2014

Micro-Data Collection and Development of Trip Generation Models of Commercial Vehicles: An Application for Windsor, Ontario

Georgiana Madar

University of Windsor

MICRO-DATA COLLECTION AND DEVELOPMENT OF TRIP GENERATION MODELS OF COMMERCIAL VEHICLES: AN APPLICATION FOR WINDSOR, ONTARIO

By

Georgiana Madar

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of **Civil and Environmental Engineering**
in Partial Fulfillment of the Requirements for
the Degree of **Master of Applied Science**
at the University of Windsor

Windsor, Ontario, Canada

2014

MICRO-DATA COLLECTION AND DEVELOPMENT OF TRIP GENERATION MODELS OF COMMERCIAL VEHICLES: AN APPLICATION FOR WINDSOR, ONTARIO

by

Georgiana Madar

APPROVED BY:

Dr. B. Anderson
Department of Political Science

Dr. C. Lee
Department of Civil and Environmental Engineering

Mr. J. Tofflemire
Department of Civil and Environmental Engineering

Dr. H. Maoh, Advisor
Department of Civil and Environmental Engineering

September 11, 2014

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

In recent years, the field of transportation research has received increased interest from researchers and modeling methods have seen significant development. This thesis presents the development of a statistical microsimulation model for trip generation in the Windsor, Ontario Census Metropolitan Area (CMA). A novel methodology was developed to collect information from local firms using a combination phone and web-survey. Establishment-level data about business characteristics and commercial vehicle activities in the study area was collected from 171 establishments for analysis. Ordered logit models were used to determine the factors affecting the number of inbound and outbound commercial vehicle trips. The modeling results show that establishment characteristics are more effective in representing outbound movements than inbound. The model results are useful in explaining some of the factors giving rise to commercial trip generation, as well as revealing areas where further research is required, for a more complete understanding of commercial behaviours.

DEDICATION

I dedicate this thesis, first and foremost to Jesus Christ, my saviour, provider, and sustainer, and to those who encouraged and supported me along this journey, especially my family and my dear Vini Bordei.

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Hanna Maoh, my advisor, for his mentorship, guidance, and patience as I progressed through this project. The encouragement I received to submit papers, attend conferences, and apply to scholarships is appreciated, as it enhanced my experience through the Master's program. I would also like to thank my committee members Dr. William Anderson, Dr. Chris Lee, John Tofflemire, as well as Dr. Rupp Carriveau, the defense chair, for taking the time to attend my defence and review my work.

Additionally, I would like to acknowledge Haibin Dong, who programmed the web-based survey used in the data collection procedure, without which this project would have been significantly more difficult to complete. I also offer thanks to Shakil Khan, who worked to distribute the survey to interested respondents and collect the results from the web survey. His aid has been greatly appreciated with the technical aspects of this project. I also acknowledge the efforts of Rahaf Hussein, who assisted in the telephone survey portion of the data collection and with whose help, the time spent recruiting survey participants was reduced.

Lastly, I thank my family and friends for their encouragement and support, especially at times where I felt overwhelmed. Their support kept me focused on the final goal and motivated to make it to the end.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

Commercial vehicle movements are the outcome of commercial vehicle trips. The latter are undertaken by a business establishment for the purpose of delivering goods or services to a customer, either another firm or a private citizen. The field of researching commercial vehicle activities is relatively new. Until recently, the majority of research efforts have been focused mostly on modeling and forecasting passenger trip generation *(Ferguson et al., 2012)*. Compared with passenger trip research, only a few studies have been published on freight trip generation and modeling in the past three decades *(Batisda and Holguin-Veras, 2009)*. Although commercial vehicle trips comprise only an estimated 10 to 15% of total urban trips *(Hunt and Stefan, 2007)*, they pose significant impacts on congestion, travel time and cost, pavement wear, noise, and air quality. In the past decade, the study of urban commercial vehicle movements has become of increasing interest and importance to researchers and policy-makers *(Hunt et at., 2006)*. Such interest has been encouraged by new methodologies for modeling the factors that give rise to freight trip generation and urban movement patterns.

Commercial vehicle activities were largely modeled and forecasted using methods primarily developed for studying private vehicle activities. The four-stage model is one such method, extensively used in modeling personal travel *(Ferguson et al., 2012)* and applied to commercial vehicle modeling by using weight factors. Despite its practicality, the explanatory ability of the four-stage model is impeded by a number of shortcomings *(Miller et al., 2004)*. Most importantly, this modeling technique lacks the behavioural realism required to accurately capture the true nature of commercial vehicle activities and the factors that give rise to commercial trips.

An extensive range of vehicle sizes are used for commercial purposes, ranging from small personal-sized vehicles to large multi-unit trucks. The smaller vehicles can easily be misrepresented in traffic count data, since they are not as easily distinguishable from passenger vehicles as the larger commercial vehicles. This is far more complex than the vehicle sizes available for passenger travel. The complexity of commercial vehicle activity is further increased by the heterogeneity in the industries involved in trading goods and the customers to whom services are provided. Due to such factors, the decision-making process for commercial vehicle trip generation is non-trivial. In recent years, these complexities have been answered through use of the microsimulation modeling paradigm *(Hunt and Stefan, 2007; Ferguson et al., 2012)*. This modeling technique analyzes the factors contributing to trip generation for each business establishment in the study area. Such a detailed study of commercial vehicle activities and their characteristics, however, requires detailed information on individual firms, which is not readily available in a majority of urban jurisdictions. Also, the acquisition of such detailed data is difficult and costly.

Within the context of the latter, a behavioural microsimulation model has been developed for the city of Calgary, Alberta in Canada *(Hunt and Stefan, 2007; Stefan et al., 2007)* and stands as a pioneering example for commercial vehicle modeling at the business establishment level. Data for this research effort was collected from individual firms through the use of surveys. Similar data collection and modeling methodologies were later implemented in the Greater Toronto Area in Ontario *(Xie and, 2009*). These successful applications of microsimulation models inspired this project and the collection of data from business establishments in the Windsor Census Metropolitan Area (CMA) to be used in a trip generation model for commercial vehicles in this study area. Such a model will form the basis of a full-fledged microsimulation model that will be developed for the region.

Individual business establishment data for the Windsor CMA forms the basis of analysis in this thesis. The data, which represents a total of over 10,000 businesses in the year 2013, were acquired from InfoCanada. The data represents the total population of business establishments in the study area. Each acquired record included the street address of the establishment, telephone number, employment size, six-digit industrial classification code, and name of the business. The telephone number was important, as it enables us to contact the establishments to recruit them to participate in the commercial vehicle movement survey. A combined telephone and web-based survey methodology was used to collect data for these individual firms about their commercial activities. The telephone-based approach used in recruiting participants for the web survey, which will be described in detail in Chapter 3, is novel and unique. The questions presented in the web survey collected data about general establishment characteristics, outbound commercial vehicle trips, and inbound commercial vehicle trips. The survey efforts yielded 171 usable observations, which were analyzed and utilized in the statistical analysis. It was decided that ordered logit models would be used to model the factors affecting the number of outgoing and incoming commercial vehicle trips in the study area, given the highly organized nature of commercial vehicle activities.

1.2 Research Objectives

A number of specific research objectives were specified at the beginning of this project, which remained fairly consistent through the duration of the work. A number of extensions to the scope of the work were suggested, however, the primary goals of this research project remained centred around the following:

1. Advancement of the knowledge on urban commercial vehicle movements with an application to the Windsor CMA,
2. Development of an effective methodology for collecting data about commercial vehicle activity at the business establishment level,
3. Explanation of the factors giving rise to the observed commercial trip productions and attractions in the study area through the use of statistical models,
4. Understanding the patterns of commercial vehicle movements, and
5. Provide a basis for commercial trip generation forecasting to improve upon current urban planning practices

1.3 Thesis Outline

This thesis is organized in the traditional format, presenting the findings from the research project in a logical progression. The following chapter outlines the existing literature on recent research efforts conducted to model commercial trips both in Canada and elsewhere. Chapter 3 describes the data used to implement the survey, the methodology used to recruit firms to partake in the survey, a preliminary analysis of the data collected from the web-based survey, and the theoretical basis for the statistical modeling approach. The results from the performed statistical analyses are presented in

Chapter 4, along with a discussion of the significant factors influencing trip generation. The final chapter outlines conclusions drawn from the results of the research project, the limitations of the conducted work, and some considerations for future research. A list of references follows Chapter 5, along with appendices containing additional information.

CHAPTER 2

LITERATURE REVIEW

2.1 Conventional Methods for Modeling Commercial Vehicle Activities

In order to have a better grasp of the theoretical basis for the data collection and modeling undertaken in this project, a literature review of existing research in the field of modeling commercial vehicle activity was conducted. It was noted that significantly more efforts have been made to research private vehicles than commercial vehicles. The study of commercial vehicles was, until recently, estimated and approximated using methods similar to those for private trips.

One of the most widely-used conventional models is the four-stage model, which studies vehicle activity at the level of aggregated traffic analysis zones. It estimates trip generation (production and attraction) for each zone, trip distribution between zones, modal split between available travel modes, and network assignment of trips to the road network. This model is commonly used to capture passenger travel and usually accounts for commercial vehicles by introducing weight factors. Doing this, however, assumes that the behaviours of private and commercial vehicle activities are the same, which is definitely not the case. The weight factors also neglect to account for the range of vehicle sizes used for commercial purposes, often only accounting for larger trucks. *Yang et al. (2009)* and *Zhou and Dai (2012)* both presented papers describing different classes of freight models. When describing the four-stage process, it is defined as modeling the movement of passengers and goods along a transportation network. Since it relates commodities and trips, shortcomings of the four-stage model include its inability to capture service trips, less-than-truckload trips, or empty truck trips. Its applicability is more related to regional contexts rather than urban settings. Four-stage models can be either commodity-based or trip-based. Data requirements for these models are costly and extensive. External data must usually be obtained, including demographic, socio-economic, land use, and employment information for the study area, as well as commodity flow or trip tables.

Another conventional method for freight modeling is by origin-destination (O-D) matrix expansion or factoring. This kind of model requires initial O-D tables for commercial vehicle trips along with supply and demand data. The *Quick Response Freight Manual (1996)* presents the general O-D modeling framework. Future flows are estimated by multiplying growth rates from economic models or traffic condition surveys into the O-D matrix *(Yang et al., 2009)*. Commodity flows or truck trips could also be calculated as percentages of the estimated passenger trips, which are determined using conventional methods *(Zhou and Dai, 2012)*.

The O-D matrix expansion method was used by *Xie and Roorda (2009)* to analyze goods movement related to the auto industry in Ontario, using data from roadside interviews, provincial input-output tables, and population and employment values. In this study, a preliminary O-D matrix was estimated by a gravity model and a travel time-sensitive impedance function was introduced to zonal commodity totals between pairs of origins and destinations. The total commodity production and attraction (consumption) rates in each geographic area were estimated based on proportions of population and employment for different industry categories. Commodity flows were then assigned to the road network using an all-or-nothing approach for the shortest travel time within O-D pairs. While this study examines only one industry type (automotive) and only one travel mode (trucks), the researchers believe this framework could be expanded to other industry types and modes.

Another O-D modeling framework was undertaken by *Nuzzolo and Comi (2014)*, who used traffic count data and roadside interviews with truck drivers and retailers to build a multi-stage model with three O-D matrix sub-systems in Rome, Italy. The three O-D matrices represented and estimated the quantity of the demand, the number of deliveries at different times of day, and the freight vehicle flows on the network. Each of these three sub-models followed a discrete choice approach. This model considered zonal economic characteristics, locations of freight centres, business characteristics, shipment size, and vehicle type. It also explored patterns of delivery of truck tours with respect to the type and size of shipped goods, accessibility to origin and destination zones, attraction of zones, and vehicle types. The calibrated models provided a good fit and a number of different calibration methodologies, such as nested logit, mixed logit, and probit models.

A third conventional type of model is the input-output or inbound-outbound (I-O) model, which makes use of I-O tables to determine zonal trip production and attraction values and commodity flows between zones. These tables record the values required and produced by each industry type as well as the values of goods that are consumed by customers. Employment and population data for the study area are then introduced in order to determine total freight flows for a particular region. This type of analysis does not consider the mode of transportation of the goods *(Jewell et al., 2007)*.

A number of other modeling methods are described in *Yang et al. (2009)*, including models that produce commodity flow estimates at facilities; models that produce estimates only for the truck mode using land use, socioeconomic, and supply and demand data; models that simulate commercial activities using economic trends; and models that study supply chains and logistics (also described in Z*hou and Dai, 2012)*.

Estimating commercial trips as a constant proportion of total traffic is incorrect, since commercial and private vehicles behave differently (*Holguin-Veras et al.,* 2013). Also, the assumption that commercial vehicles behave in a similar way as private vehicles is incorrect and using methods developed for passenger travel to model commercial trips lacks the behavioural realism required to capture the complexities of commercial movements. The way in which commercial trips are organized differs greatly from private trips. Also, the vehicle types used for commercial purposes, ranging from small passenger-sized cars to large multi-unit trucks, is much more variable than those used for passenger travel. Many of the conventional models utilize aggregate data and produce zonal estimates averaged over larger areas, neglecting the decision-making behaviours of individual businesses involved in generating the trips. The need for models that are specifically designed to model commercial activities more accurately has led to the development of a number of new approaches in recent years.

2.2 Emerging Methods for Modeling Commercial Vehicle Activities

Although the conventional methodologies are practical, since some are not too data-intensive and versatile to different regions and applications, they suffer from a number of shortcomings. The traditional four-stage model suffers from ignoring travel as a demand derived from activity decisions, ignoring spatial and temporal relationships between trips and activities, not considering the constraints that restrict choice, and neglecting various dynamic complexities by only focusing on utility maximization *(McNally and Rindt, 2007)*. Another weakness of many of the conventional models is, as previously stated, that they do not always capture all of the vehicle types that are used for commercial purposes. Many approaches study only large trucks and/or heavy vehicles, since these are much more easily observed in traffic counts than passenger-sized vehicles used for local deliveries or service calls. Using aggregate and zonal/regional data and averages will also reduce the models' ability to accurately capture some interactions, dynamic behaviours, or sensitivities of commercial vehicle movements to changes in policy.

In recent years, there has been a rise in the use of the microsimulation approaches, which use business establishments as the unit of analysis. This approach allows researchers to observe decision-making at the level of individual business establishments, allowing for a more accurate view of the behaviours that influence commercial vehicle trip generation. Here, firms are ultimately the decision makers that influence the trends in commercial vehicle behaviours. The work by *Hunt and Stefan (2007)* in Calgary, Alberta stands as a pioneering effort in the Canadian context for developing such a microsimulation framework, and served as inspiration for this research project. This framework formed the basis for some other Canadian studies (i.e. *Ferguson et al, 2012*). These, as well as some international studies, will be described in detail in the following sections. While microsimulation is a disaggregate approach that allows for more accurate commercial vehicle analysis, it is a highly involved process and quite data-intensive, as will be explained in the following section.

2.3 Data Needs and Data Collection Efforts

One of the greatest challenges to conducting a microsimulation study, which has been noted by many researchers in the field, is the lack of detailed data at the business establishment level. Many firms are reluctant to provide information about their shipping decisions. Such information may play an important role in firms' business strategies and lend them a competitive edge, which they would not want to compromise by sharing their information. There is also a higher value of time that is associated with business establishments and executives may not be able or willing to invest the time cost required to complete a detailed survey. This section outlines some of the conventional methods for dealing with the issue of data unavailability, as well as some new approaches, both in Canada and elsewhere.

2.3.1 Conventional Data Collection Efforts

Since detailed data is difficult to collect, some researchers use commercial travel data collected by different organizations that is made publicly available or that can be purchased. Examples of such databases include the Commodity Flow Survey (CFS) and the Freight Analysis Framework (FAF) available in the United States (*Yang et al., 2009*). The Commercial Vehicle Survey (CVS) of Ontario, as well as specialized tables produced by Statistics Canada and Transport Canada are used (*Xie and Roorda, 2009*).

Roadside interviews are also a conventional means for collecting data about commercial vehicle activities (*Xie and Roorda, 2009*). Some of the shortcomings of these data collection methods is that some offer averaged information, while some, such as the roadside interviews of drivers, collect information from large trucks, since these are easy to identify among the flow of traffic. The conventional data collection methodologies for commercial vehicle data are not detailed enough to be used in a microsimulation modeling framework. The following sub-sections detail some of the recent data collection procedures that have been applied to microsimulation models in Canada as well as elsewhere.

2.3.2 Recent Canadian Data Collection Efforts

In the context of Canadian research, the previously mentioned work done by John Douglas Hunt and his group of researchers in the cities of Calgary and Edmonton, Alberta *(Hunt et al., 2006)* stands as a good example. In the paper by *Hunt et al. (2006)* the design, implementation, and results from an establishment-based survey were discussed. The purpose of this project was to develop a deeper understanding of commodity movements, including the factors that influence their generation and distribution patterns, and vehicle movements that arise from the commodity movements, including delivery patterns, the use of depots, and types of vehicles used. In order to meet these objectives, the research team collected data about establishments, including employment data for different industry sectors and quantities of goods and services produced and shipped. Data about shipper behaviour was also collected, including locations of destinations for shipped goods or services sent by each establishment and characteristics about the shipments, commodity categories, and the use of depots. The research team also collected data about vehicle movements, including the types, locations, and timings of stops, the sizes, values, and types of shipments, and vehicle configurations (*Hunt et al. 2006*).

Information for this project was collected through surveys mailed to business establishments that collected data about shipments occurring on the survey day. Six different forms were sent to respondents, three of which applied to business establishments, and three for depots. The Establishment Forms collected information about the locations, employment size, and total commercial movements entering and leaving each of the establishments or depots. A Goods Depot or Goods Shipment Form collected information about the quantities of goods shipped, and origins, destinations, vehicle types, and routes for deliveries. Lastly, data about the patterns of stops made by vehicles leaving each establishments to, either, deliver goods or provide services was collected through a Goods Vehicle Form or Service Vehicle form.

Recruitment of establishments for survey participation in this study was first done through the mail, where firms were asked to identify a representative who could fill the survey. Next, the representative was contacted by telephone and asked about the employment size of the firm, whether the firm engages in any shipping or receiving and if it is a depot, and whether the firm could participate in the survey. The applicable survey forms were then forwarded to willing participants and the collected data was analyzed, verified, and used in the microsimulation modeling.

Another Canadian research effort was undertaken in the Greater Toronto and Hamilton Area (GTHA), and was based on the work done in Alberta (*Ferguson et al., 2012*). This research project used data about firms from InfoCanada, including employment size, and Standard Industry Classification (SIC) code. Data about commercial vehicles was obtained from a survey of urban commercial vehicle movements (UCVM) conducted for the Peel region between the years of 2006 and 2007, which surveyed business establishments in the Peel region to collect data about their commercial vehicle activities, including departure time, location, duration, and distance of each stop. The Commercial Vehicle Survey (CVS) conducted by the Ontario Ministry of Transportation (MTO) for the year 2006 was also used, and was a roadside survey containing information about origins, destinations, and cargo types for trucks. Traffic count data for medium and heavy vehicles, obtained from the Data Management Group (DMG) at the University of Toronto, was also utilized. All of this data was used in a microsimulation model similar to that used in Alberta and, with some modifications to suit the GTHA study area, the modeling framework was successfully transferred and implemented in the GTHA.

2.3.3 Other Recent Data Collection Efforts

Some data collection efforts to develop microsimulation models have been conducted outside of Canada as well. In the case of *Joubert and Axhausen (2013)*, for example, GPS data was used to determine locations of commercial vehicle facilities and depots. This approach is useful only in cases where firms track their vehicles using GPS transmitters, which is usually the case for longer-haul type of commercial activities and for larger vehicles. GPS data is not always available for intra-urban commercial vehicle movements and tends to exclude some vehicle and industry types.

Some studies on commercial vehicles have been conducted in New York City, such as the work carried out by *Lawson et al. (2012)*. Disaggregate data was collected for this project using an establishment-level survey sent to receivers of goods. A questionnaire was sent to collect data on establishment characteristics, operations, and patterns of freight trip generation. Land use information was combined with the geocoded locations for each of the firms. *Batisda and Holguin-Veras (2009)* also collected disaggregated data for a study in the New York City area. A survey was sent to carriers, receivers, and intermediaries (depots) in Brooklyn and Manhattan. The firms were categorized based on industry segments, determined by SIC codes. Establishment size was also collected, along with the types of commodities shipped or received, categorized in 24 different types. This data was then used in models to identify the firm attributes that influence the number of freight deliveries per day.

Another type of commercial vehicle survey was conducted by having drivers fill out travel diaries. *Wang and Hu (2012)* used this kind of data collection method for the Denver Regional Council of Governments in Denver, Colorado between the years of 1998 and 1999. The travel diaries collected information about travel purpose, tour stops, arrival and departure times at each stop, and travel times. Vehicle type characteristics and company attributes were also collected. A weighting system was implemented for the data on the basis of employment size and vehicle types. This study examined mode choice from different vehicle types used for commercial activities in the study area.

One last data collection methodology that was examined, also conducted in the USA, where it was noted that the majority of establishment surveys have low response rates was found in the work by *Samimi et al. (2013)*. The reasons for firms' reluctance to share information are, as previously discussed, the sensitivity of the information and the high value of time of the representatives filling the survey. This study compared mail-in-mail-out surveys, telephone interview surveys, and web-based surveys. It was determined that the web-based survey was the most cost-effective method. A pilot web-based survey found a response rate of about one percent. However, combining the web survey with a marketing component was more effective. Establishments were first invited to participate in the survey, then sent reminder emails at intervals of 2, 7, and 14 days after the initial contact, and finally, forwarded a report of the study's final results. This methodology found a 20% increase in the response rate. The survey collected data about establishment characteristics, attributes of recent shipments, and contact information. This study found that in order for a survey to be successful, selection bias, based on characteristics such as location, size, or industry type, must to be avoided. Contacting a biased sample has the potential of introducing a non-response bias into the survey responses. It was found that the web-based survey methodology was a viable and cost-effective methodology for data collection.

2.4 Trip Generation Modeling

2.4.1 Recent Efforts in Freight Generation Modeling

This section will highlight some recent studies done on freight generation and freight trip generation. Freight generation refers to the production and consumption of cargo, whereas freight trip generation represents the commercial vehicle trips required to transport this cargo (*Holguin-Veras et al., 2011*). The latter paper combines these two concepts to study commercial vehicle movements with respect to logistics and economics. Employment size was used as an independent variable in a linear regression model that used freight production or freight attraction as a dependent variable. The use of a constant freight trip generation rate was explored for each industry category, as determined by SIC codes corresponding to the firms. It was found that freight generation is determined by the economics of production and consumption, while freight trip generation is determined by logistic decisions. Spatial and economic aggregations and their heterogeneities were also explored. It was found that combined freight generation and freight trip generation models are better explained when the aggregations are more internally homogeneous. As mentioned before, this paper used Commodity Flow Survey (CFS) micro data for the model estimation.

Freight trip generation was also investigated by *Holguin-Veras et al. (2013)*, who explored the transferability of freight trip generation models, and the statistical significance of location variables. Three freight trip generation models were examined, namely, the National Cooperative Freight Research Program (NCFRP) Project 25, the Quick Response Freight Manual, and the Trip Generation manual by the Institute of Transportation Engineers (ITE). The first model tested SIC classifications, NAICS classifications, Land Based Classification Standards, and the City of New York Zoning resolution, which is evidently specific to New York City. The analysis was done through linear regression and multiple classification, using employment or square footage as independent variables. In the case of the QRFM, zonal employment per industry segment and SIC codes were used to estimate freight trip generation for different vehicle classes. Lastly, the ITE Trip Generation manual estimated freight trip generation as a proportion of total generated trips. Root Mean Square Error (RMSE) comparisons were made between the three methodologies. The RMSE values for the QRFM models were higher than those for Project 25, while no concrete conclusions were drawn about the ITE models. It was found that locational effects were of limited importance. A synthetic correction procedure was developed to enhance the transferability of freight trip generation models. This correction procedure was meant to extend the usefulness of constant freight trip generation rates, allowing them to be applied to any industry sectors for which freight trip generation is not assumed proportional to business size.

Freight generation models were developed for the area of New York City, as presented by *Batisda and Holguin-Veras (2009)* and *Lawson et al. (2012)*. The first paper compared the linear regression approach to a cross-classification approach for freight generation. Disaggregate establishment data was used from commercial establishments in the study area, with the number of deliveries sent or received as the independent variable. SIC codes, employment size and type, and commodity types were used in the regression modeling, in the form of categorical variables. For the cross-classification approach, continuous variables were converted to interval scales. It was found that regression models were effective in defining some relations pertaining to freight generation demand, more specifically, the commodity type was found to be a strong explanatory variable when interacting with other variables, such as employment size or revenue. It was concluded that regression models were able to forecast future freight demand and establish relationships between variables.

The other study by *Lawson et al. (2012)* examined the effects of land use and employment size on freight trip generation in New York City. This paper compared three land use classification codes, the New York City Zoning Resolution, Land-Based Classification Standards, and the Institute of Transportation Engineers manual, previously described. Establishment-level disaggregate data was collected through surveys, including information about firm characteristics, operations, and the numbers of deliveries received on a normal day. Land use designations were assigned to each firm based on geographic location. Three modeling approaches were estimated, using standard trip generation rates, an ordinary least square approach, and a multiple classification analysis. Similar to the first study, the suitability of the models was assessed using the root mean square error. The best estimates of trip generation for most models were achieved with a constant coefficient that depended on the land use designation. A smaller proportion of models were found to depend on employment size, whereas the remaining ones used a combination of a constant coefficient and employment size. It was noted by the authors that the freight trip generation models could be extrapolated and transferred to other regions with similar characteristics.

Freight generation has been analyzed using multinomial and nested logit models as shown in the study by *Wang and Hu, 2012* for Denver, Colorado. Travel related attributes, including origin, destination, trip purpose, travel distance, and travel time, as well as company attributes, land use, and cargo-related variables were obtained at a disaggregate level. The models examined vehicle type choice for commercial travel purposes using five mode categories. The statistical models found that the industry classification, location of stops, commodity characteristics such as number of cargo types carried and commodity size, as well as company related variables, including employment size, business type, and fleet size all affect the choice of vehicle types. In this case, the number of commercial trips made by a business establishment was used as an explanatory variable.

Freight generation models have also been studied based on aggregate data from the CFS, FAF, and input-output data as in *Olivera-Neto et al. (2012)*. The latter study examined the temporal effects of economic activities on variation in freight demand. Cross-classification and regression models were used in a combination to estimate freight production and attraction in some U.S. states. The parameters of the production models were estimated through a conventional linear regression model. It was found that payroll by industry sector was the only significant variable for estimating freight movements. The simplicity of this model, however, does not allow it to predict freight demands for future years.

2.4.2 Trip Generation in Microsimulation Models

Hunt and Stefan (2007) developed a methodology for studying commercial travel and its specific behaviours and characteristics. The Calgary model they proposed included a tour-based microsimulation, where commercial vehicles were considered to engage in tours with multiple stops before returning to their original location, as opposed to two-leg trips. It is argued that this approach provides a more accurate depiction of commercial vehicle patterns.

Three different components were designed and modeled separately for (1) tour-based movements, (2) fleet-allocator movements, and (3) external-internal movements. The information considered in the tour-based movement model included stop purpose, vehicle type, time of day, and location for all stops. Freight-allocator movements refer to vehicles dispatched to cover a certain area, such as newspaper delivery, waste collection, mail or courier services, taxi, and police fleets. External-internal movements refers to shipments or deliveries that have an origin or a destination that is outside of the studied urban area (*Hunt et al., 2004; Stefan et al., 2005*). The Calgary microsimulation model was calibrated using establishment-level data, as highlighted in Section 2.3.2.

The microsimulation model assigned attributes to each tour generated in a zone using Monte Carlo techniques. An aggregate trip generation model was used to determine the number of tours for each zone. A Monte Carlo process was then used to assign the vehicle type, tour purpose, and tour start time. Subsequently, characteristics for the stops of a tour, such as next stop purpose, next stop location, and stop duration, were assigned in an iterative fashion until the end of the tour. A return-to-establishment tour purpose alternative represented the end of a tour. Logit models were used to estimate the selection probabilities of trip purpose, establishment type and vehicle type combinations in the microsimulation procedure.

A description of the full modeling framework for Calgary's commercial vehicle movements is presented in the paper by *Hunt and Stefan (2007)*. Figure 2-1 shows the overall framework of the integrated system developed by *Hunt and Stefan (2007)*. The module that was of most interest to this thesis was the first, to do with tour generation. The numbers of tours were estimated as aggregate values for the different categories of establishments, for different times of day, and for each analysis zone in the study area. The aggregate values calculated were the numbers of tours generated per employee, which were described in a function of zonal attributes. When this tour generation rate was determined for a certain industry class, it was then multiplied by the total number of employees in the specific analysis zone, for the respective industry type. Thus, the total number of tours generated by each industry in one day was estimated for each analysis zone.

An exponential regression method was used for estimating the trip generation rate per employee in a certain zone and for a certain establishment type. In this calculation, the generation rate was a function of a constant term and zonal-level attributes. Employment information was obtained from the 2001 Canada Census. During the model calibration, the constant terms for specific zones were adjusted for area-specific characteristics.

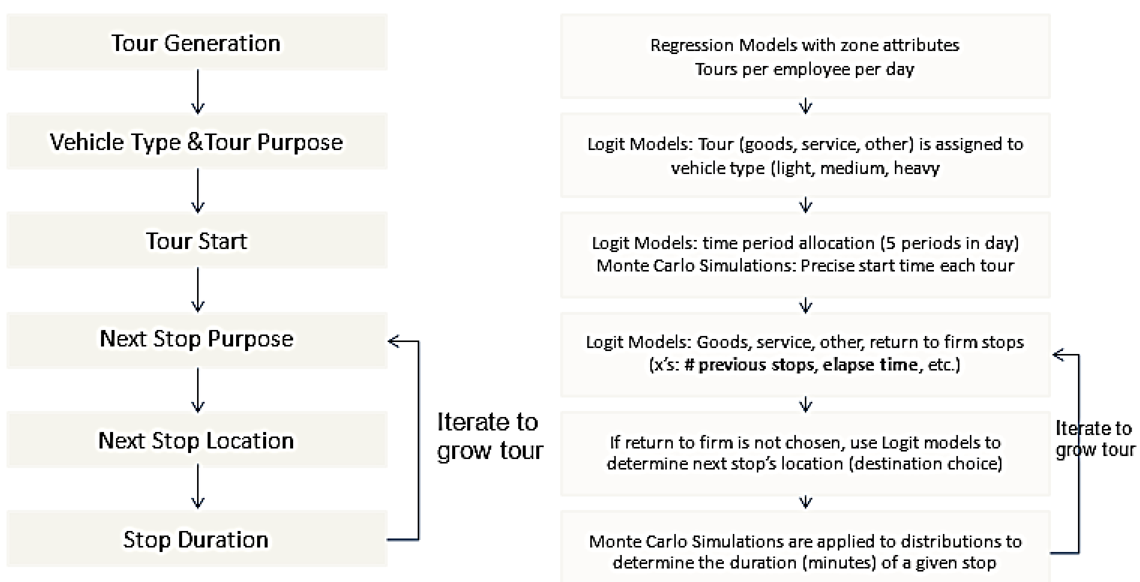


Figure 2-1: Sketch of the Tour-Based Model Developed by Hunt and Stefan (2007)

The results from this step of the estimation noted that the number of commercial trips generated for different industry categories is influenced by the land use designation and the types of activities generally associated with the respective land uses. Apart from firms in the transport industry, it was noted that fewer trips were generated by zones where a larger percentage of employment belongs to one type. With respect to accessibility, the results showed distinct differences in commercial activities compared to personal vehicle travel. The time of day for commercial tour generation was examined, calculating a utility function for five different time periods. Results here indicated that tours are generally not generated in the evening and nighttime periods. The remaining time periods were observed to have varying effects, depending on the industry category being investigated, which is to be expected, since the natures and behaviours differ among industry classes.

Some policy scenarios were tested, including the effects of increasing the money cost per unit distance, increasing the travel time, introducing or removing truck route restrictions, and introducing of a money charge per stop in the central business district. It was found that each of these affect the number of tours generated, total distances traveled, complexities of tour configurations, the number of heavy trucks present on the road network, and number of stops during a tour.

The framework developed in Alberta was transferred and applied to the study area of the Greater Toronto and Hamilton Area (GTHA) by *Ferguson et al (2012)*. The study area was split into smaller traffic analysis zones to be used in the modeling and analysis.

As previously mentioned, urban commercial vehicle movement data was obtained from a survey in the Peel region. Tour attributes, including departure time, number of stops, and tour distance, were obtained and introduced in the model estimation.

A similar framework to what had been developed in Calgary *(Hunt and Stefan, 2007)* was used in the latter study. The tour generation component, which is most closely related to this thesis, calculated the number of commercial vehicles leaving an establishment for each industry class in one day. The number of generated tours at each establishment was calculated, similarly to the Calgary study, as a function of the rate of tours per employee in the respective industry category and the total number of employees for the establishment. Random estimates were made for the number of tours per employee, based on a table of tour attributes for the Peel region. Zonal tours were then estimated by aggregating the numbers of tours generated for the different industry categories.

The GTHA study was able to successfully implement the methodologies of Hunt and Stefan (2007) and transfer the modeling framework from the Calgary to the study area. The alternative specific constants that had been adjusted for certain area characteristics, of course, had to be recalibrated when transferring the model to the new study area.

Since the trip generation model in the work presented by *Hunt and Stefan (2007)* is still an aggregate approach, the work in this thesis will follow a different, more behavioural track. It was noted that none of the existing research studies has addressed commercial trip generation at a micro-level. As such, the research presented in this thesis attempts to fill this gap by collecting data for developing such a model for a microsimulation system. The approach followed in this thesis tries to address the true behaviours of trip generation by developing microsimulation models. The variables used in previous studies were considered and hypotheses were created about the factors giving rise to commercial trip generation on the basis of the existing research. The data collection methods used in this project were inspired by some recent efforts for collecting establishment-level micro data in an efficient and effective manner, which will be described in Chapter 3. The modeling framework used and results will be detailed in Chapter 4.

CHAPTER 3

DATA AND METHODS

3.1 Windsor Commercial Survey

As discussed in Chapter 2, a microsimulation modeling effort requires detailed data at the business establishment level. In order to collect such information from individual firms in the study area, a web-based survey was designed. The process for recruiting businesses to participate in this survey was inspired by the procedure followed by *Samimi et al. (2013)*, as it involved a phone call component to recruit participants for the survey, and a web-based platform where respondents answered the survey questions over the internet. The data collection process will be described in detail in this chapter. A novel procedure was developed to determine the industry types that were most likely to be eligible and willing to partake in the survey, in order to reduce the time required for recruitment of survey respondents and eventually increase the response rate. The web survey consisted of 40 questions, grouped into three major categories: general establishment characteristics, data about inbound commercial activities on the day the survey was answered, and information about outbound commercial movements on the survey day.

3.1.1 InfoCanada Data

A dataset was acquired from InfoCanada that contained contact information for all 10,761 firms registered in the Windsor CMA in 2013. The information provided in this dataset included each firm's telephone number and mailing address (including street address, mailing suite, city, province, and postal code); executive's first and last name, title, and gender; primary North American Industry Classification System (NAICS) six-digit code and description, the Standard Industry Classification (SIC) code and description, and the InfoUSA (IUSA) code designation for each establishment. A more general two-digit NAICS code was extracted from the information provided in the dataset, which was later used to determine the industry category to which each firm belonged. A unique alphanumeric access code was generated for each business establishment and appended to the data. These access codes were used to allow respondents access to the web survey. The access code was also used as a primary key to create a relational database with the response information.

Of the list of business establishments, ten exclusive random samples were drawn, each containing approximately 10% of the entire population of businesses. The random samples were taken such that the distribution of industry types was consistent across all ten samples. The distribution of industries in each random sample was also similar to that in the entire population of businesses. These ten random samples were used in recruiting firms to the web survey.

3.1.2 Telephone survey

As a first step, recruitment of respondents for the survey was conducted through a short telephone survey. The firms were asked the following questions:

1. Whether the firm engaged in shipping or receiving of any goods or services;
2. If the answer to the first question was positive, whether a representative of the establishment would be able to partake in the online survey; and

3. If the answer to the second question was also positive, contact information was collected from the firm representative (typically an email address) in order to send a link to the web survey and an access code.

The data for each of the 10% samples was saved in a separate Access Database file. A graphical user interface form was developed to establish a computer assisted telephone interview (CATI) system to use during the telephone calls. A screenshot of this form is provided in Appendix A. The purpose of the CATI form was to display information about each contacted business and to record the representatives' responses to the three questions in an efficient manner, as well as to take note of email addresses for those interested in completing the web survey. The responses to the three questions, email addresses, and any applicable comments were updated into the Access file. Section 3.1.3 describes the analysis performed on the data about whether firms engage in shipping or receiving of goods or service and whether the firm representatives were interested in taking the web survey. A link to the web survey was sent to all establishment representatives who provided an email address.

3.1.3 Quotient Approach for Stratified Sampling

The telephone recruitment was undertaken in two phases. During the first phase, five of the ten random samples were contacted in their entirety, for a total of 4917 contacted businesses. These samples consisted of 1053, 924, 984, 982, and 974 firms, respectively. The reason that the number of firms differed in the different samples is because some of the records were corrupted, for example, not having a listed telephone number. The invalid records were removed from the dataset. Table 3-1 shows the makeup of these five samples with respect to industry category, determined by the two-digit NAICS code designation. The distribution of firms by industry classification is also given as percentage ranges in the same table. Telephone calls to these firms were conducted over the course of six weeks, between August 20th and September 26th, 2013. On average, between 150 and 200 firms were contacted each day.

Table 3-1: Makeup of Industry Categories in the First Five Ten-Percent Samples of Firms

		Sample ID					
NAICS 2-digit code	**Description**	**1**	**2**	**3**	**4**	**5**	**Total**
11	Agriculture, Forestry, Fishing and Hunting	0%	0%	0%	0%	0%	1
21	Mining, Quarrying, and Oil and Gas Extraction	0%	0%	0%	0%	0%	2
22	Utilities	0%	0%	0%	0%	0%	1
23	Construction	6%	7%	9%	8%	9%	382
31	Manufacturing	1%	1%	1%	1%	0%	37
32	Manufacturing	1%	1%	1%	1%	1%	39
33	Manufacturing	3%	3%	5%	3%	4%	184
42	Wholesale Trade	4%	2%	4%	5%	4%	183
44	Retail Trade	12%	15%	10%	10%	11%	583
45	Retail Trade	4%	4%	5%	7%	5%	235
48	Transportation and Warehousing	1%	3%	2%	2%	1%	89
49	Transportation and Warehousing	0%	0%	0%	1%	0%	14
51	Information	2%	1%	1%	1%	1%	56
52	Finance and Insurance	4%	3%	4%	4%	4%	185
53	Real Estate and Rental and Leasing	4%	3%	4%	4%	3%	166
54	Professional, Scientific, and Technical Services	10%	9%	10%	9%	8%	450
55	Management of Companies and Enterprises	0%	0%	0%	0%	0%	1
56	Administrative and Support and Waste Management and Remediation Services	2%	5%	3%	3%	4%	172
61	Educational Services	2%	2%	3%	3%	3%	132

62	Health Care and Social Assistance	12%	11%	9%	9%	9%	489
71	Arts, Entertainment, and Recreation	3%	2%	1%	2%	3%	116
72	Accommodation and Food Services	9%	9%	9%	7%	8%	408
81	Other Services (except Public Administration)	13%	12%	13%	15%	17%	700
92	Public Administration	4%	2%	2%	2%	2%	120
99	Unclassified Establishments	3%	3%	3%	5%	3%	170
(blank)	(no NAICS code assigned)	0%	0%	0%	0%	0%	2
Grand Total		100%	100%	100%	100%	100%	4917

Note: Refer to Appendix D for a detailed list of sub-industries at the 3-digit level

During the first phase of telephone recruitment, a total of 359 firm representatives expressed interest in completing the online survey and provided an email address to which a link and an access code to the survey could be sent. Given the total population size of 4917 firms for this phase, a response rate of 7.3% was calculated. It was noted during this recruitment phase that a significant number of firms reported to not engage in any shipping or receiving of goods or services, or were unwilling to partake in the web survey. This observation led to an analysis of the contacted firms, in order to determine the industry categories that were more prone to complete the survey (i.e. those who engaged in shipping or receiving of goods or services) and those who were most likely to be willing to complete the web survey, according to the responses given during the telephone recruitment. That is, the analysis aimed to devise an informed criterion to generate a stratified population of the firms that do engage in generating commercial trips and were willing to participate in the survey.

The determination of the firms was done through a Quotient analysis approach that was inspired by the Location Quotient (LQ) technique. LQ is a well-established technique that has been used in economics and economic geography to measure the degree of concentration of a certain industry in a smaller region that is part of a larger region (*Miller and Blair, 2009*). In the context of geographic areas, if $i \in I$, the size of an industry n in areas i and I are e_n^i and e_n, respectively, and the size of all industries in areas i and I are e^i and e^I, then the location quotient is defined as

$$LQ_i = \frac{e_n^i / e^i}{e_n / e^I} \ldots \text{(3-1)}$$

where $e^i = \sum_n e_n^i$ and $e^I = \sum_n e_n$.

Using Equation 3-1, industry n is said to be concentrated in area i if LQ $\geq$ 1.

The above concept was adapted to the context of this thesis to identify those industries that were more responsive to the survey. In this case, geographies i and I are replaced by populations s and S, where s is the surveyed 50% random sample and S is the total population of firms. The calculations are based on the number of firms that reported to engage in shipping or receiving of goods or services, and the number of firms who were willing to partake in the web survey, as reported during the first telephone recruitment phase.

Two Quotient (Q) values were calculated, on the basis of two-digit NAICS codes (industry classifications). One value determined which industry classifications were most likely to engage in commercial vehicle activities, and the second value determined which industry sectors were most likely to agree to complete the web survey. The first Q value (Q_S) related to the eligibility of industry sectors to participate in the survey and was calculated as described in Equation 3-2. The ratio of the number of businesses that reported to engage in shipping or receiving within a specific random sample and industry category, to the total number of firms who reported engaging in shipping or receiving, was divided by the ratio of the number of businesses in the respective sample to the number of firms in the entire population of contacted firms.

$$\boldsymbol{Q_S} = \frac{\left(\frac{F_n^S}{\sum_n F_n^S}\right)}{\left(\frac{F_n}{\sum_n F_n}\right)} \dots \text{(3-2)}$$

where the variables are defined as follows:

F_n^S: number of business in each industry category that reported shipping or receiving of goods or services

$\sum_n F_n^S$: total number of businesses in a sample that reported shipping or receiving of goods or services

F_n: total number of businesses in the respective industry category

$\sum_n F_n$: total number of responses in the sample

Similarly, the second Q values (Q_W) were calculated to determine the willingness of business establishments to complete the web survey, as given by the number of firms who provided email addresses during the recruitment telephone calls. The calculation is given by Equation 3-3.

$$\boldsymbol{Q_W} = \frac{\left(\frac{F_n^W}{\sum_n F_n^W}\right)}{\left(\frac{F_n}{\sum_n F_n}\right)} \dots \text{(3-3)}$$

where the variables are defined as:

F_n^w: number of business in each industry category that reported being willing to participate in the web survey

$\sum_n F_n^w$: total number of businesses in a sample that reported being willing to participate in the web survey

In each calculation, a Quotient value that was greater than 1 indicated that the industry category being investigated was more likely to engage in shipping or receiving of goods or services, or was more likely to agree to complete the web survey. Table 3-2 summarizes the calculated Q values with respect to firms' eligibility to partake in the web survey (whether establishments engage in any shipping or receiving). The desirable total Q values were those greater than 1, since these indicate the industry categories that were more likely to have commercial vehicle activities, and these values are bolded. Similarly, Table 3-3 shows the Quotient analysis results with respect to businesses' willingness to take the web survey (whether the representative was willing to provide an email address to which the web survey link would be forwarded). Again, a value greater than 1 was desirable for the total Q values, and the qualifying industry sectors are demarked in bold.

Table 3-2: LQ Values for Industry Sectors Engaging in Shipping/Receiving (Q_S)

		Sample					
		1	2	3	4	5	**Total**
Industry Sector (NAICS)	23	0.88	0.56	0.67	0.60	0.92	0.74
	31	***3.67***	***3.47***	***1.77***	***1.75***	***3.62***	***2.60***
	32	***1.05***	***5.01***	***2.42***	***2.62***	***3.10***	***2.77***
	33	***3.36***	***2.15***	***1.73***	***2.29***	***2.27***	***2.32***
	42	***2.38***	***2.61***	***1.11***	***1.28***	***1.86***	***1.71***
	44	***1.48***	***1.31***	***1.56***	***1.69***	***1.58***	***1.49***
	45	***1.69***	***2.03***	***1.42***	***1.35***	***1.13***	***1.48***
	48	***0.56***	***1.25***	***1.69***	***1.64***	***1.94***	***1.49***

	49	***0.00***	***3.76***	***3.99***	***2.10***	***0.00***	***2.58***
	51	0.00	1.25	0.44	0.75	2.33	0.75
	52	0.17	1.11	0.55	0.54	0.73	0.58
	53	0.00	1.16	0.28	0.90	1.16	0.65
	54	0.70	0.85	0.43	0.54	0.41	0.56
	56	0.59	0.33	0.51	0.77	0.91	0.63
	61	***1.34***	***0.36***	***0.95***	***1.52***	***1.81***	***1.32***
	62	0.98	0.52	0.75	0.80	0.79	0.76
	71	0.73	1.19	0.00	1.09	0.72	0.78
	72	***0.98***	***1.22***	***1.40***	***0.98***	***0.85***	***1.09***
	81	0.72	0.26	1.06	0.76	0.43	0.67
	92	0.73	0.44	0.97	0.95	0.57	0.75
	99	0.00	0.23	0.21	0.12	0.49	0.21

Table 3-3: LQ Values for Industry Sectors Willing to Participate in the Survey (Q_W)

		Sample					
		1	2	3	4	5	**Total**
Industry Sector (NAICS)	23	0.69	0.82	0.99	1.02	0.51	0.84
	31	***5.15***	***4.24***	***2.40***	***0.00***	***4.70***	***3.13***
	32	***0.00***	***2.30***	***2.95***	***2.85***	***6.04***	***2.97***
	33	***2.65***	***2.46***	***2.12***	***3.20***	***4.26***	***2.94***
	42	***4.59***	***4.20***	***1.26***	***1.77***	***2.41***	***2.54***
	44	***1.32***	***1.16***	***1.80***	***1.73***	***1.28***	***1.44***
	45	***0.40***	***2.61***	***1.44***	***1.38***	***0.59***	***1.32***
	48	***1.19***	***1.15***	***1.97***	***1.42***	***1.01***	***1.45***
	49	***0.00***	***6.90***	***2.70***	***2.28***	***0.00***	***2.76***
	51	0.00	0.00	0.90	0.00	2.01	0.46
	52	0.36	0.51	0.55	0.29	0.38	0.42
	53	0.00	0.53	0.00	0.98	1.51	0.54
	54	0.88	1.21	0.22	0.52	0.70	0.66
	56	0.62	0.30	0.35	1.34	1.57	0.82
	61	0.70	0.00	0.39	1.10	1.88	0.88
	62	1.33	0.41	0.35	0.80	0.68	0.71
	71	0.00	0.73	0.00	0.00	0.47	0.22
	72	0.69	0.80	1.31	0.53	0.91	0.88
	81	0.76	0.36	1.12	0.82	0.25	0.68
	92	0.77	0.81	0.49	0.00	0.00	0.43
	99	0.00	0.43	0.43	0.00	0.00	0.15

Following this analysis, the remaining five of the 10% random samples that were not contacted during the first phase were filtered on the basis of the Quotient analysis results. Only the industry classes whose Q values indicated that they were most likely to be eligible for the survey and willing to participate (where $Q \geq 1$) were selected to be contacted in the second phase of telephone recruitment. These industry sectors are summarized in Table 3-4, in which a total of 1823 firms from the non-contacted dataset fit the criteria.

Table 3-4: Industry Sectors Selected for Contact during Phase 2, Based on Quotient Analysis Results

2-Digit NAICS Code	**Description**
31	Manufacturing
32	Manufacturing
33	Manufacturing
42	Wholesale Trade
44	Retail Trade
45	Retail Trade
48	Transportation and Warehousing
49	Transportation and Warehousing
61	Educational Services
72	Accommodation and Food Services

Consequently, the second phase of recruitment phone calls was performed over the course of three weeks, between October 15th and November 4th, 2013. During this second phase, a total of 322 email addresses were collected from firm representatives who were interested in receiving the web survey (compared to 359 from phase 1). It was observed that executives of the firms contacted in phase two were more willing to cooperate and participate in the survey than those contacted during the first phase. The improvement in the response rate is evidence that the Quotient analysis and targeting method were effective. Overall, a total of 681 email addresses were collected from the two phases.

To summarize, in the first phase of phone calls, 4917 firms were contacted and 359 provided email addresses. Of these, 89 completely filled the web survey. In the second phase, 1823 firms were contacted, 322 provided email addresses, and 82 completed the web survey. These observations show that the Quotient method produced results comparable to the first, non-targeted phase of recruiting survey respondents. It can be seen that the Quotient analysis aided in reducing the number of contacted firms by approximately 64%, when comparing the first phase to the second, targeted phase of recruitment phone calls, while improving the response rate.

Further validation for the Quotient analysis method was done through statistical t-tests performed to compare the distributions of industry categories, employment size categories, and categories for the total number of vehicle owned by each firm. The industry categories were based on the two-digit NAICS designations of the firms, while the other two variables were obtained through questions in the web survey. These three characteristics have been observed in numerous articles in the literature as distinguishing the landscape of firms in a region, and were, therefore, the ones selected for comparison. The results from the validation procedures are presented in detail in Chapter 4.

3.1.5 Web-Based Survey

At the end of the second phase of telephone calls, a link to the web survey was forwarded to all firm representatives who provided an email address to receive and complete the survey. Each representative was also emailed a unique access code that was assigned to the respective firm. Subsequent to the initial informational email, three reminder emails were sent after 2, 7, and 14 days. The survey was open to respondents from August 26th until November 22nd, 2013. Of the 681 firms who provided contact information to receive the survey, 171 answered the entire survey. This represents approximately 25% of the firms who were willing to receive the survey invitation. An additional 17 firms' representatives accessed the survey and partially answered it, however, since a significant portion of the questions were left unanswered, these responses were not suitable for use in the final analysis. The partial responses were mostly in the section about establishment characteristics, while the inbound and outbound trip characteristics were left unanswered. Responses from the 171 usable observations were organized and all answers were converted to numerical values, to allow for statistical analysis that will be presented in Chapter 4.

The web survey asked a total of 40 questions, separated into three general categories: establishment characteristics, characteristics of outbound commercial vehicle activities on the survey day, and characteristics of inbound commercial vehicle activities on the survey day. A list of all of these questions is provided in Appendix B. Screenshots of the web survey platform are available in Appendix C, showing how respondents viewed the survey. Since the survey period was undertaken over the span of several months, the day on which the survey was completed by each firm was left to the representatives' discretion and time availability.

For the purpose of this project, only the data about the general establishment characteristics and the total numbers of inbound and outbound trips on the survey day were used in the statistical modeling. The establishment characteristics section was the most completely answered out of the three survey sections. The other two sections (characteristics of outbound and inbound commercial activities) were found difficult to incorporate into the models. Each firm was asked how many inbound and outbound shipments it had received or sent on the day of the survey and, on the basis of these answers, the set of questions in these two sections was asked as many times as there were reported trips. Some firms reported zero inbound or outbound trips on the survey day and, therefore, had no recorded data for these questions. Of the firms who reported multiple shipments, some representatives did not fill the survey for the total number of reported trips, thus, some values were missing. Another issue was that, in the cases where multiple trips were reported, it was not possible to average some of the answers (for example, the vehicle types used or the origin or destination of the shipment) for each respondent. The statistical model allowed for only one entry per firm; hence, questions with multiple answers for one firm were not able to be used at this time. The data, however, is quite useful and will be utilized in future modeling efforts and in the continuation and extension of this project.

3.1.6 Weighting Approach for Responses and Location Dilution

As previously mentioned, a total of 171 firms completed the survey in its entirety. In order to ensure that the survey responses were a true representation of the entire population firms, a weighting scheme was introduced. In other words, the purpose of the weights was to determine how many firms in the entire population of establishments were represented by each one of the 171 observations from the survey. The 171 observations were compared with the population of firms who reported engaging in shipping or receiving of goods or services during the telephone survey. The weights were assigned on the basis of industry types (two-digit NAICS codes) and employment size.

All of the firms who had indicated that they engage in any shipping or receiving during the telephone survey were extracted from the Access database file. The sample ID (indicating to which of the ten 10% sample each firm belonged), the IUSA number, NAICS code, and unique survey access code were extracted. This database was joined to the InfoCanada data for the year 2013 to obtain the employment sizes for this population of firms. They were matched by the IUSA number, since each firm has a unique IUSA designation. There were a total of 1461 establishments who reported engaging in shipping or receiving of goods or services. These records were updated with the InfoCanada employment size values. A comparison was made for the 171 survey respondents to determine if the employment values reported in the survey matched those from InfoCanada. In the cases where there was a difference of more than 10 employees, the values from InfoCanada replaced the reported survey value, since it was certain that the InfoCanada data was verified and accurate, whereas the survey responses were subjective to the judgment of the respondents.

Once all of the values for employment size were assigned to each firm, categories were created for ranges of numbers of employees. Next, the data for the entire population of shippers (1461 firms) and the 171 survey respondents was extracted to an Excel file where cross tables were generated for both groups of firms using the pivot table tool. These cross tables were organized separately for respondents from each of the two phases, with employment categories as the row labels and two-digit NAICS codes as the column labels. A ratio of the counts in each category in the sample to the whole population of shippers was taken. Thus, a weight factor was produced for each combination of industry type and employment size category. These values are presented in Tables 3-5a and 3-5b for the respective phone survey Phase 1 and Phase 2, respectively, and were then added into the dataset of variables for survey respondents for use in the statistical modeling. The fields that had no calculated values correspond to industry and employment categories that were not present in the survey responses or were not contacted in Phase 2.

Table 3-5a: Calculated Weights for Respondents from Phase 1

		NAICS 2-digit code																	
Employment Categories	Employment Size Ranges	23	31	32	33	42	44	45	48	49	51	52	53	54	56	61	62	72	81
0	0																		
1	1 to 10	18.5	13	9	4.9	4.3	8.2	10.2	17	5		11	14	19	16	8	19.7	23	6
2	11 to 20	8			7	1.8	11.5		3					3		2			7
3	21 to 30			3		1								1	1				
4	31 to 40				1.5														
5	41 to 50				2														
6	51 to 60				1.5														
7	61 to 70																		
8	71 to 80					1													1
9	81 to 90																		
10	> 90																		2

Table 3-5b: Calculated Weights for Respondents from Phase 2

		NAICS 2-Digit Code																				
Employment Categories	Employment Size Ranges	23	31	32	33	42	44	45	48	49	51	52	53	54	56	61	62	71	72	81	92	99
0	0																					
1	1 to 10			13	5.3	8	9.3	19.7	7.5	1									16			
2	11 to 20		1.5	8	4.3	3.3	9	8	2.7										17			
3	21 to 30				2.8	3	7			1									9			
4	31 to 40						2															
5	41 to 50				4	3	5															
6	51 to 60				1	2	1.5															
7	61 to 70						1															
8	71 to 80				3																	
9	81 to 90																					
10	> 90				6				1													

Figures 2-2 and 2-3 show Kernel density maps comparing the density of the locations for the 171 survey respondents to the density of all the contacted firms, which were geocoded and diluted. The general patterns for the densities of the two compared populations appear consistent, thus providing further confidence in the collected data.

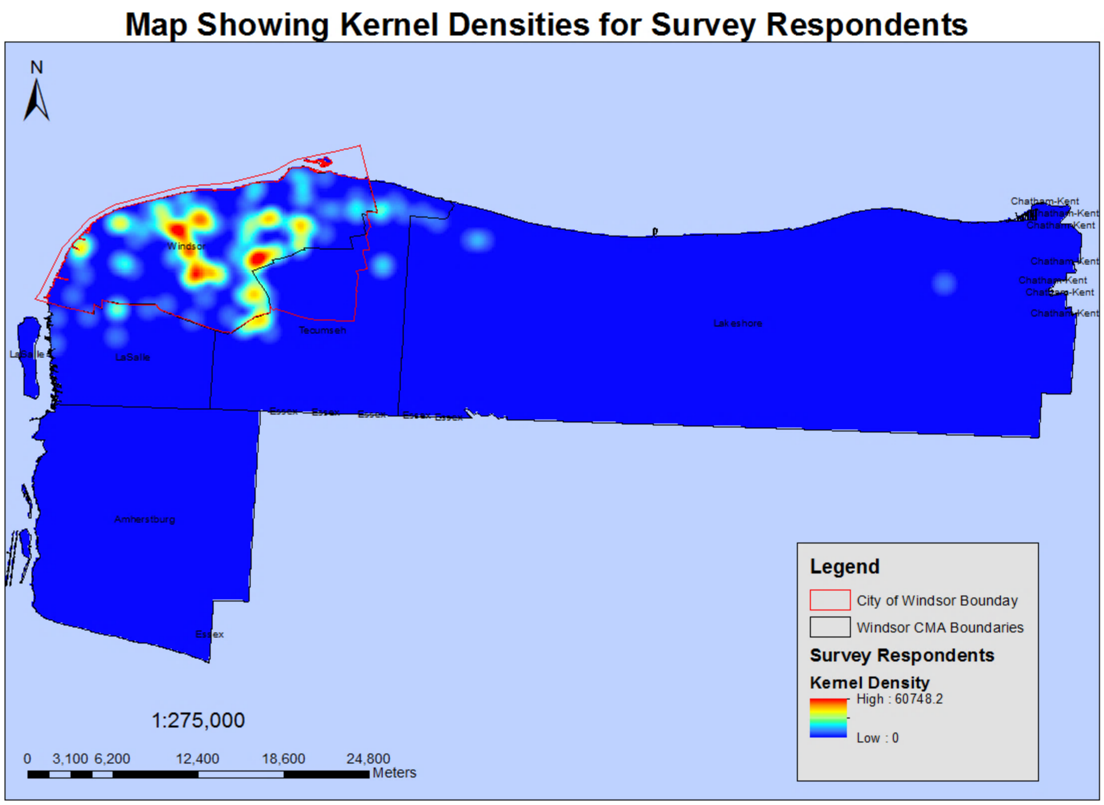


Figure 2-2: Kernel Density Map for Survey Respondents

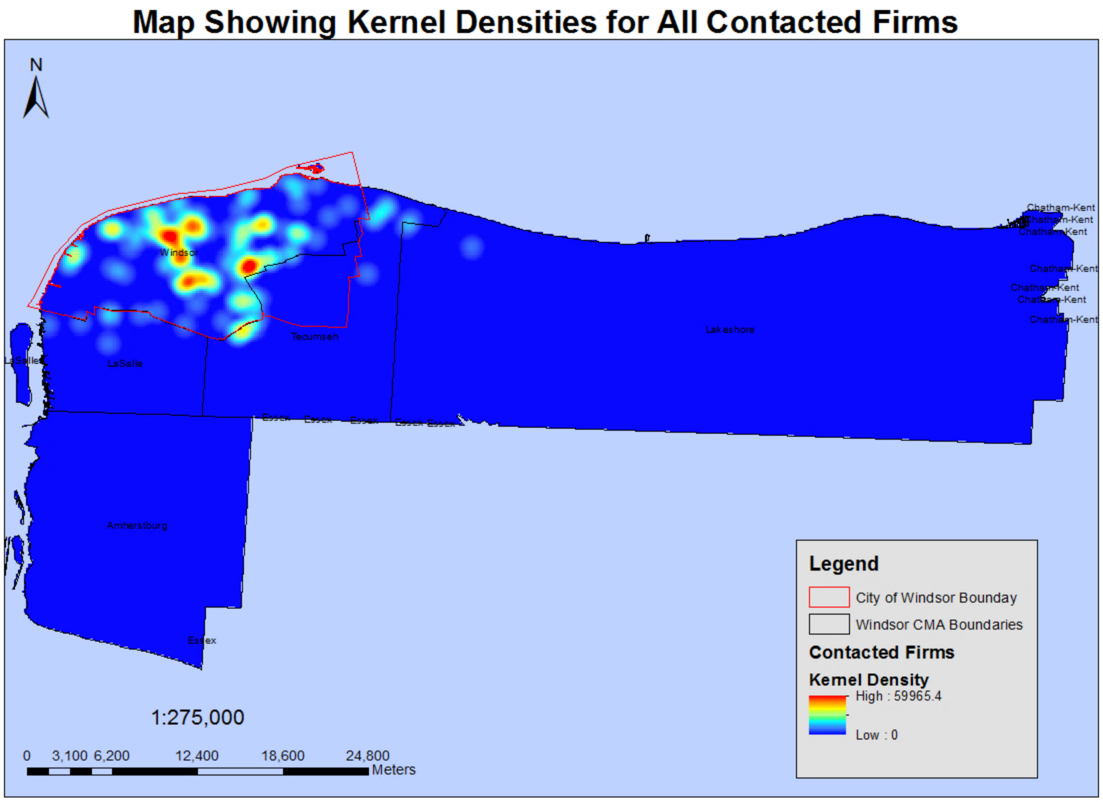


Figure 2-3: Kernel Density Map for Contacted Establishments

The respondent records were further processed to ensure confidentiality of the responding firms. Aside from removing all personal information, through which a firm could be identified, from the database of contacted firms and from the survey submissions, the geographic locations of all contacted firms and each survey respondents were diluted. This procedure was undertaken using ArcMap GIS software. Shapefiles for the road network and municipal boundaries for the Windsor CMA were downloaded from the Desktop Mapping Technologies Inc. (DMTI) spatial database for the year 2013. A network dataset and address locator were created on the basis of the roads shapefile. The address locator was then used to geocode the addresses, in three different groups: all contacted firms, all survey respondents, and remaining non-contacted firms. For the list of survey respondents, the ArcGIS online locator was also used, for comparison of the accuracy of location. This online service, however, was no longer available after December, 2013. In all cases, there were some unmatched addresses when using the address locator. Many of these were able to be repaired in the Excel database, where the city name or spelling of street addresses were repaired. Some addresses had to be manually selected on the map because the road network was not accurate in some areas due to new developments, road extensions, or missing address numbers. Validation for the addresses was done using Google Maps.

Of the 6740 total addresses contacted, 6480 were geocoded, representing approximately 96% of the addresses. All 189 survey respondents who completely or partially filled the survey were geocoded. Of these, 188 (99%) were successfully geocoded. Finally, the remaining addresses that were not contacted during the telephone survey portion, a total of 4309 firms, were mapped. Of these, 2900 addresses were successfully geocoded using the address locator. Another 933 addresses were manually corrected for city names or street address spelling, after which they were also mapped successfully. A list of 36 addresses were removed because the given addresses were located outside of the study area or, where only a postal code or street name was given, the region was too large to estimate an exact location. Another 476 firms in the database had no address or postal code specified and could not be mapped, and therefore, were also removed. Of the non-contacted firms, a total of 3797 was able to be geocoded.

The geocoding process was done in the Universal Transverse Mercator (UTM) World Geodetic Survey of 1984 (WGS84) projected coordinate system. After the geocoding procedure was complete, the point shapefiles for all geocoded firms were projected in the UTM North American Datum of 1983 (NAD83) projected coordinate system, for Zone 17N, which is most accurate for the Windsor area. The reason for this change was because the WGS84 system gives coordinates in decimal degrees, while the NAD83 system is able to measure coordinates in metres. In the property tables of each point shapefile of addresses, a field was added and calculated for the latitude (x-coordinate) and longitude (y-coordinate) of each point. These values were then extracted to an Excel file and the dilution of the coordinates was done within a radius of 100 metres. A different random value between 0 metres and 200 metres (as a dilution diameter) was both added to and subtracted from each latitude and longitude value. The diluted coordinates were then determined as the average of the two latitude and the two longitude values obtained from adding and subtracting the random values. The new coordinates were then imported into ArcMap again, converted back to the WGS84 coordinate system, and the Reverse Geocode and address locator functions were used to assign the diluted points to the nearest address on the map, within 50 metres. This procedure was undertaken to ensure that the firms could not be identified on the basis of their locations.

A preliminary analysis of the telephone survey responses and the web survey results, as well as detailed analysis of the modeling efforts will be presented in detail in Chapter 4. The next section outlines the theoretical background for the ordered modeling methodologies used in the analysis.

3.2 Ordered Choice Modeling

A few potential methods were available for modeling trip generation using the collected data. Since commercial vehicle trip generation is not a random process, normal regression and Poisson regression models were ruled out. A normal regression model considers its dependent variable to be continuous, which is not consistent with commercial trip generation behavior. Poisson regression assumes a random distribution of values, which is also not consistent with commercial vehicle activities, which are usually well planned and organized.

The modeling method that was selected as most appropriate to conduct the analysis was ordered choice modeling. The outcome of ordered logit or probit models is to predict the probability of the occurrence of 0, 1, 2, …, k events, in this case commercial vehicle trips. The explanatory variables were defined according to the business establishment characteristics obtained through the web survey. Following *Paez et al. (2007)*, an establishment will derive utility from the generated commercial trips such that 0 commercial trips will be generated when the utility U is below a give threshold μ_1 , 1 commercial trip will be generated when the utility is between μ_1 and μ_2 and k trips if the utility $U > \mu_k$. Although the utility U is unobservable, *Paez et al. (2007)* note that:

"the analyst can measure the outcome of the decision-making process (i.e. the number of trips made) and typically also other variables that relate to the level of utility". Accordingly, the utility of business establishment i is formulated as the sum of an observed utility and an unobserved error term, as per Equation 3-4.

$$U_i = V_i + \varepsilon_i \dots (3\text{-}4)$$

where

V_i: observable utility for establishment i, defined as the following linear in parameter function

$$V_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_n X_{ni} \dots (3\text{-}5)$$

where

β_0: alternative specific constant, accounting for unobserved effects

$\beta_1, \beta_2, \dots, \beta_n$: coefficients to be estimated

$X_{1i}, X_{2i}, \dots, X_{ni}$: variables pertaining to establishment i

ε_i: unobservable error term

The distribution of the unobservable error term is the basis on which an ordered model is defined as either logit or probit (*Green and Hensher, 2009*). The probability of generating a certain number of trips by establishment i can be derived from the utility function shown in equation (3-4). For instance, the probability of generating 0 trips is given as follows:

Pr (0) = Pr ($U > \mu_1$) = Pr ($V_i + \varepsilon_i > \mu_1$) $= \Pr(\varepsilon_i < \mu_1 - V_i)$

On the other hand, the probability of generating 1 trip is

Pr (1) = Pr ($\mu_1 < U < \mu_2$) = Pr ($\mu_1 < V_i + \varepsilon_i < \mu_2$)

Pr (1) = $\Pr(\varepsilon_i < \mu_2 - V_i) - \Pr(\varepsilon_i < \mu_1 - V_i)$

In the case of this project, the utility values are used to calculate the probability that a certain number of commercial trips would be generated (either produced by or attracted to a certain business establishment). The numbers of outgoing and incoming commercial vehicles were collected for each business establishment through the web survey, for the day each of them completed the survey. In each case, the numbers of incoming and outgoing commercial trips were categorized as either 0, 1, 2, or 3 or more trips. Categories 0, 1, and 2 were assigned those same numerical values in the model thus allowing for the prediction of the probabilities Pr(0), Pr(1) and Pr(2). The category representing three or more trips was designated as category 3 in the models to allow the calculation of Pr(3). This was important since the ordered model requires discrete values when estimated. The sum of the above four probabilities must add up to 1. Threshold values were determined on the basis of these categories. From a practical perspective, the predicted number of generated trips for a given establishment *i* can be calculated using the four probabilities Pr(0), Pr(1), Pr(2) and Pr(3) as follows:

$$O_i = 0 \times \Pr(0) + 1 \times \Pr(1) + 2 \times \Pr(2) + q \times \Pr(3)$$

where

q = 6.50 (outbound trips)

q = 5.19 (inbound trips)

The two q values used in the weighted probabilities represent the average number of trips that fall in the 3+ generated trips category for inbound and outbound movements. The outbound q value was determined by considering the average of all establishments that generated 3 or more trips except for two outlying observations from firms who reported having 80 and 99 trips, respectively, which is much higher than the rest of the reported values. However, the inbound q value was based on averaging the number of incoming trips from all the observations that reported to receive 3 or more inbound trips.

Table 3-5: Number of Firms Falling Under Each Ordered Category for Outbound and Inbound Movements

Number of Generated Trips	Outbound Trips	Inbound Trips
0	73	78
1	40	40
2	18	22
3+	40	31
Total	171	171

Table 3-6 also shows that the majority of businesses produce and attract between zero and two commercial trips, while there are fewer establishments who generate more than three trips. It was decided that the ordered categories for this project would be assigned as 0, 1, 2, and 3+ trips so that the most commonly occurring numbers of trips could be studied in detail. The establishments that reported producing or attracting three or more commercial trips on the day of the survey were grouped together, since these trips were sparsely distributed among a wide range of values.

Both logit and probit models were developed for both inbound and outbound vehicles, and the results were compared to determine which type of model provided the better fit for the data. As previously mentioned, the variables describing general firm characteristics were used in the analysis and modeling. The data obtained from the web survey was used to create categorical variables for the models. The previously described weight factors were introduced into all of the models. Chapter 4 provides a detailed description of the variables used, the modeling efforts, and an analysis of the results.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Phone Survey Results

After the telephone survey was completed, the results from all respondents were collected, analyzed, and used in the ordered modeling. As stated in Chapter 3, there were 189 firms who responded to the survey, 17 of which only partially completed the web survey, while the remaining 171 firms filled the survey entirely. Figure 4-1 below shows the geocoded and diluted locations of the survey respondents. The geocoding procedure was described in Chapter 3.

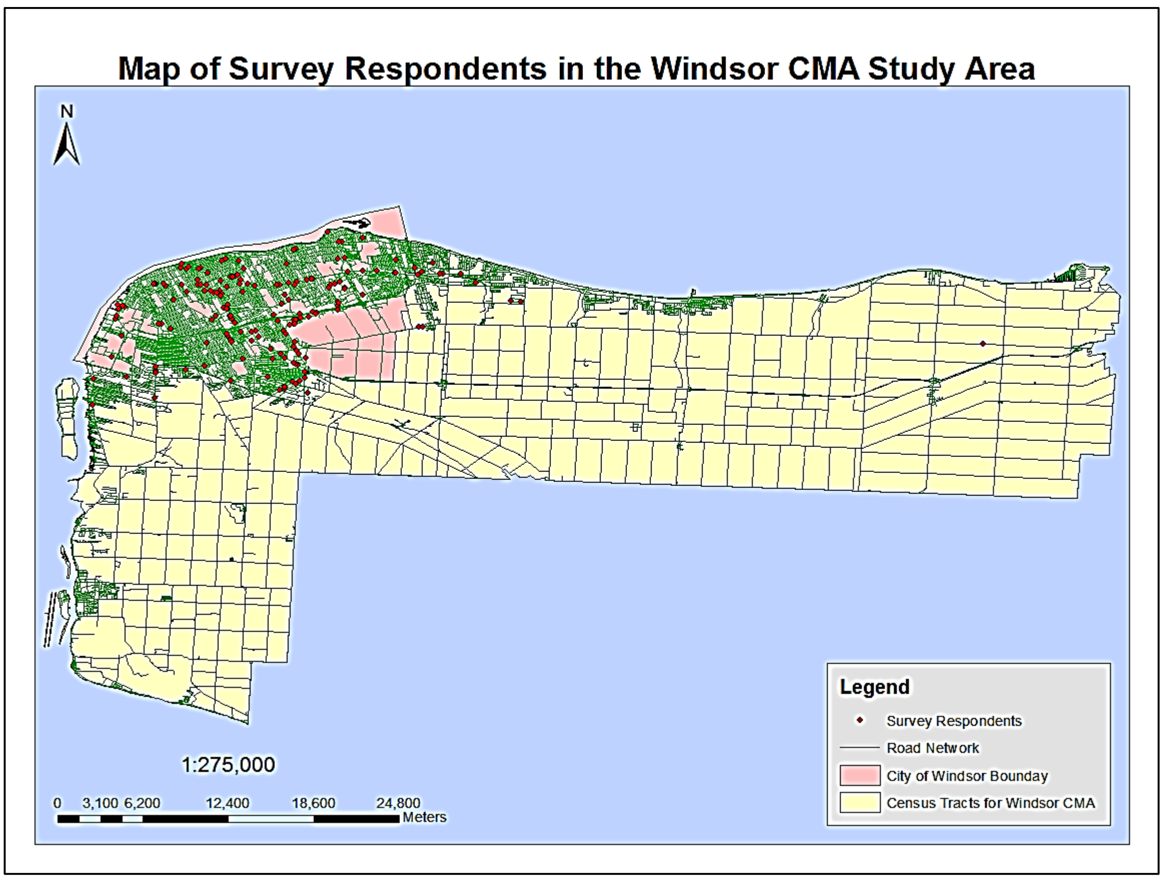


Figure 4-1: Map Showing Windsor CMA Study Area and Locations of Survey Respondents

The variables extracted from the section of the web survey about establishment characteristics, the following 18 variables, were used in the analysis.

- whether the establishment engages in shipping or receiving or goods or services
- the total number of employees at the establishment
- the number of employees in each occupation type at the establishment
- the number of years at the establishment`s current location
- the establishment's total square footage
- the number of employees who make on-call service trips
- the number of employees who make truck deliveries
- whether the establishment has a supply chain/logistics operator
- whether the establishment has a supply chain/logistics specialist
- whether the establishment uses logistics software
- the number of commercial vehicles of each type owned by the establishment
- the industry that best characterizes the establishment
- whether the number of commercial trips made by the establishment varies by day of the week
- the busiest days for delivering goods or services
- whether the number of commercial trips made varies by month
- the busiest months for delivering goods or services
- the total number of vehicles leaving the establishment on the day of the survey
- the total number of vehicles arriving at the establishment on the day of the survey

These variables were converted to numerical values for statistical analysis. All continuous variables retained the values reported in the survey. In the cases where categories were present for a variable, a unique numerical value was assigned to each category. Lastly, any yes/no answer was assigned a binary value, where a value of 1 was assigned to a response of "yes," and a value of 0 was assigned to a response of "no." These variables were then analyzed for correlations, to ensure that there wouldn't be any overlap in the explanatory powers of the variables used in the ordered models. The variables that were tested for correlation were those that would describe the size of the establishment, namely, employment, square footage, and total number of vehicles owned. These variables were the most likely to describe the same effects in the models and their correlations were taken into consideration in order that no conflicting results would be obtained from the models. The correlation between the continuous variables for employment and square footage of the establishment was 0.031, which indicates a weak correlation between these two variables. A high correlation was also found when comparing the continuous variables for employment and total number of vehicles owned by the establishments, with a correlation coefficient of 0.83. Lastly, when comparing square footage and total number of vehicles owned, no significant correlation was found, with a 0.081 correlation coefficient. Section 4.2 will explain which of the variables were used in the modeling efforts.

4.1.1 Validation of Quotient Approach for Stratified Sampling

While the Quotient method was efficient in cutting down the time required to spend contacting firms to participate in the web survey by targeting those firms that were most likely to be eligible and willing to participate, it was important to ensure that the targeting method did not introduce any bias into the data. To perform the validation, the group of survey respondents that belonged to the first phase of telephone recruitment calls was compared to the group of respondents that were contacted during the second phase. As mentioned in Chapter 3, there were 89 firms who filled the survey completely out of the firms who provided email addresses during phase one, and 82 who completed the survey out of those contacted in phase two. The distributions of these two groups with respect to several variables was compared.

The comparison variables were industry type (given by NAICS classification), employment size categories, as described in Table 3-5 of Chapter 3, and the interactions between NAICS code and employment size, and between NAICS code and total number of vehicles owned by each firm. The total number of vehicles owned was calculated as the sum of vehicles of all types that each establishment reported to own during the survey. Categories were created for this variable as well, where the number of vehicles was 1, 2, …, 9, 10 or more, and other. The "other" option was present because a number of firms did not provide any information about the number of vehicles they own.

Validation of the Quotient procedure was done through t-statistics tests. The null hypothesis tested in each case was that the mean of the distribution of each respective variable for the two groups of respondents was not different. Rejecting the null hypothesis would indicate that the distribution of the variables between the two samples was different, which would mean that there is a bias in the data. Tables were created for each of the variables tested. Table 4-1 shows the counts of firms in each of the two respondent groups who belonged to each industry classification, given by NAICS 2-digit codes. Table 4-2, similarly, shows the counts of firms in each respondent group who belonged to each employment size category.

Table 4-1: Number of Firms Belonging to Each Industry Category

NAICS-2D	Phase 1 Respondents	Phase 2 Respondents
23	3	0
31	1	2
32	2	2
33	15	20
42	16	13
44	15	27
45	5	4
48	2	6
49	1	2
52	1	0
53	1	0
54	4	0
56	2	0
61	2	0
62	3	0
72	2	6
81	14	0

Table 4-2: Number of Firms Belonging to Each Employment Size Category

Employment Category	Phase 1 Respondents	Phase 2 Respondents
1	63	43
2	13	18
3	5	9
4	2	1
5	1	3
6	2	4
7	0	1
8	2	1
9	0	0
10	1	2

The paired t-statistics tests for two samples were run on the basis of these tables using Excel's t-statistics data analysis tool. Table 4-3, below, shows the results of the t-statistics test for industry category. Among the resulting outputs, it can be seen that the critical t value for this type of two-tailed test was 2.12 and the t-statistic value obtained for the two samples tested was well below this critical value. Since this was the case, the null hypothesis could not be rejected in this case, indicating that the distribution of the two samples with respect to industry category is the same, at a 95% confidence level. This result is favourable, since it indicates that the survey respondents obtained during the first phase of telephone recruitment belonged to similarly distributed industry classes as those who responded after the second phase. The value for Pearson Correlation observed for the two respondent groups in the case of industry classification is also rather high, indicating a correlation between the NAICS distribution in the two samples. All of these results show that the two sample groups are similarly distributed, which acts to validate the Quotient procedure, showing that the targeting technique did not introduce a bias in the firms that were recruited for the web survey.

Table 4-3: Results from Paired T-Test for the Number of Firms by Industry Classification for the Two Respondent Groups

	Phase 1 Respondents	**Phase 2 Respondents**
Mean	5.24	4.82
Variance	32.44	62.65
Observations	17	17
Pearson Correlation	0.74	
Hypothesized Mean Difference	0	
df	16	
t Stat	0.32	
P(T<=t) one-tail	0.38	
t Critical one-tail	1.75	
P(T<=t) two-tail	0.75	
t Critical two-tail	2.12	

The results from the t-test run on the categorized employment size variable are similarly presented in Table 4-4. Upon examining the observed and critical t-values for this two-tailed test, it can be seen that a t-statistic value well below the critical is obtained in this case as well. This means that the null hypothesis, that the distributions of these two samples are not different, cannot be rejected. The distribution of employment categories, therefore, are considered to be the same at a 95% level of confidence. Examining the Pearson correlation measure for this test, it can be seen that the value is very close to 1, indicating that the two samples tested are highly correlated. These results show, once again, that the distribution of firms with respect to employment size was the same for firms who responded after the first phase of telephone calls as for the firms who responded after the second phase.

Table 4-4: Results from Paired T-Test for the Number of Firms by Employment Size Categorization for the Two Respondent Groups

	Phase 1 Respondents	**Phase 2 Respondents**
Mean	8.90	8.20
Variance	376.10	179.29
Observations	10	10
Pearson Correlation	0.97	
Hypothesized Mean Difference	0	
df	9	
t Stat	0.31	
P(T<=t) one-tail	0.38	
t Critical one-tail	1.83	
P(T<=t) two-tail	0.76	
t Critical two-tail	2.26	

A table was created for each of the two respondent groups to examine the number of firms in each industry category that belonged to each employment size category. This was done through pivot tables, where the columns were designated as employment size categories and the rows corresponded to NAICS codes. A table was also created for the two groups to examine the number of establishments in each industry that reported to own each amount of vehicles. Similarly, pivot tables were created with the number of vehicles owned as the column headings and the NAICS codes as the rows. The counts in each case were obtained and the data was arranged into a single column. T-tests were run for the two groups of respondents for the interactions of NAICS codes with employment categories, as well as for NAICS codes interacted with the number of owned vehicles.

Table 4-5 shows the t-stat test results comparing the distributions of the interaction of NAICS codes with employment categories for the two respondent groups. It can be seen that, similarly to the first two t-tests, the observed t-statistic value was well below the critical value, indicating that the null hypothesis was accepted and the distributions of the two samples were not different at a 95% level of confidence. The correlation between the two samples is also high, further proving that the two samples are similarly distributed.

Table 4-5: Results from T-Test for Industry Classification and Employment Categories for the Two Respondent Groups

	Phase 1 Respondents	**Phase 2 Respondents**
Mean	0.52	0.48
Variance	2.90	3.03
Observations	170	170
Pearson Correlation	0.75	
Hypothesized Mean Difference	0	
df	169	
t Stat	0.44	
P(T<=t) one-tail	0.33	
t Critical one-tail	1.65	
P(T<=t) two-tail	0.66	
t Critical two-tail	1.97	

Table 4-6 shows the results from the last t-test that was run, for the interaction between industry classification and the number of vehicles owned. As was the case with the previous three statistical tests, the t-stat value obtained here was below the critical value, meaning that the null hypothesis was not rejected, and the distributions of the two samples were considered the same at a 95% confidence level. The Pearson correlation value for this case was the lowest out of all four t-tests, however, it still indicates correlation between the distributions for the two groups of respondents. This fourth test, as with the previous three, demonstrates that the group of respondents who were recruited during the first round of telephone calls was similarly distributed to those recruited in the second phase, with respect to characteristics that describe the establishments.

Table 4-6: Results from T-Test for Industry Classification and Number of Vehicles for the Two Respondent Groups

	Phase 1 Respondents	**Phase 2 Respondents**
Mean	0.44	0.40
Variance	0.95	1.02
Observations	204	204
Pearson Correlation	0.53	
Hypothesized Mean Difference	0	
Df	203	
t Stat	0.51	
P(T<=t) one-tail	0.30	
t Critical one-tail	1.65	
P(T<=t) two-tail	0.61	
t Critical two-tail	1.97	

These results all validate the use of the LQ analysis and targeting method. The simple observation that the numbers of respondents from each of the two phases, 89 from phase one and 82 from phase two, were close. Based on these t-test results, it can be seen that the characteristics of the firms who responded out of phase one were similar to those of firms who responded from phase two. This validation was important because, following the LQ analysis, only certain industry categories were selected for contact by telephone and recruitment to the web survey. Using this targeting method had to be done with care to ensure that the firms who responded from the targeted sample were not biased toward certain industry categories or sizes (given by number of employees and number of owned vehicles), especially since only certain industry categories were selected for contact during Phase 2. The t-statistics and correlations from comparing these variables between the two groups of respondents show that no major bias was introduced. The LQ targeting method was, therefore, efficient in reducing the time spent recruiting firms, and also accurate in providing results similar to those that were obtained from contacting establishments in all categories.

4.1.2 Web Survey Results

The results from the web survey were analyzed before being introduced into the ordered models. The variables that were used for the purpose of this project were those collected from the establishment characteristics section of the web survey, namely, the eighteen variables listed in section 4.1. The current section will present a number of frequency tables created on the basis of those variables, for a preliminary analysis of the existing conditions and characteristics of the respondent firms, who engage in shipping or receiving of goods or services in the study area. The effects of each of these variables on commercial trip generation will be examined in detail in Section 4.2.

Figure 4-2 below shows a graph representing the number of firms having each reported number of employees. It can be seen that a significant portion of the respondent firms (approximately 54%) have a total of eight or fewer employees. This indicates that many of the local businesses are of smaller size, which could mean that they receive smaller shipments by smaller vehicles. It could also be an indicator that these smaller businesses receive less frequent deliveries.

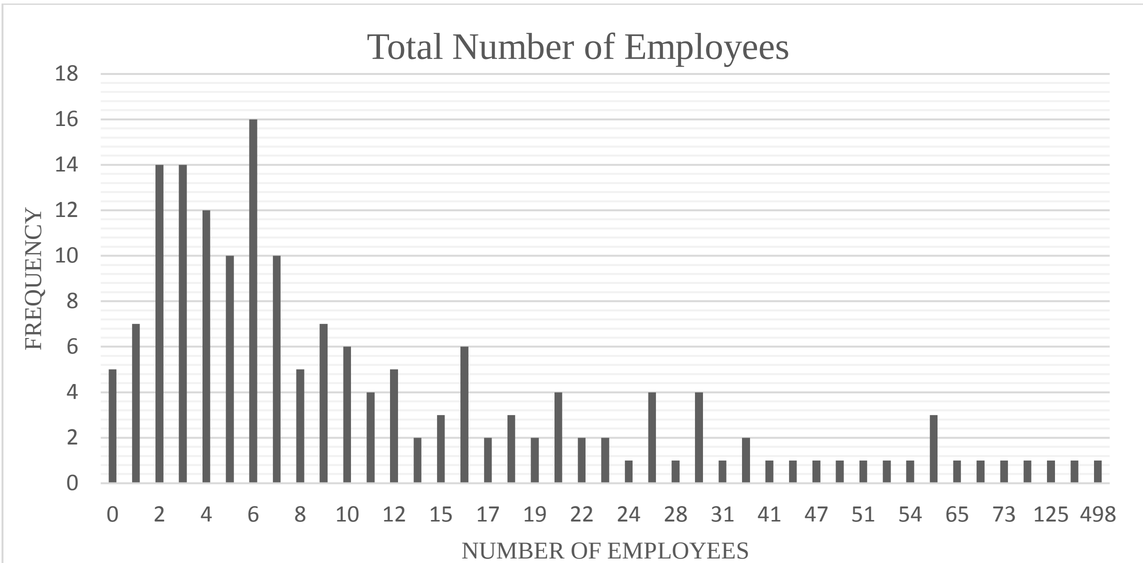


Figure 4-2: Frequency of Reported Total Number of Employees for Survey Respondents

One of the questions in the web survey asked respondents to choose one of fifteen industry classification that best described their firms. While these classifications were based on NAICS categories, it was considered more accurate to use the actual NAICS code for the firms, rather than the subjective responses of the representatives answering the survey. Table 4-7 lists all of the NAICS categories to which survey respondent firms belonged, along with a description of each category. Following this table, Figure 4-3 shows the frequency distribution of these industry classes among the respondent firms. It can be seen that a significant number of firms belongs to NAICS categories for "manufacturing," "wholesale trade," and "retail trade." Intuitively, this makes sense, since it would be these firms who receive regular shipments of goods to be sold.

Table 4-7: NAICS Codes and Descriptions Corresponding to Survey Respondents

NAICS	**Description**
23	Construction
31	Manufacturing
32	Manufacturing
33	Manufacturing
42	Wholesale Trade
44	Retail Trade
45	Retail Trade
48	Transportation and Warehousing
49	Transportation and Warehousing
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
56	Administrative and Support and Waste Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
72	Accommodation and Food Services
81	Other Services (except Public Administration)

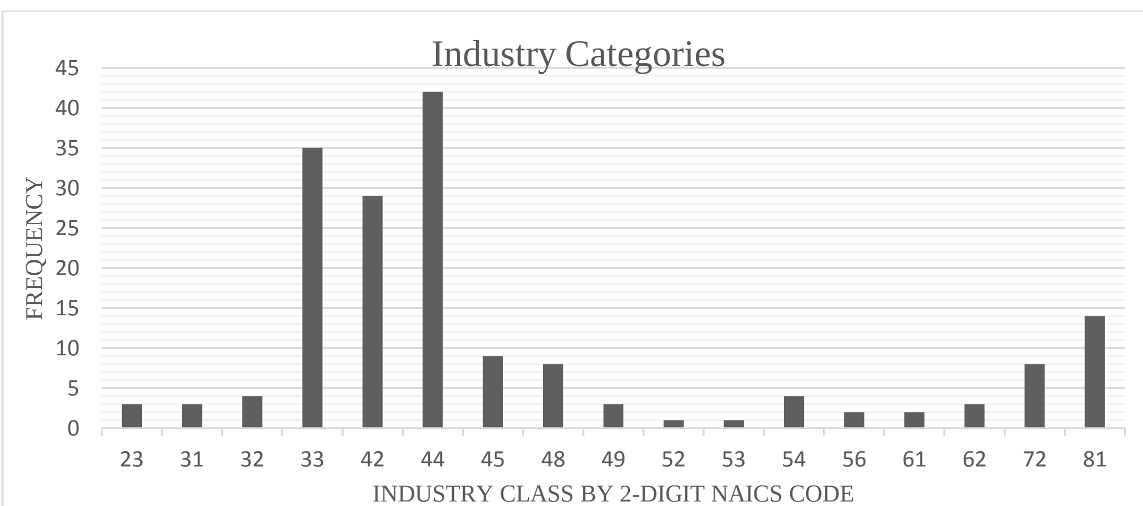


Figure 4-3: Frequency of Industry Classifications of Survey Respondents

The number of years that each respondent firm has been at its current location is examined in Figure 4-4. These values are based on each firm's response to this question in the web survey. It can be seen that the majority of firms in the study area are newer firms, having been at their current locations for under 20 years, with a large portion having fewer than eight years at their locations. The implications of this factor will be discussed further in the next section.

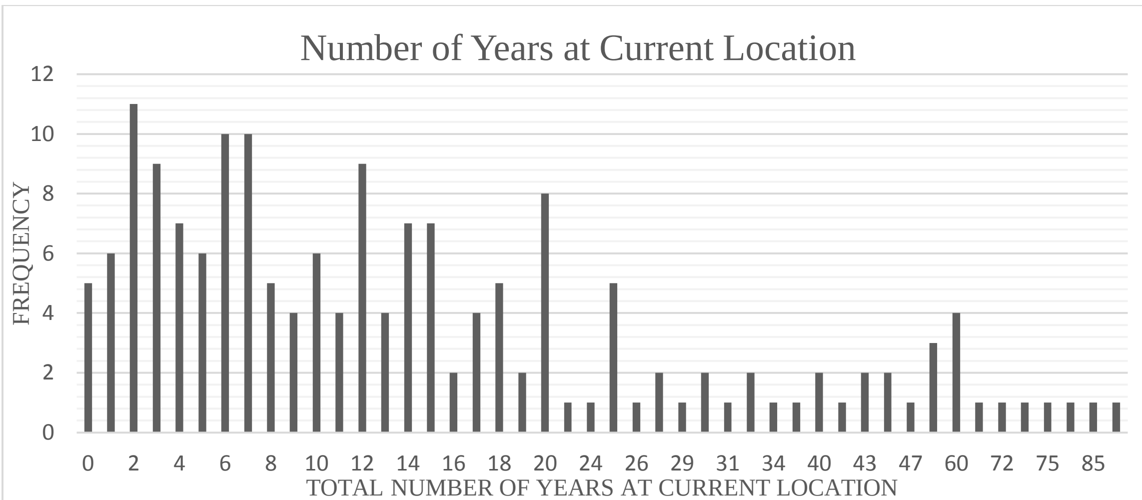


Figure 4-4: Frequency of Number of Years at Current Location for Survey Respondents

The web survey collected information about the number of employees who make on-call service trips for each firm. These values are examined in Figure 4-5, where it can be seen that the majority of firms have three or fewer employees who make on-call service trips, with a large proportion who answered that they have no employees in this function. This finding is consistent to other results found in the literature.

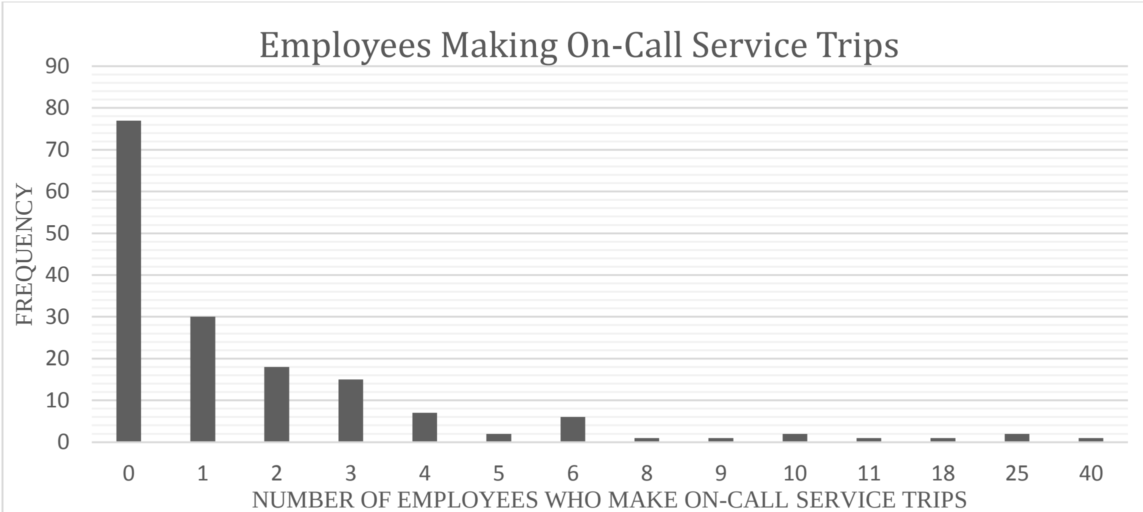


Figure 4-5: Frequency of Number of Employees Who Make On-Call Service Trips for Survey Respondents

Similarly, the number of employees who make truck deliveries for each firm was also collected through the web survey and is presented in Figure 4-6, below. It can be seen that the results are similar to the previous figure, where the majority of firms have two or fewer employees in this function, with a large portion reporting zero such employees. This is, again, consistent with previous findings in the literature.

Figure 4-6: Frequency of Number of Employees who Make Truck Deliveries for Survey Respondents

The survey asked three questions which were left unanswered by a portion of the firms. These questions asked whether the firms had a supply chain or logistics operator, whether the firms had a supply chain or logistics specialist, and whether the firm utilized logistics software. The reason that some firms did not provide an answer to these questions could be that the representatives who completed the survey were unsure of the definition of a supply chain or logistics operator or specialist. The representatives may also not have had full knowledge of this aspect of the firm. Table 4-8 below shows the frequency of the responses to these three questions. It can be seen that a negative response was given in the majority of cases for all three questions. If the size of the establishments is considered, given by the total number of employees, this result makes sense. It is usually larger establishments, who engage in moving larger quantities of goods regularly, who are the ones employing supply chain or logistics personnel and who require the organizational capabilities of logistics software to organize their trips. Since many local firms are smaller in size, they would not necessarily behave this way.

Table 4-8: Frequency of Responses to Survey Questions Regarding Having a Supply Chain/Logistics Operator or Specialist, and Use of Logistics Software

	No answer	**No**	**Yes**	**Total**
Have supply chain/logistics operator?	17	128	26	171
Have supply chain/logistics specialist?	15	133	23	171
Use logistics software?	20	132	19	171

Information was collected through the survey about the number of vehicles of each type that was owned by each establishments. Eight categories of vehicle types were available and respondents were able to respond to as many categories as applied to their establishments. Table 4-9 shows the aggregated responses to this question, displaying the number of vehicles of each kind by each industry classification. It can be seen that the most vehicles of all kinds are owned by the three prevalent industry classifications, which were discussed earlier. Again, the firms provided answers for only those vehicle categories that applied to their establishments. Not all establishments own every kind of vehicle type and some establishments reported not owning any vehicles of any kind.

Table 4-9: Number of Vehicles Owned By Vehicle Type and Industry Category

	Vehicle Types							
NAICS Code	Regular passenger cars	Pickup/ cube vans	Single-unit trucks	Multi-unit trucks	Tractors	Containers	Trailers	Other
23	2	4	4	0	0	0	0	0
31	2	0	5	0	0	0	0	0
32	5	17	3	0	0	0	0	0
33	36	63	23	1	13	200	28	0
42	21	22	19	1	6	30	13	0
44	39	59	39	0	4	5	8	1
45	1	4	0	0	0	0	0	0
48	0	22	5	0	106	0	534	0
49	4	21	2	0	0	0	0	0
52	0	0	0	0	0	0	0	0
53	0	1	0	0	0	0	0	0
54	6	11	4	0	0	0	0	0
56	3	1	8	0	0	0	0	0
61	2	0	0	0	0	0	0	0
62	1	0	0	0	0	0	0	0
72	4	2	0	0	0	0	0	0
81	18	14	1	0	2	0	4	0

The total number of vehicles was determined for each firm by adding the reported numbers of vehicles of all types. Figure 4-7 shows the distribution of the reported total number of vehicles owned by each firm. From this graphic, it can be observed that most firms own four or fewer vehicles (of any type). A significant number of firms do not report owning any vehicles of any type. This could be related with the establishment sizes, since it has been observed that the study area has a large portion of smaller establishments who may not own their own vehicles but receive deliveries or hire vehicles for their commercial activities.

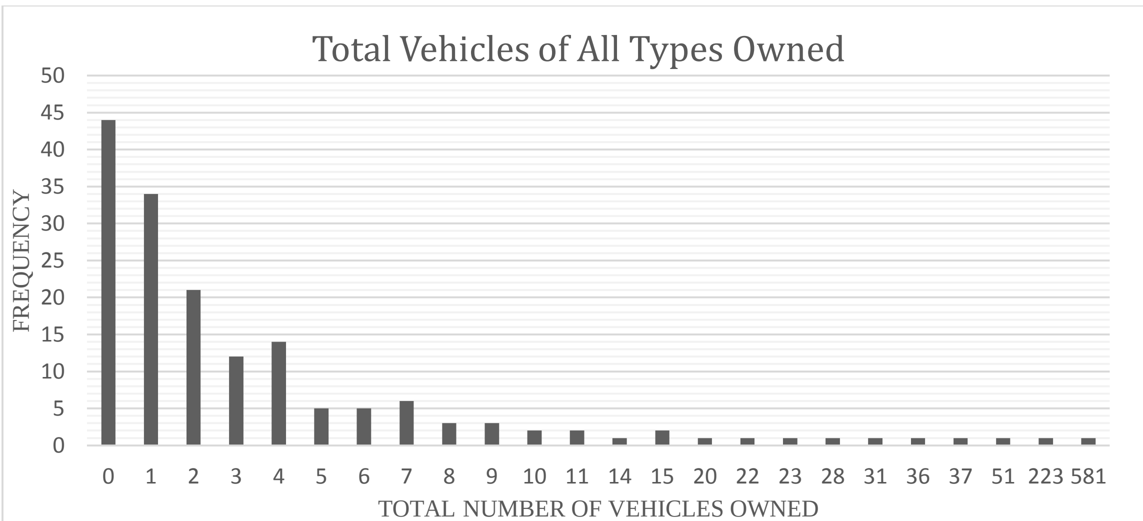


Figure 4-7: Frequency of Total Number of Vehicles of All Vehicle Types Owned by Survey Respondents

The survey asked respondents whether the number of commercial vehicle trips for their establishment varied by day of the week and by month throughout the year. As can be seen in Table 4-10, the majority of firms answered affirmatively to both of these questions. Based on these responses, firms were then asked to provide which days of the week and months throughout the year were considered the busiest for commercial vehicle activities at their establishments. In order to simplify the responses received for these two questions, which were extremely varied due to the fact that respondents could select as many or as few days or months as their busiest, some categories were created. The days of the week and the months of the year were each separated into four quadrants. The answers of each firm were then examined and it was determined under which quadrant (or quadrants) they fell. The permutations resulting from all possible combinations of these four quadrants yielded 15 categorical variables, as described in Table 4-11. The frequencies of responses falling under each category for busiest days and busiest months are subsequently given in Figure 4-8 and Figure 4-9, where a category of 0 was assigned to those who did report variability in the number of trips but did not provide the actual busiest days and months.

Table 4-10: Frequency of Responses to Survey Questions Regarding Busiest Days and Months for Deliveries

	No answer	**No**	**Yes**	**Total**
Number of Trips Varies by Day?	16	40	115	171
Number of Trips Varies by Month?	19	36	116	171

Table 4-11: Description of Categories for Busiest Days of the Week and Busiest Months

Category	**Days of the week**	**Months**
1	Monday, Tuesday	January-March
2	Wednesday	April-June
3	Thursday, Friday	July-September
4	Saturday, Sunday	October-December
5	Monday-Wednesday	January-June
6	Monday, Tuesday, Thursday, Friday	January-March, July-September
7	Monday, Tuesday, Saturday, Sunday	January-March, October-December
8	Wednesday-Friday	April-September
9	Wednesday, Saturday, Sunday	April-June, October-December
10	Thursday-Sunday	July-December
11	Monday-Friday	January-September
12	Monday-Wednesday, Saturday-Sunday	January-June, October-December
13	Monday, Tuesday, Thursday-Sunday	January-March, July-December
14	Wednesday-Sunday	April-December
15	Monday-Sunday	January-December

These two variables for the categorized busiest days of the week and months of the year were introduced as control variables in the model, rather than explanatory variables for trip generation. The complexities of these combinations should be studied further in future research.

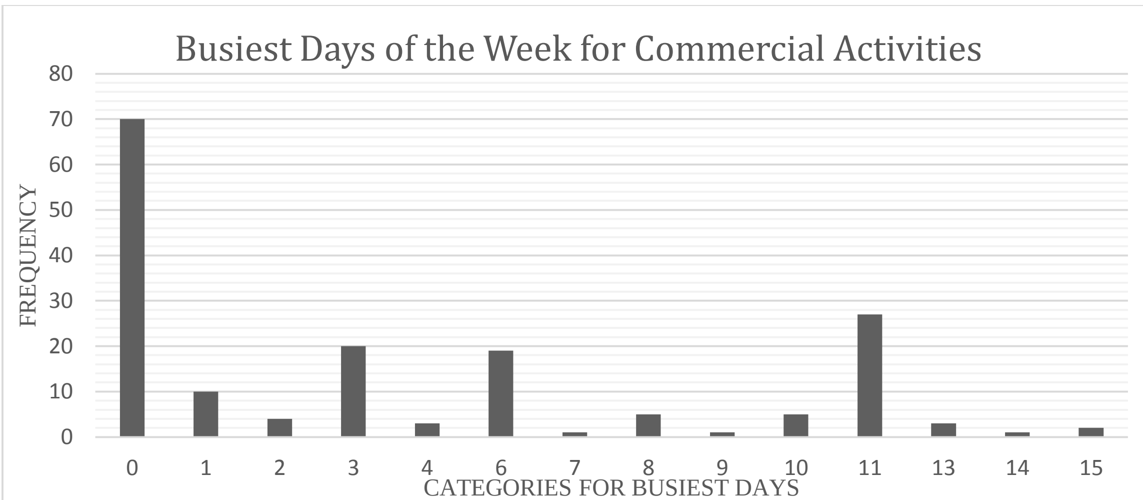


Figure 4-8: Frequency of Busiest Days for Commercial Activity for Survey Respondents

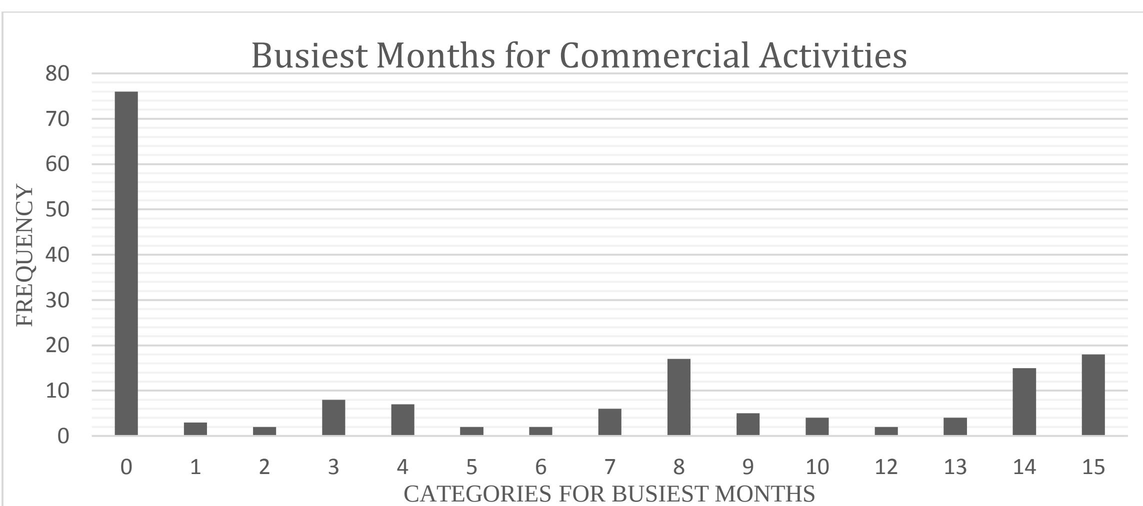


Figure 4-9: Frequency of Busiest Months for Commercial Activity for Survey Respondents

From the two preceding figures, it can be observed that the days of the week that are considered to be the busiest with respect to commercial trips belong to the categories for Thursday-Friday (right before the weekend), for Monday-Tuesday-Thursday-Friday (right before and right after the weekend), and Monday-to-Friday (during the work week). When considering that the retail trade, wholesale trade, and manufacturing industries are the most prevalent, this makes sense, as these would be times when products and materials can be restocked. When examining the busiest months, it is observed that the categories for April-to-September (spring/summer months) April to December (spring/summer/autumn months) and January-to-December (year-round) are the busiest. This also makes sense, since businesses are often more busy and receive increased patronage during the warmer months of the year. For the large proportion of retail and wholesale trade, and manufacturing industries, it also makes sense that shipments would be sent out from and arrive at the establishments year-round. It should be noted that there were also a large number of firms who neglected to provide answers to these two questions, although they did indicate that some days or months were busier (designated by a value of 0 in the Figures). This could be attributed to the assumption that the representative filling the web survey was not knowledgeable about these patterns in the establishment or because they were unable to spend the time to analyze their commercial vehicle patterns.

The survey collected information about the physical size of each establishment as well, in the form of the square footage of each firm's floor area. Since these values were also quite widely varied, categories were created to represent this data, described in Table 4-12 below. Analyzing the data obtained from this question through Figure 4-10, it can be seen that the majority of firms have smaller floor areas, of up to 2,000 square feet, or mid-sized floor areas, of between 5,000 and 20,000 square feet. This fits with the other measures of establishment size, by employment size and number of vehicles owned, which indicated that the majority of local businesses were smaller. It should be noted that a number of firms either reported a value of zero square feet or did not provide an answer to this question. The responses of zero square feet could be considered as businesses who do not have an actual storefront, such as online businesses. They could also be attributed to businesses that are run out of an individual's home.

Table 4-12: Categories for Square Footage Values

Category	Range (square ft.)
1	0
2	1 – 1,000
3	1,001 – 2,000
4	2,001 – 3,000
5	3,001 – 4,000
6	4,001 – 5,000
7	5,001 – 10,000
8	10,001 – 20,000
9	20,001 – 30,000
10	30,001 – 50,000
11	50,001 – 100,000
12	> 100,000

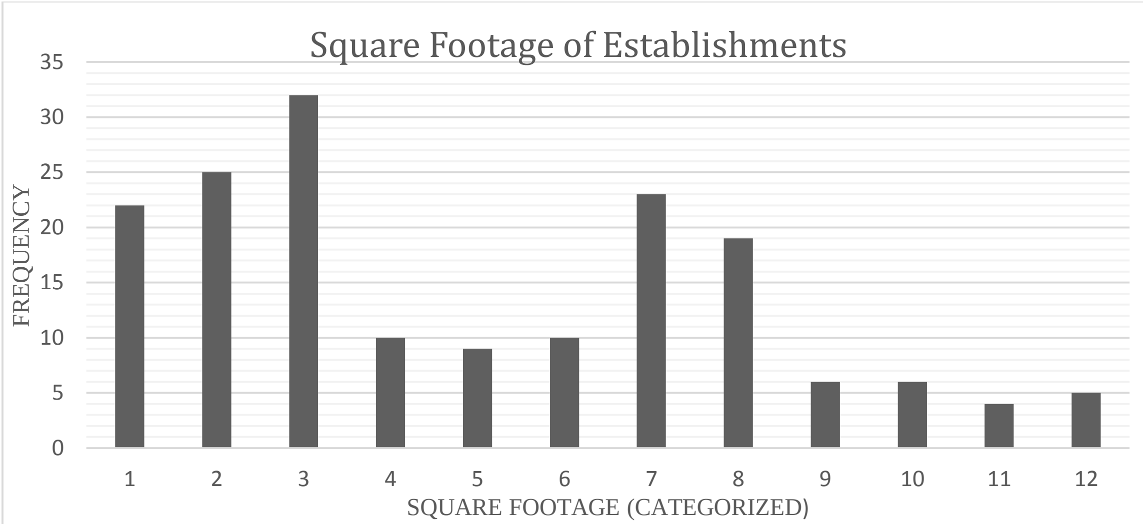


Figure 4-10: Frequency of Categorized Square Footage Values for Survey Respondents

Lastly, the survey asked respondents to provide the numbers of commercial vehicles that had left and that had arrived at the establishment on the day that the survey was completed, to deliver goods or services. The frequencies of these responses are summarized in Figures 4-11 and 4-12 shown below.

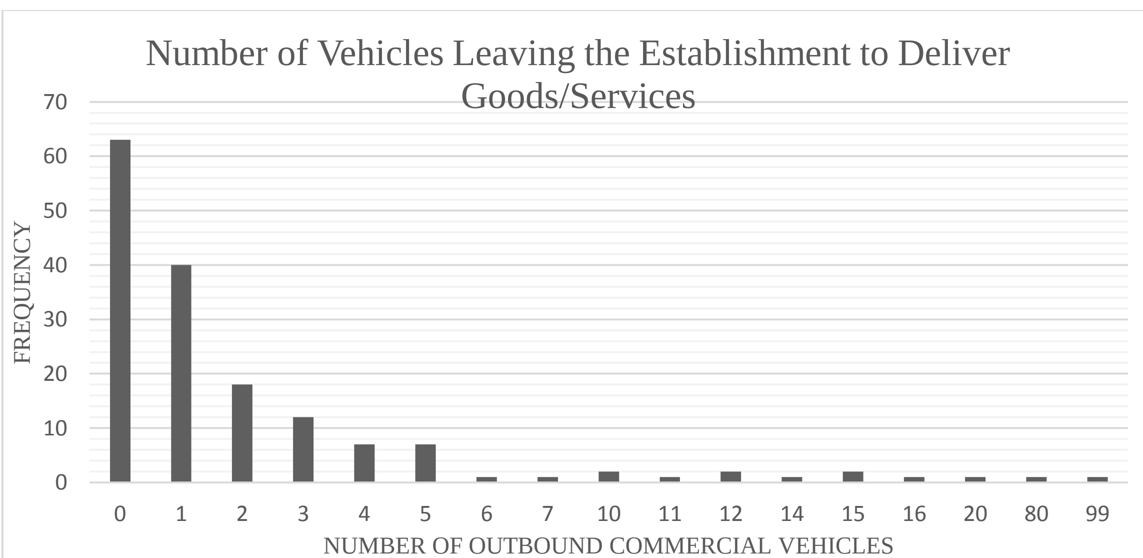


Figure 4-11: Frequency of Number of Vehicles Leaving the Establishment to Deliver Goods or Services

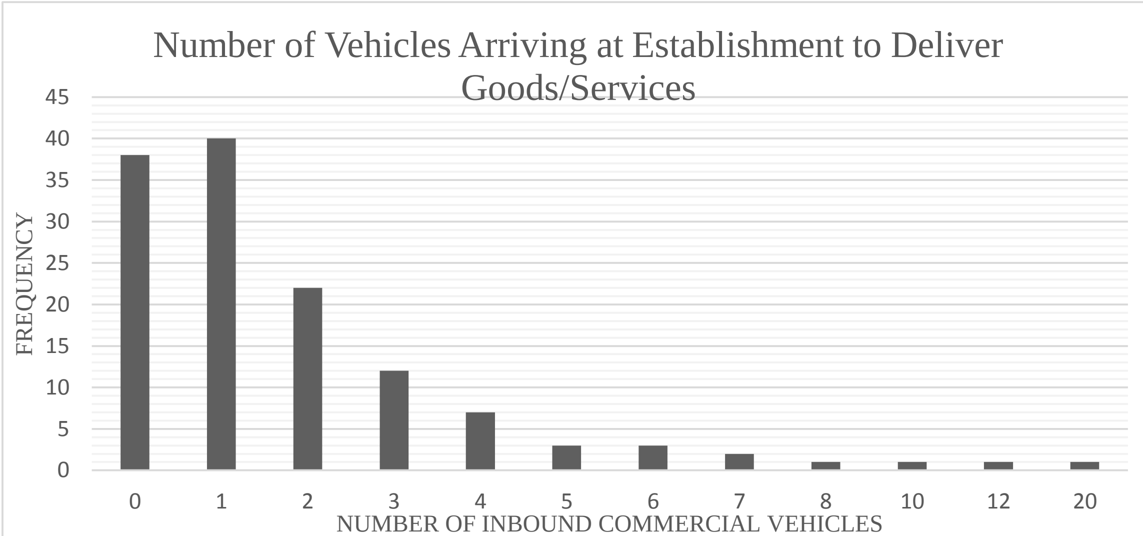


Figure 4-12: Frequency of Number of Vehicles Arriving at Establishment to Deliver Goods or Services

Upon examining these final two figures, the results seem to be in agreement with findings from the literature, where the majority of firms have between zero and three inbound or outbound commercial vehicle trips on the day the survey was completed. It can be seen that a very small number of firms reported a very large number of outbound trips. Given that the Windsor study area is home to a number of large automotive manufacturers, this result makes sense. For the purpose of the ordered model, the numbers of inbound and outbound trips for each establishment were each categorized into one of four categories, namely 0 (for no trips occurring on the survey day), 1 (corresponding to one incoming or outgoing commercial trip), 2 (representing two trips), and 3 (which included all reported values of three or more commercial trips). The continuous variables for the numbers of outbound and inbound commercial vehicles have a correlation value of 0.35, while the categorized variables for outbound and inbound commercial vehicles resulted in a correlation value of 0.17. These variables, specifically the categorical variables for inbound and outbound movements, were used as the dependent variables in the models and, therefore, were not used simultaneously. It can be seen from Table 3-6 in the previous chapter that a significant number of firms report having no outbound and no inbound commercial trips on the day of the survey (73 and 78 firms, respectively).

As mentioned earlier, the web survey consisted of three portions. The results described in here are those extracted from the first of these portions, representing establishment characteristics. The other two sections collected data about specific inbound and outbound commercial trips that each firm experienced on the day that the survey was completed. These two sets of questions were asked to some respondents multiple times. Depending on the number of incoming and outgoing vehicles that they reported, the set of questions was asked once for each incoming or outgoing trip. For this reason, some firms have multiple corresponding entries of data, which are not able to be analyzed in the same manner as the variables at hand. The questions and possible answers are available in Appendix B and it can be seen that, while some can be categorized numerically for statistical analysis, others, such as the location of the origin or destination of the trip, cannot be assigned numerical values. Due to such difficulties in categorizing the data, these two sections of the web survey were not introduced into the present trip generation model. It is advisable, however, for future research to be undertaken using these variables, since they provide important information about the nature of the commodities carried, and the behaviour of the commercial vehicles with respect to their origin or destination and the time of day when the trips were undertaken. The current modeling effort is outlined in detail in the following section.

4.2 Modeling of Trip Generation

The modeling and analysis of variables affecting inbound and outbound commercial trip generation was conducted using NLOGIT 4.0 software (2007). As described in Chapter 3, ordered logit models were selected and developed separately for inbound movements and outbound movements, to determine the factors giving rise to each, for a certain business establishment. The ordered models would determine the probability that a certain establishment would produce zero, one, two, or three or more incoming or outgoing commercial vehicle trips on a given day. Both logit and probit models were considered for the outbound and inbound movements. The calculated weight factors were introduced into all tested models, since they improved the fit of the models and increased the number of variables that became statistically significant. This section will describe the development and results of the ordered models.

As previously mentioned, the variables used in the analysis were those collected from the web survey pertaining to general establishment characteristics. All of these variables were categorized and assigned numerical values. Eight sets of categorical variables resulted, described in detail in Table 4-13. The first group of categorical variables was determined directly from the web survey responses and contained data about the number of employees in each occupation type at the respective business establishment. This categorical variable was named OCC_k, where *k* represented different occupation types. Also collected directly from the survey was the variable VEH_k, which described the number of vehicles of each type owned by the respective firms. As previously stated, the industry classification of survey respondents was done by using the official two-digit NAICS codes assigned to the firms. This information was categorized under the variable IND_k, which represented each different industry class. The self-designated industry classifications collected during the web survey were not used here, since they were subjective and not necessarily accurately reflecting the type of industry to which a firm belonged.

The categories described in the previous section for the busiest days and busiest months for deliveries to and from each establishment that were described in the preceding section were also introduced into the ordered modeling, under the categorical variables BDY_k for the busiest days and BMT_k for the busiest days. Three additional categorical variables were created on the basis of continuous variables obtained from the web survey. The variable EMP_k corresponds to the categories developed for intervals for the number of employees at each establishment, as previously explained. Categorical variable YAL_k describes ranges for the number of years that a respective establishment has been at its current location. Lastly, SQF_k describes ranges for the square footage of each firm's floor area, which was also described in the previous section. It is noteworthy that some of these categories, BDY_5, BDY_{12}, BMT_{11}, and EMP_9, are empty, since no survey answer fell under them.

Table 4-13: Descriptions of Categorical Variables Created for Use in the Analysis

Variable	**Description**
Occupation types ***OCC_k*** **(k = 1, 2, …, 11)**	Number of employees in each occupation type k at each establishment. Available types are: k = 1 for management occupations; 2 for business, finance, and administration occupations; 3 for natural and applied sciences, and related occupations; 4 for health occupations; 5 for social science, education, government service, and religion occupations; 6 for art, culture, recreation, and sport occupations; 7 for sales and service occupations; 8 for trade, transportation, and equipment operation occupations; 9 for primary industry occupations; 10 for processing, manufacturing, and utilities; 11 for other occupation types.
Vehicle types ***VEH_k*** **(k = 1, 2, …, 8)**	Number of vehicles of each type k owned by each establishment. Available types are: k = 1 for regular passenger cars; 2 for pickup trucks and cube vans; 3 for single-unit trucks; 4 for multi-unit trucks; 5 for tractors; 6 for containers; 7 for trailers; 8 other vehicle types.
Employment size class ***EMP_k*** **(k = 1, 2, .., 9)**	1 if the establishment belongs to employment size class k, 0 otherwise. Available classes are: k = 1 for size class 1 to 10 employees; 2 for 11 to 20 employees; 3 for 21 to 30 employees; 4 for 31 to 40 employees; 5 for 41 to 50 employees; 6 for 51 to 60 employees; 7 for 61 to 70 employees; 8 for 81 to 90 employees; and 9 for 91 or more employees.

Table 4-13(cont'd): Descriptions of Categorical Variables Created for Use in the Analysis

Variable	Description
Industry types **IND_k** **(k =1, 2, …, 17)**	1 if the business establishment specializes in industry type k, 0 otherwise. Available types are: k = 1 for construction industry (NAICS 2-digit code 23); 2 for manufacturing industry (NAICS 31); 3 for manufacturing industry (NAICS 32); 4 for manufacturing industry (NAICS 33); 5 for wholesale trade industry (NAICS 42); 6 for retail trade industry (NAICS 44); 7 for retail trade industry (NAICS 45); 8 for transportation and warehousing industry (NAICS 48); 9 for transportation and warehousing industry (NAICS 49); 10 for finance and insurance industry (NAICS 52); 11 real estate rental and leasing (NAICS 53); 12 for professional, scientific, and technical services industry (NAICS 54); 13 for administrative and support and waste management and remediation services (NAICS 56); 14 for educational services industry (NAICS 61); 15 for administrative and support and waste management and remediation services industry (NAICS 62); 16 for accommodation and food services industry (NAICS 72); 17 for other services, except public administration (NAICS 81).
Busiest days **BDY_k** **(k = 1, 2, …, 15)**	1 if busiest day of commercial movement activities is k, 0 otherwise. Available days are: k = 1 for Monday-Tuesday; 2 for Wednesday; 3 for Thursday-Friday; 4 for Saturday-Sunday; 5 for Monday-Wednesday; 6 for Monday-Tuesday, Thursday-Friday; 7 for Monday-Tuesdays, Saturday-Sunday; 8 for Wednesday-Friday; 9 for Wednesday, Saturday-Sunday; 10 for Thursday-Sunday; 11 for Monday-Friday; 12 for Monday-Wednesday, Saturday-Sunday; 13 for Monday-Tuesday, Thursday-Sunday; 14 for Wednesday-Sunday; 15 for Monday-Sunday.

Table 4-13(cont'd): Descriptions of Categorical Variables Created for Use in the Analysis

Variable	Description
Busiest months ***BMT_k*** **(k = 1, 2, …, 15)**	1 if busiest month of commercial movement activities is *k*, 0 otherwise. Available months are: k = 1 for January-March; 2 for April-June; 3 for July-September; 4 for October-December; 5 for January-June; 6 for January-March and July-September; 7 for January-March and October-December; 8 for April-September; 9 for April-June, October-December; 10 for July-December; 11 for January-September; 12 for January-June and October-December; 13 for January-March and July-December; 14 for April-December; 15 for January-December
Years at location ***YAL_k*** **(k = 1, 2, …, 8)**	1 if the number of years the business establishment has been at the current location is *k*, 0 otherwise. Available classes are: k = 1 for reported values of 0; 2 for 1 to 10 years at the firm's current location; 3 for 11 to 20 years; 4 for 21 to 30 years; 5 for 31 to 40 years; 6 for 41 to 50 years; 7 for 51 to 80 years; 8 for 81 or more years.
Square footage ***SQF_k*** **(k = 1, 2, …, 13)**	1 if the square footage of the business establishment belongs to class k, 0 otherwise. Available classes are: k = 1 for reported values of 0; 2 for 1 to 1,000 sq. ft.; 3 for 1,001 to 2,000 sq. ft.; 4 for 2,001 to 3,000 sq. ft.; 5 for 3,001 to 4,000 sq. ft.; 6 for 4,001 to 5,000 sq. ft.; 7 for 5,001 to 10,000 sq. ft.; 8 for 10,001 to 20,000 sq. ft.; 9 for 20,001 to 30,000 sq. ft.; 10 for 30,001 to 50,000 sq. ft.; 11 for 50,001 to 100,000 sq. ft.; 12 for 100,001 sq. ft. or more.

Four models were tested, a logit and a probit model for the number of outgoing commercial vehicles and for the number of incoming commercial vehicles. Each of the eight groups of categorical variables above was introduced into each of the four models tested, in order to assess their individual effects on trip generation and their explanatory strengths. The McFadden Pseudo R-squared value from each categorical variable's introduction into each of the four models is presented in Table 4-14. A higher pseudo R-squared value indicates that a variable has a stronger effect. The values in brackets in this table represent the ranks of the pseudo R-squared values among the eight categorical variables within each of the four types of ordered models. The rank of 1 was assigned to the variable in each of the four models that resulted in the highest pseudo R-squared value out of the eight variables for that model. These ranks were used when building the final ordered models, where the categorical variables with the highest pseudo R-squared values were added in first, followed by the rest of the variables, in order of decreasing pseudo R-square value. Statistically insignificant variables at a 90% confidence level were removed and the models were re-tested until only statistically significant variables were present. Correlated variables were not introduced into the specified models.

Table 4-14: McFadden Pseudo R-squared Values for Each Group of Categorical Variables, for each of the Four Ordered Models (Values in Parentheses Represent the Ranking of the Categories)

Model	**Variable category**							
	OCC_k	**VEH_k**	**IND_k**	**EMP_k**	**BDY_k**	**BMT_k**	**YAL_k**	**SQF_k**
OB Logit	0.163 *(5)*	0.304 *(1)*	0.237 *(2)*	0.148 *(8)*	0.201 *(3)*	0.160 *(6)*	0.159 *(7)*	0.180 *(4)*
OB Probit	0.151 *(7)*	0.305 *(1)*	0.245 *(2)*	0.144 *(8)*	0.220 *(3)*	0.171 *(5)*	0.167 *(6)*	0.176 *(4)*
IB Logit	0.017 *(7)*	0.048 *(3)*	0.049 *(2)*	0.019 *(6)*	0.015 *(8)*	0.028 *(5)*	0.037 *(4)*	0.053 *(1)*
IB Probit	0.014 *(7)*	0.065 *(1)*	0.046 *(2)*	0.013 *(8)*	0.018 *(6)*	0.027 *(5)*	0.033 *(4)*	0.040 *(3)*

After a period of testing the four different models, it was decided that the two logit models, for outgoing and incoming commercial trips, would be used as the final specification. This was because the logit models showed a better fit than the probit models, through higher pseudo R-squared values. Tables 20 and 21 summarize the results from the individual modeling of the categorical variables in the two logit models. The estimation values and t-statistics are given, along with the threshold values, the pseudo R-squared value, and the name of the reference variable in each case.

Table 4-15: Results from Logit Models of Categorical Variables, Outbound Case

Occupation types		Vehicle types		Industry types		Employment size		Busiest days		Busiest months		Years at location		Square footage	
Var.	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**
Const.	-0.580 *(-2.70)*	**Const.**	-0.884 *(-4.236)*	**Const.**	-1.427 *(-1.861)*	**Const.**	-0.446 *(-2.621)*	**Const.**	-0.994 *(-3.958)*	**Const.**	-0.650 *(-3.012)*	**Const.**	-0.385 *(-0.493)*	**Const.**	-0.184 *(-0.463)*
OCC_2	0.250 *(1.215)*	**VEH_2**	0.752 *(5.451)*	**IND_1**	0.098 *(0.075)*	**EMP_2**	1.525 *(3.472)*	**BDY_1**	1.410 *(2.318)*	**BMT_1**	0.374 *(0.249)*	**YRS_2**	0.036 *(0.044)*	**SQF_2**	-0.609 *(-1.176)*
OCC_3	-1.910 *(-1.375)*	**VEH_3**	0.896 *(4.528)*	**IND_2**	2.306 *(1.784)*	**EMP_3**	1.253 *(1.565)*	**BDY_2**	0.514 *(0.568)*	**BMT_2**	-28.215 *(0.0)*	**YRS_3**	-0.126 *(-0.153)*	**SQF_3**	-0.284 *(-0.559)*
OCC_4	0.020 *(0.22)*	**VEH_4**	0.923 *(0.485)*	**IND_3**	2.884 *(2.608)*	**EMP_4**	1.181 *(0.541)*	**BDY_3**	0.742 *(1.573)*	**BMT_3**	0.750 *(0.928)*	**YRS_4**	0.924 *(0.973)*	**SQF_4**	-0.879 *(-1.073)*
OCC_5	-0.067 *(-0.406)*	**VEH_5**	0.944 *(0.833)*	**IND_4**	2.408 *(2.804)*	**EMP_5**	2.976 *(2.225)*	**BDY_6**	2.346 *(4.724)*	**BMT_4**	0.736 *(1.35)*	**YRS_5**	1.528 *(1.343)*	**SQF_5**	0.030 *(0.036)*
OCC_6	0.609 *(0.378)*	**VEH_6**	-0.921 *(0.0)*	**IND_5**	1.653 *(1.858)*	**EMP_6**	-0.270 *(-0.134)*	**BDY_7**	57.503 *(0.0)*	**BMT_5**	-0.408 *(-0.237)*	**YRS_6**	0.555 *(0.522)*	**SQF_6**	-0.550 *(-0.743)*
OCC_7	0.0005 *(0.012)*	**VEH_7**	0.149 *(0.191)*	**IND_6**	1.521 *(1.861)*	**EMP_7**	28.163 *(0.0)*	**BDY_8**	0.807 *(0.69)*	**BMT_6**	1.996 *(1.749)*	**YRS_7**	2.922 *(2.691)*	**SQF_7**	1.417 *(2.417)*
OCC_8	0.156 *(2.433)*	**VEH_8**	0.477 *(0.339)*	**IND_7**	-0.063 *(-0.062)*	**EMP_8**	1.952 *(0.937)*	**BDY_9**	-25.234 *(0.0)*	**BMT_7**	0.896 *(1.112)*	**YRS_8**	3.135 *(2.174)*	**SQF_8**	0.841 *(1.293)*
OCC_9	0.002 *(0.236)*			**IND_8**	2.971 *(2.741)*	**EMP_{10}**	2.121 *(1.592)*	**BDY_{10}**	0.953 *(1.411)*	**BMT_8**	1.553 *(3.414)*			**SQF_9**	2.418 *(1.965)*
OCC_{10}	0.045 *(0.983)*			**IND_9**	59.664 *(0.0)*			**BDY_{11}**	2.329 *(4.731)*	**BMT_9**	1.662 *(2.369)*			**SQF_{10}**	2.527 *(2.109)*
OCC_{11}	0.173 *(2.42)*			**IND_{10}**	-29.775 *(0.0)*			**BDY_{13}**	4.867 *(2.057)*	**BMT_{10}**	-0.114 *(-0.112)*			**SQF_{11}**	3.583 *(1.668)*

Table 4-15 (cont'd): Results from Logit Models of Categorical Variables, Outbound Case (Continued)

Occupation types		Vehicle types		Industry types		Employment size		Busiest days		Busiest months		Years at location		Square footage	
Var.	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**
				IND_{11}	2.127 *(1.607)*			**BDY_{14}**	90.068 *(0.0)*	**BMT_{12}**	-0.994 *(-0.34)*			**SQF_{12}**	2.421 *(1.821)*
				IND_{12}	-0.729 *(-0.453)*			**BDY_{15}**	1.990 *(1.168)*	**BMT_{13}**	2.788 *(2.03)*				
				IND_{13}	6.311 *(2.198)*					**BMT_{14}**	1.327 *(2.224)*				
				IND_{14}	2.127 *(1.429)*										
				IND_{15}	-28.204 *(0.0)*										
				IND_{16}	0.464 *(0.507)*										
μ_1	1.261 *(7.311)*	**μ_1**	1.737 *(7.719)*	**μ_1**	1.399 *(7.673)*	**μ_1**	1.212 *(7.355)*	**μ_1**	1.370 *(7.629)*	**μ_1**	1.253 *(7.473)*	**μ_1**	1.274 *(7.317)*	**μ_1**	1.308 *(7.473)*
μ_2	1.944 *(8.404)*	**μ_2**	2.826 *(8.697)*	**μ_2**	2.080 *(8.939)*	**μ_2**	1.857 *(8.499)*	**μ_2**	2.020 *(8.933)*	**μ_2**	1.859 *(8.702)*	**μ_2**	1.953 *(8.396)*	**μ_2**	1.999 *(8.653)*
ρ^2	0.163	**ρ^2**	0.304	**ρ^2**	0.237	**ρ^2**	0.148	**ρ^2**	0.201	**ρ^2**	0.160	**ρ^2**	0.159	**ρ^2**	0.180
Ref.	OCC_1	**Ref.**	VEH_1	**Ref.**	IND_{17}	**Ref.**	EMP_1	**Ref.**	BDY_4	**Ref.**	BMT_{15}	**Ref.**	YRS_1	**Ref.**	SQF_1

Note: Variables in bolded boxes denote statistically significant variables at a 90% confidence level.

The ordered models consisted of four levels, based on the four categories for the number of inbound and outbound commercial trips. The ordered levels were categorized as 0, 1, 2, and 3, according to the number of commercial trips generated, where the last category represents 3 or more trips. Since there are four ordered levels in this model, the output would have three corresponding threshold values. The NLOGIT software automatically sets the first threshold, μ_0, to zero. The remaining two thresholds, μ_1 and μ_2, are presented in the output results for each model. Tables 20 and 21 show that a large number of categorical variables are not statistically significant on their own, according to the t-statistics obtained for each. These two tables do reveal some important trends.

The case of outbound trips is considered first, with the results from Table 4-15. The variable with the highest pseudo R-squared value was the one representing the number of vehicles of each type owned by the firms. The reference variable in this case was regular passenger cars. The only two variables that were statistically significant in this individual test were those corresponding to pickup trucks/cube vans and to single-unit trucks. This makes sense, since intra-urban commercial trips will be made mostly with small or medium commercial vehicles. When looking at the coefficients obtained, the positive values indicated that a higher number of vehicles owned of a specific type would result in a higher number of generated outbound trips compared to passenger cars. Conversely, a negative coefficient would result in a lower number of outbound trips. It is expected that these vehicle types would appear statistically significant in the final models. Since the majority of respondent firms were smaller firms (according to their employment sizes and square footage areas), the coefficients for larger commercial vehicles (such as containers) are lower. Such a result indicates that these vehicle types generate fewer trips than the reference regular passenger cars. However, the total number of vehicles was not used in the final model, since it was highly correlated to these categories as well as the employment size.

The second ranked variable, with the second-highest pseudo R-squared value was the industry classification. In this case, the class for other services, except public administration (NAICS 81) was selected as the reference. The statistically significant variables were those corresponding to manufacturing (NAICS 31, 32, and 33); wholesale trade (NAICS 42); retail trade (NAICS 44); transportation and warehousing (NAICS 48); and administrative, support, waste management and remediation industry classifications (NAICS 56). These industry categories are ones that would be intuitively expected to generate more commercial activities. It is expected that at least some of these industry classes would be present in the final model.

The categorical variable for busiest days of the week had the next highest pseudo R-squared value and the reference was set to BDY_{15}, which represents deliveries made every day of the week. The statistically significant categories in this case were those corresponding to Monday-Tuesday, Monday-Tuesday-Thursday-Friday, Monday-to-Friday, and Monday-Tuesday-Thursday-to-Sunday. This observation is sound, since these show that the most deliveries occur right before or right after the weekend. This also confirms what was observed in the frequency table for this variable.

Square footage was the next categorical variable, with the reference set as SQF_1, corresponding to a reported value of zero. It is observed here that larger establishments have larger and positive coefficients, meaning that they generate an increased number of outgoing commercial vehicles, which is intuitively correct. The significant variables are those for establishments having a square footage between 5,000 and 10,000 square feet, and those higher than 20,000 square feet. The three categories that encompass the latter range of areas were observed to have similar coefficients, meaning that they have a similar effect on the number of generated trips. In the final model, these three categories were combined into a single one representing the given range of areas. A hypothesis made prior to the final model was that businesses with larger floor areas would be more likely to generate commercial trips. The categorical variables were tested in the final model, but not the continuous square footage variable, again, to avoid correlations.

The categorical variable for the number of vehicles in different occupation types was next, with the reference set as the variable corresponding to the number of employees in management occupations. Only two of these variables were found statistically significant in this model. The numbers of employees in trade, transportation, and equipment operation occupations and other occupation types were statistically significant. This is intuitively correct, since an establishment that has a greater number of employees in these occupation types are more likely to engage in generating commercial trips to deliver goods or services.

The sixth ranked categorical variable corresponds to the busiest months for deliveries. The reference in this case was the category for year-round shipments. Significant categories in this case corresponded to April-to-June and October-to-December; January-to-March and July-to-December; and April-to-December. This indicates that the spring and autumn months are the busiest months for deliveries. This observation is consistent to the results from the frequency table for this variable.

When examining the categorical variable for the number of years of each establishment at its current location, the reference variable was set to that corresponding to zero, namely, establishments who were less than one year old. It can be seen that older businesses have a higher positive coefficient, indicating that they generate more commercial trips than newer businesses. This makes sense, since newer firms may not be able to afford sending as many shipments, or may not yet own enough vehicles to generate more trips. Similarly to the variable for square footage, the last two categories, corresponding to 51 years or more spent at each firm's current location, had similar coefficients. Since this was the case, these two categories have a similar effect on the number of outgoing commercial trips that are generated, and were combined into a single variable for the final analysis. Similarly to the square footage and employment variables, only one of the continuous or categorical variables were tested in the final model.

Lastly, employment size showed the lowest pseudo R-squared value when tested individually. The reference was set as EMP_1, corresponding to a reported number of employees ranging between one and ten. The statistically significant categories in this case were EMP_2 and EMP_5. It is expected that smaller businesses, those with fewer employees, would generate fewer commercial vehicle trips. The effects of this variable will be explored in the final model. Since this categorical variable is based on a continuous variable for the total number of employees, care needed to be taken that a correlation would not be introduced. Therefore, only the continuous variable or the categories were used in the final model. The two were not used together.

The results from the same analysis of the eight categorical variables with respect to the number of inbound commercial vehicles are presented in Table 4-16 below.

Table 4-16: Results from Logit Models of Categorical Variables, Inbound Case

Occupation types		Vehicle types		Industry type		Employment size		Busiest days		Busiest months		Years at location		Square footage	
Var.	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**
Const.	0.140 *(0.692)*	**Const.**	0.533 *(3.217)*	**Const.**	0.516 *(0.904)*	**Const.**	0.298 *(1.866)*	**Const.**	0.417 *(1.983)*	**Const.**	0.447 *(2.369)*	**Const.**	2.745 *(3.194)*	**Const.**	0.911 *(2.366)*
OCC_2	0.015 *(0.083)*	**VEH_2**	-0.034 *(-0.408)*	**IND_1**	-0.156 *(-0.146)*	**EMP_2**	-0.149 *(-0.342)*	**BDY_1**	0.341 *(0.638)*	**BMT_1**	1.491 *(0.968)*	**YRS_2**	-2.529 *(-2.86)*	**SQF_2**	-1.390 *(-2.796)*
OCC_3	0.781 *(0.646)*	**VEH_3**	-0.064 *(-0.447)*	**IND_2**	-2.723 *(-1.148)*	**EMP_3**	2.186 *(2.532)*	**BDY_2**	0.136 *(0.166)*	**BMT_2**	0.085 *(0.075)*	**YRS_3**	-2.061 *(-2.303)*	**SQF_3**	-0.439 *(-0.929)*
OCC_4	0.123 *(1.529)*	**VEH_4**	-1.476 *(-1.264)*	**IND_3**	0.523 *(0.486)*	**EMP_4**	1.487 *(0.758)*	**BDY_3**	0.106 *(0.271)*	**BMT_3**	0.515 *(0.729)*	**YRS_4**	-3.187 *(-2.977)*	**SQF_4**	0.492 *(0.725)*
OCC_5	0.014 *(0.138)*	**VEH_5**	-0.274 *(-0.764)*	**IND_4**	-0.079 *(-0.113)*	**EMP_5**	0.439 *(0.291)*	**BDY_4**	-0.348 *(-0.389)*	**BMT_4**	-0.550 *(-0.968)*	**YRS_5**	-2.822 *(-2.426)*	**SQF_5**	0.108 *(0.127)*
OCC_6	-0.680 *(-0.486)*	**VEH_6**	0.027 *(0.811)*	**IND_5**	-0.716 *(-0.947)*	**EMP_6**	-0.084 *(-0.045)*	**BDY_5**	-0.802 *(-1.448)*	**BMT_5**	-1.170 *(-0.661)*	**YRS_6**	-3.306 *(-2.661)*	**SQF_6**	0.908 *(1.285)*
OCC_7	0.083 *(1.938)*	**VEH_7**	0.036 *(0.39)*	**IND_6**	0.213 *(0.339)*	**EMP_7**	1.360 *(0.353)*	**BDY_8**	-0.522 *(-0.525)*	**BMT_6**	-1.269 *(-0.687)*	**YRS_7**	-3.078 *(-3.078)*	**SQF_7**	-0.792 *(-1.391)*
OCC_8	-0.003 *(-0.193)*	**VEH_8**	1.796 *(1.541)*	**IND_7**	-0.134 *(-0.189)*	**EMP_8**	-0.374 *(-0.183)*	**BDY_9**	-27.019 *(0.0)*	**BMT_7**	1.558 *(1.837)*	**YRS_8**	-0.950 *(-0.673)*	**SQF_8**	-0.418 *(-0.635)*
OCC_9	0.002 *(0.178)*			**IND_8**	1.283 *(1.259)*	**EMP_{10}**	0.428 *(0.305)*	**BDY_{10}**	-0.198 *(-0.278)*	**BMT_8**	-0.320 *(-0.658)*			**SQF_9**	-2.748 *(-1.328)*
OCC_{10}	-0.018 *(-0.358)*			**IND_9**	-28.770 *(0.0)*			**BDY_{11}**	0.056 *(0.099)*	**BMT_9**	-0.631 *(-0.827)*			**SQF_{10}**	-1.440 *(-1.142)*

Table 4-16 (cont'd): Results from Logit Models of Categorical Variables, Inbound Case (Continued)

Occupation types		**Vehicle types**		**Industry type**		**Employment size**		**Busiest days**		**Busiest months**		**Years at location**		**Square footage**	
Var.	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**	**Var.**	**Coef. (t-stat)**
OCC_{11}	-0.043			**IND_{10}**	0.172			**BDY_{13}**	-2.229	**BMT_{10}**	-0.425			**SQF_{11}**	1.376
	(-0.647)				*(0.128)*				*(-0.944)*		*(-0.492)*				*(0.763)*
				IND_{11}	-30.104			**BDY_{14}**	-32.018	**BMT_{12}**	-28.080			**SQF_{12}**	-0.472
					(0.0)				*(0.0)*		*(0.0)*				*(-0.35)*
				IND_{12}	-0.965					**BMT_{13}**	1.542				
					(-1.011)						*(0.964)*				
				IND_{13}	-3.234					**BMT_{14}**	-0.517				
					(-1.145)						*(-0.883)*				
				IND_{14}	-29.516										
					(0.0)										
				IND_{15}	-0.427										
					(-0.524)										
				IND_{16}	0.072										
					(0.103)										
μ_1	1.317	**μ_1**	1.370	**μ_1**	1.377	**μ_1**	1.320	**μ_1**	1.318	**μ_1**	1.348	**μ_1**	1.372	**μ_1**	1.415
	(8.69)		*(8.936)*		*(8.898)*		*(1.320)*		*(8.762)*		*(8.811)*		*(8.771)*		*(8.942)*
μ_2	1.984	**μ_2**	2.023	**μ_2**	2.048	**μ_2**	1.996	**μ_2**	1.962	**μ_2**	2.018	**μ_2**	2.066	**μ_2**	2.129
	(10.158)		*(10.51)*		*(10.466)*		*(10.084)*		*(10.29)*		*(10.326)*		*(10.234)*		*(10.499)*
ρ^2	0.017	**ρ^2**	0.048	**ρ^2**	0.049	**ρ^2**	0.019	**ρ^2**	0.015	**ρ^2**	0.028	**ρ^2**	0.037	**ρ^2**	0.053
Ref.	OCC_1	**Ref.**	VEH_1	**Ref.**	IND_{17}	**Ref.**	EMP_1	**Ref.**	BDY_7, BDY_{15}	**Ref.**	BMT_{15}	**Ref.**	YAL_1	**Ref.**	SQF_1

Note: Variables in bolded boxes denote statistically significant variables at a 90% confidence level.

The top ranked categorical variable in the ordered model for the number of inbound commercial vehicles was square footage. In this case, only one variable, SQF_2, was statistically significant, having a negative coefficient. This indicates that small business establishments do not attract as many commercial vehicle trips. Again, this could be attributed to the assumption that these firms do not receive deliveries as often, or belong to industry classes that do not engage in commercial vehicle activities. Square footage was also tested using only the categorical variables in the final mode, and not the continuous variable, to avoid the effects of a correlation.

Industry category was the next ranked categorical variable, with the category for NAICS 81 set as the reference. None of these categories provided significant results in the individual test. The manufacturing and retail trade industry categories resulted in positive coefficient values, indicating that they would attract a larger number of commercial trips, although this effect is statistically insignificant. This makes sense, since these types of businesses require stock or raw materials for their respective operations.

The third categorical variable was the one describing the number of vehicles owned of each type. Similarly to the outbound case, the reference was set as the number of regular passenger cars. The category for the number of other vehicle types owned was the only significant variable in this case. The majority of variables resulted in negative coefficients, excepting the larger vehicle types, containers and trailers, which is opposite to what might be expected. It may be assumed that establishments who own larger vehicle types are larger firms, such as manufacturers. These types of establishments, as previously explained, would require inbound shipments of resources for their operations.

The categorical variable for the number of years at an establishment's current location was tested next with the reference set as YAL_1, for reported values of zero, similarly to the outbound model. This model showed all but the last category to have statistically significant results. All of these coefficients were negative, indicating that newer businesses generate higher numbers of inbound trips, opposite to the effect observed in outbound activities. This could be attributed to the assumption that newer establishments rely more on deliveries from others. Similarly to other categorical variables that were derived from a continuous variable, only the categorical variables were used in the final model, in order to explore the effects more in-depth and to avoid correlations.

Fifth ranked was the categorical variable for busiest months for deliveries. As before, BMT_{15} for year-round deliveries was set as the reference. Only one variable was statistically significant in this case as well, namely the category for January-March and October-December. The coefficient for this variable had a positive value. These months are during the winter and it makes sense that an increased number of incoming commercial trips would be made during these months, when it is more convenient for establishments to have their supplies delivered to their location, due to the harsher weather conditions during these seasons.

The next ranked variable was for the categories of employment size and the reference was the category corresponding to establishments having one to ten total employees. Only one category was statistically significant in this case, EMP_3. A hypothesis that can be made about this group of categorical variables is that larger establishments, which would have a higher number of employees, could attract a higher number of commercial vehicle trips. The effects of this variable, similarly to the outbound model, were tested using only the categorical variables or only the continuous variable for employment size, to avoid a correlation that would lead to resulting coefficients that would cancel each other out.

The categorical with the seventh ranked pseudo R-squared value was the one for the number of employees at each firm in different occupation types. The reference was set as the management occupations category. Statistically significant variables were those for health occupations and sales and service occupations. Intuitively, this makes sense, since these occupations require supplies and product for their operations. These materials would be delivered to the establishments.

Lastly, the categories for busiest days of the week for deliveries had the lowest pseudo R-squared value in the individual test for the number of inbound commercial vehicles. The two variables BDY_7 and BDY_{15} were taken as references in this case. The model only ran without errors when two variables were extracted as references. None of the categories provided statistically significant results in this model.

It can be seen in Tables 20 and 21 that a large number of categorical variables are not statistically significant when tested on their own. The final models for outbound and inbound movements were estimated using a combination of these categorical variables. They were introduced in the order of the rankings listed in Table 4-14. Only the categories that were statistically significant were kept in the models. When insignificant variables were removed, the remaining variables were introduced one by one, to test if the new combinations yielded statistically significant results. Only variables that were statistically significant at a 90% or higher level of confidence were considered for the final ordered models.

A number of additional variables about establishment characteristics, also obtained from the web survey, were tested in the final models. The continuous variables for total number of employees (EMP_T) and total years at a firm's current location (YAL_T) were considered and, as explained, were only introduced if the corresponding categorical variables were not statistically significant. Other continuous variables tested were the number of employees who make on-call service trips (EOC), the number of employees who make truck deliveries (ETD), the number of commercial vehicles leaving the establishment on the survey day to deliver goods or services (CVL), and the number of vehicles arriving at the firm on the survey day (CVA). Binary variables, corresponding to yes/no answers in the web survey, were introduced. These variables dealt with whether the firm had a supply chain or logistics operator (LSO), whether it had a supply chain or logistics specialist (LSP), whether the firm used logistics software (LSW), whether the number of commercial trips undertaken by the firm varied by day of the week (VDY), and whether the number of trips varied by months (VMT). The total number of vehicles, VEH_T, was not introduced in the final model, since it was found to be highly correlated with the total number of employees at a firm.

The final results from the ordered logit models for outbound and inbound commercial vehicles are presented in Table 4-17. The coefficients for each statistically significant variable are provided, along with each coefficient's t-statistic. The resulting thresholds for each model were listed, along with the pseudo R-squared value.

Table 4-17: Results from Final Ordered Logit Models for Outbound (OB) and Inbound (IB) Commercial Vehicle Activities

Var	OB Coef. (t-value)	IB Coef. (t-value)
CONST	-5.036 *(-5.662)*	4.220 *(5.139)*
EMP_T	0.158 *(3.555)*	
ETD	1.050 *(5.205)*	
LSP	-0.009 *(-3.104)*	
LSW	0.009 *(3.428)*	
CVA	-0.004 *(-5.506)*	
OCC_1		0.450 *(3.727)*
OCC_2	0.760 *(2.267)*	
OCC_7	-0.346 *(-3.611)*	
OCC_8	-0.389 *(-4.352)*	
OCC_9	-0.215 *(-4.292)*	
OCC_{11}		-0.156 *(-2.195)*

Var	OB Coef. (t-value)	IB Coef. (t-value)
VEH_1	0.939 *(3.736)*	-0.462 *(-2.418)*
VEH_2	1.247 *(5.867)*	-0.458 *(-3.598)*
VEH_3	0.936 *(3.401)*	
VEH_4		-2.206 *(-1.754)*
VEH_5		-2.554 *(-3.261)*
VEH_6		0.079 *(2.242)*
VEH_7		0.946 *(2.378)*
VEH_8		4.669 *(3.628)*
IND_3	4.962 *(3.155)*	
IND_5	1.503 *(1.806)*	
IND_6	1.832 *(2.958)*	
IND_8	4.798 *(3.608)*	4.122 *(3.838)*

Var	OB Coef. (t-value)	IB Coef. (t-value)
IND_{11}	4.844 *(2.821)*	
IND_{16}	-2.884 *(-2.213)*	
EMP_3		3.703 *(3.438)*
EMP_7		11.667 *(2.481)*
BDY_3	-1.313 *(-1.688)*	
BDY_6	2.670 *(3.500)*	
BDY_{15}	-4.830 *(-1.767)*	
BMT_3		1.385 *(1.784)*
BMT_4	2.401 *(1.940)*	
BMT_6	4.844 *(2.916)*	
BMT_7		2.429 *(2.688)*
YAL_2	1.119 *(1.900)*	-4.393 *(-5.148)*

Var	OB Coef. (t-value)	IB Coef. (t-value)
YAL_3		-3.352 *(-3.929)*
YAL_4		-6.114 *(-5.004)*
YAL_5		-3.276 *(-2.558)*
YAL_6		-6.173 *(-4.498)*
YAL_7		-4.594 *(-3.265)*
SQF_4	-2.851 *(-1.956)*	3.403 *(4.786)*
SQF_6	-2.325 *(-2.203)*	2.739 *(3.881)*
SQF_8	1.815 *(2.094)*	
SQF_{11}		7.701 *(3.300)*
μ_1	3.971 *(9.886)*	2.166 *(9.788)*
μ_2	6.132 *(11.560)*	3.317 *(11.273)*
ρ^2	0.646	0.262
Correct	76.7%	59.1%

The ordered logit model for the number of outbound vehicle trips will be examined in detail first. It can be seen that there are 27 statistically significant variables. The constant term obtained in this model is somewhat larger than most of the coefficient values obtained. This indicates that there are still unobserved effects influencing the number of outbound commercial vehicle trips that are not explained by the variables at hand. Since the categorical variables for employment size were not significant, the continuous EMP_T was introduced instead. This variable has a positive coefficient, which acts to prove the hypothesis that a larger firm, given by employment size, is more likely to produce a larger number of outgoing commercial vehicle trips. The next variable, ETD, which represents the number of employees who make truck deliveries for each firm, also had a positive coefficient. This result also makes sense, since these employees would evidently be making commercial vehicle trips from the firm. Therefore, having a larger number of employees who make deliveries is likely to result in a larger number of outbound commercial trips.

The variable describing whether a firm has a supply chain or logistics specialist has a negative coefficient, which may seem odd, since it indicates a lower number of outgoing commercial trips. The variable for whether a firm uses logistics software, however, is positive, indicating an increased number of trips for firms who do use logistics software. The correlation between these two variables is 0.19 and not significantly high, so there should not be a negative interaction between the two variables. These results may be explained by assuming that a supply chain or logistics specialist may organize incoming commercial activities for a firm, rather than outgoing trips, whereas logistics software may be used by a firm to organize all of their outbound movements. These variables are not statistically significant in the inbound model. The number outbound commercial activities is also modeled as lower when a firm has a higher number of commercial activities arriving for deliveries. This can be seen from the negative coefficient for the CVA variable.

With respect to the number of employees in different occupation types, outbound commercial activities were shown to be influenced by four occupation types. These are business, finance, and administration; sales and service; trade, transportation, and equipment operation; and primary industry occupations. A positive coefficient was achieved for the number of employees in the first occupation type listed, OCC_2. This could be due to the assumption that such occupations would use couriers or similar services to send paperwork and documents to other locations. Therefore, the more employees in these positions that exist at a firm, the higher the number of outbound trips will be. The other three occupation types resulted in negative coefficients, indicating that a higher number of employees in these three occupation types at a firm will act to reduce the number of outbound commercial trips. The reason for this behaviour could be that these occupations may be responsible for acquiring merchandise or materials for the establishment’s operations, whether they be sales or manufacturing. The reason that a higher number of employees in these occupation types leads to a decrease in the number of outbound trips may be that, since these firms accept deliveries and generate more inbound trips, they would then generate fewer outbound trips.

The next categorical variable introduced was the number of vehicles of each type owned by a specific establishment. Statistically significant categories of vehicle types included regular passenger cars, pickup trucks/cube vans, and single-unit trucks. All of these variables resulted in positive coefficients, meaning that an establishment who owns a higher number of vehicles in each of these categories is more likely to produce a higher number of outgoing commercial vehicle trips. It is also observed that the category for the number of pickup trucks/cube vans is higher than the other two, which are very similar to each other. The category with a higher coefficient is the more influential to the ordered model. These results make sense when considering the commercial activities undertaken within the urban study area. Many intra-urban commercial trips are made using smaller commercial vehicle types. This observation aids to further the behaviour of commercial activities with respect to service trips, since the majority of these are undertaken using smaller vehicles such as trucks or vans.

A number of industry classifications were significant in the ordered logit model for the number of outbound trips. These correspond to manufacturing (NAICS 32), wholesale trade (NAICS 42), retail trade (NAICS 45), transportation and warehousing (NAICS 48), real estate rental and leasing (NAICS 53), and accommodation and food services (NAICS 62). All of these industry classes, save the last one in the list, had positive coefficient results. This finding reinforces the hypothesis that firms belonging to these industry classes act to increase the number of commercial vehicle trips leaving the establishments. This is due to the nature of these industry categories, which often have various products to sell or offer transportation services.

The significant variables related to busiest days of the week for deliveries were BDY_3 (Thursday-Friday), BDY_6 (Monday-Tuesday, Thursday-Friday), and BDY_{15} (Monday-Sunday). The variables BDY_3 and BDY_{15} resulted in negative parameters, meaning that firms for which these days are considered the busiest will have a lower probability of producing outgoing commercial vehicles, while BDY_6, has the opposite effect with a positive coefficient. With respect to the categorical variables for the busiest months, two significant variables were found, both having positive coefficients and, thus, contributing to increasing the number of outgoing commercial trips. The two significant variables are for the categories of October-to-December and January-to-March and July-to-September. These values indicate that firms who reported being the busiest during these respective time periods during the year will be more likely to have higher numbers of outgoing commercial trips.

Only one categorical variable representing the number of years of each firm at its current location was statistically significant, namely, that for one to ten years. The positive coefficient attained for this variable contributes to an increase in the number of commercial trips leaving the establishment. Lastly, square footage variables SQF_4, SQF_6, and SQF_8 were statistically significant. The model gave negative parameter values for the first two categories, and a positive coefficient for the third. This indicates that larger establishments are more likely to generate outbound commercial vehicle trips, which makes sense.

Now, in examining the ordered logit model for inbound commercial trips, it was observed that 23 variables were statistically significant. Again, the constant term shows that there are still some unobserved factors not explained by the current variables. Only variables belonging to the categorical groups were statistically significant in this case. Two variables were significant from the group representing the number of employees by occupation type at each establishment, namely, that representing management occupations and other occupations. The parameter estimated for the first variable was positive, indicating that a firm with more employees in this occupation would be more likely to receive a higher number of commercial vehicles. The variable for other occupations was negative, having the opposite effect on the number of inbound commercial trips.

Upon examining the categorical variables about the number of vehicles owned of different types, it can be seen that all vehicle types except single-unit trucks are significant in influencing the number of incoming commercial trips. The first four variables, VEH_1, VEH_2, VEH_4, and VEH_5 resulted in negative parameters, while the remaining three were estimated with positive parameters. These results show that firms who own larger vehicles would be more likely to attract commercial vehicles. Usually, establishments who own these large vehicles are larger manufacturers, transportation firms, or other such industries would require products to be delivered to the establishments. It could also indicate that the firms who own larger vehicles receive deliveries more often than others. Since the majority of the surveyed firms were of smaller size (considering the total number of employees and the floor area), it can be assumed that smaller vehicle types are more commonly used. Since the proportion of firms using small vehicles is larger in the context of the study area, more commercial trips would be generated by firms who own more of these vehicles.

Only one industry category was statistically significant in this model. A positive parameter was obtained for the transportation and warehousing industry type, which is intuitively correct. This industry type, obviously, deals with transportation services, having a larger number of incoming vehicles. The businesses classified under warehousing would also contribute to a higher number of inbound trips, since patrons would arrive to deposit products for storage.

The categorized employment size variables were introduced here and EMP_3 and EMP_7 were found statistically significant. The coefficients for these variables both had positive values, indicating that they both lead to an increase in the number of inbound trips. It was also noted that the parameter for the EMP_7 variable is larger than the other, indicating that it has a more significant impact on the end result of the model. This reinforces the hypothesis that larger businesses, as determined by employment size, contribute to increased commercial vehicle trips.

There were no statistically significant variables related to the busiest days of the week for deliveries, however, two of the variables describing the busiest months were significant in the inbound model. Both variables, corresponding to July-September, and January-March and October-December, were estimated with positive parameters. These results show that a higher number of incoming commercial vehicles arise during the summer months. This could be attributed to the assumption that some establishments receive increased business during these months, therefore, the firms would need to restock their products more often. The number of incoming commercial trips is also increased during the winter months, which could be attributed to the assumption that businesses may prefer to have their supplies delivered to the establishment during the cold winter months.

With respect to the number of years of establishments at their current locations, all of the categorical variables except for YAL_1 were statistically significant. These variables all held negative parameter values, which may be an indicator that newer establishments rely more on deliveries from others. Thus, the negative parameters show that a business that has been at its current location is less likely to attract inbound commercial vehicle trips.

Lastly, three categorical variables describing the square footage of firms were statistically significant in this ordered model, namely, SQF_4, SQF_6, and SQF_{11}. The parameters for all three variables were positive, and the parameter for SQF_{11} was larger than the others. These observations indicate that larger establishments attract more inbound trips. This could be because firms with larger surface areas have greater capacities for storing goods and may use greater quantities of products that need to be replenished.

The explanatory power of these two models is indicated by the McFadden pseudo R-squared values. The ordered logit model for outbound commercial activities yielded a value of 0.65, which is a fairly good fit. Although there are effects that are yet unobserved and not explained by the currently used variables, the model that was developed describes a fair portion of the behaviour of commercial vehicles. The pseudo R-squared value for the inbound ordered model, however, was considerably lower, at 0.26, but it is still considered a good fit. This shows that the factors influencing the number of inbound commercial trips are not explained as well by the variables used in the present model. The reason for this could be that outbound trips are coordinated by the establishments themselves, are highly organized, and are influenced by the characteristics of the business establishments. On the other hand, firms do not have as much control over the incoming commercial vehicles. Those trips are organized outside of the firms and are not as dependent on the characteristics of the establishment receiving the commercial vehicles. Further investigation is required to better describe inbound commercial vehicle movements.

Another analysis of the effectiveness of the models can be done through cross-tabulation analyses. Cross-tabulations show the amount of results in each category that are predicted correctly and those that are predicted erroneously, assigned to a different category than the one they actually belong to. In the case of the ordered logit model for the number of outbound trips, the proportion of observations that were predicted correctly was 76.7%. A similar cross-tabulation was generated for the inbound ordered logit model, where only 59.1% of the results were predicted correctly. The explanatory ability in the case of the inbound model is seen, again, as inferior to the ordered logit model for outbound commercial activities.

From a model application perspective, the type of detailed micro data collected and analyzed in this thesis is typically not readily available to use in transportation and policy planning exercises. However, firm data that is usually available and can be acquired from companies like InfoCanada relates to the employment sizes and industry classifications. As such, a second pair of ordered logit models was generated for the number of inbound and outbound commercial trips as a subset of the full models above, using only these two groups of variables. The results are shown in Table 4-18. It can be seen that the results for the outbound model, once again, show a better explanatory power than the inbound model. The McFadden pseudo R-squared values are lower than in the full models, which is to be expected, since only a subset of the variables was used.

Table4-18: Results from Ordered Logit Models for Outbound (OB) and Inbound (IB) Commercial Vehicle Activities Using a Subset of Variables (Employment Size and Industry Classes)

Variable	OB	IB
	Coefficient (t-value)	Coefficient (t-value)
Const.	-1.477 *(-4.764)*	0.094 *(0.534)*
EMP_T	0.023 *(1.995)*	
EMP_3		2.353 *(2.718)*
IND_2	2.100 *(1.939)*	
IND_3	2.764 *(3.291)*	
IND_4	1.955 *(3.875)*	
IND_5	1.495 *(2.771)*	
IND_6	1.384 *(3.382)*	0.592 *(1.864)*
IND_8	2.814 *(3.447)*	1.704 *(1.971)*
IND_{11}	2.151 *(1.919)*	
IND_{13}	6.110 *(2.194)*	
μ_1	1.394 *(7.614)*	1.350 *(8.712)*
μ_2	2.083 *(8.864)*	2.054 *(10.146)*
ρ^2	0.218	0.032
Correct	49.7%	45.03%

With respect to the subset model for outgoing commercial vehicles, it can be seen that the continuous variable for employment is statistically significant at a 90% confidence level. The coefficient for this variable is positive, indicating that a larger number of employees at a firm will lead to an increase in the number of outgoing commercial trips. Since this continuous variable was statistically significant in this model, the categorical variables for employment size were not introduced, since they would have had clashing effects. The continuous variable for total employment was not statistically significant in the case of the inbound subset model. Therefore, the categorical employment size variables were tested. As in the case of the full inbound model, it was found that employment category 3, corresponding to firms having between 21 and 30 employees, was statistically significant. The positive coefficient, again, indicates that firms of this size by employment are likely to attract more inbound commercial trips.

All of the coefficients for the statistically significant industry classifications were positive, indicating that firms belonging to these industry classes were more likely to generate more outbound and inbound trips. The number of outbound trips is influenced by a larger number of industry classes. These are firms belonging to the manufacturing; wholesale trade; retail trade; transportation and warehousing; real estate rental and leasing; and administrative, support, waste management, remediation service industries. The coefficients for most of these variables are fairly close to each other, indicating that they have similar effects on the number of outgoing commercial trips. The last listed industry type has the highest coefficient, indicating that firms belonging to this industry class have the highest influence on the number of incoming vehicles. These would include fleets such as those used for trash, recycling, and yard waste collection.

With respect to the industry categories that affect the number of inbound commercial trips, only two showed statistically significant results. These two industry classes represent firms belonging to the retail trade and transportation and warehousing industries. Both of the coefficients associated with these variables were positive, indicating that firms belonging to these industry categories will generate a higher number of incoming trips, which makes sense given the nature of the respective industries. The transportation and warehousing industry category resulted in a higher coefficient, indicating that this industry type is more influential in attracting commercial vehicles.

When examining the fit of the models, the McFadden pseudo R-squared values appear lower than those attained by the full models. The value pertaining to the outbound model is, again, higher than the inbound model, indicating that the explanatory power of the former is greater than that of the latter. The proportions of observations predicted correctly in the subset models were not as different from each other as those in the full models. Fewer observations were predicted correctly in both the outbound and inbound cases in the subset models than in the full models.

CHAPTER 5

CONCLUSIONS

5.1 Summary of Data Collection Methods

The purpose of this thesis was to develop a statistical model for determining the factors giving rise to commercial vehicle trip generation in the Windsor CMA study area. A microsimulation approach was implemented, in order to analyze commercial travel behaviours at the level of individual business establishments. To this end, a combined telephone and web-based survey instrument was implemented. A novel methodology was introduced for identifying the firms belonging to industry categories more likely to be eligible and willing to partake in the web survey, in order to reduce the time required to recruit survey respondents. The resulting data collected from completed web surveys was then used to create ordered logit models analysing the factors giving rise to the production and attraction of commercial vehicle trips in the Windsor CMA study area.

The telephone portion of the survey effort was conducted in two phases, the first of which involved contacting five samples, comprising 50% of the population of firms in the region, in their entirety. The telephone survey inquired whether the firm engaged in shipping or receiving of goods or services, whether a representative would be willing and able to complete the web survey, and collected email addresses from interested parties where the web survey link would be sent. On the basis of the response rate from this phase, a Quotient analysis was conducted to determine which industry categories (by NAICS classification) were more likely to engage in shipping or receiving of goods or services, as well as which industry classes were more likely to provide an email address to receive the web survey. A list of NAICS codes was identified and the remainder of the population of firms was filtered on the basis of these industry categories, such that only those identified in the Quotient value procedure were contacted during the second phase for recruitment to the web survey.

The Quotient analysis method was proven effective in increasing the response rate and reducing the time required to recruit survey participants. This procedure was validated when comparing the distributions of the survey respondents who participated from the first phase and from the second phase, with respect to industry classification, employment size, and total number of vehicles owned. The latter two factors represented the size of the business establishments. It was found that the distributions of each of these characteristics and combinations thereof were not statistically different between the two contacted samples. This indicated that the targeted recruitment technique arising from the LQ analysis did not introduce a bias in the types of firms who responded to the web survey.

A total of 171 firms completed the web survey and these observations were weighted in order to ensure that they would more accurately represent the entire population of businesses in the study area. The weighting was done on the basis of employment size and industry classification by NAICS codes, taking ratios between the survey respondents and the population of all establishments contacted during the telephone survey portion. Introduction of the weights into the statistical models improved the fit of these models, increasing the pseudo R-squared parameter that indicates the goodness-of-fit of the models, and identifying a larger number of statistically significant variables.

5.2 Summary of Modeling Results

The ordered models were built using the variables pertaining to establishment characteristics collected through the web-based survey. Prior to statistical analysis, the variables were all assigned numerical values and categorized. Correlations between certain variables were explored and, where there was a high correlation coefficient, care was taken that the two correlated variables were not used in the same model together. This ensured that there would be no clash in the variables, leading to results that canceled each other out.

The variables under analysis related to the size of the establishments were the number of employees, the number of vehicles of each type owned, the number of employees in each occupation type, and the square footage of the establishments. Variables describing establishments' commercial vehicle activities included the number of employees making on-call service trips and truck deliveries, whether the firm employs a supply chain or logistics operator or specialist, whether the firm uses logistics software, whether the number of commercial trips varies by day of the week or by month, the busiest days and months for deliveries, the numbers of vehicles leaving and arriving at the establishment on the survey day for deliveries. Other establishment characteristics were described by the industry classification (by NAICS 2-digit codes), and the number of years of each firm at its current location.

Examining the frequency of responses to each of the survey questions pertaining to these variables revealed some trends of commercial activities in the study area. On the basis of these observations, some a priori hypotheses were developed regarding the influence that these factors would pose on the number of incoming and outgoing commercial vehicle trips.

The statistical modeling followed an ordered structure with four levels for the numbers of incoming and outgoing commercial trips, namely 0, 1, 2, and 3 or more on the day the survey was completed. Both ordered logit and ordered probit models were tested for the numbers of both incoming and outgoing commercial vehicles. Categorical variables were created for the numbers of employees in different occupation types, the numbers of vehicles owned of different types, the industry classes, ranges of employment sizes, combinations for days and months considered the busiest for deliveries, ranges of number of years at the firm's current location, and ranges for square footages. Each of these eight categories was introduced individually into the four ordered models. Trends were observed and hypotheses generated about the effects of each of these factors on commercial trips.

The ordered logit model was found to have a better fit in both the inbound and outbound cases than the ordered probit model, and was selected for development of the final models. Two models were created, one exploring the factors giving rise to outgoing commercial vehicles and the other for incoming trips. The outbound trip model showed a relatively good fit, with a McFadden pseudo R-squared value of 0.65 and approximately 77% of the observations predicted correctly by the model. The outbound movements appear to be most influenced by the numbers of employees in different occupation types, the number of smaller commercial vehicles owned, the industry classification, day of the week and month of the year, and a firm's square footage. The inbound trip generation model had a lower fit, with a pseudo R-squared value of only 0.26 and 59% of the observations predicted in the appropriate category. The inbound model was most influenced by the types of vehicles owned by firms, firms in the transportation and warehousing industry class, employment size, the years of a firm at its current location, and the firm's square footage.

Two models, one for outbound movements and one for inbound, were generated using a subset of the variables, which would be readily available for practical purposes, such as policy planning. Although the fit and explanatory power of these subset models were lower than the full models, the subset models would be practical for use by engineers or policy makers. The number of employees at each firm and the industry classification of each establishment influence the numbers of outbound and inbound commercial trips that are generated in the study area.

5.3 Recommendations

It can be seen that the models do show some of the effects of important factors influencing commercial trip generation in the context of the Windsor study area. It was also observed, however, that the performance of the model describing the number of incoming commercial vehicles was lower than that for outgoing commercial vehicles. The reason for this is not completely known. Part of it can be attributed to the assumption that, while business establishments have control over organizing their outbound commercial activities, not as much control exists over incoming vehicles. The latter is also heavily influenced by the firms from which they are dispatched. Not enough information is available to explain the factors influencing the number of inbound commercial vehicle activities.

One recommendation that arises from the results of this project is to study incoming commercial vehicle activities more in-depth. The remaining survey data can be used for this purpose, pertaining to the characteristics of actual inbound and outbound shipments observed by the respondent firms. These two sections of the web survey collected data regarding the types of outbound activities (service or shipment), location information for the origin or destination of the trip, the vehicle type used, the number of deliveries undertaken in each trip, the time of day that the commercial vehicle left or arrived at the firm, the value, weight, and type of commodities shipped, or the type of service provided. Expanding the current models to include this data has the potential to provide a stronger explanatory power, especially in the inbound case.

A second method by which the models can be extended in future efforts is through the introduction of interaction terms. These terms would explore the effects of variables in combination, providing further insight into the influences of each variable. Interactions could be tested between the industry classification and any of the following: number of employees in different occupation types, number of vehicles owned of different types, square footage of the business establishment, employment size, and the days and months considered to be busiest for deliveries. Of course, other interactions can and should also be explored. Addition of these terms would resolve some of the hypotheses that were tested, which were only partially answered in this model.

As previously discussed, the categorical variables created for the busiest days of the week and months of the year for commercial activities were introduced as control variables, rather than explanatory variables. In future research and modeling efforts, the complexities of the combinations for the different days and months should be studied.

The busiest days of the week and months of the year could be used as interaction terms to describe the days of the week and the months through the year when more commercial vehicles are generated.

With respect to practical verification of the model, the Rhodes Drive area in Windsor, Ontario could be studied. This region is populated largely by commercial and industrial establishments and the majority of the traffic that passes through this area consists of commercial vehicles. A cordon count of the commercial vehicle volumes passing through could be conducted. The number of vehicles of each vehicle type could be observed in this area.

Future research on this subject should differentiate between firms who deliver goods, those who deliver services, and those who engage in both of these activities. In the web survey used for data collection for this thesis, respondents were only asked whether they engage in shipping or receiving of goods or services, but were not asked to further elaborate on their shipping activities. Obtaining this information would be beneficial for a future study, as it would allow further investigation into the factors giving rise to commercial trips undertaken for goods delivery purposes and those undertaken for service calls individually. These two activities do differ from each other and it is important that an attempt is made in future work to differentiate between them and the effects influencing each.

REFERENCES/BIBLIOGRAPHY

Bastida, C. & Holguin-Veras, J. 2009, "Freight Generation Models: Comparative Analysis of Regression Models and Multiple Classification Analysis", *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 2097, pp. 51 – 61

Ferguson, M., Maoh, H., Ryan, J., Kanaroglou, P., Rashidi, T. 2012, "Transferability and enhancement of a microsimulation model for estimating urban commercial vehicle movements", *Journal of Transport Geography*, vol. 24, pp. 358-369

Greene, W. & Hensher D. 2010, "A Model for Ordered Choices", *Modeling Ordered Choices: A Primer.* Cambridge: Cambridge UP, 99-110. Print

Holguin-Veras, J., Jaller, M., Destro, L., Ban, X., Lawson, C., & Levinson, H. 2011, "Freight Generation, Freight Trip Generation, and Perils of Using Constant Trip Rates", *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 2224, 68-81

Holguin-Veras, J., Sanchez-Diaz, I., Lawson, C., Jaller, M., Campbell, S, Levinson, H., & Shin, H-S. 2013 "Transferability of Freight Trip Generation Models", *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 2379, pp. 1-8

Hunt, J. D., et al. 2004, "A commercial movement modelling strategy for Alberta's major cities." *Proceedings of the 2004 Annual Conference of the Transportation Association of Canada.* Quebec

Hunt, J., Stefan, K & Brownlee, A. 2006, "Establishment-Based Survey of Urban Commercial Vehicle Movements in Alberta, Canada: Survey Design, Implementation and Results", *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 1957, pp. 75-83

Hunt, J.D. & Stefan, K.J. 2007, "Tour-based microsimulation of urban commercial movements", *Transportation Research Part B*, vol. 41, pp. 981-1013

Jewell, L., et al. 2007 *Commercial Vehicle Usage and Forecasting-Stage 2: National Freight Traffic*. Land Transport New Zealand

Joubert, J. & Axhausen, K. 2013, "A Complex Network Approach to Understand Commercial Vehicle Movement", *Transportation*, vol. 40(3), pp. 729-750

Lawson, C., Holguin-Veras, J., Sanchez-Diaz, I., Jaller, M., Campbell, S., & Powers, E. 2012, "Estimated Generation of Freight Trips Based on Land Use", *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 2269, pp. 65-72

Madar, G. & Maoh, H. 2014, "A micro-approach to understanding urban commercial vehicle movements: Preliminary results from a web-based establishment survey."

Proceedings of the Annual Canadian Transportation Research Forum Conference, Windsor

Maoh, H., Ryan, J. Kanaroglou, P. 2009, “The Impacts of Future Economic Growth on Urban Commercial Vehicle Movement in Hamilton, Ontario”, *Proceedings of the Annual Canadian Transportation Research Forum Conference*, Victoria

McNally, M. G.; & Rindt, C. 2008, “The Activity-Based Approach. Center for Activity Systems Analysis.” *Center for Activity Systems Analysis.* Retrieved from: https://escholarship.org/uc/item/86h7f5v0

Miller, E., Hunt, J., Abraham J. & Salvini, P. 2004, “Microsimulating urban systems”, *Computers, Environment and Urban Systems*, vol. 28, pp. 9-44

Miller, R. E. & Blair P. D. 2009 “Nonsurvey and Partial-Survey Methods: Extensions” *Input-output analysis: foundations and extensions*. Cambridge University Press, 374-359. Press

Nuzzolo, A. & Comi, A. 2014, “Urban freight demand forecasting: A mixed quantity/delivery/vehicle-based model”, *Transportation Research Part E*, [online] Available at: < http://dx.doi.org/10.1016/j.tre.2013.12.014>

Oliveira-Neto, F., Chin, S., & Hwang, H. 2012, “Aggregate Freight Generation Modeling: Assessing Temporal Effect of Economic Activity on Freight Volumes with Two-Period Cross-Sectional Data”, *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 2285, pp. 145 – 154

Paez, A., Scott, D., Potoglou, D., Kanaroglou, P., & Newbold, K. B. 2007, "Elderly mobility: Demographic and spatial analysis of trip making in the Hamilton CMA, Canada", *Urban Studies*, vol. 44(1), pp. 123 – 146

Roorda, M., McCabe, S., & Kwan, H. 2007a, "A shipper-Based Survey of Goods and Service Movements in the Greater Golden Horseshoe: Report I Survey Design and Implementation", *Final Report Submitted to the Ministry of Transportation of Ontario and the Region of Peel*. Sept 14, 2007

Roorda, M., McCabe, S., & Kwan, H. 2007b, "A shipper-Based Survey of Goods and Service Movements in the Greater Golden Horseshoe: Report II Preliminary Analysis of Results and Comparison of Survey Techniques", *Draft Report Submitted to the Ministry of Transportation of Ontario and the Region of Peel*. Sept 20, 2007

Samimi, A., Mohammadian, A., & Kawamura, K. 2010, "A behavioural freight movement microsimulation model: Method and data", *Transportation letters: The International Journal of Transportation Research*, vol. 2(1), pp. 53-62

Samimi, A., Mohammadian, A., & Kawamura, K. 2013, "A nationwide web-based freight data collection", *Canadian Journal of Civil Engineering*, vol. 40(2), pp. 114-120

Stefan, K., McMillan, J. & Hunt, J. 2005, "Urban Commercial Vehicle Movement Model for Calgary, Alberta, Canada", *Transportation Research Record, Journal of the*

Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C., no. 1921, pp. 1-10

Stefan, K., Hunt, J., McMillan, J. & Farhan, A. 2007, "Development of a fleet allocator model for Calgary", *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 1994, pp. 89-93

Wang, Q. & Hu, J. 2012, "Behavioral Analysis of Decisions in Choice of Commercial Vehicular Mode in Urban Areas", *Transportation Research Record, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*, no. 2269, pp. 58-64

Xie, W. & Roorda, M. J. 2009, "Commodity-based goods movement model for the Ontario auto industry", *Proceedings of the Annual Canadian Transportation Research Forum Conference*, Victoria

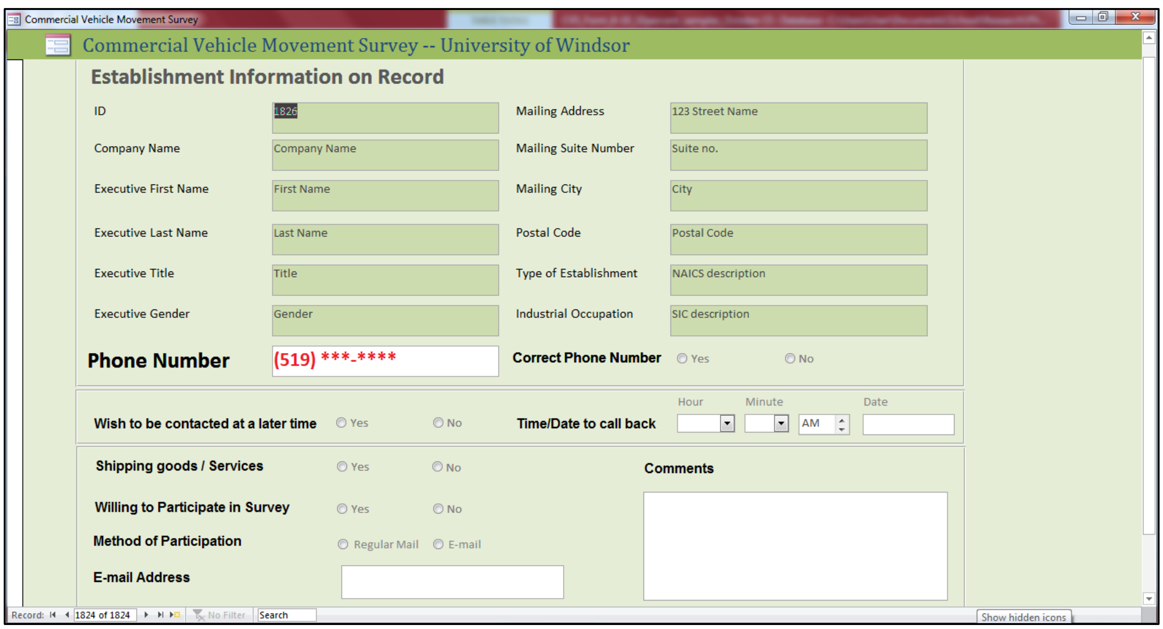
Yang, C., Regan, A., & Son, Y. 2010, "Another view of freight forecasting modeling trends", *KSCE Journal of Civil Engineering*, vol. 14(2), pp. 237-242

Zhou, J. & Dai, S. 2012 "Urban and Metropolitan Freight Transportation: A Quick Review of Existing Models", *Journal of Transportation Systems Engineering and Information Technology*, vol. 12(4), pp. 106-114

APPENDICES

Appendix A: Screen Capture of Telephone Survey

This appendix contains a screen capture of the platform used during the telephone recruitment phases. As mentioned in the Data Collection chapter of the thesis, an Access database was used to organize the responses and, within this database, the following form was created to aid in the telephone survey. Each establishment's contact information was displayed in turn, and the form allowed for input of the respondent's answers to the phone survey questions. The inputted responses were then updated in the Access database and subsequently used in the LQ and frequency analyses.



Appendix B: Web Survey Questions

This appendix contains the entire list of questions asked in the web-based survey, along with the possible answers to each question. In some cases, only a yes/no answer was required. In other cases, the respondent had to manually input a value or answer (such as the number of employees, or the address to which a shipment was destined). Still other responses were given through selection from a dropdown list, or through checkboxes where respondents selected as many (or as few) options as were applicable to their respective business establishment.

Establishment Characteristics	**Does your establishment engage in delivering goods or services to clients or establishments in the region?** (Y/N)
	The total number of employees who work at this location is: (input value)
	The number of employees by occupation type in my establishment are as follows: (input value for each category, as applicable: *Management occupations; Business, finance and administrative occupations; Natural and applied sciences and related occupations; Health occupations; Occupations in social science, education, government service and religion; Occupations in art, culture, recreation and sport; Sales and service occupations; Trades, transport and equipment operators and related occupations; Occupations unique to primary industry; Occupations unique to processing, manufacturing and utilities; Other*)
	My establishment has been at this location for _ years. (input value)
	The total square footage of my establishment is _ square feet. (input value)
	My establishment currently has _ employees who make on-call service trips. (input value)

Establishment Characteristics (cont'd)	**My establishment currently has _ employees who make truck deliveries.** (input value)
	My establishment has a Supply-Chain/Logistics operator. (Y/N)
	My establishment has a Supply-Chain/Logistics specialist. (Y/N)
	My establishment makes use of Logistics Software to manage trips. (Y/N)
	My establishment owns the following commercial vehicles: (input value for each category, as applicable: *Regular passenger cars, Pickup or cube van, Single unit trucks, Multiple unit trucks, Tractors, Containers, Trailers, Other – specify)*
	The industry which best characterizes my establishment is: (choose from list: *Retail trade; Transportation and warehousing; Information and cultural industries; Finance and administration; Real estate and rental and leasing; Professional, scientific, and technical services; Management of companies and enterprises; Administration and support, waste management and remediation services; Educational services; Health care and social assistance; Arts, entertainment and recreation; Accommodation and food services; Other services (except public administration); Public administration; Other – specify)*
	The number of commercial trips made by my establishment varies by day of the week. (Y/N)
	The busiest days for delivering goods or services are. (select as many as apply from: *Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday)*
	The number of commercial trips made by my establishment varies by month. (Y/N)
	The busiest months for delivering goods or services are: (select as many as apply from: *January, February, March, April, May, June, July, August, September, October, November, December)*

Outbound Commercial Trips	**How many commercial vehicles left your establishment today to deliver goods or services?** (input value)
	Type of outbound activity: (select one from: *shipment only, service only, shipment and service)*
	Destination location information. (select one from: *Street address of destination receiving shipment/service, postal/zip code of destination receiving shipment/service,* input address/postal code and city)
	Destination location information. (select one from: *Street address of destination receiving shipment/service, postal/zip code of destination receiving shipment/service*, input address/postal code and city)
	Select all means of transportation used to make this outbound trip: (select as many as apply from: *regular passenger car, pickup truck, minivan or cube van, single unit truck, tractor with one trailer, tractor with two or more trailers, rail, air, marine, other – specify)*
	If the outbound trip involved deliveries, how many shipments were delivered on this trip? (input value)
	The time vehicle left my establishment: (input *hour* and *minute*; select *am* or *pm)*
	Was this a Just In Time (JIT) delivery? (select one from: *yes, no, not sure)*
	Estimated value of all shipments/service (Canadian dollars): (input value)
	What was the weight of the commodities shipped out on this shipment: (input value; select one from: *kg, tonnes, lb, other – specify)*
	Select the type of commodity shipped out on this shipment: (select from dropdown list)
	Select the exact type of commodity shipped out on this shipment: (select from dropdown list – specific to selection in previous question)
	Select the type of service for this outbound commercial trip: (select from dropdown list)

Inbound Commercial Trips	**How many commercial vehicles arrived at your establishment today to deliver goods or services to your establishment?** (input value)
	Type of outbound activity: (select one from: *shipment only, service only, shipment and service)*
	Origin location information. (select one from: *Street address of destination receiving shipment/service, postal/zip code of destination receiving shipment/service,* input address/postal code and city)
	Origin location information. (select one from: *Street address of destination receiving shipment/service, postal/zip code of destination receiving shipment/service*, input address/postal code and city)
	Select all means of transportation used to make this inbound trip: (select as many as apply from: *regular passenger car, pickup truck, minivan or cube van, single unit truck, tractor with one trailer, tractor with two or more trailers, rail, air, marine, other – specify)*
	If the inbound trip involved deliveries, how many shipments were delivered on this trip? (input value)
	The time vehicle arrived at my establishment: (input *hour* and *minute*; select *am* or *pm)*
	Was this a Just In Time (JIT) delivery? (select one from: *yes, no, not sure)*
	Estimated value of all shipments/service (Canadian dollars): (input value)
	What was the weight of the commodities received on this shipment: (input value; select one from: *kg, tonnes, lb, other – specify)*
	Select the type of commodity received on this shipment: (select from dropdown list – specific to selection in previous question)
	Select the exact type of commodity received on this shipment: (select from dropdown list – specific to selection in previous question)
	Select the type of service for this inbound commercial trip: (select from dropdown list)

Appendix C: Screen Capture of Telephone Survey

This appendix contains screenshots of each page of the web-based survey, in the order that respondents would see them.

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Survey

Welcome Message

Dear Respondent,

Welcome to the official webpage of the Business Establishment Commercial Travel Survey. This survey is being conducted by Dr. Hanna Maoh and his research group of the Department of Civil and Environmental Engineering at the University of Windsor. Data collection is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC). Please note the following:

1. The purpose of this online survey is to collect data on trips resulting from the delivery of goods and services within and across key urban areas in Canada. The collected data will be used study the patterns of freight transportation in and between Canadian markets. It will also be used to propose a sustainable goods movement strategy for Canada.
2. The objective of the study is to improve the mobility of goods and services in the studied regions in future years. Upon completion of the study, the results regarding the patterns of freight transportation and the proposed goods movement strategy will be available to you.
3. This survey does not collect specific personal information about individuals. The collected information will be solely used in academic research work and will not be sold or shared with any third party. Furthermore, published research will be in the form of summary tables and/or statistics that cannot identify the characteristics of your establishment and the transportation activities associated with your business.
4. Your participation in this survey is voluntary. You can stop and exit the survey at any time.
5. Your cooperation and completion of the survey is highly appreciated.

For further details and to view or print the Letter of Information for Consent to Participate in Research please click **here**.

To complete the survey, please visit the following link: **Business Establishment Commercial Travel Survey**

Sincerely,
Dr. H. Maoh
Associate Director
Cross Border Institute
University of Windsor

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

Dear Respondent,

Thank you for taking the time to participate in the **Business Establishment Commercial Travel Activities** online questionnaire. Your survey code is a uniquely assigned code, which will enable you to access the survey at any time. Please make sure you keep the code secure while completing the survey.

Sincerely,
Dr. H. Maoh
University of Windsor

Please enter the Access Code you received with the survey invitation e-mail then click the button below to start

Click here to start the survey

Click here to exit this page

Note: During the survey please do not use the Next/Previous buttons of your web browsers, instead use the buttons provided at the end of each survey form.

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

0%

1. Does your establishment engage in delivering goods or services to clients or establishments in the region? (•) Yes ( ) No

2. The total number of employees who work at this location is [5]

3. The number of employees by occupation type in my establishment are as follows:

Management occupations	2
Business, finance and administrative occupations	1
Natural and applied sciences and related occupations	1
Health occupations	1
Occupations in social science, education, government service and religion	0
Occupations in art, culture, recreation and sport	0
Sales and service occupations	0
Trades, transport and equipment operators and related occupations	0
Occupations unique to primary industry	0
Occupations unique to processing, manufacturing and utilities	0
Other	0

4. My establishment has been at this location for [5] years

5. The total square footage of my establishment is [50000] square feet

Cancel   Next >>

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

17%

Establishment Characteristics Information

1. My establishment currently has 2 employees who make on-call service trips
2. My establishment currently has 2 employees who make truck deliveries
3. My establishment has a Supply-Chain/Logistics operator ◉ Yes ○ No
4. My establishment has a Supply-Chain/Logistics specialist ◉ Yes ○ No
5. My establishment makes use of Logistics Software to manage trips ○ Yes ○ No
6. My establishment owns the following commercial vehicles:

1	Regular passenger cars	0	Pickup or cube van
1	Single unit trucks	0	Multiple unit trucks
0	Tractors	0	Containers
0	Trailers	0	Other, specify

7. The industry which best characterizes my establishment is [Select from the list, or choose 'Other, Specify ...' to type in your]
8. The number of commercial trips made by my establishment varies by day of the week ◉ Yes ○ No
9. The busiest days for delivering goods or services are: ☑ Monday ☑ Tuesday ☐ Wednesday ☐ Thursday ☐ Friday ☐ Saturday ☐ Sunday
10. The number of commercial trips made by my establishment varies by month ◉ Yes ○ No
11. The busiest months for delivering goods or services are: ☑ Jan ☐ Feb ☐ Mar ☐ Apr ☐ May ☑ Jun ☑ Jul ☐ Aug ☐ Sept ☐ Oct ☐ Nov ☑ Dec

Cancel | << Previous | Next >>

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

33%

Outbound Trips

1. How many commercial vehicles left your establishment today to deliver goods or services? 1

Cancel | << Previous | Next >>

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

50%

Outbound Trip #1

1. Type of outbound activity: ◉ Shipment only ○ Service only ○ Shipment and Service

2. Destination location information
   - ◉ Street address of destination receiving shipment/service:
   - ○ Postal/Zip code of destination receiving shipment/service:

   City: Province: Please Choose:

3. Select all means of transportation used to make this outbound trip:
   - ☐ Regular passenger car
   - ☐ Pickup truck
   - ☐ Mini van or cube van
   - ☐ Single unit truck
   - ☐ Tractor with one trailer
   - ☐ Tractor with two or more trailers
   - ☐ Rail
   - ☐ Air
   - ☐ Marine
   - ☐ Other, specify

4. If the outbound trip involved deliveries, how many shipments were delivered on this trip? 0

5. The time vehicle left my establishment: Hour 00 : Minute 00 ○ AM ○ PM

6. Was this a Just In Time (JIT) delivery? ○ Yes ○ No ○ Not sure

7. Estimated value of all shipments/service (Canadian dollars): $0

8. What was the weight of the commodities shipped out on this shipment: 0 ○ kg ○ tonnes ○ lb ○ other, specify

9. Select the type of commodity shipped out on this shipment: Select from the list, or choose 'Other, Specify ...' to type in your

10. Select the exact type of commodity shipped out on this shipment: Select from the list, or choose 'Other, Specify ...' to type in your

11. Select the type of service for this outbound commercial trip: Select from the list, or choose 'Other, Specify ...' to type in your

Cancel << Previous Next >>

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

67%

Inbound Trips

1. How many commercial vehicles arrived at your establishment today to deliver goods or services to your establishment? 1

Cancel | << Previous | Next >>

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

83%

Inbound Trip #1

1. Type of Inbound activity: ◉ Shipment only ○ Service only ○ Shipment and Service

2. Origin Location Information
   - ◉ Street address of origin where shipment/service was initiated:
   - ○ Postal/Zip code of origin where shipment/service was initiated:

   City: Province: Please Choose:

3. Select all means of transportation used to make this inbound trip:
   - ☐ Regular passenger car
   - ☐ Pickup Truck
   - ☐ Mini Van or Cube Van
   - ☐ Single Unit Truck
   - ☐ Tractor with one trailer
   - ☐ Tractor with two or more trailers
   - ☐ Rail
   - ☐ Air
   - ☐ Marine
   - ☐ Other, specify

4. If the inbound trip involved deliveries, how many shipments were delivered on this trip? 0

5. The time vehicle arrived at my establishment: Hour 00 : Minute 00 ○ AM ○ PM

6. Was this a Just In Time (JIT) delivery? ○ Yes ○ No ○ Not sure

7. Estimated value of all shipments/service (Canadian dollars): $0

8. What was the weight of the commodities received on this shipment: 0 ○ kg ○ tonnes ○ lb ○ other, specify

9. Select the type of commodity received on this shipment: Select from the list, or choose 'Other, Specify ...' to type in your

10. Select the exact type of commodity received on this shipment Select from the list, or choose 'Other, Specify ...' to type in your

11. Select the type of service for this inbound commercial trip: Select from the list, or choose 'Other, Specify ...' to type in your

Cancel | << Previous | Finish

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

Dear Respondent,

Congratulations, you have successfully completed the survey. Thank you for your effort.

Your help is highly appreciated.

Best regards,
Survey Team

Close Window

University of Windsor

Natural Sciences and Engineering Research Council of Canada

Business Establishment Commercial Travel Activities

Dear Respondent,

You have not successfully completed the survey. However, you can return and complete the survey at your convenience.

Your help is highly appreciated.

Best regards,
Survey Team

Close Window

Appendix D: Descriptions of 3-Digit NAICS Codes

This appendix shows more detailed descriptions for the NAICS codes that shared the same general description at the 2-digit code level. The 3-digit NAICS codes and corresponding descriptions are given to differentiate between these industry classes.

3-Digit NAICS Code	**Description**
311	Manufacturing: Food manufacturing
312	Manufacturing: Beverage and tobacco product manufacturing
313	Manufacturing: Textile mills
314	Manufacturing: Textile product mills
315	Manufacturing: Clothing manufacturing
316	Manufacturing: Leather and allied product manufacturing
321	Manufacturing: Wood product manufacturing
322	Manufacturing: Paper manufacturing
323	Manufacturing: Printing and related support activities
324	Manufacturing: Petroleum and coal product manufacturing
325	Manufacturing: Chemical manufacturing
326	Manufacturing: Plastics and rubber products manufacturing
327	Manufacturing: Non-metallic mineral product manufacturing
331	Manufacturing: Primary metal manufacturing
332	Manufacturing: Fabricated metal product manufacturing
333	Manufacturing: Machinery manufacturing
334	Manufacturing: Computer and electronic product manufacturing
335	Manufacturing: Electrical equipment, appliance and component manufacturing
336	Manufacturing: Transportation equipment manufacturing
337	Manufacturing: Furniture and related product manufacturing
339	Manufacturing: Miscellaneous manufacturing

441	Retail trade: Motor vehicle and parts dealers
442	Retail trade: Furniture and home furnishings stores
443	Retail trade: Electronics and appliance stores
444	Retail trade: Building material and garden equipment and supplies dealers
445	Retail trade: Food and beverage stores
446	Retail trade: Health and personal care stores
447	Retail trade: Gasoline stations
448	Retail trade: Clothing and clothing accessories stores
451	Retail trade: Sporting goods, hobby, book and music stores
452	Retail trade: General merchandise stores
453	Retail trade: Miscellaneous store retailers
454	Retail trade: Non-store retailers
481	Transportation and warehousing: Air transportation
482	Transportation and warehousing: Rail transportation
483	Transportation and warehousing: Water transportation
484	Transportation and warehousing: Truck transportation
485	Transportation and warehousing: Transit and ground passenger transportation
486	Transportation and warehousing:: Pipeline transportation
487	Transportation and warehousing: Scenic and sightseeing transportation
488	Transportation and warehousing: Support activities for rail transportation
491	Transportation and warehousing: Postal service
492	Transportation and warehousing: Couriers and messengers
493	Transportation and warehousing: Warehousing and storage

VITA AUCTORIS

NAME:	Georgiana Madar
PLACE OF BIRTH:	Timisoara, Romania
YEAR OF BIRTH:	1990
EDUCATION:	University of Windsor, BASc in Civil Engineering, Windsor, ON, 2012
	University of Windsor, MASc in Civil Engineering, Windsor, ON, 2014