Understanding the Travel Behaviour of Families with Dependent Children Within the Context of Activity Based Modelling

by

James Lamers

A thesis submitted in conformity with the requirement for the degree of Master of Applied Science, Graduate Department of Civil Engineering, in the University of Toronto

© Copyright by James Lamers 2017
Understanding the Travel Behaviour of Families with Dependent Children Within the Context of Activity Based Modelling

James Lamers

Master of Applied Science

Graduate Department of Civil Engineering
University of Toronto

2017

Abstract

This thesis presents an investigation into several inter-related travel decisions made by families with children, a topic which has been generally overlooked by major travel demand models. The primary goal of this investigation was to develop an empirical understanding of these decisions which would allow the behaviour of families with children to be accurately represented in an activity scheduling model. The first of these decisions is the joint choice of whether a family will use an out-of-home childcare service, and the location of that service, and the thesis proposes an econometric model for this nested choice. A novel method for generating the daycare location choice set using a stochastic frontier model is proposed and tested. Finally, the effect of needing to drop off a child at daycare or school in the morning on an adult’s choice of travel mode and departure time is analyzed.
Acknowledgements
The process of writing these acknowledgements has been a humbling one. Reflecting upon the past few years of developing this thesis and thinking of everyone involved led me to realize how hopeless this endeavour would have been for me if I was working on it alone.

I will first thank my supervisor, Khandker Nurul Habib. Habib, your excellent advice and guidance – from technical to professional, and personal – have contributed tremendously to my growth throughout this degree. I can’t count the number of times I entered your office feeling lost, anxious, and stumped, and left feeling sure and confident about the project and myself.

My colleague and friend Adam Weiss’s technical expertise and creative ideas contributed to almost all the chapters of this thesis, and I’m very proud of everything we’ve created together. I thank all my classmates that contributed in their own way to making the lab such a welcoming and engaging place to work, notably Sami, Yunfei, Wafic, Adam R., Teddy, Mehdi, Kathryn, Julian, Chris, Moein, Soheil and Graeme.

Thank you to my family for your unquestioning love and support when at times this degree became too much for me to get through on my own. Thank you, Mom, Dad, Nona, Uncle Mike, Ray, Adam, and Mariah.

Finally, I acknowledge the importance of my friends over the past years, on whom I’ve relied in different ways to be a successful student and person. To everyone at Chestnut, the Oxford program, the Cobras, and particularly, Amanda, Dan, and Victoria, I am thankful for your invaluable combination of advice, criticism, support, confidence, laughs, and love.
Table of Contents

1 Introduction...................................................................................................................... 1

2 Literature Review ........................................................................................................... 5
   2.1 Travel Behaviour of Families with Children ............................................................. 5
   2.2 Location Choice Modelling ...................................................................................... 7

3 Conceptual Framework .................................................................................................. 10
   3.1 Research Motivation in the Context of Activity Based Modelling ......................... 10
   3.2 Daycare Utilization and Location Choice .................................................................. 13
      3.2.1 Stochastic Frontier Model of Daycare Location Choice Set Generation .......... 13
      3.2.2 Nested Model of Daycare Utilization and Location Choice .............................. 16
   3.3 Morning Mode and Departure Time Choice .............................................................. 21
   3.4 Overview .................................................................................................................. 22

4 Data.................................................................................................................................. 24
   4.1 Study Area ................................................................................................................ 24
   4.2 Transportation Tomorrow Survey .......................................................................... 24
   4.3 Traffic Assignment Model ...................................................................................... 28
   4.4 Enhanced Points of Interest .................................................................................... 28
   4.5 CanMap Land Use ................................................................................................. 30

5 Empirical Investigation of Daycare Utilization and Location Choice ...................... 31
   5.1 Stochastic Frontier Model of Daycare Location Choice Set .................................... 31
   5.2 Nested Logit Model of Daycare Utilization and Location Choice ............................ 32

6 Empirical Investigation of Morning Mode Choice and Departure Time ............... 37
   6.1 Mode choice ........................................................................................................... 38
   6.2 Departure time ....................................................................................................... 41
      6.2.1 TTS Observations ............................................................................................ 41
      6.2.2 Multinomial logit model .................................................................................... 44

7 Conclusions and Future Work ..................................................................................... 48

8 References .................................................................................................................... 50
List of Tables

Table 1 Stochastic Frontier Model of Daycare Location Choice Set Generation ................ 31

Table 2 Nested Logit Model of Daycare Utilization and Location Choice ....................... 33

Table 3 Multinomial Logit model of Morning Departure Time ..................................... 45
List of Figures

Figure 1 CUSTOM’s treatment of time-budget considerations ........................................... 11

Figure 2 Nesting structure of CUSTOM............................................................................ 12

Figure 3 Graphical representation of zones included in choice set using the stochastic frontier model........................................................................................................ 16

Figure 4 Nesting structure of the daycare utilization and location choice model.............. 18

Figure 5 Framework of the daycare utilization and location choice model...................... 19

Figure 6 Nesting structure when drop off task allocation is included in the model ........ 20

Figure 7 Choice sets when both drop off and pick up are considered.......................... 21

Figure 8 TTS Study Area.................................................................................................. 25

Figure 9 Employment density calculated from 2011 TTS.............................................. 27

Figure 10 Population density calculated from 2011 TTS.............................................. 27

Figure 12 Schools in the EPOI dataset ........................................................................... 29

Figure 13 Childcare location in the EPOI dataset............................................................ 29

Figure 14 Land use polygons in the study area................................................................. 30

Figure 15 Mode split for trips from home to work/school by adults with children........... 39

Figure 16 Mode split for work/school chains with a daycare drop off............................ 40

Figure 17 Mode split for work/school chains with a school drop off............................... 40

Figure 18 Departure time distribution for trips from home to work/school by adults with children .................................................................................................................. 43

Figure 19 Departure time distribution for work/school chains with a daycare drop off... 43

Figure 20 Departure time distribution for work/school chains with a school drop off..... 43
1 Introduction

It is widely viewed that activity-based travel demand models, which simulate the trips and activities of a population over the course of an entire day, are more behaviourally accurate than their trip-based and tour-based counterparts. This can be explained by the concept that people plan their days by scheduling activities, and the trips in-between are only a way to access those activists. And so, from a behavioural standpoint, it makes more sense to model a person’s choices related to the activities they want to conduct, and not the trips, which are only a means to access their activities.

Activity based models use many different frameworks, and the differences and advantages and disadvantages of each will be discussed in more detail in the following chapter of this thesis. However, they are similar in that they all seek to reproduce daily travel patterns, including those made by various modes, at all times of the day, to many different activity types, including work and non-work, in an effort to capture the complex nature of modern urban travel. As a result, the conceptual development and estimation of these models are much more complex than those of traditional four-step trip-based models.

A complete activity based model will simulate, for each agent, a complete list of all activities undertaken within a given time period, usually 24 hours. For each activity, the type, duration, and location must be modelled, along with eventually the access mode and route. The model would be estimated to fit a local travel survey that collects a complete travel diary from the respondent over the corresponding time period, whether over the phone, online, by mail, or in person. In the Greater Toronto and Hamilton Area (GTHA), most travel demand modelling is based on the Transportation Tomorrow Survey (TTS) (Ministry of Transportation of Ontario, 2016), which is a household-based travel survey conducted by landline for a 5% sample of the population every 5 years. It collects a full travel diary for all members of the respondent household over age 12 for the day before the survey is completed.
Each trip recorded in the TTS is classified as one of seven trip purposes: work, school, home, shopping, facilitate passenger, daycare, and other. While shopping and facilitate passenger trips happen with a high degree of randomness and are difficult to predict, work, and school trips tend to be regular and easy to predict on a day-to-day basis. The survey collects data about the location of the respondent’s work, school, and home location, and these locations are used as fixtures for any modeling investigation using the dataset. Daycare trips are somewhere in the middle of the spectrum for randomness and repetitively. If a respondent makes a daycare trip on survey day, it’s likely that they repeat that trip regularly, but the location is less permanent than work or school for most people.

In many travel demand models, work, school, and home activities are simulated to take place at a predetermined location, while the other activity types have some randomness in the location choice, consistent with scheduling behaviour. However, daycare location choice is a longer-term decision that is not made on a day-to-day basis, and should be treated as such in travel demand models. Whether a given household will even use a daycare service at all is a decision of similar timescale.

For many families, dropping off and picking up their children at daycare, school, and other activities comprises a significant portion of the daily travel of both working and non-working parents, and the scheduling of these trips can be a determining factor in the daily activity patterns of all household members. From a travel modelling perspective, to look at parents as agents that schedule their day based on only their own work, shopping, and other discretionary activities would be an oversimplification of the scheduling complexity caused by having dependent children in the household with their own activities but no means to transport themselves. However, many travel demand models either completely ignore or oversimplify this idea. The specific treatment of children in various contemporary activity-based models will be detailed in the subsequent chapter, but all of them fall short of simulating regular attendance at a daycare facility, which is the core daily activity for many children and the fixture for pick up and drop off activities for their parents.
The main goal of this thesis is to explore a conceptual framework that could provide a more behavioural representation of household decision making related to childcare and travel than what has previously appeared in the literature, and report on an empirical investigation of that framework in the GTHA. This research, while on its own is a study of a few specific household decisions related to childcare and travel, it meant to be a key building block in understanding daily activity scheduling, and implementing an operational activity based travel demand model that treats all activity types with behavioural realism.

At the University of Toronto, the CUSTOM (K. M. N. Habib, 2011) model is currently under development. It is an activity-based model that avoids the use of deterministic rules for activity generation and scheduling, and is based on a three-level nested structure to model the type, location, and duration of each activity of the day. More details on the background and implementation of CUSTOM will be given in the two subsequent chapters, including how the novel childcare travel research presented in this thesis fits into its overall structure. The motivation for this research arose through discussions of CUSTOM’s implementation, and a recognition that very little was understood about how to treat children within the framework of scheduling daily activities. After some exploration, it became clear that this lack of understanding was shared throughout the literature, and this gap needed to be addressed before feeling confident in the way the CUSTOM handles childcare pick up and drop off.

This thesis presents three contributions to the understanding of childcare-associated travel within the demand modelling context. The first is a novel approach to location choice set generation using a stochastic frontier model, a structure first published in 1977 and widely used as a way to assess productivity and efficiency of a firm (Aigner, Lovell, & Schmidt, 1977). More background of the stochastic frontier model will be presented in chapter 2, and details its framework for choice set generation and it’s empirical testing will be given in chapters 3 and 5, respectively. In all activity based models, location choice set generation is an important step, and all researchers aim to find an approach that balances computational complexity with a need to fit the data, while avoiding over fitting.
The methodology presented in this thesis has shown good performance in all of these criteria and has the potential for further development and implementation. In this research, the framework was tested on household childcare location choice set generation, but it demonstrates potential to be applied in various location choice models.

The second contribution is a nested logit model of daycare utilization and location choice. The model is representative of a household’s joint decision of whether they will bring their children to an out of home daycare, and where in the region they will do so. The outputs of such a model could be used as inputs to an activity based travel demand model, providing realistic spatial fixtures for the daily activity scheduling of parents with children. Although this model is the first to explicitly model daycare location choice, a broader discussion of the literature concerning daycare choices and travelling with children will be presented in chapter 2. The formulation of the nested model and its place within the context of activity scheduling will be detailed in chapter 3, and the procedure and results of a case study using the model in the GTHA will be presented in chapter 5.

The final contribution is an investigation of mode choice and departure time choices and how they are affected by the presence of children and the need facilitate the travel of children to school and daycare in the morning. While morning mode choice and departure time choices have been investigated frequently and with depth in the literature, these choices are less understood when children need to be dropped off, and activity scheduling models which ignore these effects may cause errors in forecasting. The procedure for, and results of this investigation are presented in chapter 6.
2 Literature Review

2.1 Travel Behaviour of Families with Children

Research on the topic of daycare activities in the field of activity-based modelling has been scarce. Some major operational models consider the activity schedules of children and their need for facilitated travel and supervision with varying degrees of complexity. However, none of them have explicitly considered the household need to bring their child to a regular daycare facility daily.

The simplest consideration of children is the approach taken by both FAMOS (R. M. Pendyala, 2004) and ALBATROSS (Arentze & Timmermans, 2004), which include the number of children as a variable in the models of adults’ activity choices. A somewhat more direct approach is taken by TASHA (Roorda, 2005), which notes if a child is being accompanied during a given activity episode. In addition, it is noted in the documentation that the drop off at daycare and pick up from day care can have an important effect on the activity schedule of adults, but this effect is not explicitly modelled.

ADAPTS (Auld, 2011), CEMDAP (Bhat, Guo, Srinivasan, & Sivakumar, 2008), the Jakarta Model (Yagi & Mohammadian, 2010), the Portland Model (Bowman, Bradley, Shiftan, Lawton, & Ben-Akiva, 1998), the New York Best Practice Model (Vovsha, Petersen, & Donnelly, 2002), SACSIM (Bradley, Bowman, & Griesenbeck, 2010), SimAGENT (Goulias et al., 2012), SimTRAVEL (R. Pendyala et al., 2012), etc. all model the activity schedules of children and how they affect the adults in the household, but none explicitly model daycare as an activity type. Although these models recognize the need for children to have facilitated travel and chaperoned activities, their attendance at a regular daycare facility, like a student’s school activity or a worker’s work activity is not modelled.

Outside of the major operational activity based-models, it has been shown in an investigation of non-work activity location choice that married people with children less than 16 years of age prefer to conduct shopping activities in the zones with daycare
facilities, presumably due to an effort to seek shopping locations close to their child’s daycare to minimize trip chain travel time (Sivakumar & Bhat, 2007).

Because the subject of household daycare choices has been neglected by travel demand modelling researchers, it is worth reviewing related literature in urban geography, social science, public policy, and operations research. (Holmes, Williams, & Brown, 1972) approached the issue from a system planning perspective and used the example of Columbus, Ohio to plan the location of new daycare facilities across the city to increase accessibility while being constrained by a limited budget. (Hodgson, 1981) used a similar approach to plan the locations of daycare facilities in Edmonton, Alberta, but recognized that there are three types of accessibility which might be desired by households for their daycare location: residential accessibility, workplace accessibility, and journey-to-work accessibility.

While falling short of doing any econometric modelling of household daycare choices, social scientists have reported valuable insights until the factors involved in such decisions. It has been found that families with higher incomes (Davis & Connelly, 2005) (Hofferth & Wissoker, 1992) and parents with higher education (Hofferth, Chaplin, Wissoker, & Robins, 1996) are more likely to place their children in centre-based care over other modes such as parental care or relative care (with a non-parent family member). In addition, households with teenagers are less likely to use centre-based care (Davis & Connelly, 2005). It has also been shown that ethnicity is a significant determinant of a household’s use of centre care (Early & Burchinal, 2002) (Radey & Brewster, 2007). The distinction between centre-based and parental-based or relative-based care is significant for travel demand modelling because in contrast with centre-based care, parental-based is not likely to require any travel. For these reasons, it is possible that any policies which encourage parental care may be beneficial from a travel demand management perspective.
2.2 Location Choice Modelling

None of the above-mentioned studies investigated the spatial location choice issues involved in out-of-home daycare location choices by the households. However, it can be speculated that availability and/or difficulties of choosing a daycare location that is feasible within the daily activity-travel context of working parents may preclude the choice of daycare utilization. So, it is important to consider the spatial location choice within the context of daycare utilization choice. In absence of any empirical evidence on this issue, we now focus on reviewing literature related to discrete choice application in spatial location choice, in general.

The conditional logit formulation introduced by (McFadden, 1973) provides the framework for understanding how locational characteristics affect spatial behaviour. When researchers use this framework to model location choice, the most naïve approach is to assume that the decision maker considers all choices in the universal choice set. However, this practice has the potential to introduce bias into the model, and when random sampling of alternative from the universal choice set is used to reduce computational complexity, overfitting is likely.

In some studies of destination choice, researchers have proposed criteria defined by distance or travel time for generation choice sets in situations where the universal choice set is too large to be modelled as the consideration set (Scott, 2006) (Termansen, McClean, & Skov-Petersen, 2004) (Hicks & Strand, 2000). This deterministic approach is based on the recognition that individuals will not consider extremely distant destinations for choices in which feasible alternatives are nearby. The deterministic set generation methodology is attractive because it is easy to implement and is seemingly behavioural accurate: it is intuitive that individuals do not collect information about alternative destinations which are distant from their origin beyond some threshold.

However, further examination of the deterministic approach reveals some shortcomings. The first is that the coefficients for the destination choice model are sensitive to the time or distance threshold selected by the researchers for the choice set (Termansen et al.,
Thus, the success of this methodology is dependent on the judgement of the investigators in choosing an appropriate threshold. In addition, the size of the consideration set varies across individuals, dependent on their available travel modes and time budget, and these characteristics should be considered in order to generate an accurate choice set. One alternative approach was proposed by (Rashidi, Auld, & Mohammadian, 2012), who used a hazard model to determine a travel time threshold. While the approach employed within this work is intuitive, we argue that the subsequent model formulation is a more appropriate than the hazard model for the determination of a travel time threshold.

To account for the uncertainty in determining a decision maker’s choice set, researchers have developed econometric structures that use a stochastically generated choice set. There have been many applications in different forms of the two-stage choice framework adopted by (Manski, 1977), in which the choice set is first explicitly modelled using characteristics of the decision maker and the transportation network, and location choice is then modelled from that set. The first stage, choice set generation, has taken many different forms in the literature within and outside of spatial decision making, including capturing travel mode captivity (Swait & Ben-Akiva, 1987), the limited ability of individuals to gather information about many alternatives (Meyer, 1980), or as a distance threshold with a probabilistic function (Zheng & Guo, 2008).

The stochastic frontier model was originally formulated by (Aigner et al., 1977) as a way to model the maximum productivity of a firm, which is dependent on measurable inputs such as the characteristics of its employees and technology, but also subject to random effects such as climate and equipment breakdowns. The stochastic frontier is suitable for modelling transportation behaviour because the costly trip inputs of travel time and cost which result in utility-generating outputs are analogous to the resource costs required to create products in a production setting. However, the application of stochastic frontier models in transportation research is limited.

In (Kitamura, Yamamoto, Kishizawa, & Pendyala, 2000), (R. M. Pendyala, Yamamoto, & Kitamura, 2002), and (Yamamoto, Kitamura, & Pendyala, 2004) a stochastic frontier
model was used to model the location of origin and terminal vertices of space-time prisms on the time axis based on socioeconomic attributes and household structures. The application of stochastic frontiers in this setting is appropriate, because although analysts can observe the actual times of departure from home in the morning and returning home in the evening, the true original and terminal vertex of the space-time prism is unknown. Any deviation of the trip start and end time from the true original and terminal vertices can be thought of as inefficiencies in individual’s use of time.

Finally, (Scrogin, Hofler, Boyle, & Milon, 2004) used a stochastic frontier to model location choice set generation for recreational fishing trips in a two-stage model of location choice. Fishing sites were rejected from an individual’s choice set if they did not meet an acceptable level of efficiency based on minimizing travel time and maximizing catch potential at the site.
3 Conceptual Framework

This chapter introduces the CUSTOM framework which is the activity based platform on which this thesis is based. The novel concepts in daycare and school drop off behaviour, which are the core of this thesis are then detailed, including some theoretical structures for which no empirical work has yet been done. Finally, these conceptual frameworks are then tied together in one unified structure.

3.1 Research Motivation in the Context of Activity Based Modelling

As discussed in the introductory chapter, the research presented in this thesis related to childcare travel planning and household decision-making was developed with the intention of producing results that can serve as inputs for an activity-scheduling model. Specifically, the models that will be described in subsequent sections were designed around the framework of the Comprehensive Utility System of Travel Options Modelling (CUSTOM), which is being developed at the University of Toronto.

The reader is referred to (K. M. N. Habib, 2011) and (K. Habib, 2015) for a detailed description of CUSTOM, but an overview is given here so that the context in which the rest of the thesis was developed can be properly understood. CUSTOM uses a random utility maximization approach to model the activity scheduling process, including the types of activities that are conducted, their location, and how much time is allocated to each one. Time is considered to be a continuous, and the duration of each activity is chosen sequentially throughout the day while considering the remaining time budget until the end of the day. Furthermore, the choices for activity type and location are dependent on the remaining daily time budget, taking into account the endogeneity between the choice of activity type and time expenditure.

The CUSTOM framework considers that individuals will schedule their activities throughout a 24-hour time period sequentially, starting the choice of when to leave home in the morning. Following that, the type, location, and duration of each activity will be chosen to maximize the utility of said activity, considering the diminishing time budget
as the day progresses. At any point, the time budget is the total amount of time in a day minus the total time expended on previously performed activities. The sequential process is depicted in Figure 1 and Figure 2. The activity type choice is based on various household and personal socioeconomic characteristic of the individual, in addition to the time of day that activity is taking place, and the nature of the activities that have been performed earlier in the day. The location choice is based on maximizing utility while taking into account land use characteristics of the city, transportation level of service, and the time available for travel based on remaining daily time budget.

Figure 1 CUSTOM's treatment of time-budget considerations (K. M. N. Habib, 2011)
When CUSTOM is used to simulate a daily activity schedule, the locations of shopping and other discretionary activities are chosen randomly based on available land use and transportation level of service data, in addition to the activities that have occurred earlier in the day and the remaining time budget. This is representative of the choice process used in reality with which we decide on many of our shorter activities of the day. This is unlike the choice of home, work, and school activity locations, which are inputs into the model, whether taken directly from the survey sample or from a population synthesis, representing the long-term choice of home, work, or school location, certainly not a part of everyday decision making and planning.

Details about the data used in the development of this model will be given in chapter 4, but a critical point is discussed here because one feature of the local travel survey motivated the development of the models presented in the subsequent sections. The Transportation Tomorrow Survey, which is the major travel survey in the GTHA, would
be the basis for a local implementation of CUSTOM. Each trip that makes up the TTS dataset is classified as one of seven purposes: home, school, work, shopping, facilitate passenger, daycare, and other. For each respondent, the survey collects the location their home, in addition to work and/or school, if applicable, and these locations could be directly fed into any simulation or forecasts using CUSTOM. And as mentioned above, the locations of shopping and other trips would be simulated randomly within the CUSTOM framework. The trip purpose that had not yet been assessed with much rigor was daycare.

It is clear that like the choice of home location, daycare location is not a decision made by households on a day-to-day basis. A family’s choice of whether or not they will bring their child to a daycare centre and if so, where they will do that is a long term one, although usually less permanent than the choice of work, school, and home. The decision is often made with the characteristics of the adult household members’ home and work taken into consideration. Thus, the framework of CUSTOM, or any other major activity based model is not adequate for representing a household’s decisions related to childcare and how those decisions affect their daily travel patterns. The research presented in the remainder of this thesis seeks to improve the understanding of these processes and guide their implementation in activity scheduling models, with the CUSTOM framework in mind.

3.2 Daycare Utilization and Location Choice

3.2.1 Stochastic Frontier Model of Daycare Location Choice Set Generation

As discussed in chapter 2, using the correct choice set when modelling location choice behaviour is an important step to achieving behaviourally accurate results. In this project, a novel choice set generation approach was implemented for modelling the choice of daycare location. When considering how to narrow down the consideration set for daycare choice from the entire study area, the behaviour and planning strategies of families with children was considered, despite the lack of data about these strategies.
It was assumed that for families planning their daily activities and travel schedule, their home and location would be considered fixed, and all other decisions, including the choice of daycare location, would be made with respect to these fixed locations as constraints. More specifically, the choice set generation method which is used in this research reflects the idea that commuters will bring their child to daycare in a place which is convenient for their home-work commute. While most location modelling frameworks would include this strategy as a parameter in the location choice model, making it more likely for an agent to choose a location closer to home and work, it’s important that it’s used in the choice set stage to reflect the reality of the decision. The reality is that not only would a daycare location which is across the entire study area from a respondent’s home and work commute be undesirable and an improbable choice, it would be a completely unfeasible choice and not even considered due to the scheduling constraints associated with travelling across the entire study area for pickup and drop off.

A stochastic frontier approach was adopted to reflect this planning behaviour. As discussed in chapter 2, the stochastic frontier model has been used to model efficiency maximization, where the maximum efficiency can be calculated based on the observable characteristics of an individual or firm, but their actual efficiency is subject to various unobservable and random factors. These random effects can be categorized into those that affect the production strictly negatively, and random effects that may affect the production either positively or negatively. The analog in this context is that, using travel survey data, the maximum allowable travel time for the home-daycare-work work can be estimated as a function of observable household and personal characteristics, with the actual commute time being a function of this maximum and some random effects. Similarly to the original intent of the stochastic frontier model, some random effects such as a commuter’s laziness or the reliability of their vehicle will affect their maximum travel time strictly negatively, other random effects may be positive or negative, such as the commuter’s ability to gather information about daycare locations.
The SFM approach simply states that the maximum possible production of a good (or minimum expenditure) for a firm or individual is a latent variable. As such the model can be formulated as follows: (Aigner et al., 1977)

\[ y = \beta'x + v - u \]

Where: \( y \) is the observed output (in this case travel time from the individuals home to the daycare and then to work), \( \beta'x \) is a set of estimated parameters including a constant times a set of independent variables, \( v \) is a stochastic normally distributed error component with mean zero and unknown variance \( \sigma_v^2 \) and \( u \) is a positive inefficiency component, which is commonly treated as the absolute value of a normal distribution with mean zero and variance \( \sigma_u^2 \). It should also be noted that the literature referenced in chapter 2 typically advises the use of a logarithmic model whereby the dependent and continuous independent variables are logged.

Through simple arithmetic and noting that the expected value of \( v \) is zero, the above equation can be rearranged to show that the linear in parameter systematic component \( \exp(\beta'x) \) is a prediction of the maximum an individual would be willing to travel to make a daycare trip before continuing to work. This model can then be applied to each respondent in the sample to generate a feasible set of locations given the home and work location of each individual. This process is represented graphically in Figure 3.
16

In the hypothetical example demonstrated in Figure 3, the three zones shown are included in the data set because the travel time from the respondent’s home to that zone in addition to the travel time from that zone to the respondent’s work or school location is less than the estimated maximum travel time. A spatial frontier can be drawn as an ellipse with the respondent’s home and work or school locations at the focal points, and the distance from home to the curve and then to work is the estimated maximum travel time.

3.2.2 Nested Model of Daycare Utilization and Location Choice

Given the highly plausible hypothesis that the choice not to use daycare is not correlated with the choice of a given location for a daycare trip, a nested logit formulation is proposed with an upper level choice of using daycare or not and a lower level choice of
location conditional on the choice to use daycare. As a result, the model structure can be thought of as using the usual nested logit formulation, constructed as follows:

\[ P_{\text{Location } M | \text{Use}} = \frac{\exp(\mu_{\text{use}} V_m)}{\sum_{n \in C} \exp(\mu_{\text{use}} V_n)} \]

\[ P_{\text{Use}} = \frac{\exp(V_{\text{use}} + (I V_{\text{locations}})/\mu_{\text{use}})}{\exp(V_{\text{use}} + (I V_{\text{locations}})/\mu_{\text{use}}) + \exp(V_{\text{no daycare}})} \]

\[ I V_{\text{locations}} = \ln \left( q \sum_{n \in C} \exp(\mu_{\text{use}} V_n) \right) \]

Where \( q \) is a sampling correction factor to account for the sampling of locations. This factor is calculated simply as the greater value between the ratio of the number of alternatives in the consideration choice set generated by the SFM approach divided by the number of alternatives sampled and 1 (as is the case when the consideration set size is less than the number of alternatives sampled). The original intent of this method was to reduce the choice set to one that is computationally manageable. However, in the local context, some respondent’s home and work location were so far from each other, that the size of their choice sets made model estimation prohibitively time consuming, and so the sampling correction was introduced into the structure. The nesting structure is depicted graphically in Figure 4.
In any empirical application, it is necessary to generate the choice set for daycare location before this nested model can be estimated. The choice set is necessary so that the logsum of the lower location choice nest can be computed and used as an independent variable for the daycare utilization choice. This reflects the reality that a household’s decision of whether to use daycare may be affected by the quality of location choices available. The flow of information between these two models is shown in Figure 5.
Figure 5 Framework of the daycare utilization and location choice model

The two aforementioned models have been tested empirically and the results of that investigation are presented in chapter 5. One part of the proposed conceptual framework which has not been tested is the decision of which adult in the household will make the drop off. The incorporation of that decision within the existing nesting structure is shown in Figure 6.
This structure has the potential to be expanded to any number of adults to include multi-generational households, households with adult children, or any configuration of multi-adult households. Importantly, the location choice set for the household can vary greatly depending on which adult is making the strip, since the frontier is based on their work or school location. In the empirical investigation performed for this thesis, the task allocation decision was considered to be exogenous to the utilization and location choice decision, and the adult that made the drop off in the survey was used for estimation.

Another component of the decision is that of which adult will make the daycare pick up in the afternoon, since it is always not the same adult that made the morning drop off. While this thesis falls short of creating a modelling structure for this complex decision, Figure 7 gives a graphical depiction of how the choice set generation process described earlier in this section changes when pick up allocation is introduced.
3.3 Morning Mode and Departure Time Choice

The standard structure of CUSTOM, as shown earlier in this chapter, simulates activity start time and duration, in addition to location and trip mode in a standard way for the whole population. In chapter 6 of this thesis, I show that in the local context, whether an adult needs to drop off a child at daycare or school during their morning work commute has a significant effect on their departure time from home and their mode choice. Knowing this, any simulation of CUSTOM which does not explicitly consider the need to make these facilitate passenger trips will introduce error into the forecast for families with children.

To fix this issue, it’s not necessary to introduce any changes to CUSTOM’s daily modelling structure, but rather to run task allocation models for daycare and school drop off and pick up, as described in the previous section, and using these allocations as inputs to the existing nested models. Specifically, for each household with children aged 17 and
younger, models for daycare and school drop off and pick up must be run. The result of these models is to assign flags to some adults indicating whether one of these tasks has been assigned to them.

Beyond the assignment of these flags to the agents in the model, the existing CUSTOM framework for mode and departure time choice do not need to be modified. These flags simply need to be included within the model as independent variables affecting the outcome of the decisions. It should also be noted that the CUSTOM framework explicitly models activity duration, and departure time from these activities is determined from this duration.

3.4 Overview

All the concepts in this chapter can be combined into one modelling system which results in 24-hour activity schedules, with accurate representation of the facilitation of children’s trips. This system is described below:

- For all adults living in a household with children age 12 and under in the survey sample (or synthetic population), stochastic frontiers for daycare location are generated based on the home and work/school location of the adult. For adults that are neither students or employed, the frontier is a circle around their home.
- For each household, the nested daycare utilization, task allocation, and location choice model is run. The output of this is a variable for each household indicating whether they will use daycare, and if so, which adult makes the drop off and pick up, and the location of that daycare.
- For each household with children age 5-17, a school pick up and drop off model will be run to determine whether an adult will facilitate the child’s trip to and from school, and if so, which adult.
- The regular CUSTOM models can then be run, as described in (K. M. N. Habib, 2011) with the following key differences:
  - For any adult with one of these flags, the choice set for their activity type choice models (already existing in the CUSTOM framework) will be
modified to include one of the following, whichever is applicable, as determined by the task allocation model: daycare drop off, daycare pick up, school drop off, and school pick up.

- The daycare drop off and pick up activities are set to take place in the location predicted from the estimated daycare location choice model, instead of using CUSTOM’s activity location choice model. Similarly, school drop off and pick ups take place at the location specified in the survey (or population synthesis).
- Both the trip mode choice and activity duration models include dummy variables for whether the adult facilitates a child’s trip that day.
4 Data

4.1 Study Area

The study area for the empirical application of this thesis is the GTHA and surrounding area, including the City of Toronto, the regional municipalities of Durham, York, Peel, and Hamilton, Peterborough County, City of Peterborough, Victoria County, Town of Orangeville, Simcoe County, City of Barrie, Wellington County, City of Guelph, Waterloo Region, Niagara Region, Dufferin County, City of Orillia, Northumberland County, City of Brantford, and Brant County. To give the reader some context, these municipalities are all located in proximity to each other in southern Ontario, with the City of Toronto generally at the centre, as shown in Figure 8 TTS Study Area (Ministry of Transportation of Ontario, 2016). The population of the entire study area is 8,515,750 as measured by the 2011 Canadian census. (Data Management Group, 2013) It is also the Canada’s largest urban area (Metrolinx, 2012). The area is served by a network of transportation corridors with emphasis on serving Downtown Toronto, including regional rail radiating outwards from downtown, a subway system that exists entirely within the City of Toronto and a system of controlled-access expressways, many of which lead to Toronto.

4.2 Transportation Tomorrow Survey

The Transportation Tomorrow Survey (TTS) is a major travel survey conducted every five years in the GTHA since 1986(Data Management Group, 2013). While the specifics of the survey have varied from year to year, the research for this thesis is based on the 2011 survey, and this section will familiarize readers with that survey in some detail, since the TTS is the foundation of data that the research is based on. As such, the extent and limitations of the survey data are what guide the conceptual framework given in the previous chapter into empirical testing.

The 2011 survey included all the municipalities listed in section 4.1. These municipalities, in addition to the Ministry of Transportation of Ontario and local and regional transit operators jointly fund the survey.
The survey sample, 5% of the population, is reached by landline telephone. The bias that this may introduce due to systematically missing certain populations that do not have landlines has been addressed for future TTS surveys, but it remained the sole method of contact in 2011. When a respondent answers the phone and volunteers to participate in the survey, they are asked questions about their household, each person living there, and finally to provide a full travel diary for each member of the household aged 11 and older for the day before the survey is conducted. In the travel diary, details about each trip made throughout the day is collected, including the start time, mode of travel, origin and destination locations, and the activity taking place at the destination (purpose of the trip). Some households are given the option of completing the survey online using unique access codes provided to them. Further details about the data that is collected in the
survey will be given in chapters 5 and 6 when presenting the empirical application of the models using TTS data.

The TTS data is expanded to represent the total population of the survey area. The sample data was expanded within each forward sortation area (FSA) based on comparisons to the population in the 2011 Canadian Census, and adjusted to ensure that the age distributions in each municipality match the census. These expansion factors were validated and the reported difference in total population between the 2011 TTS and the Census is 0.1%. These expanded counts are also used in the empirical studies described in chapters 5 and 6. The 2011 survey recorded responses from 159,157 households, 410,404 people, 772,1245 non-transit trips, and 86,704 transit trips (Data Management Group, 2013).

Figure 9 and Figure 10 show the population and employment distribution around the study area. It’s clear that the population is most dense in downtown Toronto, with the density generally decreasing until the area is completely rural at the edges. Employment in the area is more decentralized. While the density of jobs is in fact highest in the financial district of downtown Toronto, this only represents a small number of zones, and so the red is not very visible at this map scale. Major employment centres are visible near the airport, downtown Kitchener, Hamilton and, St Catherine’s, all major commercial or industrial areas.
Figure 9 Employment density calculated from 2011 TTS

Figure 10 Population density calculated from 2011 TTS
4.3 Traffic Assignment Model

In order to define the alternatives daycare locations available to each household, the region is decomposed into a set of TAZs. This deconstruction is done in order to facilitate the development of a calibrated traffic assignment model (Travel Modelling Group, 2015), which is used by numerous municipalities and transit agencies within the region. This model is used to produce both auto and transit travel time and travel cost matrices between each of the 2298 TAZs defined within the region. The time and cost matrices generated by the assignment model can then be fed into the location choice model as alternative specific variables with generic coefficients.

The area covered by the traffic assignment model is more limited than the full TTS area. Of the 3257 households from TTS that used daycare, 2484 have all of their home, daycare, work and school locations within the traffic assignment model area. Consequently, all others were removed from the analysis.

4.4 Enhanced Points of Interest

The 2013.3 version of the Enhanced Points of Interest (EPOI) file produced by DMTI spatial was used to identify the locations of daycare facilities and schools in the study area(DMTI Spatial, 2013). This data set is a national database of over 1 million geocoded business and recreational points of interest. Within the study area, the EPOI dataset gives the location of 4296 childcare service locations and 7885 schools. We can expect that the point data for schools is very accurate, as these locations are all either accredited, or operated, by local governments and maintained in databases. The childcare dataset, however, can be expected to be missing some locations, especially some more informal ones based in homes. While this is a limitation, since the data collection and section is consistent throughout the study area, the distribution of locations throughout the area should be accurate, and that’s the important part since the empirical analysis is done at the zonal level, and only the relative differences between the zones is important.
Figure 11 Schools in the EPOI dataset

Figure 12 Childcare location in the EPOI dataset
4.5 CanMap Land Use

Land use data was taken from the 2014.2 version of the CanMap RouteLogistics (DMTI Spatial, 2014) comprehensive spatial data package. The package contains various geographic data from expressway casements to golf courses, but this research used the land use file, which split the entire country into land use polygons which are all one of the following type: government and institutional, open area, parks and recreational, residential, resource and industrial, and waterbody.

Figure 13 Land use polygons in the study area
5 Empirical Investigation of Daycare Utilization and Location Choice

5.1 Stochastic Frontier Model of Daycare Location Choice Set

The empirical application of the model described in section 3.2.1 is based on the TTS data described in section 4.2, using travel times from the traffic assignment model described in 4.3. Specifically, the estimation of the stochastic frontier model of daycare location choice is based on the survey records of all households with an adult that made a daycare trip. For simplicity at this stage, the estimation was restricted to adults that are either workers or students, and made a trip in the AM period.

The results of parameter estimation of the daycare location choice set generation model are shown in Table 1 Stochastic Frontier Model of Daycare Location Choice Set Generation, including the set of $\beta$ parameters in addition to the variances $\sigma^2_v$ and $\sigma^2_u$ which define the inefficiency distribution of the stochastic frontier. The number of records used for model estimation is 2484, the number of valid households using daycare.

### Table 1 Stochastic Frontier Model of Daycare Location Choice Set Generation

<table>
<thead>
<tr>
<th>Logarithm of maximum home-daycare-work morning travel time (minutes)</th>
<th>$\sigma^2_v = 0.1195$</th>
<th>$\sigma^2_u = 0.6253$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-likelihood</td>
<td>1791.29</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood when $\sigma_u = 0$</td>
<td>1855.78</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables of Deterministic Component</th>
<th>Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.6021</td>
<td>-52.8280</td>
</tr>
<tr>
<td>Logarithm of the number of vehicles in household</td>
<td>0.3200</td>
<td>-6.4092</td>
</tr>
<tr>
<td>Dummy, employed in professional sector</td>
<td>-0.0675</td>
<td>31.2927</td>
</tr>
<tr>
<td>Dummy, house dwelling type</td>
<td>0.1029</td>
<td>7.4778</td>
</tr>
<tr>
<td>Dummy, commute to the City of Toronto from outer suburbs</td>
<td>0.2138</td>
<td>-1.5784</td>
</tr>
<tr>
<td>Dummy, commute to planning district 1 from outer suburbs</td>
<td>0.6174</td>
<td>11.9395</td>
</tr>
<tr>
<td>Dummy, live and work in planning district 1</td>
<td>-1.4534</td>
<td>-2.3337</td>
</tr>
<tr>
<td>Dummy, live and work in City of Toronto</td>
<td>-0.5385</td>
<td>1.6130</td>
</tr>
<tr>
<td>Dummy, live and work in the same region outside of the City of Toronto</td>
<td>-0.5383</td>
<td>0.3273</td>
</tr>
</tbody>
</table>
The magnitude and sign of the $\beta$ parameters in Table 1 indicate the size of the $\nu$ and $\mu$ random error distributions, in addition to the effect of the respective variables on the maximum allowable home-daycare-work travel times for each respondent.

As shown by the table, households with a greater number of vehicles are willing to consider longer travel times to daycare and work, consistent with the increased flexibility offered by more cars. Parents employed in the professional sector are less likely to travel further to drop their children in daycare. This may be caused by the urban lifestyle of many professional workers, or reduced time availability due to longer working hours. It may also be indicative of something more complex, such as more choices for daycare location for parents with the relatively high incomes of parents in the professional sector.

Households living in detached homes are more likely to travel further for their daycare trip chain, likely due to the decreased density and longer travel times generally associated with suburban living. This may be a cause for concern because not only do suburban residents face longer home-work commutes than, but suburban parents that need to drop off and pickup their children in daycare face even further increases in everyday travel time than their urban counterparts.

The rest of the variables demonstrate the effect of commuting patterns on the size of a household’s daycare location choice set. The size of the choice set is relatively consistent with the length of a parent’s commute. Commuters going from an outer suburb all the way to downtown consider the largest number of zones as potential daycare locations, and those living and working in downtown consider the fewest. This is consistent with the graphical interpretation of the daycare location choice set being an ellipse with the adult’s work and home locations at the focal points, and when the two focal points have the highest travel times between them, the ellipse is largest.

### 5.2 Nested Logit Model of Daycare Utilization and Location Choice

The results of parameter estimation of the nested choice model of daycare utilization and location choice are shown in Table 2. The number of records used for model estimation is
23872, which is the total number of households in the 2011 TTS that have a child. 16 coefficients are alternative specific and are associated with the affirmative choice for household daycare use, and the final 7 coefficients are associated with attributes of TAZs. For location choice, a random sample of 20 feasible locations defined by stochastic frontier model was used for estimating the location choice model component. All parameters exhibit intuitive or reasonable signs and most are significant at the 95% confidence level. In addition, a McFadden’s R-squared of 0.548, which indicates a good fit to observed data.

**Table 2 Nested Logit Model of Daycare Utilization and Location Choice**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Daycare is used</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-8.7697</td>
<td>-22.604</td>
</tr>
<tr>
<td>Number of adults in the household</td>
<td>-0.3917</td>
<td>-7.778</td>
</tr>
<tr>
<td>Number of children age 0 to 5</td>
<td>0.5597</td>
<td>10.87</td>
</tr>
<tr>
<td>Number of children age 6 to 12</td>
<td>-0.7648</td>
<td>-15.556</td>
</tr>
<tr>
<td>Number of children age 13 to 17</td>
<td>-0.8839</td>
<td>-9.296</td>
</tr>
<tr>
<td>Logarithm of the number of household vehicles</td>
<td>0.7662</td>
<td>6.463</td>
</tr>
<tr>
<td>Dummy, apartment dwelling type</td>
<td>-0.2270</td>
<td>-2.155</td>
</tr>
<tr>
<td>Dummy, household is in planning district 1</td>
<td>-2.7094</td>
<td>-7.128</td>
</tr>
<tr>
<td>Dummy, household is in the City of Hamilton</td>
<td>-0.8388</td>
<td>-7.699</td>
</tr>
<tr>
<td>Dummy, household is in City of Toronto</td>
<td>-0.6757</td>
<td>-7.836</td>
</tr>
<tr>
<td>Dummy, household is in the Region of Peel</td>
<td>-0.4872</td>
<td>-5.851</td>
</tr>
<tr>
<td>Dummy, an unemployed adult lives in household</td>
<td>-1.2330</td>
<td>-12.261</td>
</tr>
<tr>
<td>Dummy, an adult part-time worker lives in household</td>
<td>-0.5609</td>
<td>-6.382</td>
</tr>
<tr>
<td>Dummy, an adult working in manufacturing lives in household</td>
<td>-0.4275</td>
<td>-4.864</td>
</tr>
<tr>
<td>Dummy, an adult working in retail lives in household</td>
<td>-0.2537</td>
<td>-3.963</td>
</tr>
<tr>
<td>Scale of daycare use nest*</td>
<td>1.4774</td>
<td>6.981</td>
</tr>
<tr>
<td><strong>Location Choice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of jobs in zone</td>
<td>1.0074</td>
<td>30.196</td>
</tr>
<tr>
<td>Number of residents in zone</td>
<td>0.2407</td>
<td>17.242</td>
</tr>
<tr>
<td>Number of schools in zone</td>
<td>0.6706</td>
<td>12.452</td>
</tr>
<tr>
<td>Number of childcare facilities in zone</td>
<td>1.1826</td>
<td>12.432</td>
</tr>
<tr>
<td>Number of shopping trips destined to zone</td>
<td>-1.2946</td>
<td>-17.616</td>
</tr>
<tr>
<td>Logarithm of home-daycare-work travel time (minutes)</td>
<td>-1.3301</td>
<td>-12.087</td>
</tr>
<tr>
<td>Logarithm of home-daycare-work travel cost (CAD)</td>
<td>-0.3398</td>
<td>-2.464</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-4956</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood (null model)</td>
<td>-71514</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood (constant-only model)</td>
<td>-15890</td>
<td></td>
</tr>
<tr>
<td>McFadden R-squared (null model)</td>
<td>0.930</td>
<td></td>
</tr>
<tr>
<td>McFadden R-squared (constant-only model)</td>
<td>0.548</td>
<td></td>
</tr>
</tbody>
</table>

*Note that the scale value here is greater than 1. This is because we use the inverse notation of what many textbooks will use.*
The estimated scale parameter of location choice nest is significantly higher than 1 that indicates the daycare location choices are more correlated than the two alternative discrete choices of choosing or not choosing an out-of-home daycare utilization. The resulting approximate correlations among alternative location choices are around 0.30. This imply the fact that spatial location choice have a strong influence on the choice of an out-of-home daycare utilization choice. In addition to this, the estimated parameters/coefficients of the variables in the model reveal many behavioural details of daycare utilization and location choice.

First, a negative coefficient for the number of adults people in a household suggests that the need for daycare is decreased, as there are more people at home that might be able to care for children. The coefficient for the number preschool aged children is positive, which is expected. Interestingly, households with teenagers are less likely to use daycare, which may also indicate that teenagers can assist in the care of small children before and after school hours, so that the use of before-and-after school childcare services can be avoided.

More vehicles in a household is associated with increased daycare use. It is hypothesized that the number of vehicles in a household is a proxy for household income, which is consistent with the findings of (Hofferth & Wissoker, 1992) and (Davis & Connelly, 2005) that higher income families are more likely to use centre-based care, compared to home-based or family-based.

Interestingly, households living in apartments are less likely to use daycare. It is hypothesized that apartment dwellings encourage home-based childcare among neighbours. It has been shown in the literature that if a friend or neighbour is available, parents are less likely to make use of centre-based care (17), and such an arrangement seems more likely to occur in a setting where more families live in close proximity. It is also possibility that once again, this variable is a proxy for lower income, which in turn causes decreased daycare use.
Households in planning district one (downtown Toronto), the City of Hamilton, the rest of the City of Toronto, and the Region of Peel are less likely to keep the children in daycare than in the rest of the GTHA, with the largest effect observed in planning district one.

As expected, households with an unemployed adult are less likely to use daycare, since that adult is more likely to have time to care for children. The same effect is observed in households with at least one part-time employee, but to a lesser magnitude.

Finally, households in which at least one adult works in the manufacturing or retail sectors have less tendency to use daycare. It is hypothesized that unlike the other TTS occupation types (office/clerical, professional/management/technical), these employees are more likely to work shifts, which means that they may be available to care for children during daytime hours, especially in households with multiple adults when the shifts don’t overlap, meaning at least one adult is always available to supervise the children.

In the lower nest of location choice, households are more likely to put their children in a daycare facility in a TAZ with a higher number of jobs, residents and schools. This makes sense because childcare facilities are more likely to locate in proximity to jobs and residential areas because these are the start and end of daycare trip chains that adults with children are making. Similarly, daycares should ideally be located conveniently close to schools if they offer before and after school care services. Obviously, households are more likely to choose a daycare centre in a zone with a higher number of daycare facilities.

Interestingly, the number of shopping trips destined for each zone, which is our proxy variable for retail density, negatively influences daycare location choice. Although we might expect parents to select daycare locations in a way which would increase the convenience of trip chaining with shopping activities, in our dataset it is possible that a high proportion of these shopping trips take place in shopping malls and suburban retail centres, which are not desirable locations for a daycare centre. The coefficients for
morning home-daycare-work travel time and cost are negative, which is to be expected, indicating that respondents seek to minimize their travel time and cost.
6 Empirical Investigation of Morning Mode Choice and Departure Time

In this chapter, the travel behaviour of adults that drop off their child at school or daycare in the morning is compared to that of adults with children in their household, but who do not make a drop off. This is done to understand how the need to make these drop offs affect travel decisions independently of other differences between commuters.

Preparing the data for this exercise was a complex process. For the models in the previous chapter, querying daycare trips from the TTS dataset was relatively straightforward, since “daycare” is one the trip purposes from the survey. This is unlike trips for making school drop offs, which have no special trip purpose value. These school drop offs and pick ups are labelled as the generic “facilitate passenger” for the adult, but as “school” for the child who is being transported to school. As a result, in order to create the dataset for this chapter, the process was to find all the school trips made by children aged 17 and under in the survey. For each of these trips, the next step was to check the trips of all the adults in the household and check if there was a facilitate passenger trip starting at the same time with the same origin and destination zones as the school trip in question. If there was, we would know that that facilitate passenger trip was in fact a school drop off or pick up.

However, the above process would only work for students aged 11-17, as travel data was not recorded for younger children. In order to find out if an adult made a school drop off or pick up for a child under 11, the first step was to query for any facilitate passenger trips made my adults with children under 11 in the household. Next, the trips of all other household members were checked to look for a trip with the same start time and origin and destination zones as the facilitate passenger trip in question. If no corresponding trip was found, it could be reasonably assumed that that trip was in fact a school drop off or pickup for a child under age 11.

Once the above data preparation processes were completed it was possible to make comparisons between the trips made by adults making daycare drop offs, school drop
offs, and those that went straight to work or school. For clarity, when I refer to an adult making a trip to work or school, when say school, am I referring the place where the adult is a student, such as university, college, or high school, if applicable. This is not to be confused with the place where their child is a student.

6.1 Mode choice

In the 2011 TTS, all trips were classified as one of the following modes:

- Public transit
- Bicycle
- Auto driver
- GO rail
- Joint GO rail and public transit
- Motorcycle
- Auto passenger
- School bus
- Taxi
- Walk
- Other

For the purposes of this empirical investigation, school bus, taxi, motorcycle, and other were all put into the same bin and labeled as “other” because of their low frequency among adults with children in the household, which is the relevant sample from TTS. In addition, Public transit, GO rail, and joint GO trail and public transit were all put into the same bin and labelled as “transit for the sake of simplicity.

Figure 14 shows the mode split for adults that have children in their household making a non-stop trip from home to their work or school in the AM period (6:00AM-8:59AM). Although the majority of commuters use the auto driver mode, the auto mode share is relatively small when we compare to Figure 15, which shows the mode split for trip chains from home to work or school in the AM period that make a stop to drop off a child.
at daycare. The increased share for auto driver, and especially decreased share for transit is large. This trend is even more marked in Figure 16, which shows the mode split for trip chains from home to work or school in the AM period that make a stop to drop off a child at school, for which approximately 100% of adults use auto modes.

Figure 14 Mode split for trips from home to work/school by adults with children
Figure 15 Mode split for work/school chains with a daycare drop off

Figure 16 Mode split for work/school chains with a school drop off

While the higher convenience of auto modes for transporting children is clear and would likely never be questioned, the statistical magnitude of the difference in mode share
suggests that not only do auto modes offer the most utility for dropping children at school in the local area, they are the only feasible option.

The policy implications of these findings are notable. Since reducing automobile trips, especially during the peak periods is generally always a local for local agencies, an important part of the solution may be to find ways for children to get to school without their parents needing to drive them. If we naively assume that all the trips represented in Figure 16 can be expanded to population levels simply based on an overall 5.01% sampling rate (Data Management Group, 2013), if the mode split for all school drop offs was made identical to the mode split for direct home-work/school trips, the number of auto trips in the AM period in the study area would be reduced by over 35,000. Although the accurate expansion process is not so naïve, the correct number is certainly on a similar order of magnitude. In light of this information, school travel planning programs such as Safe and Active Routes to School (Canada Walks, 2016) have the potential to impact the region significantly.

6.2 Departure time

The data preparation process and analysis approach for this section are similar to those in the previous section, except morning departure time from home is the focus in place of mode choice.

6.2.1 TTS Observations

Figure 17 shows the departure time distribution for adults with children in the house making a direct trip. The distribution shows very large on the :00 and :30 points of each hour, which is for two reasons. This is affected by rounding errors caused during the telephone survey. While this may cause errors in estimating precise traffic assignment models, for the purposes of this chapter, as long as the errors are consistent throughout the survey sample, the different distributions can be compared. Figure 18 and Figure 19 show the morning departure time distribution for adults making trip chains from home to work or school with a stop at daycare and school, respectively. Both of these distributions are shifted towards the later part of the AM period, more so for the school drop offs. In
addition to their overall shift, both distributions are less spread out over the period, indicating reduced flexibility when a child needs to be dropped off.

This reduced flexibility is to be expected since schools around the region start at generally the same time, as opposed to jobs that start at all times of the day. However, the policy implications of the finding are significant. Shifting automobile trips to outside of the peak periods is generally a goal for agencies in the region to alleviate rush hour traffic, but these distributions suggest that such a policy goal will be ineffective for these populations, since the timing of their commute is constrained by the starting hours of their child’s daycare or school.
Figure 17 Departure time distribution for trips from home to work/school by adults with children

Figure 18 Departure time distribution for work/school chains with a daycare drop off

Figure 19 Departure time distribution for work/school chains with a school drop off
6.2.2 Multinomial logit model

To gain a more quantitative understanding of how making a school or daycare drop off affects morning departure, a choice model was estimated. As a preliminary exploration, the AM period was discretized into 30-minute segments and a multinomial logit model (MNL) was used. Although this is an oversimplification because the MNL does not recognize the correlation between adjacent time periods or the continuous nature of time, it still gives us an opportunity to determine the effects of school and daycare drop offs on departure time, even if the model is to simple to accurately make prediction of departure time.

\[ P_{timei} = \frac{\exp(V_i)}{\sum_k \exp(V_k)} \]

The sample size for model estimation is 65087, which is the total number of trips (or chains) from home to work or school by adults in the AM period, 1801 of which are daycare drop off chains, and 4780 are school drop off chains. Results of parameter estimation are given in Table 3. Utility for the 6:00-6:29 period was normalized to 0 so that positive coefficients indicate an increased utility gained from departing later in the morning.
Table 3 Multinomial logit model of morning departure time

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>6:00 – 6:29 normalized to 0 utility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.432</td>
<td>21.47</td>
</tr>
<tr>
<td>6:30 – 6:59</td>
<td>1.870</td>
<td>28.33</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>2.636</td>
<td>41.29</td>
</tr>
<tr>
<td>7:30 – 7:59</td>
<td>2.088</td>
<td>30.39</td>
</tr>
<tr>
<td>8:00 – 8:29</td>
<td>2.032</td>
<td>28.54</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:30 - 6:59</td>
<td>-0.010</td>
<td>-18.69</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>-0.013</td>
<td>-14.66</td>
</tr>
<tr>
<td>7:30 – 7:59</td>
<td>-0.021</td>
<td>-18.92</td>
</tr>
<tr>
<td>8:00 – 8:29</td>
<td>-0.017</td>
<td>-11.77</td>
</tr>
<tr>
<td>8:30 – 8:59</td>
<td>-0.023</td>
<td>-9.15</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:30 - 6:59</td>
<td>-0.124</td>
<td>-4.04</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>-0.255</td>
<td>-7.73</td>
</tr>
<tr>
<td>7:30 – 7:59</td>
<td>-0.312</td>
<td>-9.76</td>
</tr>
<tr>
<td>8:00 – 8:29</td>
<td>-0.331</td>
<td>-9.51</td>
</tr>
<tr>
<td>8:30 – 8:59</td>
<td>-0.163</td>
<td>-4.43</td>
</tr>
<tr>
<td>Works in retail</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:30 - 6:59</td>
<td>-0.299</td>
<td>-8.03</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>-0.535</td>
<td>-14.47</td>
</tr>
<tr>
<td>7:30 – 7:59</td>
<td>-0.594</td>
<td>-9.76</td>
</tr>
<tr>
<td>8:00 – 8:29</td>
<td>-0.525</td>
<td>-13.69</td>
</tr>
<tr>
<td>8:30 – 8:59</td>
<td>-0.220</td>
<td>-5.53</td>
</tr>
<tr>
<td>Works in manufacturing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:30 - 6:59</td>
<td>-0.767</td>
<td>-17.43</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>-1.367</td>
<td>-29.04</td>
</tr>
<tr>
<td>7:30 – 7:59</td>
<td>-2.027</td>
<td>-39.96</td>
</tr>
<tr>
<td>8:00 – 8:29</td>
<td>-2.341</td>
<td>-35.68</td>
</tr>
<tr>
<td>8:30 – 8:59</td>
<td>-2.244</td>
<td>-30.79</td>
</tr>
<tr>
<td>Commutes to Toronto from outer suburbs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:30 - 6:59</td>
<td>-0.242</td>
<td>-4.01</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>-0.457</td>
<td>-7.55</td>
</tr>
<tr>
<td>7:30 – 7:59</td>
<td>-0.795</td>
<td>-13.32</td>
</tr>
<tr>
<td>8:00 – 8:29</td>
<td>-1.017</td>
<td>-15.33</td>
</tr>
<tr>
<td>8:30 – 8:59</td>
<td>-1.049</td>
<td>-14.96</td>
</tr>
<tr>
<td>Commutes to planning district 1 from outer suburbs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:30 - 6:59</td>
<td>-0.199</td>
<td>-3.09</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>-0.719</td>
<td>-10.75</td>
</tr>
<tr>
<td>7:30 – 7:59</td>
<td>-1.563</td>
<td>-22.07</td>
</tr>
<tr>
<td>8:00 – 8:29</td>
<td>-2.301</td>
<td>-24.24</td>
</tr>
<tr>
<td>8:30 – 8:59</td>
<td>-2.280</td>
<td>-22.35</td>
</tr>
<tr>
<td>Lives and works in outer suburbs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:30 - 6:59</td>
<td>-0.217</td>
<td>-5.33</td>
</tr>
<tr>
<td>7:00 – 7:29</td>
<td>-0.229</td>
<td>-5.73</td>
</tr>
</tbody>
</table>
Increasingly negative coefficients for the age and male variables indicate that men and older people are more likely to leave the house earlier in the morning than younger people and women.

Similarly, increasingly negative coefficients for the retail and manufacturing dummy variables mean that people working in these sectors are more likely to leave for work early in the morning.

The following four dummy variables are essentially proxies for morning commute length. These variables were used in place of actual commute time for two reasons. The first was to simplify the data preparation procedure, and the second was that zone-to-zone travel time is not available for all zones in the study area, meaning that if we only wanted to use respondents with home and work locations in zones covered by the traffic assignment model, our sample size would be reduced. As is intuitive, all these variables suggest people with the longest commutes are likely to leave earlier in the morning.
Most importantly, the magnitude and sign of the final two dummy variables and their t-statistics show that the need to drop off a child at school or daycare have sizeable and significant effects on the time a commuter leaves the house in the morning, even when considered independently of other explanatory variables.
7 Conclusions and Future Work

The research presented in this thesis represents progress in three significant areas. The first is a novel approach to location choice set generation using a stochastic frontier model. The methodology was not only useful for estimating a choice model about daycare location choice, but likely has the potential to be used in a variety of location choice applications. The second is a model of daycare utilization and location choice which can be used as an input to an activity scheduling model. Third, data from the 2011 Transportation Tomorrow Survey has been analyzed to demonstrate the effect that needing to drop off a child at school or daycare in the morning has on travel decisions like mode choice and travel time. It is felt that these three developments provide an increased understanding of how decisions related to transporting dependent children are made, and how those decisions impact the travel behaviour of the adults that make them.

In addition to this understanding, the models developed in this thesis provide direction about how to ensure that these decisions are accurately represented in activity scheduling models.

Several opportunities for future work emerge from the work presented in this thesis, including moving forward with the empirical work for models that have been conceptualized here.

The task allocation models described in chapter 3 could be tested empirically. This included adding another nest to the daycare utilization and location choice model to include a daycare drop off task allocation model, in addition to a separate school drop off task allocation model.

The interaction between the morning drop off and the afternoon pickup could be investigated. Namely, the pickup task can be included in the task allocations models described above. In addition, the effect of one adult making the drop off and the other making the pickup on the location choice set can be assessed.
One major behavioural limitation of the daycare location choice model was the absence of explanatory variables related to given childcare locations, such as price and quality, and specific programs offered. While it’s clear that these elements are important in the decision of parents, the data for including these variables in the model was not available in the local context. A data collection effort into at least a sample of childcare centres to collect this information would be valuable in understanding how these effects compare with those already included in the model. Alternatively, the model could be tested in any city around the world where this data is already available in some government database.

The effect of child drop off on morning mode choice could be assessed by estimating a choice model. However, it’s important that any such modal considers the captivity to auto modes that parents making drop offs face.

Finally, the morning departure time model presented in this thesis can be improved upon, either using a cross-nested logit model to properly represent the correlation between time segments, or with a continuous choice model which represents the progression through AM period as a continuous variable.
8 References


DMTI Spatial. (2014). *CanMap RouteLogistics*.


*SimAGENT* in Southern California. Paper presented at the 91st annual meeting of the Transportation Research Board, Washington, DC.


