18637	9055	7256	48178	83126
Type 3	5990	3704	4748	17513	31955
Type 4	4258	2938	1426	7024	15646
Type 5	7111	2976	1811	25309	37207
Total	40625	20932	16464	106067	184088

Table E2.2 Predicted Gap Choices (Counts) – by Day of Week

Chosen	Day of Week							
Shopping Type	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Type 1	13	7	13	16	12	16	11	88
Type 2	67	60	60	63	71	76	135	532
Type 3	39	22	19	20	25	35	45	205
Type 4	16	14	8	8	11	13	20	90
Type 5	45	31	15	38	37	37	68	271
Total	180	134	115	145	156	177	279	1186
Predicted	**Day of Week**							
Shopping Type	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Type 1	2223	819	883	878	873	986	2238	8900
Type 2	8662	6904	7035	7306	7563	7335	8595	53400
Type 3	1639	3029	3753	3403	3352	3712	1512	20400
Type 4	2178	736	880	691	830	840	2745	8900
Type 5	2913	4108	3722	3977	4128	4532	3320	26700
Total	17615	15596	16273	16255	16746	17405	18410	118300

Table E2.3 Predicted Gap Choices (Distribution) – by Time Period

Chosen	Time Period				
Shopping Type	Mor	Aft	Eve	Lat	Total
Type 1	2%	1%	0%	4%	7%
Type 2	11%	6%	4%	24%	45%
Type 3	5%	2%	1%	10%	18%
Type 4	2%	1%	1%	4%	8%
Type 5	6%	2%	1%	13%	22%
Total	26%	12%	8%	55%	100%
Predicted	Time Period				
Shopping Type	Mor	Aft	Eve	Lat	Total
Type 1	3%	1%	1%	4%	9%
Type 2	10%	5%	4%	26%	45%
Type 3	3%	2%	3%	10%	17%
Type 4	2%	2%	1%	4%	8%
Type 5	4%	2%	1%	14%	20%
Total	22%	11%	9%	58%	100%
Difference	Time Period				
Shopping Type	Mor	Aft	Eve	Lat	Total
Type 1	0%	0%	0%	0%	2%
Type 2	-1%	-1%	0%	2%	0%
Type 3	-2%	0%	1%	0%	0%
Type 4	0%	0%	0%	0%	0%
Type 5	-2%	-1%	0%	1%	-2%
Total	-4%	0%	1%	3%	0%

Table E2.4 Predicted Gap Choices (Distribution) – by Day Of Week

Chosen	Day of Week							
Shopping Type	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Type 1	1%	1%	1%	1%	1%	1%	1%	7%
Type 2	6%	5%	5%	5%	6%	6%	11%	45%
Type 3	3%	2%	2%	2%	2%	3%	4%	17%
Type 4	1%	1%	1%	1%	1%	1%	2%	8%
Type 5	4%	3%	1%	3%	3%	3%	6%	23%
Total	15%	11%	10%	12%	13%	15%	24%	100%
Predicted	Day of Week							
Shopping Type	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Type 1	2%	1%	1%	1%	1%	1%	2%	8%
Type 2	7%	6%	6%	6%	6%	6%	7%	45%
Type 3	1%	3%	3%	3%	3%	3%	1%	17%
Type 4	2%	1%	1%	1%	1%	1%	2%	8%
Type 5	2%	3%	3%	3%	3%	4%	3%	23%
Total	15%	13%	14%	14%	14%	15%	16%	100%
Difference	Day of Week							
Shopping Type	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Type 1	1%	0%	0%	-1%	0%	-1%	1%	0%
Type 2	2%	1%	1%	1%	0%	0%	-4%	0%
Type 3	-2%	1%	2%	1%	1%	0%	-3%	0%
Type 4	0%	-1%	0%	0%	0%	0%	1%	0%
Type 5	-1%	1%	2%	0%	0%	1%	-3%	0%
Total	0%	2%	4%	2%	1%	0%	-8%	0%

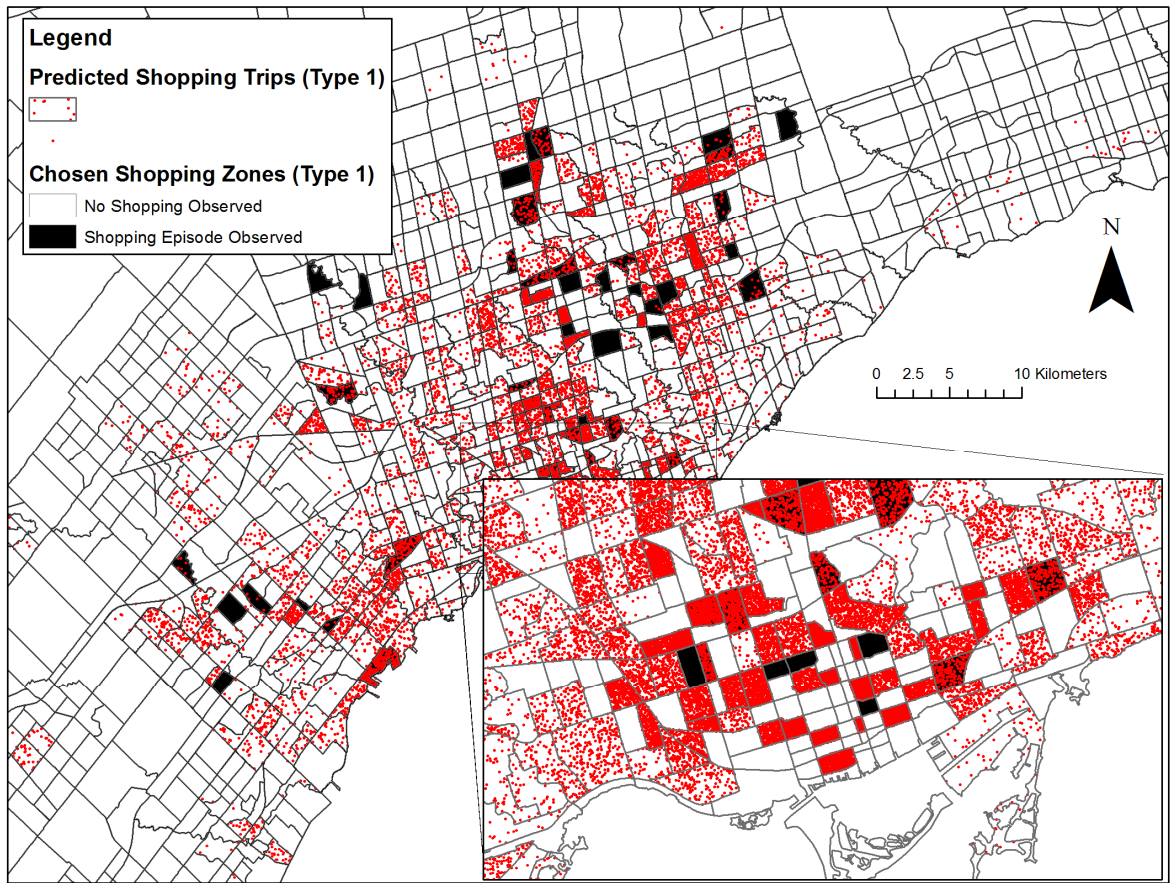**Figure E3.1 Spatial Distribution of Predicted Shopping Locations (Type 1)**

Figure E3.2 Spatial Distribution of Predicted Shopping Locations (Type 2)
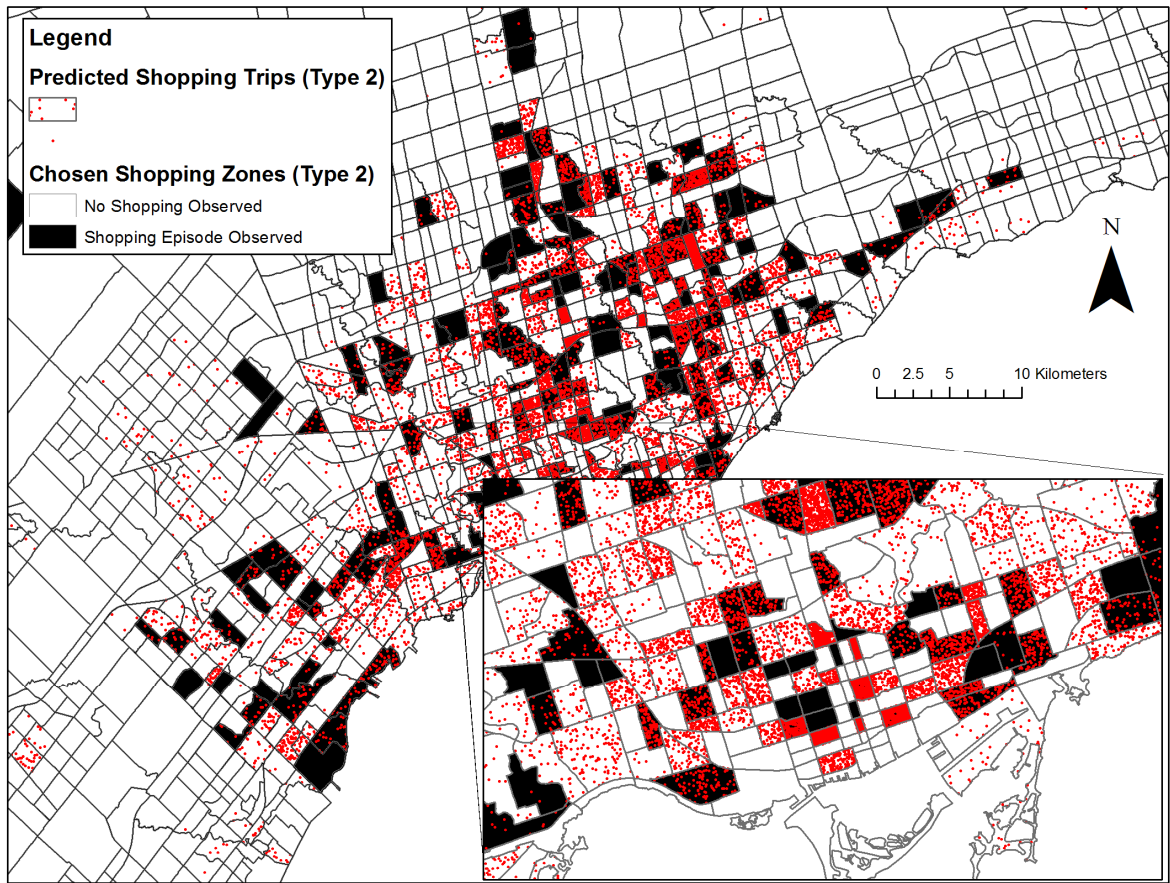
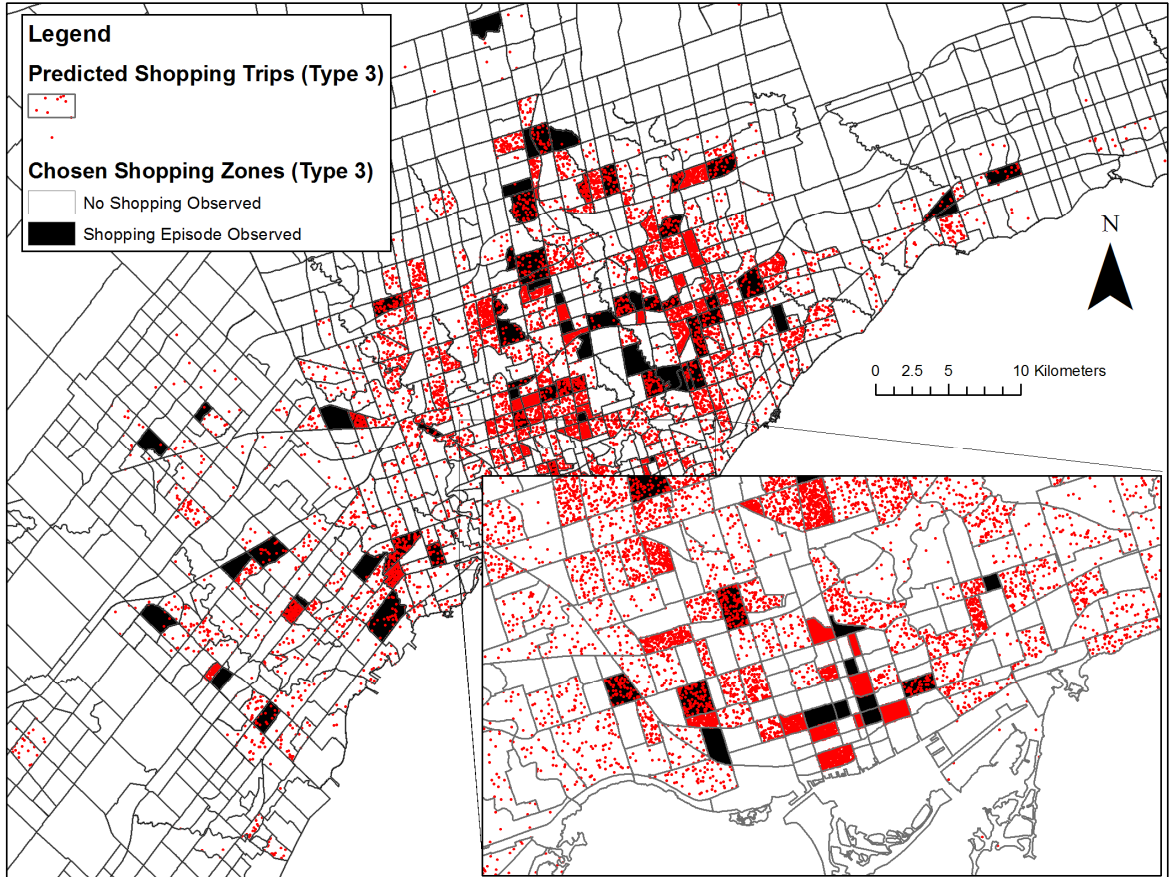**Figure E3.3 Spatial Distribution of Predicted Shopping Locations (Type 3)**

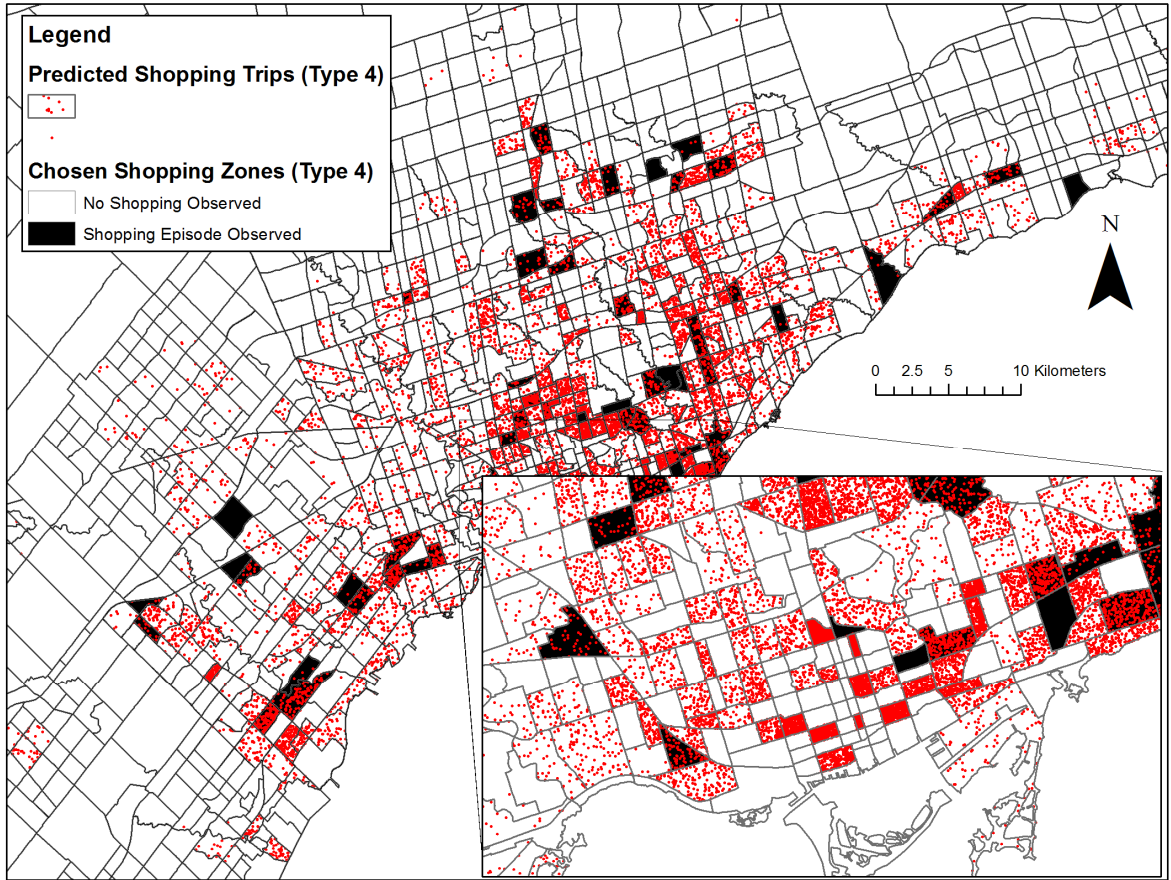**Figure E3.4 Spatial Distribution of Predicted Shopping Locations (Type 4)**
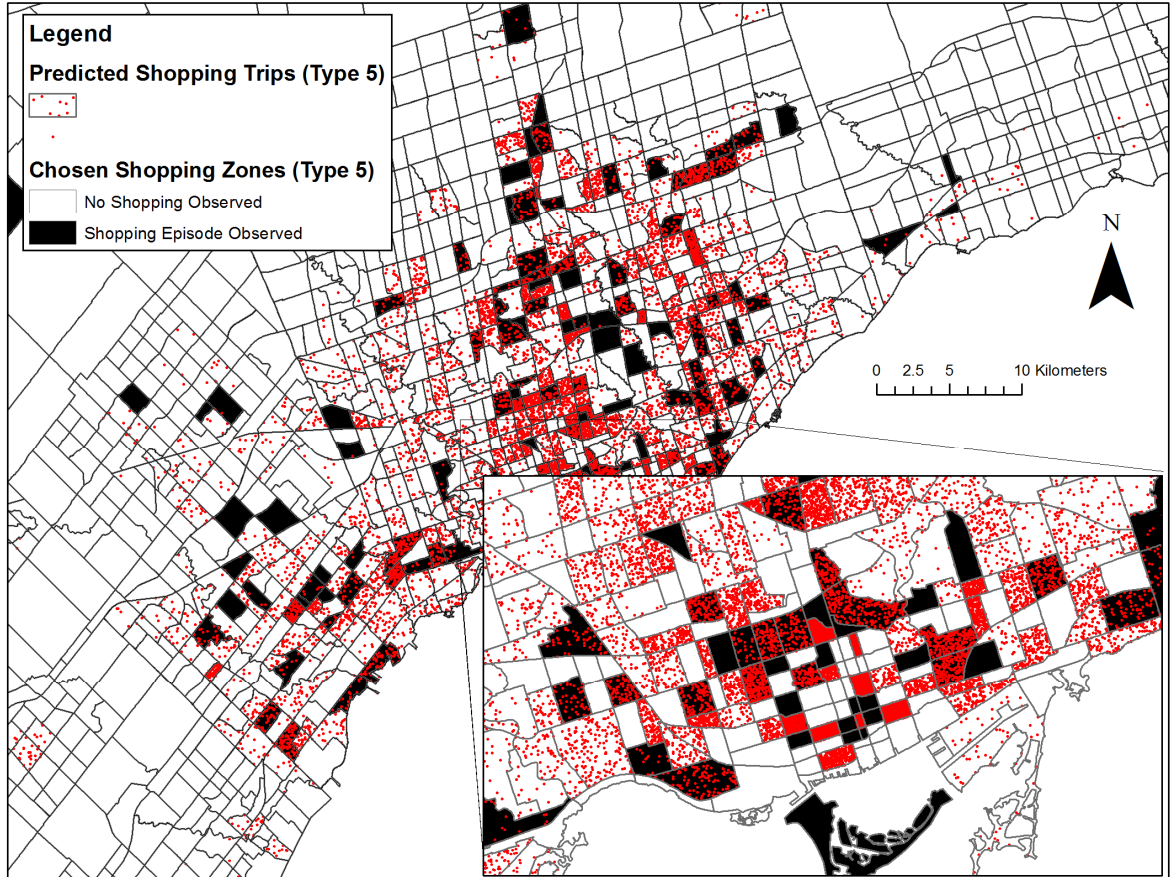
Figure E3.5 Spatial Distribution of Predicted Shopping Locations (Type 5)

Copyright Acknowledgements

A portion of this paper was adapted from a previously submitted paper:

Wang, J., E.J. Miller (2011). Choice Set Formation in Shopping Destination Choice. Presented at the *Computers in Urban Planning and Urban Management 2011 Conference*, Lake Louise, AB.