FIRM DEMOGRAPHY IN URBAN AREAS: AN AGENT-BASED APPROACH

MODELING FIRM DEMOGRAPHY IN URBAN AREAS WITH AN APPLICATION TO HAMILTON, ONTARIO: TOWARDS AN AGENT-BASED MICROSIMULATION MODEL

By

HANNA FRANCIS MAOH, B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2005)
(School of Geography and Earth Sciences)

McMaster University
Hamilton, Ontario

TITLE: Modeling Firm Demography in Urban Areas with an Application to Hamilton, Ontario: Towards an Agent-Based Microsimulation Model

AUTHOR: Hanna F. Maoh, B.Sc. (Bethlehem University, 1996)
M.Sc. (McMaster University, 1999)

SUPERVISOR: Professor Pavlos S. Kanaroglou

NUMBER OF PAGES: xi, 213

ABSTRACT

The concern about sustainable planning of cities over the past few decades has generated a considerable amount of empirical literature on the emerging urban form and its causes. Part of this literature has focused on modeling the behavior driving the location decision of households and firms, as well as the urban development practices to simulate the evolution trends of urban form and growth. Earlier efforts were based on building and using aggregated zone-based urban land use models. However, due to their aggregate nature, most of these models have been criticized for lacking the behavioral realism in imitating the complexities of the city development process. Research during the past decade and a half calls for the adoption of a micro-analytical approach, whereby change in urban form is driven by the collective behavior of individual agents (households, business establishments, developers, etc.) and their location decisions.

This dissertation is concerned with the behavior of business establishments in urban areas. Specifically, we use principles from firm demography to develop a framework for an agent-based microsimulation model for the City of Hamilton, Ontario, Canada. The framework will be used to simulate the evolution of small and medium size business establishments as an outcome of processes that related to the failure, mobility and location choice of establishments in the city. The statistical models underlying these processes are discussed and estimated using data extracted from Statistics Canada Business Register (BR) for the period 1996 – 2002. The parameters of these models attempt to identify and quantify the causes associated with firmographic processes in Hamilton and will form the basis towards implementing an agent-based model.

ACKNOWLEDGMENTS

I would like to extend my sincere appreciation and thanks to my supervisor, Dr. Pavlos S. Kanaroglou for his endless support and limitless patience and time in mentoring me during the years of my graduate studies. Thanks Pavlos for sharing your experience and knowledge with me, and for all the advice and encouragement whenever I needed them. I am also thankful to the members of my supervisory committee Dr. Darren M. Scott and Dr. Eric J. Miller for their useful input to my research. I would like also to thank Dr. John Douglas Hunt and Dr. Spiro Panagiotou for evaluating this dissertation and providing useful comments that helped to improve certain parts of it.

I would like to gratefully acknowledge the support of Statistics Canada through their 2003 – 2004 PhD Research Stipend program. Thanks to Dr. John Baldwin, Dr. Mark Brown and Mr. Desmond Beckstead for providing me with office space and access to the Business Register data to conduct my doctoral research. I am also thankful to them for their useful discussions and input to my research. Special thanks go to Dr. Mark Brown for all his support and mentoring during my tenure at Statistics Canada. I am also grateful to the Social Sciences and Humanities Research Council of Canada (SSHRC) for providing me with a Doctoral Fellowship in the fourth year of my PhD program (award number: 752-2004-1080).

There are many others who deserve an acknowledgment for their support, friendship and words of encouragement. Thank you to Dimitris Potoglou and Theodora Pouliou, Jamie Spinney, Pavlos Kanaroglou and his wife Vina, Jacqueline Binyamin and her husband Ghassan, Darren Scott and his wife Marianne, Mark Brown, Kostas

Apostolou and his wife Angeliki, Talar Sahsuvaroglu, and Maria Triantafullou for their friendship and good hospitality. Special thanks go to Dimitris and Theodora for their sincere friendship, pleasant day-to-day company, hospitality and all the fun we had mixing and matching Arabic and Greek cuisines. Also, thanks to everyone I knew at the School of Geography and Geology and the Center for Spatial Analysis (CSpA) for whatever help or words of encouragement. Thank you to my friends in Bethlehem for their encouragement and e-mail correspondence. Specifically, I would like to mention Majed AbuKubbi and Issa Zboun.

Finally, I would like to extend my deepest and warmest gratitude to my mother Georgette and my two sisters, Antoinette and Mary, for limitless patience and understanding, for unwavering support and encouragement and above all else for the unconditional love they provided all the way till this moment. The words are not sufficient to describe my gratitude and appreciation. Also, thanks to all the members of my family for caring and asking about me.

PREFACE

This dissertation is a compilation of four main research papers that have either been submitted or will be submitted for publication in peer-reviewed journals. The research papers are as follows:

Chapter 2:

Maoh, H.F and P.S Kanaroglou (2005). Agent-Based Firmographic Models: A Simulation Framework For the City Of Hamilton. For submission to *Computers, Environment and Urban Systems.*

Chapter 3:

Maoh, H.F and P.S Kanaroglou (2005). A Discrete-Time Hazard Duration Model of Survival for Small and Medium Business Establishments with an Application to Hamilton, Ontario. Submitted to *Geographical Analysis*.

Chapter 4:

Maoh, H.F and P.S Kanaroglou (2005). Business establishment mobility behavior in urban areas: An application to the city of Hamilton in Ontario, Canada. For submission to the *Professional Geographer*.

Chapter 5:

Maoh, H.F and P.S Kanaroglou (2005). The Location of Business Establishments in the City of Hamilton, Canada: A Micro-Analytical Model Approach. For submission to *Growth and Change*.

It is worth noting that some repetition of content is inherent in the different chapters, particularly in their introduction, given that they are compiled as individual research papers. The various research activities including: literature review, conceptualization, data gathering, compilation, statistical analysis, interpretation of results and writing of the papers were completed by the dissertation author. However, Dr. Pavlos Kanaroglou is co-author of the four papers. His contribution included guidance on research methods, critical appraisal of manuscripts and editorial reviews.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 1: Introduction

1.1 Justification of Research Topic

The concern about sustainable planning of cities over the past few decades has generated a considerable amount of empirical literature on the emerging urban form and its causes. Part of this literature has focused on modeling the behavior driving the location decision of households and business establishments as well as the urban development practices of planners and developers in the city to assess the evolution trends of urban form and growth. In general terms, a household makes a decision on the type and location of a dwelling based on its demographic characteristics, the number of workers in it and the place of employment of its workers. Business establishments, on the other hand, are looking for advantageous profit maximizing locations, while developers will engage in residential or commercial development in locations where they perceive high demand for their output.

All of these location decisions and land development projects have their impact on urban form since they are influenced and driven by the level of accessibility to a particular site in the city. For instance, the addition of an expressway in a city's transport network can have significant effects on the location decision of establishments or on the perceived monetary returns from the residential development of a parcel of land that is adjacent to the new expressway. Similarly, locational decisions, such as the development of a shopping mall at the outerskirts of a city, may necessitate changes in the transportation network, thereby affecting other locational decisions. Also, change in urban form can have a profound influence on the flows of goods and people within the city since they translate into motorized trips on the transportation network. The latter are crucial to the sustainability of the city, since they are responsible for traffic congestion and tailpipe emissions in all the links of the urban transport network. The key point here is that the urban land use and transportation systems interact and influence each other. This two-way relationship is the driving force that governs the sustainability of cities.

Earlier efforts to assess the impact of land use and transportation interactions were based on building and using zone-based integrated urban land use models (IUMs). Such models rely on computer simulations of urban processes. IUMs are regarded as sophisticated planning tools that could assist public officials in their decision-making by allowing them to evaluate the long-term repercussions of proposed land-use and transport network changes. A number of operational IUMs for cities around the world are in existence (see Southworth, 1995; Miller et al., 1998; Kanaroglou and Scott, 2002; Wegener, 2004 and Hunt et al., 2005 for a comprehensive review of these models). However, a wave of research is taking place to improve on these models in response to environmental concerns pertaining to green house gas emissions in cities. Identified drawbacks in existing models include: (1) Their aggregate nature; (2) their static nature and lack of behavioral realism and (3) their weak response to long-term planning policies due to their inadequate policy sensitivity. Such drawbacks are believed to impede the performance of these models (Miller et al., 2004).

Recently, many scholars propose the micro-analytical modeling framework to rectify these drawbacks. Within this framework, urban development is driven by the micro behavior of the various agents (persons, households, business establishments, developers, etc.) that interact in the city. Accordingly, a micro-analytical theory of urban development has to decompose the total processes of urban change into subprocesses and within these identify the main actors (agents) and their decision behavior (Wegener and Spiekermann 1996). The dynamic behavior within each process and for each individual agent is then explicitly simulated over both time and space to generate aggregate system behavior (Miller et al. 2004). Such approach is believed to result in agent-based microsimulation models that are truly behavioral and policy sensitive, capable of promoting a wide range of sustainability practices in urban areas.

The adoption, in recent years, of the activity-based approach for understanding and modeling the urban travel behavior of individual persons and households necessitated the development of disaggregate land use models. Within these models, the locations of individual households and business establishments are used as input to the activity-based model. Therefore, it is a chief task to model and be able to simulate the change in the location of individual households and business establishments when developing agent-based IUMs.

In accordance with this paradigm, this dissertation is concerned with the location of business establishments. Specifically, we use principles from firm demography to develop a framework for an agent-based microsimulation model for the City of Hamilton, Ontario, Canada. With few exceptions, our review of the literature suggests a clear shortage of empirical studies that employ the micro approach to analyze the urban development process in general and the firm demographic behavior in particular. The review also indicates that research to devise operational agent-based IUMs is still in its early stages (Miller et al., 2004), albeit microsimulation is not a new topic in the social sciences. A survey conducted by Merz (1991) shows that microsimulation as a tool to analyze the individual impacts of economic and social policies goes back to the early 1960s. However, the lack of spatio-temporal micro data on individual households and business establishments remains as one of the major obstacles that impede the progress in developing operational agent-based IUMs.

Firm demography or firmography is dedicated to the study of processes that relate to the establishment of new firms in a study area or to the growth, decline, failure, or migration of existing firms. It is an interdisciplinary research area in the social sciences (Van Wissen, 1997), with most of the literature focusing on studies at the national and regional, as opposed to the urban level. Like the demography of people, firmography is concerned with identifying and quantifying the causes associated with firmographic processes. It focuses on the dynamic changes that take place among the population of firms in a country or a region (Van Wissen, 1997; Dijk and Pellenbarg, 1999). It is also ideal for analyzing the structure and evolution of the space economy. This is accomplished by examining the spatial distribution of groups of firms, identified by their economic sector, and the evolution of this distribution over time. Characteristic of this approach is that it pays special attention to the spatial dimension. Therefore, the firm demographic approach is comprehensive and can be used to study the evolution in the form of contemporary cities.

Firmography is a relatively new field of research that first emerged in the early 1980s to study economic regional development. However, most of the empirical studies in this area were conducted in the 1990s (Dijk and Pellenbarg, 1999) with the aim of studying particular firmographic process, such as the formation (entry in to the market) of firms (for example Geroski, 1995; Berglund and Brannas, 2001; Van Wissen, 1997; 2000); failure (exit from the market) of firms (for example Audretsch and Mahmood, 1995; Mata and Portugal, 1994; Van Wissen, 1997; 2000; Baldwin et al., 2000; Berglund and Brannas, 2001); migration of firms (for example Dijk et al., 1999; Dijk and Pellenbarg, 2000; Van Wissen, 2000; Brouwer et al. 2004) or growth/decline of firms (for example Evans, 1987; Hart and Oulton, 1996; Audrestsch, 1995). Moreover, there were no attempts in the past to use firmography for projections in order to assess planning policies. Van Wissen (2000) notes that model specification and estimation within a firmographic framework is relatively new and received special attention by organizational ecologists, such as Hannan and Freeman (1989), for the first time in the late 1980s. Therefore, firmography is a new approach for policy analysis and makes for an interesting area of research that can advance the development of policy related decision support tools.

The few studies reported by Van Wissen (2000), Arbia (2001), Brenner (2001), Otter et al. (2001), Waddell et al. (2003) and Hunt et al. (2003) to develop agent-based firmographic models provide a positive start. However, none of the former studies provide a comprehensive operational model at the urban level. Van Wissen (2000) developed *SIMFIRMS* for the Netherlands with the objective to microsimulate the life trajectory of firms at the municipal and district levels. The model is comprehensive in the sense that it microsimulates all the firmographic processes pertaining to the formation, failure, migration and growth/decline of individual firms. The work of Brenner (2001) is focused on simulating the evolution of localized industrial clusters at the regional level as an outcome of the formation, failure and growth of firms. However, the model utilizes economic theory to evolve firms in a hypothetical situation and does not rely on real world data in its simulations. Similarly, the model devised by Otter et al. (2001) is not based on real world observations and is concerned with the behavior of firms and households at the regional level. It is also based on economic theory and aspects of complexity theory and decision rules. Arbia (2001) uses concepts from firmography to study economic clustering in the regional context. His work is based on estimating statistical models using data from the San Marino Republic. On the other hand, Waddell et al. (2003) and Hunt et al. (2003) develop firm demographic models at the urban level. However, their modeling efforts have mainly been focusing on the location choice of intra-urban mobile firms. In both cases, the location choice problem is handled via models derived from maximum utility theory.

This dissertation opts to fill the existing gap in the literature by devising a framework for an agent-based firmographic model at the urban scale. The proposed framework aims to integrate the various firmographic processes into a unified system which is capable of simulating the evolution of business establishment population as an outcome of processes that relate to the failure, mobility and location choice of individual establishments. We focus on the behavior of small and medium size (SME) establishments with less than 200 employees in the city of Hamilton in Ontario, Canada. SME establishments have a dynamic firmographic behavior compared to larger establishments. The latter have a more predictable behavior, given their small number and the high monetary value associated with their infrastructure. Therefore, large establishments will be treated exogenously in the devised agent-based model. In the chapters to follow, the statistical models underlying the former processes are discussed and estimated using data extracted from Statistics Canada Business Register (BR) for the period 1996 – 2002. To our knowledge, this is the first study that utilizes data from the BR to identify and quantify the causes associated with firmographic processes for a city in the Canadian context.

1.2 Scope of Research Topic

The need to improve upon the conventional zone-based methods used to model the location of firms in Integrated Urban Land Use and Transportation Models is the primary objective of this dissertation. To fulfill this objective, the research topic is focused on developing an agent-based microsimulation Decision Support Tool (DST). The modeling framework is based on the idea of evolving a business establishment population over time and space as an outcome of demographic events that affect the life trajectory of individual business establishments. Obviously, the successful implementation of such DST requires a critical evaluation of the factors driving the firm demographic events. Therefore, the scope of this dissertation is defined by the task of identifying and quantifying the causes associated with the firmographic processes that drive the evolution of establishments. Accordingly, three specific objectives are identified and fulfilled. First, firm demographic research is reviewed to assess the current state of knowledge in terms of the modeling methods and empirical findings. Second, a conceptual framework for modeling the evolutionary process of the small and medium size business establishments in the city of Hamilton is developed. The framework is based on three submodules that will be used to model the failure, mobility and location choice behavior of individual establishments. Finally, a number of statistical models underlying the firmographic submodules are formulated, specified and estimated. This is followed by an interpretation of the statistical results. The outcome of these results is expected to serve two purposes: (1) advance the current state of knowledge on firm demographic processes in the urban context, and (2) set the basis for implementing the firmographic DST.

1.3 Contents of Dissertation

The remainder of the dissertation is organized as follows. Chapter 2 presents the simulation framework for the agent-based firmographic model proposed in this dissertation. It starts by reviewing the progress of firm demographic research to assess the current state of knowledge in terms of modeling methods and empirical findings. Accordingly, an overview of the firmographic concept is presented to highlight the different firmographic processes and the associated factors that affect the life trajectory of firms. The review also sheds light on the microsimulation approach in firmography to identify the different modeling frameworks used to date in developing agent-based firmographic models. Next, a conceptual framework for modeling the evolution of small and medium size business establishments is proposed. The framework is comprehensive and assumes that the evolutionary process of establishments can be modeled as an outcome of subprocesses that relate to failure, mobility and location choice events. These three subprocesses will be handled via three inter-linked submodules. The chapter also emphasizes the nature of the firm micro data that are extracted from the Statistics Canada Business Register (BR) and used in model estimations. Given the confidential nature of the BR data, the chapter also emphasizes the means of creating a synthetic population that can be used along with the estimated parameters to implement the proposed framework. It also summarizes the methods of analysis that are used to model the failure, mobility and location choice processes. Finally, a discussion about the model development issues is provided to highlight the state of the art technologies that can be used to facilitate the implementation of sophisticated agent-based microsimulation models.

Chapter 3 presents the empirical findings for the model underlying the establishment failure submodule. First, an overview of factors and determinants influencing the survival and failure of firms is provided based on the most recent empirical literature. This is followed by a non-parametric survival analysis to explore the determinants that affect the survival of business establishments in the City of Hamilton. Finally, the chapter discusses the formulation of the discrete-time hazard duration model, specified to handle the failure process of business establishments. Unlike cross-sectional models, the devised model utilizes a longitudinal dataset from the BR to analyze the annual life trajectory of each establishment in the 1996 cohort till 2002. Such data enrich the analysis since the temporal dimension is explicitly incorporated in the model. The estimation results suggest that the business failure process in Hamilton can be explained by the establishment characteristics, the industry it belongs to, the geographic location it is at and the macro-economic conditions in the city. Among the above factors, establishment characteristics and macro-economic conditions appear to be the most influential in explaining the failure of business establishments.

Chapter 4 presents the empirical results for the model underlying the establishment mobility submodule. The chapter starts with a review of the theoretical frameworks used to explain the mobility behavior of establishments. It also attempts to shed light on the determinants that drive this behavior. However, the review concludes that firm mobility studies are scarce and not well addressed, especially in the urban context. To this end, the chapter explores the establishment mobility trends in the city of Hamilton. It also models the mobility behavior based on a micro dataset extracted from the BR for 1996 and 1997. Accordingly, we estimate multinomial logit models by economic sector to predict the probability of staying, moving or leaving the city. Estimation results suggest that the willingness to move can be explained by age, growth rate, size and the industry of the establishment. We also find that as stress from local competition increases, an establishment will tend to move a longer distance. On the other hand, agglomeration economies increase the propensity of staying at the current location. These results highlight the existence of market power and market share in Hamilton. The estimated models suggest that mobility is more pronounced among the Central Business District (CBD) establishments.

Chapter 5 presents the empirical findings for the model underlying the establishment location choice submodule. First, a literature review is conducted to highlight the intra-metropolitan location factors that affect the location choice behavior of firms. Next, econometric models are formulated and specified to analyze the location choice behavior of establishments. Unlike earlier studies, the proposed framework models the location choice of intra-urban migrating, newly formed and in-migrating establishments simultaneously. It also analyzes the consistency in location choice behavior over time by modeling the process for the 1996 – 1997 and 2001 – 2002 periods. The modeling framework divides the developed land in the city into grid cells of 200 by 200 meters and predicts the location choice of each establishment at this spatial resolution. Land use, population census and business establishment data are geo-coded to grid cells of the developed land to create the explanatory variables of the econometric models. Estimation results suggest that CBD, highway and mall proximities, population density, new residential development and agglomeration economies influence the location decision of business establishments. The results also depict the existence of land use specialization and economic clustering in the region. The estimation for the second period shows a consistency in the location behavior over time as depicted from the estimated parameters of the 2001 - 2002 models.

Finally, Chapter 6 discusses the major finding of the dissertation and suggests directions for future research.

1.4 References

Arbia, G. (2001) Modeling the geography of economic activities on a continuous space. *Papers in Regional Science,* **80,** 411-424.

Audretsch, D. and Talat, M. (1995) New firm survival: new results using a hazard function. *Review of Economics and Statistics,* **77,** 97-103.

Audretsch, D. (1995) Innovation, growth and survival. *International Journal of Industrial Organization,* **13,** 441-457.

Baldwin, J., Bian, L., Dupuy, R. and Gellatly, G. (2000) *Failure rates for new Canadian firms: new perspectives on entry and exit.* Statistics Canada, Ottawa.

Berglund, E. and Brannas, K. (2001) Plants' entry and exit in Swedish municipalities. *The Annals of Regional Science,* **35,** 431-448.

Brenner, T. (2001) Simulating the evolution of localised industrial clusters - An identification of the basic mechanisms. *Journal of Artificial Societies and Social Simulation,* **4,** Retrieved February 2, 2005, from http://www.soc.surrey.ac.uk/JASSS/4/3/4.html.

Brouwer, A., Mariotti, I. and Ommeren, J. (2004) The firm relocation decision: an empirical investigation. *The annals of regional science,* **38,** 335-347.

Dijk, J. and Pellenbarg, P. (1999) The demography of firms: progress and problems in empirical research. In Dijk, J. v. and Pellenbarg, P. H. (eds.), *Demography of Firms: Spatial Dynamics of Firm Behaviour.* Netherlands Geographical Studies, KNAG, RUG, pp. 325-337.

Dijk, J., Pellenbarg, P. and Steen, P. (1999) Determinants of firm migration in the Netherlands: An exercise in the demography of firms approach. In Dijk, J. and Pellenbarg, P. (eds.), *Demography of Firms: Spatial Dynamics of Firm Behaviour.* Netherlands Geographical Studies, KNAG, RUG, pp. 87-121.

Dijk, J. and Pellenbarg, P. (2000) Firm relocation decisions in the Netherlands: an ordered logit approach. *Papers in Regional Science,* **79,** 191-219.

Evans, D. (1987) The relationship between firm growth, size, and age: estimates for 100 manufacturing industries. *Journal of Industrial Economics,* **35,** 567-581.

Geroski, P. (1995) What do we know about entry? *International Journal of Industrial Organization,* **13,** 421-440.

Hannan, M. and Freeman, J. (1989) *Organizational Ecology.* Harvard University Press, Cambridge.

Hart, M. and Scott, R. (1993) Measuring the effectiveness of small firm policy: Some lessons from Northern Ireland. *Regional Studies,* **28,** 849-858.

Hunt, J., Khan, J. and Abraham, J. (2003) Microsimulating firm spatial behavior. In *Eighth International Conference on Computers in Urban Planning and Urban Management "CUPUM, "* Sendai, Japan.

Hunt, J., Kriger, D. and Miller, E. (2005) Current operational urban land-use-transport modelling frameworks: a review. *Transport Reviews,* **25,** 329-376.

Kanaroglou, P. and Scott, D. (2002) Integrated urban transportation and land-use models for policy analysis. In Dijst, M. (eds.), *Governing Cities on the Move: Functional and Management Perspectives on Transformations of Urban Infrastructure in European Agglomerations.* Ashgate, pp. 43-75.

Mata, J. and Portugal, P. (1994) Life duration of new firms. *Journal of Industrial Economics,* **42,** 227-246.

Merz, J. (1991) Microsimulation - A survey of principles, developments and applications. *International Journal of Forecasting,* **7,** 77-104.

Miller, E., Kriger, D. and Hunt, J. (1998) Integrated urban models for simulation of transit and landuse policies. *TCRP Web Document 9,* **Project H-12,** 1-252.

Miller, E., Hunt, J., Abraham, J. and Salvini, P. (2004) Microsimulating urban systems. *Computers, environment and urban systems,* **28,** 9-44.

Otter, H., Veen, A. and Vriend, H. (2001) ABLOOM: Location behavior, spatial patterns, and agent-based modelling. *Journal of Artificial Societies and Social Simulation,* **4,** Retrieved February 2, 2005, from http://www.soc.surrey.ac.uk/JASSS/4/4/2.html.

Southworth, F. (1995) A technical review of urban land use-transportation models as tools for evaluating vehicle travel reduction strategies. Oak Ridge National Laboratory, Tennessee, pp. 1-114.

Waddell, P., Borning, A., Noth, M., Freier, N., Becke, M. and Ullfarsson, G. (2003) Microsimulation of urban development and location choices: design and implementation of UrbanSim. *Networks and spatial economics,* **3,** 43-67.

Wegener, M. and Spiekermann, K. (1996) The potential of microsimulation for urban models. In Clarke, G. P. (eds.), *Microsimulation for urban and regional policy analysis.* London: Pion, pp. 149-163.

Wegener, M. (2004) Overview of land use transport models. In Hensher, D., Button, K., Haynes, K. and Stopher, P. (eds.), *Handbook of transport geography and spatial systems.* Elsevier, pp. 127-146.

Van Wissen, L. (1997) Demography of the firm: modelling birth and death of firms using the concept of carrying capacity. In Brekel, H. and Deven, F. (eds.), *Population and Families in the Low Countries 1996/1997.* Den Haag, Brussel: NIDI/CBGS, pp. 219-244.

Van Wissen, L. (2000) A micro-simulation model of firms: applications of concepts of the demography of the firm. *Papers in Regional Science,* **79,** 111-134.

CHAPTER 2: Agent-Based Firmographic Models: A Simulation Framework for the City of Hamilton

2.1 Introduction

Many North American cities have changed in form over the past few decades. The observed changes were, to a large extent, the outcome of increasing automobile dependency along with the continuous investment in road infrastructure and land use development projects. Changes in urban form include the outward expansion of the city boundary, known as urban sprawl, and the layout of the internal city structure that contributes to the green house gas emissions, threatening the sustainability of the natural habitat. Such environmental concerns have put the future of urban development at the top of the political agenda in North America (Dieleman *et al.*, 2002). Academics, practitioners and politicians are interested in developing a planning process that will provide a path to a sustainable city future.

A large body of research for more than three decades has concentrated on understanding the interrelation between urban land-use and travel behaviour to assess the sustainability of cities. Part of this research effort has focused on modeling the dynamics of urban land-use and the transportation system, as well as the feedback loops between them. Models that address this issue are known as integrated urban land use and transportation models (IUMs). They constitute a family of computer simulation models that allow decision makers to assess the impact of land-use and transportation infrastructure projects on the evolution of urban areas. In this respect, IUMs are sophisticated Decision Support Systems (DSS) that allow planners and other decision makers to design policies that will lead to sustainable cities. A number of operational IUMs for cities around the world are in existence (see Southworth, 1995; Wegener 1998; Miller *et al.* 1998; Kanaroglou and Scott, 2002; Wegener, 2004; Hunt *et al.*, 2005 for a review of these models). While these operational models have been used in the planning process of cities, a wave of research is taking place to improve the capabilities of these models. Identified drawbacks include: (1) Their aggregate (macro) nature; (2) their static nature and lack of behavioural realism and (3) the weak response to long-term planning policies due to their inadequate policy sensitivity (Miller *et al.*, 1998).

Over the last decade and a half conceptualization within an IUM has evolved from macro to micro. Recent thinking calls for understanding the complexities of the micro-behaviour of agents and the interactions among them in the urban system. According to Wegener and Spiekermann (1996), a microanalytical theory of urban development has to decompose the total process of urban change into subprocesses and within these identify the main actors and their decision behaviour. Therefore, there is a need to identify the subprocesses that affect the dynamic changes in the spatial distribution of households and firms in an urban area. Thus, the focus of current research is on developing disaggregated models that identify individual persons, households, developers and business establishments as the units of analysis. It is believed that such research will lead to models that are behavioural and policy sensitive, capable of promoting a wide range of sustainability practices in urban areas (Miller *et al.*, 2004).

This paper is concerned with the behaviour of firms in urban areas. Specifically, we use principles from firmography to develop a framework for an agent-based microsimulation model. The model will be used to simulate the evolution of business establishments as an outcome of processes that are related to the failure, mobility and location choice of business establishments in the city. While the framework we discuss has general applicability we focus our attention to an application for the City of Hamilton, Canada. A review of the literature indicates the absence of studies on operational agent-based firmographic models in the urban context. We attempt to fill this gap by proposing a modeling framework that can be used as a prototype for the development of operational agent-based firmographic models. Such a model can be used to assess the interplay between the local economy and urban form.

Here is a brief description of the remainder of this paper. In the next section we provide a background about the concept of the demography of firms and review the recent work that has been done in this field. We also shed light on the microsimulation approach in firmography. Section three describes the modeling framework we are proposing and highlights the nature of firm micro data suitable for developing agent-based firmographic models. We also discuss the analytical methods that were used to formulate the econometric models of the microsimulation system. Section four provides a discussion on the methods that can be used to develop and implement microsimulation firmographic models. Finally, we conclude with a summary of the paper and highlight the future direction of this research.

2.2 Background

2.2.1 The Demography of Firms

A growing body of literature, characterized by the term "firm demography" or "firmography", is dedicated to the study of processes that relate to the establishment of new firms in a study area or to the growth, decline, failure, or migration of existing firms. This is an interdisciplinary research area in the social sciences (Van Wissen, 1997), with most of the literature focusing on studies at the national and regional, as opposed to the urban level. Like the demography of people, firmography is concerned with identifying and quantifying the causes associated with firmographic processes. It focuses on the dynamic changes that take place among the population of firms in a country or a region. Birch (1979) was a pioneer in using the firm demographic approach as a tool to study economic development in the regional context. He argued that the job generating process could be analyzed as an outcome of the basic firm demographic processes we listed above.

According to Van Wissen (1997) firm demography is used to analyze the dynamics of sectors in an economy, conceptualized as the outcome of the behaviour of individual firms. Thus, the unit of analysis in such a framework is the individual firm, which is subjected to the potential events of birth (formation), growth/decline, migration (intra-urban relocation, in-migration and out-migration) and death (failure). The firm demography approach is therefore comprehensive and can provide a behavioural view of economic development at the national or regional scale (Van Wissen, 1997; Dijk and Pellenbarg, 1999). It is ideal for analyzing not only the firm population size but also the structure and evolution of the space economy. This is accomplished by examining the spatial distribution of groups of firms, identified by their economic sector, and the evolution of this distribution over time. Characteristic of this approach is that it pays special attention to the spatial dimension.

Dijk and Pellenbarg (1999) note that most of the firm demographic research since the early 1980s has evolved from the organizational ecology and industrial organization schools of thought. The study of entry and exit rates of particular industries by ecologists such as Hannan and Freeman (1989) and Carroll and Hannan (2000) is at the heart of the organizational ecology approach. They rely heavily on the number of firms (density) in the different industries to explain the entry and exit rates. Therefore, they use the population as their level of analysis to statistically examine the birth (entry) and mortality (exit) of organizations and organizational forms within the population over long periods. The industrial organization approach by contrast to the organizational ecology approach pays more attention to the behaviour of individual firms (Dijk and Pellenbarg, 1999). As a result, the industrial organization approach is deemed more appropriate for the development of agent-based simulation models (known as microsimulation models).

The industrial organization approach is mainly based on the work of economists who studied three main firm demographic processes. These are: (a) firm exit or the failure process (for example Audretsch and Mahmood, 1995; Mata and Portugal, 1994; Van Wissen, 1997; 2000; Baldwin et al., 2000; Berglund and Brannas, 2001), (b) firm entry or the birth process (for example Geroski, 1995; Berglund and Brannas, 2001; Van Wissen, 1997; 2000), and (c) firm growth/decline (for example Evans, 1987; Hart and Oulton, 1996; Audrestsch, 1995). Caves (1998) provides an extensive review on the topic based on the existing body of empirical industrial organization literature. Geographers and spatial scientists, on the other hand, contributed to the industrial organization approach by focusing on the migration behaviour of individual firms (for example Dijk et al., 1999; Dijk and Pellenbarg, 2000; Van Wissen, 2000 and Khan et al., 2002; Brouwer et al. 2004).

The birth of a firm is referred to as the firm entry to the spatial market. Firm entry is associated with the ability to penetrate the market in an attempt to retain a profit higher than the nominal profit (Berglund and Brannas, 2001). There are a number of factors that contribute to the explanation of firm entry. Firm entry is central to the theory of entrepreneurial choice. Here, the formation of the firm is motivated by the decision of an entrepreneur to change status from being an employee to being self-employed. In this context, the propensity for firm entry will be higher if the expected income from entry into the market is higher than the expected income from employment. Studies have shown that regions with a large fraction employed in small firms are more fertile. This is the case since employees can gain the experience they need to run a small firm and eventually will start their own business if they find it more fruitful. This suggests that the number of small firms in a region has a positive effect on firm entry (Berglund and Brannas, 2001; Johnson and Cathcart, 1979). Regions with a specialized and well-trained work force are more attractive for entry than regions that need to recruit employment; therefore, employment structure can be used to explain firm entry (Berglund and Brannas, 2001; Garfoli, 1994).

Not surprisingly, access to capital is crucial for establishing a new firm. Therefore, rich regions have higher entry level to the market. Studies also concluded that government incentives play a major role in the birth of firms, especially R&D firms (Hart and Scott, 1993; Berglund and Brannas, 2001; Dijk and Pellenbarg, 2000; Frenkel, 2001). Geroski (1995) provides a rather interesting review on firm entry. He summarizes most of the empirical work that has been conducted on entry as stylized facts that report what drives firm entry and to what extent firm entry influences the market. Among the stylizations that he draws from empirical research and were not mentioned above is that high rates of entry are often associated with high rates of innovation and increases in efficiency.

The failure of firms, or what is known in the literature as firm exit from the spatial market, is the outcome of two main causes: bankruptcy or involuntary liquidation, and voluntary exit (Dietmar et al., 1998). Bankruptcy occurs when the firms' profit declines and becomes less than the nominal profit. Berglund and Brannas (2001) suggest that involuntary exit is influenced by local market conditions rather than national conditions. The local market condition is dependent on the level of competition a firm is facing. In this respect, it is suggested that high level of firm entry triggers competition resulting in a high level of firm exit among the new and incumbent firms. On the other hand, studies have found that agglomeration economies have negative effect on the exit rate of firms (Berglund and Brannas, 2001; Love, 1996). Such a result is anticipated since agglomeration economies are associated with a number of externalities that can provide benefits to the firm. Empirical studies have found that the survival rate of the firm is positively related to its age and size (Mata and Portugal, 1994; Audretsch and Mahmood, 1995; Geroski, 1995; van Wissen, 1997; Baldwin et al., 2000). Small firms are more likely to be born and die in a short period as compared to large companies and enterprises. Also, old firms have a lower propensity of failure compared to young firms. The technological changes in a given industry might lead the firm into bankruptcy and exiting from the market if the firm cannot accommodate for the change in technology or is faced with a lack of specialized labor.

Firm relocation or migration is the movement of the firm from one location to another in order to retain an optimal level of profit. According to Bade (1983), firms have a strong preference to remain in situ and will only move in situations of high locational pressure caused by location deficiencies. In order to understand the relocation process we adopt the scheme used in many studies, which classify factors that influence the decision to relocate into two; push factors which explain the reasons to leave the present location, and pull factors which highlight the forces that attract a firm to the new location. The optimal location a firm might have when it is formed may become less optimal in terms of the profit due to location deficiencies. Van Wissen (2000) notes that location deficiencies (push factors) may arise as a result of changing market orientation, space requirements, technological change, location costs, accessibility problems, local policy and labour market mismatch. Such factors are also highlighted elsewhere in the literature (See for example Dijk and Pellenbarg, 2000; Brouwer et al. 2004). Once a moving decision has been taken, the firm will search for a new location. The search process is known in the literature as the location choice problem. The problem entails evaluating and ranking different locations based on their attractiveness (pull factors). Van Wissen (2000) argues that pull factors are to some extent the mirror of the push factors but with a positive content. Therefore, the attractiveness of the site is a function of a number of factors that may include: locational quality, better market orientation, higher accessibility, better labour market, more space, governmental incentives, etc.

The firm growth/decline process determines the change in firm size and consequently the spatial distribution of jobs in a region or a country. Growth generates more job opportunities as a result of expansion and new recruitment. In some cases, the expansion associated with growth can lead the firm to have more than one branch, thus contributes directly to more births. On the other hand, decline is associated with the loss of jobs and in certain circumstances can lead to the failure of the firm. Based on the existing literature, a number of hypotheses with respect to firm growth can be postulated.

One of the well-known firm growth hypotheses in the economic literature is the Gibrat's hypothesis, which is also known as the Gibrat's law of proportionate. This law states that the proportionate change in the size of a firm is independent of its initial size. However, the validity of this assumption has been questioned by some scholars such as Evans (1987) and Hart and Oulton (1996). Evans (1987) findings suggest that firm growth decreases with firm age and that firm growth decreases with firm size. He also found that the relationship between firm growth and firm size is highly non-linear so that the growth-size relationship varies over the size distribution of firms. Hart and Oulton (1996) obtained similar results to those achieved by Evans (1987). Their results show that for the sample of firms as a whole, the estimated model shows that growth is negatively related to initial size. However, when they split the sample by size group, they found that small firms with 8 employees or less are generating more jobs (high growth) whereas for larger firms, there is essentially no relation between growth and size. Van Wissen (2000) argues that the life cycle of the firm, in which life shocks such as relocation or merger, may trigger new growth. He also argues that market conditions determine growth to a large extent.

While firm demographic studies have had the interest in quantifying the causes that drive the evolution of firms via statistical modelling, there were no attempts in the past to use these models for projections in order to assess planning policies (Van Wissen, 2000). Therefore, firmography is a relatively new approach for policy analysis and makes for an interesting area of research in the social sciences. In this context, we believe that the use of the firmographic technique to develop microsimulation models will bridge many of the gaps that currently exist.

2.2.2 *The Microsimulation Approach in Firmography*

2.2.2.1 Simulating Firms in Conventional IUMs

The application of land use simulation models to study the evolution of urban areas can be traced to early 1960s. The model introduced by Ira Lowry in 1964 was the first of a family of urban land use models that has evolved over the past few decades (Horowitz, 2004). The principal use of a Lowry-type model is to allocate a fixed amount of population and employment to zones of a region for the purpose of land use planning or travel forecasting. The allocation process is based on a recursive procedure that attempts to predict the change in population as a function of available employment and vice versa. Three schools of thought have influenced the development of Lowry-type firm location models. The first school extended the conceptual framework proposed by Lowry (1964) to generate a family of operational firm location models that are formulated as spatial interaction models. Putman (1983) was a pioneer in developing this class of gravity-based models since the early 1970s. Compared to other existing models, gravity-based firm location models are simplistic in their theoretical foundation for at least two reasons. First, they do not account for the linkages between the different sectors of the economy. Second, they lack the mechanism to simulate the land market clearing processes. The theoretical foundation of such models is based on the assumption that the change in the employment of a zone is based on an attractiveness measure in the zone, the travel time between the place of residence zone and place of work zone and the current location of the region's residents and employment. Therefore, such models lack the behavioural realism needed to simulate urban complexities. However, the gravity-based modeling approach has the advantage of relying on minimal data sources compared to other modeling approaches.

The second school of thought pays more attention to linkages between the different economic sectors and the land market clearing process. Macgill (1977) conducted early experiments to account for the economic linkages by formulating the Lowry model as an input-output model that derived its final demand according to entropy maximization. De La Barra (1984) adopted the same approach in developing the *TRANUS* land use model. However, *TRANUS* is formulated within the framework of a random utility and input-output model. A similar modeling approach was adopted to develop the *MEPLAN* model (Hunt and Simmonds 1993; Echenique and Owers 1994; Abraham and Hunt 1999). However, *MEPLAN* appears to offer the most experience with the input-output approach to date (Southworth, 1995).

MEPLAN treats land use and transportation as two parallel and interacting markets. Therefore, the land market clearing process is explicitly simulated in the model under the assumption of system equilibrium. Behavior in each market is modeled as a response to money-price and price-like signals (including travel disutility). For instance, in the land market, production, consumption and the location decision of households and firms are driven by price signals. The land market model is based on a spatially disaggregate input-output social accounting matrix (SAM) that is expanded to include variable technical coefficients and zonal land use information such as the building floor space and type in a given zone. The volume of firms from the different economic sectors as defined in SAM is assigned to geographic zones using a multinomial logit (MNL) model. The MNL works as a distributing mechanism to allocate the productions from all the sectors that should be consumed by a firm of type n in zone j. The utility function of the MNL is a composite of four elements that include: the total cost of producing unit n in zone j; zone specific disutility associated with producing n in a supplying zone i; the cost of transporting unit n from i to j and the floor space size available to unit n in zone j. The disutility of transporting the firm production to the consumer suggests that customers are more likely to purchase from close zones or equivalently firms are more likely to supply to close zones. Moreover, firms will locate in zones that are cheaper since production cost includes the cost of purchasing and transporting to various inputs to production. Also, firms will want to locate so that they are closer to their suppliers and their labor. To this end, *MEPLAN* is capable of representing the bidding process for land and the classic trade offs in location choice. That is: (1) the trade off between proximity to suppliers and being closer to customers and (2) the trade off between transport costs and land costs. Therefore, *MEPLAN* provides a significant theoretical advancement compared to the gravity-based models. However, the implementation of *MEPLAN* requires extensive data that are not always available.

The third school of thought is based on ideas that were originally generated in the urban economic literature to model the location of firms (for example Hansen, 1987; Shukla and Waddell, 1991; Waddell and Shukla, 1993). This literature utilizes the concept of maximum utility to model the location choice of firms in the city. Within this framework, the firm is assumed to be a profit maximizer and typically will search the metropolitan area for the location that will maximize its profit. Following Martinez (1992), the problem can be modeled using the bid-choice theory. Here, the city is divided into a set of mutually exclusive zones and the location of firms is modeled as the problem of choosing the zone with the maximum utility. Accordingly, the probability of choosing a particular zone can be predicted via a multinomial logit model. The estimation of such model will depend on formulating a utility function that incorporates factors relating to the attributes of the zone and the characteristics of the firm. Such approach has been used to model the location of firms in the *MUSSA* model (Martinez, 1996). The utility of the firm location model in *MUSSA* is a composite of: attractiveness measure reflecting the number of trips with purpose of interest for firm type f in the zone, the average households' income in the zone, the floor-space by type of building in the zone, lot size in location option type v in the zone, a dummy equal to zero if v is a flat, and a dummy equal one if f stands for retail and service.

While the idea of enhancing land use models remained under development in academic research for a long time; recently, there has been a renewed interest in urban land-use models as potential tools for policy analysis. This renewed interest presented new challenges to the urban-modeling community due to the need for more detailed information on household demographics and workplace characteristics, which are required as input for the new generation of activity-based travel demand models (Wegener and Spiekermann, 1996). The latter are believed to be more appropriate than the four-stage model when studying complex urban transportation systems. Therefore, the Lowry-type zone-based models are not suitable for these emerging micro-level travel demand models. This is the case because zone-based models do not take account of topological relationships and ignore the fact that socio-economic activities are continuous in space. They also assume a static system that is in equilibrium state and to a large extent they assume homogeneity among the actors that drive the evolution of the system. This suggests a need to shift to urban models that are based on examining the dynamic behavior of individual households and firms over both time and space to generate aggregate system behavior. Such approach gives rise to the so-called agent-based microsimulation land use models.

2.2.2.2 Agent-Based Microsimulation Models

A review of the recent literature indicates that research to devise agent-based microsimulation land use models is still in its early stages (Miller et al., 2004), albeit microsimulation is not a new topic in the social sciences. A survey conducted by Merz (1991) shows that microsimulation as a tool to analyze the individual impacts of economic and social policies goes back to the early 1960s. In the case of firm demography, the very few studies found in literature (Van Wissen, 2000; Khan et al., 2002; Waddell et al., 2003) provide a positive start on the topic. Van Wissen (2000) developed the SIMFIRMS model to microsimulate the evolution of the Netherlands firm population as an outcome of the four firmographic processes highlighted in the previous section. Despite that SIMFIRMS is a regional firmographic model, the ideas used in its development are interesting and could be applied to study the evolution of the firm population in the city. Van Wissen (2000) utilizes the concept of carrying capacity, which has been used before by organizational ecologists such as Hannan and Freeman (1989), to model the individual behavior of the firm. The individual behavior is determined by a market stress or pressure factor $C = K(t) - X(t)$, which measures the difference between market supply $X(t)$ and market capacity or demand $K(t)$ at time t.

A negative C suggests an increase in the production level whereas a positive C suggests a decrease in the production level. Since $X(t)$ relates to the size of the population of firms, measuring $K(t)$ will determine a statistic like C, which can be used as a covariate of the functions determining birth, growth and death of firms. Market demand $K_{is}(t)$ of economic sector s at location i is modeled in a multi-sectoral and spatial input-output framework[1]; that takes into account the inter-industry relations between economic sectors and the spatial separation between suppliers and producers. The outcome market demand is then compared to the level of production $X_{is}(t)$ so that the relative difference between the two is used as a predictor of the economic demographic behavior at the micro-level.

[1] See Van Wissen (1997) for a detailed description of input-output framework.

Van Wissen (2000) utilizes the devised market stress covariates along with firm, industry and location specific covariates to model the birth, death, growth/decline and migration of firms. This is done through the use of binomial logit formulation to model the birth and death of firms by economic sector, an exponential growth model to examine the growth and decline of firms and a multinomial logit models to simulate the migration of firms.

The work of Khan et al. (2002) focused on the firmographic behavior of business establishments in urban areas. The modeling framework they present to microsimulate the evolution of firms is influenced by the ideas originally used to develop the firm location component of the *MEPLAN* and *PECAS* land use models[2]. Accordingly, business establishments are described primarily by their business transactions between the different locations (markets) in the city. The city is divided into a number of discrete zones that constitute the locations (markets) at which the transactions take place. In a multisectoral market, the probability of choosing one particular market for one particular transaction is modeled using a nested logit model. This modeling approach gives rise to the spatial input-output model, in which, the probability of having a transaction from zone i and to market m in zone j is modeled in a two level nested logit model. Therefore, choosing a certain market m will depend on choosing a destination zone for the transaction to take place. The utility of each market m in a given zone is a function of three variables. These are (1) the size of market m (S_m), (2) the cost of transporting the commodity to market m from its original location z (T_{mz}) and (3) the price of the

[2] See Abraham and Hunt (2000) and Hunt and Abraham (2003) for a description of the MEPLAN and PECAS models, respectively.

commodity at market m (P_m). Therefore, the utility of choosing a certain market m within zone z (call it Zutility) will depend on the expected maximum value of the utility of all the possible markets.

The model of Khan et al. (2002) starts with a synthetic population of discrete business establishments that is generated for the base year. Business establishment mobility is modeled as the outcome of three decisions that include: to stay at the existing location, to move to a new location within the city or to leave the city. If the move decision is made then the choice of a destination (new zone) will be based on maximizing the utility of all possible locations. Therefore, the mobility is modeled using a nested logit model. The logsum utility of choosing a certain zone is a function of the vacancy rate, floor space price and the Zutilities described above. As can be seen, firm mobility is modeled with a three-level nested logit model. Firmography in the model is simulated on a year-by-year basis but without resorting to a full equilibrium with market clearing. The dynamic disequilibrium is maintained in the model by allowing floorspace rents to be updated more frequently within one year. This enables the system to remain in disequilibrium in reaction the supply-demand mismatches.

Waddell et al. (2003) developed the *UrbanSim* land use model to microsimulate the change in location choices of households and firms as well as the change in the real estate markets. The model is regarded as one of the first and few operational microsimulation land use models. Three main innovative features differentiate *UrbanSim* from existing operational land use models. First, as in the case of Khan et al. (2003), *UrbanSim* operates as a dynamic disequilibrium in each year. Waddell et al. (2003) argue that the assumptions of the equilibrium conditions are over-simplifications when considering the complexities of the urban system. Examples of over-simplified equilibrium assumptions include: actions of any individual cannot affect prices, the products of all firms in the market are homogenous, resources are perfectly mobile, present and future prices and costs are perfectly known to all market participants and the agendas of all buyers and sellers in all markets are being coordinated simultaneously. In their defence, there are at least three different time scales that violate the assumptions of equilibrium when addressing the interactions between the land use and transportation systems. These are: (1) short-term changes that happen within the scope of a single day such as the change in travel behavior, (2) mid-term changes that occur over the span of months such as the change in the location choices of households and firms and (3) long-term changes that take place over a longer time frame of several years such as the change in the real estate development process. Therefore, the disequilibrium approach adopted in *UrbanSim* attempts to dynamically evolve the urban system over a series of finer time steps to account for the different time scales that drives the evolution of the urban system. A second feature of *UrbanSim* is its highly disaggregated nature relative to most other existing operational models. The model microsimulates the annual evolution in locations of individual households and jobs, and the development practices of real estate developers using grid cells of 150 × 150 meters in size. The development environment of the model adopts a modular object-oriented programming approach, which sets another remarkable feature of the model. The developers of *UrbanSim* adopts the Java programming language and develop the system as an Open Source project using the GNU General Public License. They argue that such an approach will encourage synergy among researchers and will speed the evolution and robustness of the simulation system.

On the firm side, *UrbanSim* handles the change in the job population rather than change in the business establishment population. Waddell et al. (2003) argue that the main advantage of this approach is that it affords greater capacity for modeling the location of jobs in large businesses. However, we believe that this is an unrealistic assumption since the firm and not the job is what drives the location choice behavior. The modelling process starts with an economic transition model, which determines the jobs that will be lost and those that will be created per economic sector. Therefore, this transition model is a database manager that determines job losses and removes them from the database for the sectors that are declining. It also place newly formed jobs for sectors that experience growth in a queue that will be processed by the employment location choice model. Once the transition rules are set, an employment mobility model is used to predict the mobility likelihood of jobs from their current location during a particular year. This is done using a Monte Carlo simulation, which calculates the probability of a job moving from its current location. The Monte Carlo sampling procedure is achieved using the annual mobility rates for the different sectors of the economy. In the case of move or stay, the procedure evaluates whether the generated random number from the Monte

Carlo simulation is greater than the observed mobility probabilities (rates) in which case the move outcome is chosen.

Upon determining the mobility status of jobs, an employment location choice model is then used to distribute jobs that moved within the region along with the newly formed jobs and assigns them to the different grid cells that comprise the study area. The location choice model is based on a multinomial logit model that draws heavily on the urban economics literature in its specification. Therefore, the estimated probability that the current job will move to each of the alternative job spaces use factors like accessibility to population, agglomeration economies measures, age of the building, proximity to amenities (malls, green space and parking areas), proximity to highways and CBD. Once the location choice probabilities are calculated, Monte Carlo simulation is used to generate a final decision of assigning the job to a particular grid cell. The alternative grid cells for the location choice of jobs are based on the available vacant floor space, which is determined using a real estate development model.

2.3 Methodology

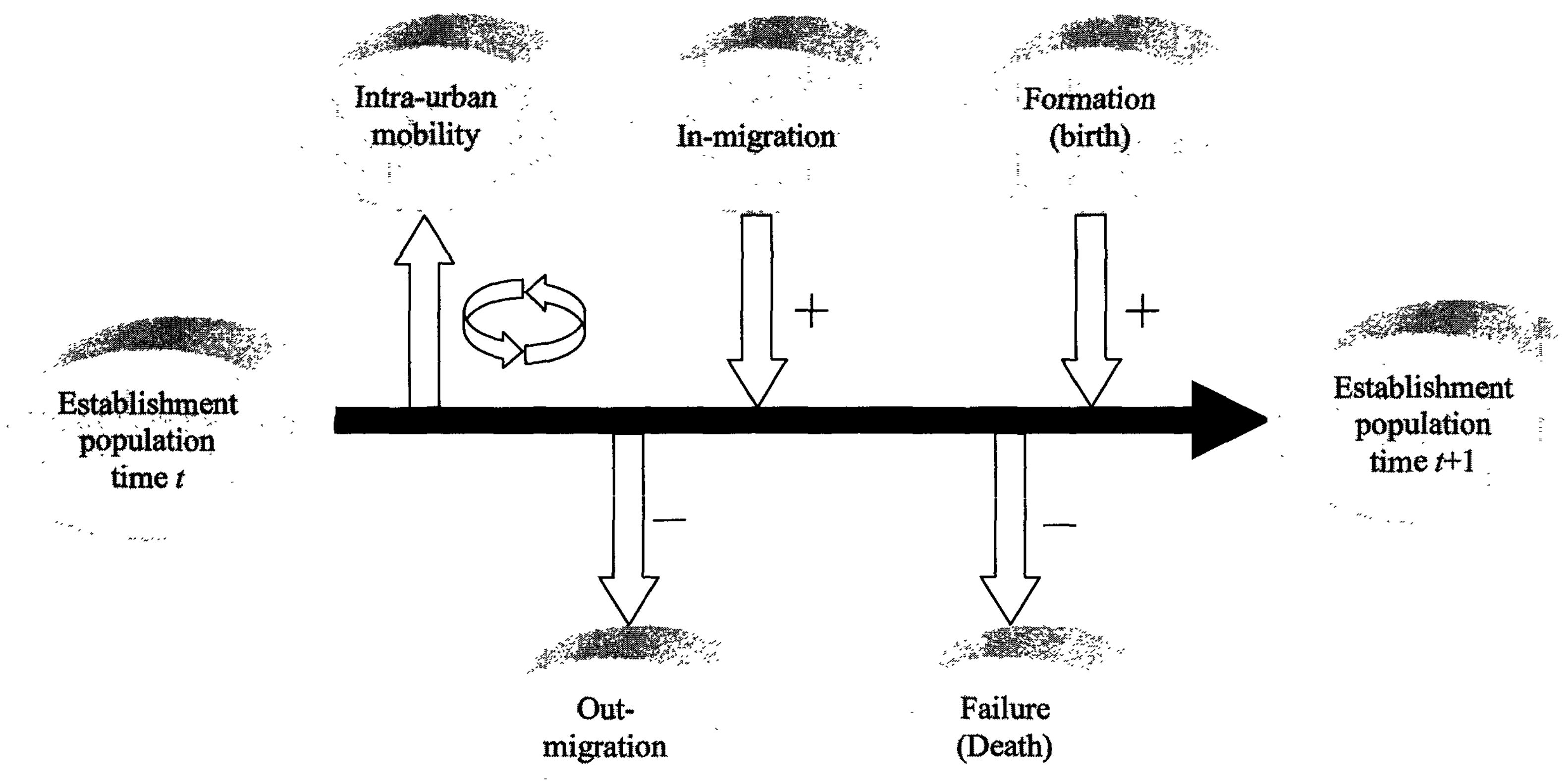
2.3.1 Modeling Framework

The modeling framework we are proposing will be used to simulate the evolution of individual firms over space and time. The firm in our context refers to a single location business establishment. Dijk and Pellenbarg (1997) note that the research entity in firmographic studies is not a well-clarified term. They argue that the term "firm" in firmographic research can stand for a single establishment or a company with multi branch plants. It can also refer to the entrepreneur or in the case of large firms can resemble the business unit or marketing division. This, as they note, will depend on the aim of the conducted research. Given that the aim of our research is to use firmography to study the urban development process, it is critical to focus on the firmographic processes at the business establishment level. Therefore, the business establishment, which occupies a certain amount of floorspace at a fixed location in space and is engaged in the production of goods and services, is considered as the unit of analysis.

The proposed model will focus on the firmographic behavior of small and medium (SME) size establishments with an application to the city of Hamilton in Ontario, Canada. SME establishments have a dynamic firmographic behavior compared to larger establishments. According to the literature, the mobility, entry and exit of large establishments is not commonplace given the amount of invested capital and the size of the labor population that are associated with these establishments. Van Wissen (1997) notes that capital and labor constitute what is called "exit barriers". Therefore, the developed model will handle the firmography of large establishments exogenously. The evolutionary process of the SME establishment population in the proposed framework is driven by a number of firmographic processes. These processes will affect the way business establishments are distributed over space. Intuitively, we assume that the change in the establishment population over time is driven by the formation of new establishments as well as the failure and migration (in-migration, out-migration and intramigration) of incumbent establishments. We also assume that the change in the job population (i.e.: size of each establishment) between time t to $t+1$ is based on growth/decline process. Mergers and spin-offs of establishments are not explicitly specified in our paradigm since the available data to develop the system do not include explicit information on these firmographic processes. Figure 2-1 highlights the evolutionary process of business establishment population over time

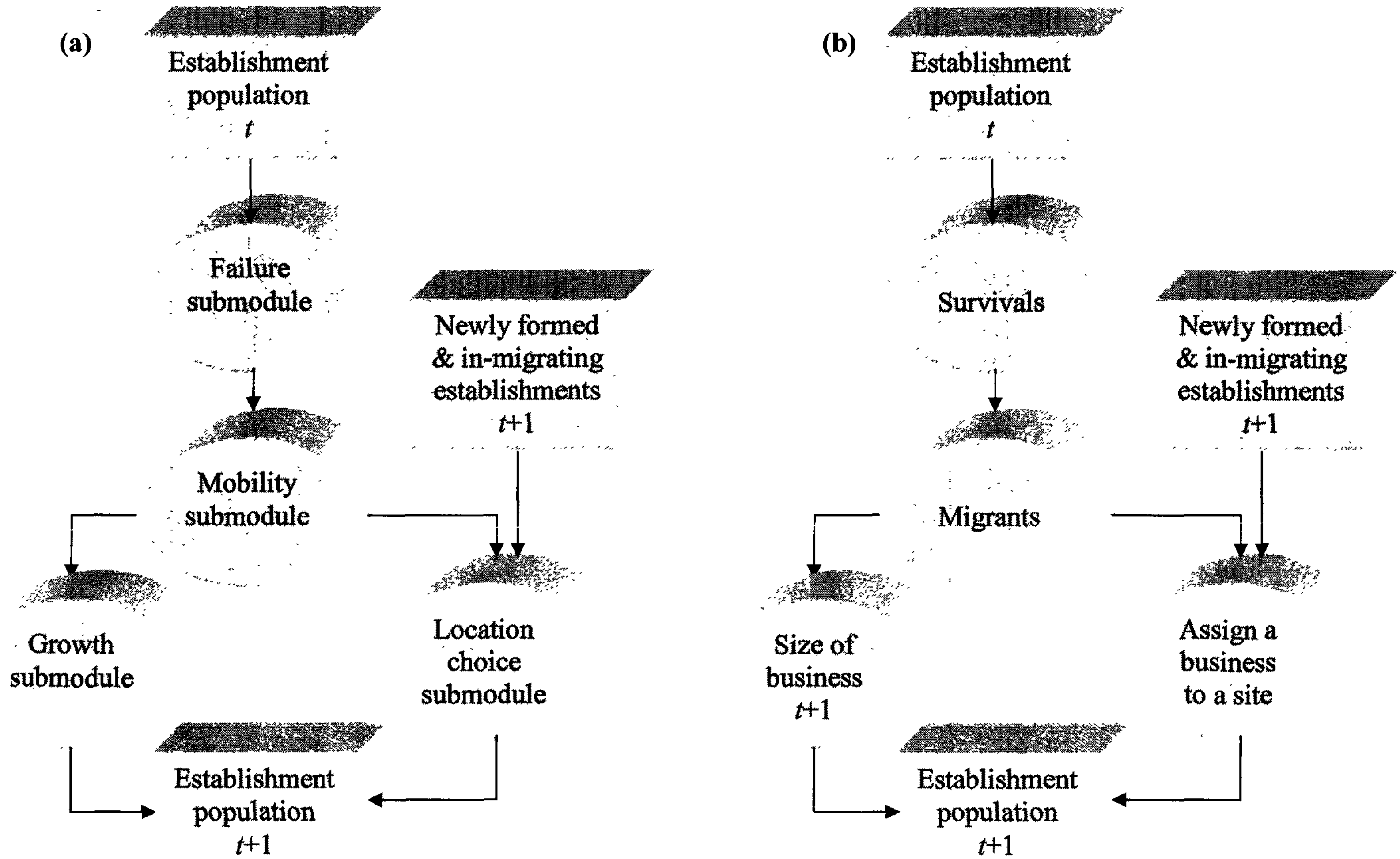
A market for industrial/commercial floor space at the parcel level drives the framework we propose. Firms are considered consumers of space. Demand for floor space is generated by the demographic events that operate on the firm population over the study or simulation period. The establishment of a new business (birth) or the relocation of an incumbent establishment from outside into the study area (in-migration) creates demand for floor space. The relocation of an establishment from one location within the city to another (intra-migration) frees up space at one place but creates demand for space in another. Also, the failure (death) or the relocation of an establishment outside the study area (out-migration) frees up floor space. Growth and decline as well as the merging or splitting of establishments may also contribute to changes in demand for space. The total available floor space in the city constitutes the supply side, which, in addition to firmographic processes, is affected by new development or redevelopment that leads to alteration in land uses. The latter includes converting industrial/commercial floor space into residential and vice-versa.

Figure 2-1: The evolutionary process of business establishment population over time



Simulations will be carried out to evolve the establishment population in one-year increments. During a given simulation period t to $t+1$, establishments are subjected to potential demographic events with the help of estimated econometric models. The parameters of these models identify and quantify the causes that govern the failure, mobility (intra-migration and out-migration), location choice and growth/decline of business establishments. A number of transition probabilities will be calculated and used along with Monte Carlo simulation to determine the future state of the business establishment at time $t+1$. The simulation starts by executing a failure submodule, which splits the establishment population into survivors and failures. This is achieved by calculating the probability of failing (exiting from the market) between t and $t+1$ for each individual business establishment. Following, a mobility submodule is executed to determine the mobility status of the surviving establishments. Three complementary probabilities for each establishment are calculated to determine if the establishment will stay at the current location, intra-migrate or out-migrate from the city. Intra-urban migrating establishments are placed in a searching queue along with an exogenous list of newly formed and in-migrating establishments. Finally, a location choice submodule is executed to determine the location of establishments in the search queue. At the end of the time period $t+1$, the business establishment population is updated. On the job side, a growth/decline model is executed to update the change in employment size of surviving establishments. Figure 2-2 presents the architecture of the devised model. The business establishment population will be modeled by the main economic

Figure 2-2: Model architecture - (a) Firmographic processes; (b) Process output



sectors to account for heterogeneity in the firm population.

2.3.2 Firm Micro-Data

2.3.2.1 *Statistics Canada Business Register (BR)*

The estimation of the econometric models that make the basis of the proposed firmographic system is based on individual business establishment data that were extracted from Statistics Canada Business Register (BR). The BR is a structured list of businesses engaged in the production of services and goods in Canada. It is a repository of the universe of Canadian business establishments developed to provide the framework for the production of coherent statistics for National Accounts. It is also used as a sampling frame for the different business surveys conducted by Statistics Canada. The BR retains information about all Canadian businesses at the business establishment level since 1990. Each business establishment is attributed a unique Establishment Number (EN), created at the time the establishment is registered in the BR. The micro records of the BR are confidential and can only be accessed at the headquarters of Statistics Canada to conduct statistical analysis. Therefore, the data extracted from the BR are used to obtain parameter estimates for the different econometric models we specified in our modelling framework.

We used the EN to create a longitudinal firmographic database of SME establishments in the city of Hamilton for the time period 1996 to 2002. The devised model will be used to microsimulate the firmographic events on a year-by-year basis.

Therefore, annual information is required to estimate the econometric models. However, such information is only available for self-owned non-incorporated (NIP) establishments. Accordingly, the generated database stores annual information about Hamilton's individual NIP business establishments. Incorporated (IP) businesses are not included in the database because information on them is not updated annually. Consequently, we only study the firmographic behaviour of self-owned single establishments in the city of Hamilton. Preliminary analysis of the extracted data reveals that the NIP population accounted for 86% of the SME establishments in 1996. The exploration also suggests that 79% of SME establishments had employment size less than or equal to 10, 93% of which are NIP establishments. Accordingly, the NIP population is deemed appropriate for our objectives since it accounts for the majority of the SME establishments. However, the absence of the IP establishments from the extracted sample will undermine our knowledge of what role organizational structure has on the firmographic behaviour of establishments. Unlike the IP establishments, NIP establishments are simple in their organizational structure and each establishment corresponds to a unique firm in the city.

The first step for creating the firmographic database was based on extracting a set of annual files each with the basic attributes that the BR retains. Typically, an annual dataset for a particular year in the BR includes all the establishments that existed between January 1st and December 31st of that year. Each establishment has a list of attributes that include: (1) establishment number (2) postal code address in year t (3) paid worker jobs in year t (4) total operating revenue in year t (5) 1980 Standard Industrial Classification

(SIC) in year t (6) 1986 SGC Census Division in year t (7) operating name in year t (8) street name and number in year t (9) city/municipality in year t. These files were used to measure the annual firmographic events by comparing the BR records for each of three consecutive years. The comparison between years $t-1$, t and $t+1$ was mainly based on the establishment number (EN) and Standard Geographical Classification (SGC) code. A number of firmographic events were measured and assigned to each establishment in the year t file. These included: continuer, birth, failure, in-migration and out-migration events. A Continuer event is coded if the establishment has the same EN and Hamilton SGC code in years $t-1$ and t. A Birth event is coded if the establishment is not listed under any SGC in year $t-1$ but existed in year t. A Failure event is coded if the establishment existed in year t in the Hamilton SGC but are no longer in any SGC in year $t+1$. An In-migration event is coded if the establishment has the Hamilton SGC code in year t and a different SGC code in year $t-1$. An Out-migration event is coded if the establishment has the Hamilton SGC code in year t and a different SGC in year $t+1$. Continuing establishments were reclassified as non-movers and intra-urban migrants. Such a distinction was possible by comparing the postal code address for each continuing establishment between time t and $t+1$. An intra-urban migration event was coded if the establishment's postal code and associated coordinate was different in years $t-1$ and t, otherwise, the establishment is considered a non-mover.

We linked the 1996 file to the files for the successive years from 1997 to 2002 upon measuring the annual firmographic events for each business establishment. The outcome is a database with annual profiles of firmographic events for each establishment. These profiles enable the modelling of the firmographic behaviour of individual business establishment between the years 1996 and 2002. We also applied a backward link to create a new variable that depicts the age of each establishment in the 1996 SME cohort. This was achieved by linking the 1996 file to the annual records of the BR for the successive years from 1990 to 1995. The generated age variable took on ordered discrete values of (1, 2, …, 6 and 7+) depending on the earliest year the establishment was observed in the annual records.

2.3.2.2 *Synthetic Establishment Population*

The generation of a synthetic population is required to obtain a base year list of establishments that can be used to conduct simulations. The objective is to generate individual establishment observations for 1996 to replace the confidential establishment list used in the estimation of the econometric models of the system. Therefore, the procedure works on creating attributes similar to those found in BR dataset for each synthesized establishment. Methods to generate a synthetic population from aggregate data can be found in Wilson and Pownall (1976), Beckman et al. (1996), Wegener and Spiekermann (1996), Miller (1996) and Voas and Williamson (2000). In all these methods, the synthetic population is driven from cross-tabulations that can be generated from publicly available census data. Some methods also utilize 5% representative samples of publicly available micro data to improve the quality of the synthesized information.

For illustration, we follow the example used in Miller (1996) to synthesize an individual establishment h with three attributes X_1, X_2 and X_3. The procedure starts by forming probability distributions from available aggregate cross-tabulations for the three attributes in question. Accordingly, we can define $P(X_1 = x_1)$ as the marginal probability of X_1, and $P(X_2 = x_2|x_1)$ and $P(X_3 = x_3|x_2)$ as joint probability distributions for X_1 and X_2 and for X_2 and X_3, respectively. A uniform random number u_{1h} on the range [0,1] is generated using the Monte Carlo simulation. This number is used to determine x_{1h} from the distribution $P(X_1=x_{1h})$. Once x_{1h} is determined, a uniform random number u_{2h} is generated to determine x_{2h} from the distribution $P(X_2 = x_{2h}|x_{1h})$. Finally, a uniform random number u_{3h} is generated and used with the predicted x_{2h} value to determine x_{3h} from the distribution $P(X_3 = x_{3h}|x_{2h})$.

While the above method has been the common practice in creating micro-data from aggregate zonal information, there have been concerns about its performance (Miller, 1996). The issue of the two-way ordering of variables in the conditional probability remains ambiguous when structuring these probabilities. Therefore, the procedure tends to ignore the potential for significant multi-way correlations among the variables to be synthesized. The work conducted by Beckman et al. (1996) has directly addressed this issue. Their procedure utilizes 5% samples of almost complete census records for collections of census tracts from Public Use Microdata Sample (PUMS) files along with the aggregate cross-tabulations. Adding up the records in a PUMS provides an estimate of the full multi-way distribution across all attributes for the collection of census tracts. The PUMS multi-way distribution provides important additional information to the synthesis process under the assumption that each census tract has the same correlation structure as its associated PUMS. The primary tool used to complete the multiway table for each zone is the Iterative Proportional Fitting (IPF) procedure, which is similar to the fratar method commonly used in transportation studies. A number of steps are followed to generate the synthetic individual observations. These are: (1) generating summary tables for a selected set of demographics for each zone or census tract; (2) constructing multi-way tables of the selected demographics from the PUMS for the whole study area; and (3) applying the IPF to estimate a multi-way table for each census tract such that the marginal totals in the constructed table match the marginal totals from step 1 and the correlation structure of the multi-way table from step 2 is maintained. Individual synthesized observations are then generated by selection of observations from the associated PUMS according to the proportions generated by the IPF.

Voas and Williamson (2000) criticize the IPF procedure and argue that it is not well suited to handling variables from different levels (eg. households and the individual) when estimating joint distributions. They note that the IPF becomes cumbersome when a large number of estimated cross-tabulations are required. Therefore, they propose a new procedure, which they call the combinatorial optimization. In this procedure, a combination of observations from a sample of microdata is selected to reproduce the characteristics of a chosen zone. The process starts from an initial set of observations that are chosen randomly from the sample of microdata where the effects of replacing one of the selected observations with a new observation from the microdata sample are considered. If the replacement improves the fit between the actual data and the selected sample, then the observations are swapped, otherwise they are not swapped and the procedure is iterated. Voas and Williamson (2000) use the normal Z score statistics to measure the goodness-of-fit of the generated microdata. A critical review of different statistical methods to evaluate the goodness-of-fit of synthesized microdata can also be found in Voas and Williamson (2001).

2.3.2.3 *Methods of Analysis*

The development of the proposed firmographic model depends on formulating, specifying and estimating econometric models for the four submodules shown in Figure 2. The failure submodel is based on a discrete-time hazard duration model that predicts the probability of failure and the likelihood of survival for each establishment. Hazard based duration models represent a class of analytical methods, which are appropriate for modeling data that have as their focus an end-of-duration occurrence such as a failure state, given that the duration has lasted to some specified time (Kiefer, 1988; Bhat, 2000). Mata and Portugal (1994) and Audretsch and Mahmood (1995) were pioneers in utilizing hazard duration models to study firm exit or failure. The development of our hazard duration model is based on analyzing the annual life trajectory of each establishment in the 1996 cohort till 2002 to record the year when the establishment shutdown. We use a discrete-time hazard duration model to formulate the failure probability as a function of factors that relate to: (1) establishment characteristics, (2) establishment industry (3) establishment geographic location and (4) macro-economic conditions in the city. Our analysis suggests that while all types of specified factors contribute to explaining the failure of business establishments in Hamilton, firm specific and macro economic conditions are the most influential. Further details about the model and the estimation results can be found in Maoh and Kanaroglou (2005b).

The mobility submodule is based on a multinomial logit model that predicts the probability of staying at the same location, intraurban-migrating or out-migrating from the city. Examples of utilizing the logit model to study the mobility behavior of firms can be found in Van Wissen (2000) and Brouwer et al. (2004). We model the mobility behavior based on the concept of maximum utility. We assume that each of the three alternatives "stay, intra-migrate or out-migrate" is associated with a certain utility and that the status of the establishment will change when it chooses the alternative with the highest utility. The utilities are formulated as functions of firm internal characteristics and location factors (Maoh and Kanaroglou, 2005c). The estimation results from logit models by economic sector suggest that the willingness to move can be explained by age, growth rate, size and the industry of the establishment. We also find that as stress from local competition increases, an establishment will tend to move a longer distance. On the other hand, agglomeration economies increase the propensity of inertia. The estimated models also suggest that mobility is more pronounced among the Central Business District (CBD) establishments.

The Location choice submodule is based on a multinomial logit model that predicts the probability of an establishment choosing a vacant location, from a list of available locations in the city, as a destination. The literature suggests that the firm location choice problem has been modeled in two different ways. In the first approach, the analyst tries to model the probability of a firm's occupying a given site. Here, the concept of bid rent theory is applied such that firms are assumed to out-bid each other for the site and the site will go to the highest bidder. Lee (1982) was a pioneer in applying this concept to study the firm location choice problem using a discrete choice model. The second approach is directly related to our location choice problem. In this case, the firm is assumed to face the decision of finding a site that will maximize its profit. Shukla and Waddell (1991) show that profit maximization can be achieved when the firm chooses the site with the maximum utility. Therefore, the problem can be modeled via a multinomial logit model. Later work by Martinez (1992) shows that the two approaches are equivalent and complement each other. To this end, the formulation of our location choice model is based on the assumptions of profit maximization and out-bidding theory. The location choice problem is modeled by specifying and estimating multinomial logit models by economic sector. Accordingly, the set of alternative developed locations are represented by grid cells of 200 × 200 meters. Following the urban economic literature, the utility associated with each location is formulated as a function of establishment attributes and location characteristics. Estimation results suggest that CBD, highway and mall proximities, population density, new residential development and agglomeration economies influence the location decision of business establishments in the city of Hamilton. The results also point to the existence of land use specialization and economic clustering in the city (Maoh and Kanaroglou, 2005a).

Work to estimate the models for the growth/decline submodule is in progress. The purpose of the models will be to determine the change in the size of surviving establishment during the simulations. The existing literature indicates that regression models are the common statistical method used to model firm growth and decline. Examples of modelling firm growth/decline can be found in Evans (1987); Hart and Oultman (1996) and Van Wissen (2000).

On the supply side of the model, we propose a real estate development model to predict the change in industrial and commercial floor space as an outcome of the development practices in the city. The heart of the process would be a nested decision structure with regards to any given parcel of land. With that, developers decide at the higher level of the model to operate on a given parcel of land or not. The lower level pertains to the decision of adding new development to a vacant parcel, intensifying the existing development or altering the land-use of a parcel. Discrete choice models of the logit family (MNL or nested) were used before to estimate the probability of a certain development event (Waddell et al. 2003; Hunt et al., 2003). Potential variables that make up the utility function of such models include price, vacancy and age of the structures as noted by Hunt et al. (2003), as well as site characteristics and spatial proximities, regional accessibilities and vacancy rates as noted by Waddell et al. (2003). Macro-economic indicators such as interest rates are also important. The model will provide a set of transition probabilities for the different development choices and Monte Carlo simulation could be used to generate a final development decision.

2.4 Model Development Issues

Early development of IUMs and other spatial decision support tools followed a procedural-programming approach. Programs ran from top to bottom and program designs revolved around procedures to perform tasks and manipulate data. However, the extensibility of these procedural programs became cumbersome and in some cases extremely difficult when different types of data were introduced to expand the model. Object-oriented programming approach tried to provide a remedy for the problems inherited in the procedural programming approach. The object-oriented approach focus on real world objects (agents) such as people, households, firms, etc. and tries to model the state and behaviour of these objects. It can be argued here that the object-oriented approach is a natural tool for the development and implementation of agent-based microsimulation models. Also, the adoption of the object-oriented approach to develop agent-based systems is more efficient since, unlike the procedural approach, it affords enhancing the extensibility and maintainability of the modeled system.

Spatial agent-based decision support tools are heavily dependent on the databases that are used and updated during the microsimulation. Therefore, managing and maintaining these databases during the microsimulation is very important. The "geodatabase data model" of the ESRI ArcGIS 9.x platform is an appropriate tool for handling and maintaining spatial databases (Zeiler, 1999). This object-oriented data model allows the user to make the features of the GIS datasets smart by endowing them with natural behaviours, and allows any sort of relationship to be defined among features. This geodatabase data model brings a physical data model closer to its logical data model. A powerful feature of the geodatabase data model is the ability to implement the majority of custom behaviors without writing any code. Here, most behaviors are implemented through domains, validation rules and other functions of the framework provided in ArcGIS.

Visual programming via the unified modelling language (UML) leverages the development of geodatabase data models that are suitable for microsimulation models. Using the CASE tools, one can create custom features that extend the data model of ArcGIS. The CASE tools consist of two major activities: code generation and schema generation. Code generation is used to create the behaviour while the schema generation is used to create schemas in geodatabases. Object oriented design tools can be used to create object models that represent the design of the applications' custom features. These tools use the UML to create the design in an interactive fashion. Based on these models the CASE tools will help creating Component Object Model (COM) classes that implement the behaviour of the custom features and database schemas in which custom feature properties are maintained.

The Arc-objects module of ArcGIS enables the customization of ArcGIS interface while maintaining many of the build-in spatial analysis tools that could be used in designing policy scenarios once the system is implemented. Special attention will be paid to the design of the scenario dialogue boxes to include a wide range of functionalities that can be used to devise and simulate scenarios. The purpose is to design a friendly Graphical User Interface (GUI) that can be used to test the interplay between the local economy and urban form in Hamilton. The choice of ArcGIS as the benchmark technology will enable the transferability of the proposed system to be used in planning, research and teaching.

2.5 Summary

This paper has presented an overview of a modeling framework proposed to develop an agent-based firmographic microsimulation model. While the framework we discuss has general applicability, we focus our attention to an application for the city of Hamilton, Canada. We use principles from firmography to identify and quantify the causes associated with firmographic processes such as the failure, mobility and location choice of small and medium size business establishments. Our framework is based on a set of integrated submodules that are used to evolve a base year business establishment population over time. Initially, the system starts with a failure submodule that determines the surviving establishments between t and t+1. Next, a mobility submodule based on information produced by the failure submodule, determines the establishments that will relocate within the city (intra-migrate) as well as those that will out-migrate. Intra-migrants are added to a searching pool of an exogenous list of in-migrating and newly formed establishments. The location of establishments in the pool is determined with a location choice submodule. Finally, a growth/decline submodule is used to update the change in establishment employment size in the time period t to $t+1$.

The implementation of the above submodules is based on a number of econometric models that were estimated using micro-data extracted from the Statistics Canada Business Register. The failure submodule is based on a discrete time hazard duration model that predicts the probability of business establishment failure at a certain point in time between 1996 and 2002. The mobility submodule is a multinomial logit model that predicts the probability of remaining at the same location, relocating within the city or out-migrating from the city between 1996 and 1997. The location choice submodule is a multinomial logit model that is based on the profit maximization concept. Here, an establishment is assumed to choose among the available locations in 1997, the one that will maximize its profit. The growth/decline submodule is a regression model that predicts the growth or decline of an establishment's employment size between 1996 and 1997.

In terms of spatial resolution, our first intention was to use the delineation of developed land parcels to present the location of the firm. However, the idea was dismissed since the geographic coordinates (Longitude/Latitude) for each establishment in the BR data are driven from the postal code addresses. These addresses have a one-to-many relation with the land parcels. It was, therefore, impossible to assign a firm to one land parcel. Consequently, we opted for dividing the developed land available to firms into grid cells of 200 × 200 meters. The layer of grid cells was then superimposed onto the event layer of establishments. As a result, an establishment is matched to the grid cell it fell into. The 200 x 200 meters resolution is deemed sufficient for an agent-based model as indicated by earlier studies (Wegener and Spiekermann, 1996; Waddell et al., 2003; Miller et al., 2004). Grid cells provide a stable presentation of space in a microsimulation model. They also avoid the aggregation bias problems inherited in the zone-based presentation of space when defined on a fine spatial scale. This spatial scale is a better abstraction of the real world processes since, in the limit, as the cell size goes to zero the raster cells becomes individual points in space and the spatial representation becomes fully continuous (Miller el al., 2004).

While the above econometric models have already been estimated, the work to implement the agent-based firmographic model is still a work in progress. A number of research initiatives are set forward to fulfill the overall objectives of this project. Short term initiatives involve: (1) generating a baseline synthetic population that can be used as the basis for any simulation as described in section 2.3.2.2, (2) implementing and developing the system as an ArcGIS application following an object-oriented approach, and (3) experimenting with the devised system by designing and simulating a number of future planning scenarios. Part of the experimentation will also involve comparing the results from the devised agent-base system to those produced by a conventional employment location model (Maoh et al., 2005). This will provide novel results about value added from using agent-based firmographic models. Mid-term research initiatives will involve the implementation of a real estate development model that will simulate the addition of new floor space either through the construction of new development, intensification (add more to the same) or conversion (redevelopment) of existing development. Finally, a long-term research initiative would be to test land use and transportation policy scenarios by integrating the devised system with a microsimulation urban land use and transportation model.

2.6 References

Abraham, J.E. and J.D. Hunt (2000). Firm location in the MEPLAN model of Sacramento *Transportation Research Record,* 1685, 187-198.

Audretsch, D.B. (1995). Innovation, growth and survival *International Journal of Industrial Organization,* 13, 441-457.

Audretsch, D.B. and M. Talat (1995). New firm survival: new results using a hazard function *Review of Economics and Statistics,* 77, 97-103.

Bade, F.J. (1983). Locational behavior and the mobility of firms in West Germany *Urban Studies,* 20, 279-297.

Baldwin, J., L. Bian, R. Dupuy and G. Gellatly (2000). *Failure rates for new canadian firms: new perspectives on entry and exit,* Statistics Canada, Ottawa.

Beckman, R.J., K.A. Baggerly and M.D. McKay (1996). Creating synthetic baseline populations *Transportation Research A,* 30 (6), 415-429.

Berglund, E. and K. Brannas (2001). Plants' entry and exit in Swedish municipalities *The Annals of Regional Science,* 35, 431-448.

Bhat, C.R. (2000). Duration Modeling. In: *Handbook of Transport Modelling*, Vol. 6 (Eds, Hensher, D. and Button, K.J.), Elsevier Science Ltd, pp. 91-111.

Birch, D. (1979). *The Job Generation Process,* MIT Program on Neighborhood and Regional Change, Cambridge, MASS.

Brouwer, A.E., I. Mariotti and J.N.v. Ommeren (2004). The firm relocation decision: an empirical investigation *The annals of regional science,* 38, 335-347.

Carroll, G. and M.T. Hannan (2000). *The demography of corporations and industries,* Princeton University Press, Princeton, NJ.

Caves, R. (1998). Industrial organization and new findings on the turnover and mobility of firms *Journal of Economic Literature*, 36 (4), 1947-1982

Dieleman, F., M. Dijst and B. Guillaume (2002). Urban form and travel behavior: micro-level household attributes and residential context *Urban Studies,* 39 (3), 507-527.

Dietmar, H., S. Konrad and M. Woywode (1998). Legal form, Growth and Exit of West Germany Firms-Empirical Results for Manufacturing, Construction, Trade and Service Industries *The Journal of Industrial Economics,* XLVI, 453-488.

Dijk, J.v. and P. Pellenbarg (1999). The demography of firms: progress and problems in empirical research. In: *Demography of Firms: Spatial Dynamics of Firm Behaviour*, Vol. 17 (Eds, Dijk, J.v. and Pellenbarg, P.H.), Netherlands Geographical Studies, KNAG, RUG, pp. 325-337.

Dijk, J.V. and P.H. Pellenbarg (2000). Firm relocation decisions in the Netherlands: an ordered logit approach *Papers in Regional Science,* 79, 191-219.

Dijk, J.v., P.H. Pellenbarg and P.J.M.v. Steen (1999) Determinants of firm migration in the Netherlands: An exercise in the demography of firms approach. In: *Demography of Firms: Spatial Dynamics of Firm Behaviour,* Vol. 5 (Eds, Dijk, J.v. and Pellenbarg, P.H.), Netherlands Geographical Studies, KNAG, RUG, pp. 87-121.

Dijk, J.v. and P.H. Pellenbarg (1999) The demography of firms: progress and problems in empirical research. In: *Demography of Firms: Spatial Dynamics of Firm Behaviour*, Vol. 17 (Eds, Dijk, J.v. and Pellenbarg, P.H.), Netherlands Geographical Studies, KNAG, RUG, pp. 325-337.

Echenique M and J. Owers (1994) Research into practice; the work of the Martin Center in urban and regional modeling, *Environment and planning B: Planning and design*, 21 (5), pp. 513-650

Evans, D.S. (1987). The relationship between firm growth, size, and age: estimates for 100 manufacturing industries *Journal of Industrial Economics,* 35 (4), 567-581.

Frenkel, A. (2001). Why high technology firms choose to locate in or near metropolitan areas *Urban Studies,* 38 (7), 1083--1102.

Garfoli, G. (1994). New firm formation and regional development: The Italian case *Regional studies,* 28, 381-393.

Geroski, P.A. (1995). What do we know about entry? *International Journal of Industrial Organization,* 13 (4), 421-440.

Hannan, M.T. and J. Freeman (1989). *Organizational Ecology,* Harvard University Press, Cambridge.

Hansen, E. (1987). Industrial location choice in Sao Paulo, Brazil: a nested logit model, *Regional Science and Urban Economics*, 17, 89-108.

Hart, M. and R. Scott (1993). Measuring the effectiveness of small firm policy: Some lessons from Northern Ireland *Regional Studies,* 28, 849-858.

Hart, P.E. and N. Oulton (1996). Growth and size of firms *Economic Journal,* 106 (3), 1242-1252.

Horowitz, A.J. (2004). Lowry-type land use models. In: *Handbook of transport geography and spatial systems* (Eds, Hensher, D., Button, K., Haynes, K. and Stopher, P.), Elsevier, pp. 167-182.

Hunt, J., and J. Abraham (2003). Design and application of the PECAS land use modeling system. *Eighth International Conference on Computers in Urban Planning and Urban Management "CUPUM",* Sendai, Japan.

Hunt, J., J. Khan and J. Abraham (2003). Microsimulating firm spatial behavior. *Eighth International Conference on Computers in Urban Planning and Urban Management "CUPUM",* Sendai, Japan.

Hunt J and D. Simmonds (1993) Theory and application of an integrated land use and transport modeling framework, *Environment and Planning B: Planning and Design*, 20, pp. 221-224

Johnson, P.S. and D.G. Cathcart (1979) New manufacturing firms and regional development: Some evidence from the Northern Region *Regional Studies,* 13, 269-280.

Kanaroglou, P.S. and D.M. Scott (2002) Integrated urban transportation and land-use models for policy analysis. In: *Governing Cities on the Move: Functional and Management Perspectives on Transformations of Urban Infrastructure in European Agglomerations.* (Ed, Dijst, M.), Ashgate, pp. 43-75.

Khan, A.S., J.E. Abraham and J.D. Hunt (2002). A system for microsimulating business establishments: analysis, design and results. *Access to Activities and Services in Urban Canada: Behavioural processes that condition equity and sustainability,* Laval University, Quebec City.

Kiefer, N.M. (1988). Economic duration data and hazard functions *Journal of Economic Literature,* XXVI, 646-679.

Lee, K. (1982). A model of intraurban employment location: an application to Bogota, Colombia *Journal of Urban Economics,* 12, 263-279.

Love, J.H. (1996). Entry and exit: A county-level analysis *Applied Economics,* 28, 441-451.

Lowry, I. (1964). *A model of metropolis*. Santa Monica, Calif., Rand Corp

Maoh, H. and P. Kanaroglou (2005a). *The Location of Business Establishments in the City of Hamilton, Canada: A Micro-Analytical Model Approach,* Working paper, Center for Spatial Analysis, School of Geography and Geology, McMaster University, Hamilton.

Maoh, H. and P. Kanaroglou (2005b). *A Discrete-Time Hazard Duration Model of Survival for Small and Medium Business Establishments with an Application to Hamilton, Ontario,* Working paper, Center for Spatial Analysis, School of Geography and Geology, McMaster University, Hamilton.

Maoh, H. and P. Kanaroglou (2005c). *Business establishment mobility behavior in urban areas: an application to the city of Hamilton in Ontario, Canada,* Working paper, Center for Spatial Analysis, School of Geography and Geology, McMaster University, Hamilton.

Maoh, H., P. Kanaroglou and R. Buliung (2005). *Modeling the location of firms within an integrated transport and land use model for Hamilton, Ontario,* Working Paper No. 6, Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University, Hamilton.

Martinez, F. (1992). The bid-choice land use model: an integrated economic framework *Environment and Planning A,* 24, 871-885.

Martínez. F. (1996). MUSSA: a land use model for Santiago City. *Transportation Research Record,* 1552: 126-134.

Mata, J. and P. Portugal (1994). Life duration of new firms *Journal of Industrial Economics,* 42 (3), 227-246.

Macgill S (1977) The Lowry model as an input-output model and its extension to incorporate full intersectoral relations, *Regional Studies*, 11, 337-354

Merz, J. (1991). Microsimulation - A survey of principles, developments and applications *International Journal of Forecasting*, 7, 77-104.

Miller, E., J. Hunt, J. Abraham and P. Salvini (2004). Microsimulating urban systems *Computers, environment and urban systems*, 28, 9-44.

Miller, E.J. (1996). Microsimulation and activity based forecasting *presented at the TMIP conference on activity-based travel forecasting*, 1-30.

Miller, E.J., D.S. Kriger and J.D. Hunt (1998). Integrated urban models for simulation of transit and landuse policies *TCRP Web Document 9*, Project H-12, 1-252.

Putman, S. (1983). *Integrated Urban Models: Analysis of Transportation and Land Use.* London: Pion.

Shukla, V. and P. Waddell (1991). Firm location and land use in discrete urban space: A study of the spatial structure of Dallas-Fort Worth *Regional Science and Urban Economics*, 21, 225-253.

Southworth, F. (1995). *A technical review of urban land use-transportation models as tools for evaluating vehicle travel reduction strategies,* Oak Ridge National Laboratory, Tennessee.

Voas, D. and P. Williamson (2000). An evaluation of the combinatorial optimisation approach to the creation of synthetic microdata *International Journal of Population Geography*, 6, 349-366.

Voas, D. and P. Williamson (2001). Evaluating goodness-of-fit measures for synthetic microdata *Geographical and Environmental Modeling*, 5 (2), 177 - 200.

Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke and G. Ullfarsson (2003). Microsimulation of urban development and location choices: design and implementation of UrbanSim *Networks and spatial economics*, 3, 43-67.

Waddell, P. and V. Shukla. (1993). Manufacturing location in a polycentric urban area: a study in the composition and attractivness of employment subcenters. *Urban Geography*, 14, 277-296.

Wegener, M. (2004). Overview of land use transport models. In: *Handbook of transport geography and spatial systems* (Eds, Hensher. D., Button, K., Haynes, K. and Stopher, P.), Elsevier, pp. 127-146.

Wegener, M. and K. Spiekermann (1996). The potential of microsimulation for urban models. In: *Microsimulation for urban and regional policy analysis* (Ed, Clarke, G.P.), Pion, London, pp. 149-163.

Wilson, A.G. and C.E. Pownall (1976). A new representation of the urban system for modelling and for the study of micro-level interdependence *Area,* 8, 246-254.

Van Wissen, L. (1997). Demography of the firm: modelling birth and death of firms using the concept of carrying capacity. In: *Population and Families in the Low Countries 1996/1997* (Eds, Brekel, H.v.d. and Deven, F.), NIDI/CBGS, Den Haag, Brussel, pp. 219-244.

Van Wissen, L. (2000). A micro-simulation model of firms: applications of concepts of the demography of the firm *Papers in Regional Science,* 79, 111-134.

Zeiler, M. (1999). *Modeling our world: the ESRI guide to geodatabase design,* Environmental Systems Research Institute (ESRI), Inc, Redlands

CHAPTER 3: A Discrete-Time Hazard Duration Model of Survival for Small and Medium Business Establishments with an Application to Hamilton, Ontario

3.1 Introduction

Most North American cities witnessed dramatic changes in their spatial structure over the past century. These changes were the outcome of car dependency and the various locational decisions made by individual households, commercial and industrial firms along with the development practices imposed by developers in the city. A household makes a decision on the type and location of a dwelling based on its demographic characteristics, the number of workers in it and the place of employment of its workers. Firms, on the other hand, are looking for advantageous profit maximizing locations for their business establishments, while developers will engage in residential or commercial development in locations where they perceive high demand for their output.

All of these location decisions and land development projects have their impact on urban form since all are influenced and driven by the level of accessibility to a particular site in the city. For instance, the addition of an expressway in a city's transport network can have significant effects on the location decision of firms or on the perceived monetary returns from the residential development of a parcel of land that is adjacent to the new expressway. Similarly, locational decisions, such as the development of a shopping mall at the outerskirts of a city, may necessitate changes in the transportation network, thereby affecting other locational decisions. Over the past three decades, scientists have focused on understanding the behaviour driving the location decision of households and firms as well as the urban development processes to depict the evolution trends of urban form and growth. A number of Integrated Urban Models (IUMs) have been developed for cities around the globe to address the issue.

IUMs are a family of computer simulation programs that allow decision makers to assess the impact of land use and transportation infrastructure projects on the evolution of urban areas. They are planning tools that are used to evaluate the long-term repercussions of proposed land-use or/and transport network changes. These include the differential estimation of land development, location decision of households and firms and the generated trips between the different parts of the city. These trips are then used to estimate congestion and associated tailpipe emissions over all the links of the urban transport network. The results achieved from simulating the urban system will enable decision makers to accept, rectify or abandon certain aspects of their infrastructure projects. A number of existing operational IUMs were developed since the early 1980s and are being used as Decision Support Systems (DSS) in the planning process of many cities around the globe. A review of these models can be found in Southworth (1995); Miller et al. (1998); Kanaroglou and Scott (2002); and Wegener (2004).

Research to develop and improve the performance of IUMs received more attention in recent years in response to the environmental concerns raised by the different governments on the green house gas emissions problem (Dieleman et al. 2002). The research community recognizes a number of drawbacks that makes the existing IUMs models ineffective for long-term policy making and planning. Major drawbacks include: (1) the aggregate nature of the existing models which considers the traffic zone or census tract as the unit of analysis; (2) the static nature and lack of behavioral realism as a result of dealing with time lags of 5 years or more to model land use changes and (3) the weak response to long-term planning policies due to their weak policy sensitivity (Miller et al. 2004).

The proposed solution to these drawbacks is to adopt a micro-analytical approach to handle urban processes via micro-simulation models. Within this framework, urban development is driven by the micro behavior of the various agents (persons, households, trips, business establishments, etc.) that interact in the city. Accordingly, a micro-analytical theory of urban development has to decompose the total processes of urban change into subprocesses and within these identify the main actors (agents) and their decision behavior (Wegener and Spiekermann 1996). The dynamic behavior within each process and for each individual agent is then explicitly simulated over both time and space to generate aggregate system behavior (Miller et al. 2004).

On the firm side, we believe that firm demography can be effective in informing the changes in distribution of business establishments over time and space. Firm demography or firmography is the study of processes that relate to the establishment of new businesses in a study area or to the growth, decline, shutdown, or migration of existing business establishments. This is an interdisciplinary research area (Van Wissen 2000), with most of the literature focusing on studies at the national and regional, as opposed to the urban level. It is however ideal for the development of agent-based or microsimulation IUMs because it adopts a bottom-up approach that starts with the individual establishment as the unit of analysis. Firmography can provide an accurate assessment of the dynamic change of the firm population. It can be used to analyze not only the firm population size but also the structure and evolution of firms over space and time (Dijk and Pellenbarg 2000).

A review of the literature suggests that attempts to develop firmographic microsimulation models are rare and still emerging at both the regional (Van Wissen 2000) and urban (Waddell et al. 2003; Hunt et al. 2003) levels. Consequently, our objective is to develop an agent-based firmographic DSS for the City of Hamilton in Ontario, Canada. This paper reports on the results we achieved from specifying and estimating the firm survival submodel. Our goal here is to fill the gap in the literature on intra-urban business failures in general and for the Canadian context in particular[1]. We make use of a longitudinal business establishment database extracted from Statistics Canada Business Register (BR) to determine the duration of survival for the 1996 small and medium (SME) size establishment population. SME establishments are more dynamic in their firmographic behavior as compared to large establishments. The later have a more predictable firmographic behavior in the urban context and their population

[1] A survey of the literature suggest a scarcity in studies that investigated the closure and survival of business establishments in urban areas (Lerman and Liu 1984; Miron 2002)

is relatively small. These could be handled exogenously in an agent-based firmographic model.

We follow the annual life trajectory of each establishment in the 1996 cohort till 2002 and record the year when the establishment shutdown. We use a discrete-time hazard duration model to formulate the failure probability as a function of a firm's characteristics, the industry it belongs to, its geographic location and the macro-economic conditions in the city. An innovative feature of our study is the application of hazard duration modeling techniques[2] to explain and eventually simulate the intra-urban failure of establishments.

The remainder of the paper is organized as follows. Section two reviews some of the recent mainstream research on firm survival and failure to highlight the factors and determinants that drive this firmographic process. Section three describes the Statistics Canada Business Register (BR) to highlight the properties of the data employed in our empirical analysis. It also presents the results we achieved from the non-parametric survival analysis. Section four discusses the formulation of the discrete-time hazard duration model, specified to handle the failure process of business establishments in Hamilton. Section five presents the estimated models. The final section provides a conclusion and direction for future research.

---

[2] While there has been a handful of studies that utilized hazard duration modeling to investigate the survival of firms, all of them were conducted at the national level and were concerned with new manufacturing firms as oppose to the whole establishment population (for example Baldwin and Gellatly 2004; Baldwin et al. 2000; Audretsch and Talat 1995; Mata et al. 1995; Mata and Portugal 1994)

3.2 Firm Survival and Failure: an Overview of Factors and Determinants

The future of contemporary cities depends on their economic development. This in turn is dependent on the existence of the various types of firms and their demographic behavior in the city. For instance, the decline, closure or emigration of existing firms from the city could lead its economy to decline, leaving a large portion of the labor force unemployed. On the other hand, the growth of existing firms, the establishment of new firms or the arrival of migrant firms into the city could lead to economic growth that would create new job opportunities. We thus contend that firm demography is central to the study of dynamic change in urban areas.

Firm demography like the demography of people studies the basic demographic processes of firm death, birth and migration (Van Wissen 2003). Dijk and Pellenbarg (1997) note that the research entity in firmographic studies is not a well-clarified term. They argue that the research entity depends on the aim of the conducted research. For instance, the term "firm" in firmographic research can stand for a single establishment or a company with multi branch plants. It can also refer to the entrepreneur or in the case of large firms can resemble the business unit or market division. In the urban context, we believe firmography should be concerned with the spatial distribution of individual business establishments since these establishments have a direct influence on urban form. Therefore, firmography can be very effective in assessing the interplay between the dynamics of a city's economy and the evolution of urban form when analyzing the firmographic processes that relates to individual business establishment.

The failure of business establishments is one of the basic firmographic processes. Failure of a business can be an involuntary liquidation or a voluntary exit from the city's market. The shutdown of one or more establishments from a certain location has a number of long run ramifications that may influence the spatial distribution of establishments in the city. Most geographic studies have focused on explaining change in the space economy by analyzing the intra-urban location behavior of business establishments. In contrast, very little has been done in analyzing firm survival and failure in the urban context. The work of Mirron (2002) and Lerman and Liu (1984) to model the failure of retail trade firms in the City of Toronto and City of Boston, respectively, fill some of the gap in the literature.

Most of the existing studies on firm survival and exit are conducted at the national level as opposed to the urban level (for example Baldwin and Gellatly 2004; Baldwin et al. 2000; Holm et al. 2000; Fotopoulos and Louri 2000; Caves 1998; Audretsch and Talat 1995; Mata et al. 1995; Mata and Portugal 1994). The objective in these studies is to examine how the exit process influences the regional or the national economy by determining when firms exit from the market. According to the reviewed literature, the variables used to explain firm exit could be summarized into the following generic factors: (1) firm specific, (2) industry specific, (3) macro-economic and (4) spatial or geographic factors. In the remainder of this section, we will draw on these groups of factors to provide an overview of the determinants influencing firm survival and failure.

Firm-specific factors reflect the individual characteristics of the business establishment. The literature usually highlights the following individual characteristics when studying survival and failure: establishment age, employment size, establishment growth, organizational structure and adaptability to new technologies. Studies have shown that young and small establishments are the most vulnerable in the firm population and their exit rates are the highest as compared to large establishments (Baldwin et al. 1997; Fotopoulos and Louri 2000; Holm et al. 2000; Audretsch and Talat 1995; Geroski 1995; Mata and Portugal 1994). Van Wissen (1997) notes that this relation is known as the "*liability of newness*" in the firmographic literature. Baldwin et al. (2000) show that 23% of new Canadian firms are likely to exit during the first year of entry. A new firm has a 0.42 probability of surviving past its 4^{th} year and a 0.20 likelihood of completing its first decade. They also show that small firms in their first years of entry have a higher hazard of exit when compared with medium and large size firms. Mata and Portugal (1994) achieved comparable results for studying the survival of new manufacturing firms in Portugal. Their results show that one fifth of these firms failed during the first year of entry and that only 50% survived for four years.

Organizational structure of the firm refers to its ownership structure. Studies show that multi establishment firms are less likely to exit as compared to single establishment firms (Holm et al. 2000; Audretsch and Talat 1995; Mata and Portugal 1994). Such a result is expected since multi establishment firms have access to large capital that can immunize them and lengthen their life span in the market. Change in technology can also affect the survival of business establishments. Mata and Portugal (1994) show that manufacturing firms that were not able to accommodate for the change in technology and did not have access to specialized labor were forced to shutdown and exit.

Other firm specific factors that are less highlighted in the literature but appear to have a significant influence on the survival of firms relate to the internal factors of the firm. Managerial experience and financial management are vital in determining the success or failure of the firm. Baldwin et al. (1997) conducted a survey to study the causes of firm bankruptcy in Canada. Their findings suggest that factors internal to the firm are instrumental in explaining Canadian firm bankruptcy. They found that almost half of the firms in Canada go bankrupt due to internal problems caused by managerial inexperience and knowledge as well as lack of vision to run the business. They also found that 71% of firms fail because of poor financial planning that occurred due to unbalanced capital structure, inability to manage working capital and undercapitalization. Their findings suggest that even firms that failed due to external events, such as economic downturn or increased competition, suffered from internal managerial deficiencies that did not allow them to react to the external shocks. These deficiencies included lack of vision, initiative, flexibility, adaptability and marketing competencies. While internal firm specific factors appear to be very important in explaining the failure process, most of the available datasets lack such exhaustive information on individual firms.

Industry specific factors refer to variables measured at the industry level to examine the role of the industry environment on firm survival. Studies have shown that high industry growth can lead to firm exit (Baldwin et al. 2000; Audrestch and Talat 1995; Mata and Portugal 1994). Baldwin et al. (2000) use the average firm size of an industry to measure the performance of new entrants. They argue that establishments usually enter at a suboptimal level and their likelihood to survive will decrease if they are not able to pull their size to match the average firm size of the industry. Many studies have also shown a variation among the survival rates of business establishments by industry (Agarwal and Audretsch 2001; Baldwin et al. 2000; Holmes et al. 2000; Harhoff et al. 1998; Audretsch et al. 1999; Mata et al. 1995; Dunne and Hughes 1994; Lerman and Liu 1983). The various studies use the Standard Industrial Classification (SIC) coding scheme to categorize firms and business establishments by type of industry. Other industry related factors that were used in previous studies include industry concentration as a surrogate measure of competition. A concentrated industry suggests low levels of competition since most of the output is generated by a relatively small number of firms. Baldwin et al. (2000) found that new firms in more concentrated industries are less likely to fail. They argue that new start-ups benefit from lower levels of competition and do not face much incumbent response.

Macro-economic conditions have also been addressed in the firm survival and exit literature (Baldwin et al. 2000; Holmes et al. 2000; Audretsch and Talat 1995). Studies have used a number of macro-economic indicators to account for the change in business cycle over time. The most common indicators are unemployment rate, interest rate, exchange rate and growth rate in real GDP. High unemployment rates due to economic downturn and high levels of interest rates or exchange rates are expected to increase the likelihood of exit from the market (Holmes et al. 2000; Audretsch and Talat, 1995). On the other hand, high growth in real GDP can prolong survival as suggested by Baldwin et al. (2000).

Spatial or geographic factors are variables that reflect the individual characteristics of an establishment or the industry it belongs to in a spatial setting. An example of an individual characteristic is the postal code address of the establishment, which identifies the geographic location of the business in the regional context or within a particular region. Some studies made use of such information to examine if survival varies by geographic location (Baldwin et al. 2000; Fotopoulos and Louri 2000). Baldwin et al. (2000) show that exit of the Canadian firms vary by province and this variation is dependent on the age of the firm. Their results suggest that Ontario hosts the most successful young firms when compared to the other provinces. Such patterns tend to dissolve for firms older than five years. Fotopoulos and Louri (2000) show that young Greek firms locating in Athens exhibit higher survival rates when compared to their counterparts elsewhere in Greece. The variation in firm exit over space suggests the need to control for geography when modeling survival and failure.

The levels of local competition and agglomeration economies are spatial measures at the industry level for a business establishment. These measures reflect the market power and market share effects, respectively. Pinkse and Slade (1998) use the Hotelling model of product differentiation to explain these local market effects. They argue that if an establishment is very similar to its competitors then this will allow it to steal some of its rivals' customers. Consequently, establishments will tend to cluster and agglomeration economies will lead to the market share effect. Studies have shown that agglomeration economies have negative effect on the exit rate of establishments (Berglund and Brannas 2001; Love 1996). On the other hand, when establishments become more distinct than their rivals, they enjoy a higher degree of market power since the demand for their products becomes less elastic, which leads to the market power effect. While the clustering of incumbent establishments can trigger local competition or increase the benefits of agglomeration economies, high levels of establishment entry could be an additional catalyst. The literature notes that high levels of firm entry are usually associated with high levels of firm exit among new and incumbent establishments (Baldwin et al. 2000; Geroski 1995).

Geographic factors are influential when studying establishment survival and failure in the urban context. An example of introducing spatial variables that reflect the geography of the city can be found in Lerman and Liu (1984). They study the closure of retail establishments in the city of Boston using a logistic regression model. Among other factors, they define several locational attributes to reflect the geography of the city. These include distance to the CBD and proximity to main routes in the city as measured by a set of categorical variables. The objective of these variables is to examine if establishment closure is sensitive to variation in location within the metropolitan areas. When addressing local competition, Lerman and Liu (1984) devised a ratio measure of the total number of employees in similar types of retail store in a given zone to the total number of employees at the store of interest. Mirron (2002), on the other hand, adopted a different approach to handle local competition in his survival model of retail establishments in Toronto. He defines local competition by counting the number of rival establishments within a 2 km search radius from a given establishment. We believe that the approach proposed by Lerman and Liu (1984) is more appropriate to capture the effect of local competition since it also accounts for the size of the rival firms in terms of employment.

3.3 Data and Non-Parametric Survival Analysis

3.3.1 Study Area

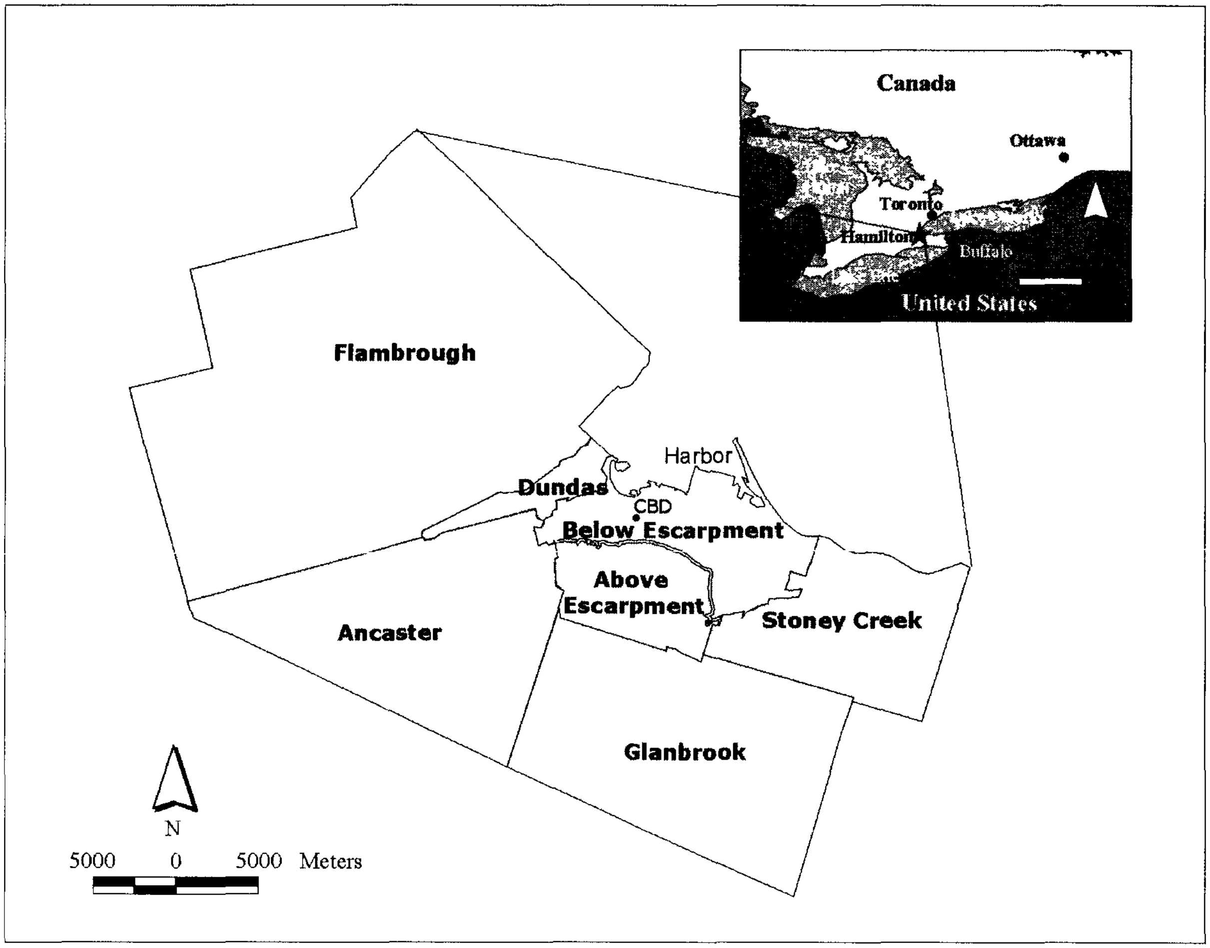
The focus of our study is the City of Hamilton, Ontario, located south west of Toronto. Topography is shaped by the Niagara escarpment, which separates the city's geography into a lower coastal area between the escarpment and the harbor to the north and a highland beyond the escarpment to the south. While early urban development took place in the lower city, the continuous population and economic growth over time led to urban development above the escarpment. Recently, the city expanded its boundary and annexed the municipalities of Ancaster, Dundas, Flamborough, Glanbrook and Stoney Creek as shown in Figure 3-1. The total population in Hamilton has increased from 465,153 in the year 1996 to 504,513 in the years 2001 (Statistics Canada, 1996 and 2001).

3.3.2 Hamilton's Firmographic Database

The empirical analysis is based on individual business establishment data extracted from Statistics Canada Business Register (BR). The BR is a structured list of businesses engaged in the production of services and goods in Canada. It is a repository of the universe of Canadian business establishments developed to provide the framework for the production of coherent statistics for National Accounts. It is also used as a sampling frame for the different business surveys conducted by Statistics Canada. The BR retains information about all Canadian businesses at the business establishment level. Each business establishment is attributed a unique Establishment Number (EN), created at the time the establishment is registered in the BR.

We used the EN to create a longitudinal firmographic database of small and medium size (SME) establishments in the city of Hamilton for the time period 1990 to 2002. SME are the target of our analysis since they account for the majority of business establishments in the city of Hamilton. Our analysis shows that more than 94% of the Hamilton's business establishment population was SME in the years 1990, 1996 and 2002. These accounted for 36%, 37% and 35% of the total jobs in the three years, respectively. While larger establishments contribute the majority of the jobs in the city, their firmography is more predictable from the micro-analytical point of view. Therefore, the firmographic behavior of those establishments will be handled exogenously in the devised agent-based model. Studies have shown that large establishments are less likely

Figure 3-1: The City of Hamilton in the regional context



to exit the market due to the high cost associated with their entry (Baldwin et al. 1997). In contrast, exit among the small and medium size establishments is commonplace since their cost of entry is relatively low (Lerman and Liu 1984).

The need for longitudinal information to conduct survival analysis constrained the nature of the extracted sample from the BR. The generated database stores annual information about Hamilton's individual non-incorporated (NIP) business establishments between 1990 and 2002. Incorporated (IP) businesses are not included in the database because the BR does not retain annual information on them. Consequently, we analyze the survival of the NIP business establishment population in the city of Hamilton. This is a reasonable proposition when analyzing the survival of SME business establishments. The figures in Table 3-1 show that the NIP population constitutes the majority of Hamilton's small and medium size establishments in 1990, 1996 and 2002. Further investigation suggests that 83%, 82% and 81% of the SME establishments in 1990, 1996 and 2002 were small establishments with employment size less than or equal to 10. Therefore, the absence of the IP population from the analyzed sample will not impair the overall objectives of the study. It will, however, undermine our understanding of what role organizational structure has on the survival of business establishments.

The first step in the creation of the Hamilton's firmographic database was to extract a set of annual files each with the basic attributes that the BR retains. Typically, an annual dataset for a particular year in the BR includes all the establishments that existed between January 1st and December 31st of that year. Each establishment has a list of attributes that include: (1) establishment number (2) postal code label and geographic coordinates (longitude/latitude) in year t (3) paid worker jobs in year t (4) total operating revenue in year t (5) 1980 Standard Industrial Classification (SIC) in year t (6) 1986 SGC Census Division in year t (7) operating name in year t (8) street name and number in year t (9) city/municipality in year t. These files were used to measure the annual firmographic events by comparing the BR records for each of three consecutive years. The comparison between years $t-1$, t and $t+1$ was mainly based on the establishment number and Standard Geographical Classification (SGC) code. The following firmographic events were measured and coded in the year t file: continuer, birth, failure, in-migration and out-migration. The following rules were applied when measuring a firmographic event:

1- *Continuer event*: if the establishment has the same EN and Hamilton SGC code in years $t-1$ and t.
2- *Birth event*: if the establishment is not listed under any SGC in year $t-1$ but existed in year t.
3- *Failure event*: if the establishment existed in year t in the Hamilton SGC but are no longer in any SGC in year $t+1$.
4- *In-migration event*: if the establishment has the Hamilton SGC code in year t and a different SGC code in year $t-1$.
5- *Out-migration event*: if the establishment has the Hamilton SGC code in year t and a different SGC in year $t+1$.

Table 3-1: Size distribution of IP and NIP SME business establishments in 1990, 1996 and 2002

Size Category	Size	1990 IP (%)	1990 NIP (%)	1990 Total (%)	1996 IP (%)	1996 NIP (%)	1996 Total (%)	2002 IP (%)	2002 NIP (%)	2002 Total (%)
Small	**1 – 10**	5	95	100	7	93	100	9	91	100
Medium	**11 – 20**	17	83	100	28	72	100	27	73	100
	21 – 50	34	66	100	50	50	100	43	57	100
Large	**51 - 199**	69	31	100	74	26	100	68	32	100
Total SME	**1 - 50**	8	92	100	12	88	100	14	86	100

Source: Special tabulation of the Statistics Canada Business Register, 2004

Continuing establishments were reclassified as non-movers and intra-urban migrants. Such a distinction was possible by comparing the postal code labels and geographic coordinates for each continuing establishment between time t and $t + 1$. An intra-urban migration event was coded if the establishment's postal code label and geographic coordinates were different in years $t - 1$ and t, otherwise, the establishment is considered a non-mover.

The second step in the development of the database was to use the annual files to create a firmographic profile for each business establishment in the 1991 and 1996 populations. While both cohorts were used to explore business establishment survival via the non-parametric survival analysis, the 1996 cohort was also used to model the survival process. The creation of the firmographic profile for the two cohorts was based on the Establishment Number (EN). Accordingly, we linked the 1991 file to the files for the successive years from 1992 to 2002. The outcome is a database with annual profiles of firmographic events for each business establishment. These profiles were then used to depict the duration of survival for each business establishment between the years 1992 and 2002. The same method was applied to create the firmographic profiles for the 1996 population. We also applied a backward link to create a new variable that depicts the age of each establishment in the 1996 cohort. This was achieved by linking the 1996 file to the annual records of the BR for the successive years from 1990 to 1995. The generated age variable took on ordered discrete values of (1, 2, ..., 6 and 7+) depending on the earliest year the establishment was observed in the annual records.

3.3.3 Non-Parametric Survival Analysis

We explored the survival of Hamilton's business establishment population using the life table method (Allison 1995). The method provides estimates for the survival *S(t)* and hazard *h(t)* probability functions that are used to gauge the time and rate of establishment survival and failure for a particular cohort. The survival function measures the probability that an individual establishment from a particular cohort will survive until time *t*. According to Kiefer (1988), the survival or endurance probability *S(t)* is the counterpart of a distribution function *F(t)* used to specify the probability that an individual establishment from a particular cohort will have a lifetime less than or equal to *t*. The life table method formulates the survival probability as the following non-parametric function:

$$S(t) = \prod_{j=1}^{t}\left(1 - \frac{d_j}{n_j}\right)$$

where d_j is the number of establishments that failed at time j and n_j is the population of business establishments at risk of failure at time j. The hazard function *h(t)* provides another useful measure in survival analysis to emphasize the instantaneous risk of failure faced by an individual business establishment from a particular cohort at time *t* given that the risk of failure still exits at time *t*. Intuitively, the hazard rate at time *t* can be calculated by taking the ratio of the number of failed establishments at time *t* to the number of establishments that are at risk of failure at time *t*.

We used Hamilton's firmographic database to calculate the survival and hazard rates for the 1991 and 1996 cohorts. We considered the two cohorts to examine the consistency of survival and hazard rates over time. The survival and hazard rate estimates for the two cohorts are reported in Table 3-2. The figures indicate consistent trends in the survival and hazard of failure over time. As a result, we decided to consider the 1996 cohort to model the establishment failure process in the city of Hamilton. Doing so will enable us to utilize the age variable we measured for the 1996 establishments in the explanatory analysis. Such a variable is absent and cannot be measured for the 1991 cohort.

The figures in Table 3-2 suggest that establishments in the 1991 cohort have a 0.84 likelihood of surviving beyond the first year of observation, a 0.54 probability of surviving past the fifth year from the time of observation and a 0.40 likelihood of surviving a decade from 1991. The hazard rates on the other hand suggest that the likelihood of failure after the first year of observation for both cohorts is 0.17. The estimates also suggest that the likelihood of failure will decline as time passes. For instance, the likelihood of failure after the fifth year of observation is 0.10 and 0.07 for both cohorts, respectively. These rates continue to decline with time as can be seen from the 1991 hazard rates.

We further explore the survival of business establishment by size class, industrial sector, age groups and geography to examine if these factors influence the survival of establishments. Figure 3-2 presents the estimated survival curves for the 1991 cohort by employment size class. The results indicate that the survival rate for small and medium size business establishments is virtually the same over time.

Table 3-2: Survival and hazard rates, 1991 and 1996 SME cohorts

	Survival *S(t)*		Hazard *h(t)*	
Time t	**1991 cohort**	**1996 cohort**	**1991 cohort**	**1996 cohort**
[1-2)	0.84	0.85	0.17	0.17
[2-3)	0.73	0.74	0.14	0.14
[3-4)	0.65	0.67	0.11	0.10
[4-5)	0.59	0.62	0.09	0.07
[5-6)	0.54	0.58	0.09	0.07
[6-7)	0.49	0.54	0.10	0.07
[7-8)	0.45		0.09	
[8-9)	0.42		0.06	
[9-10)	0.40		0.05	
[10-11)	0.38		0.06	
[11-12)	0.36		0.05	

By comparison, large establishments have lower survival probabilities. For instance, while small establishment have 0.84, 0.54 and 0.40 likelihood of surviving after the first, fifth and tenth year of observation, large establishments have only a 0.77, 0.46 and 0.23 per cent probability of surviving. This counter intuitive result about the relation between size and survival for those large establishments is a result of right censoring. That is, NIP establishments have low survival rates because they are becoming incorporated (IP) and not because they actually failed. However, we were not able to control for this form of right censoring because the annual failure event is based on a comparison between the NIP establishment populations for two successive years. Therefore, the firmographic event for a NIP establishment in year t that became incorporated (IP) by $t+1$ will be coded as a "failure" event. This situation is more likely to be prevalent for establishments with size greater than 50. Another less appealing explanation for the low survival rates of large establishments could be that large self-owned NIP establishments are facing high levels of competition from their IP counterparts who normally have a bigger market share due to access to more capital. Table 3-1 indicates that the majority of large size establishments are part of the incorporated businesses (IP) in Hamilton. This suggests that large NIP establishments have more IP rivals compared to small NIP establishments, which make them more susceptible to failure. However, the validity of this hypothesis requires further investigation.

The estimated survival curves by industry in Figure 3-3 suggest that health and social service establishments have the highest survival rates among the different types of businesses. By comparison, finance insurance establishments have the lowest survival rates. A close look at the estimated curves indicates that other services and manufacturing establishments have virtually the same survival curves. A ranking comparison of the survival rates for the different types of industries to the overall aggregate survival suggests that other industries, manufacturing, other services, real estate, education and businesses are well above the average survival. On the other hand, finance insurance, construction, communication, retail trade, transportation, accommodation food and beverages, and wholesale trade have lower survival rates as compared to the average survival of all establishments in Hamilton.

Figure 3-2: Survival rates, by size class, 1991 cohort

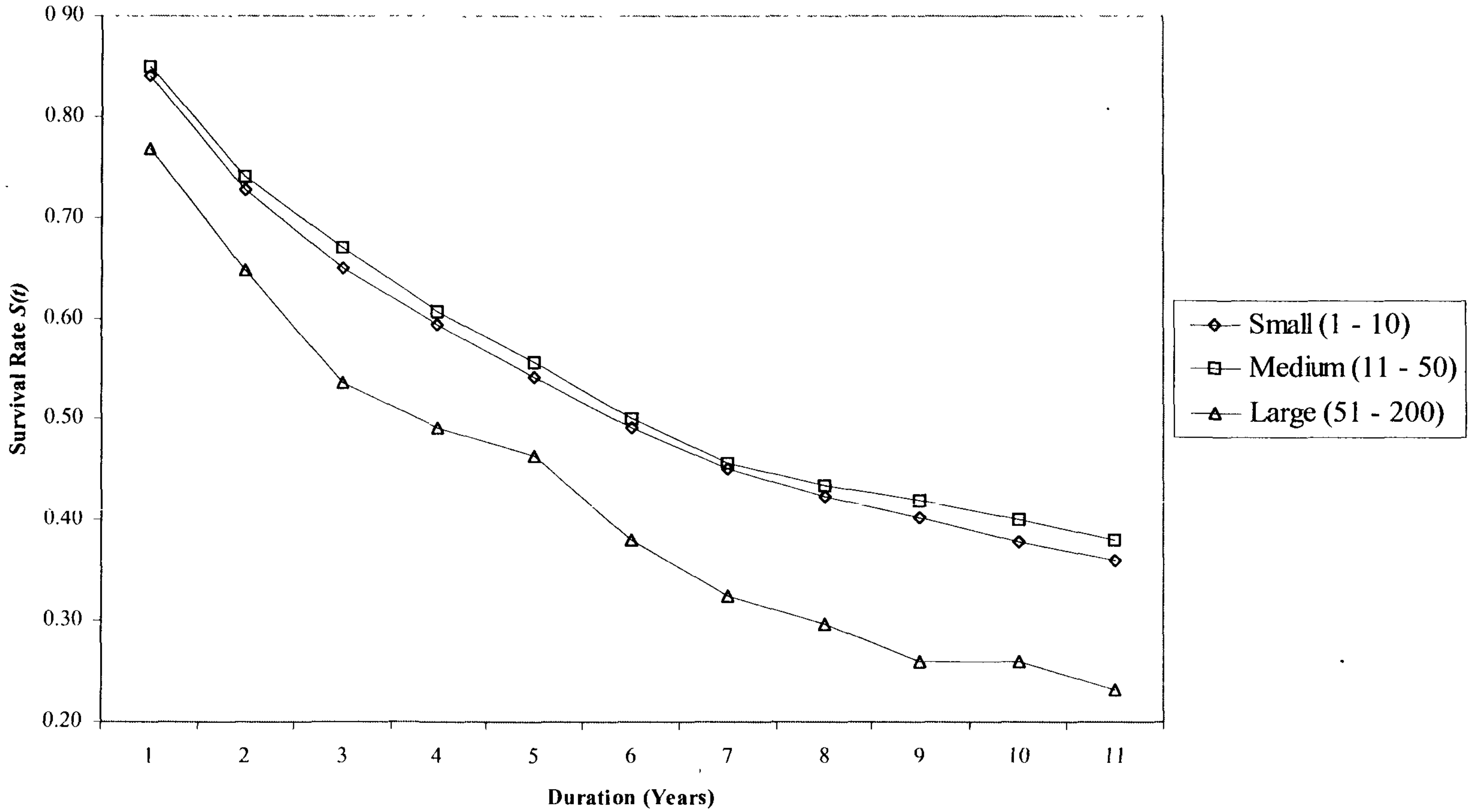


Figure 3-3: Survival rates, by industry, 1991 cohort

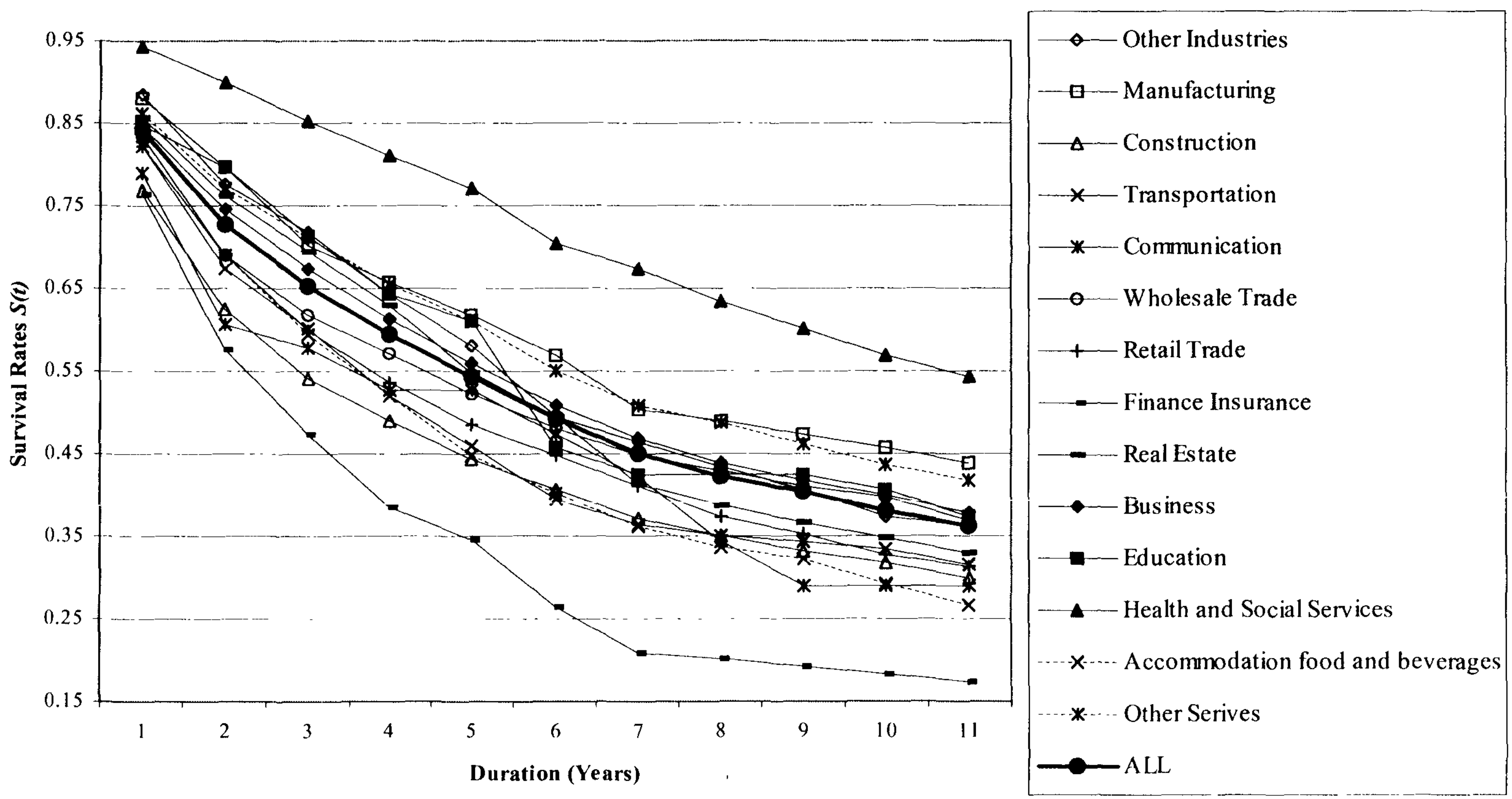


Figure 3-4 shows the survival curves for the 1996 cohort by age class. The resulting curves indicate a ladder effect between age and survival that is consistent over time. Accordingly, young establishments have a lower survival rate as compared to older establishments. For instance, a new establishment has a 0.78 likelihood of surviving its first year. This is compared to 0.90 if the establishment is more than six years old. These results coincide with the findings reported elsewhere in the literature about the relation between age and survival of SME business establishments.

The results from exploring survival by geography suggest that establishments located in Dundas, the CBD and Flamborough have higher survival rates compared to establishments elsewhere in the city. As shown in Figure 3-5, the survival curves for these three areas are above the aggregate survival curve achieved for all establishments in the city. While the CBD establishments had higher survival rates in the first year of observation, they rank second in their survival for later years. Dundas establishments exhibited the highest survival rates between the second and fifth year of observation while Flambrough establishments had the highest survival rates after the fifth year. On the other hand, the survival curves for Stoney Creek, land above and below the escarpment, Glanbrook and Ancaster are below the aggregate survival curve. An examination of these survival curves indicates that Stoney Creek establishments have the lowest survival rates during the first five years of observation, while establishments located above the escarpment tend to have lowest survival rates after the fifth year of observation.

Figure 3-4: Survival rates, by age groups, 1996 cohort

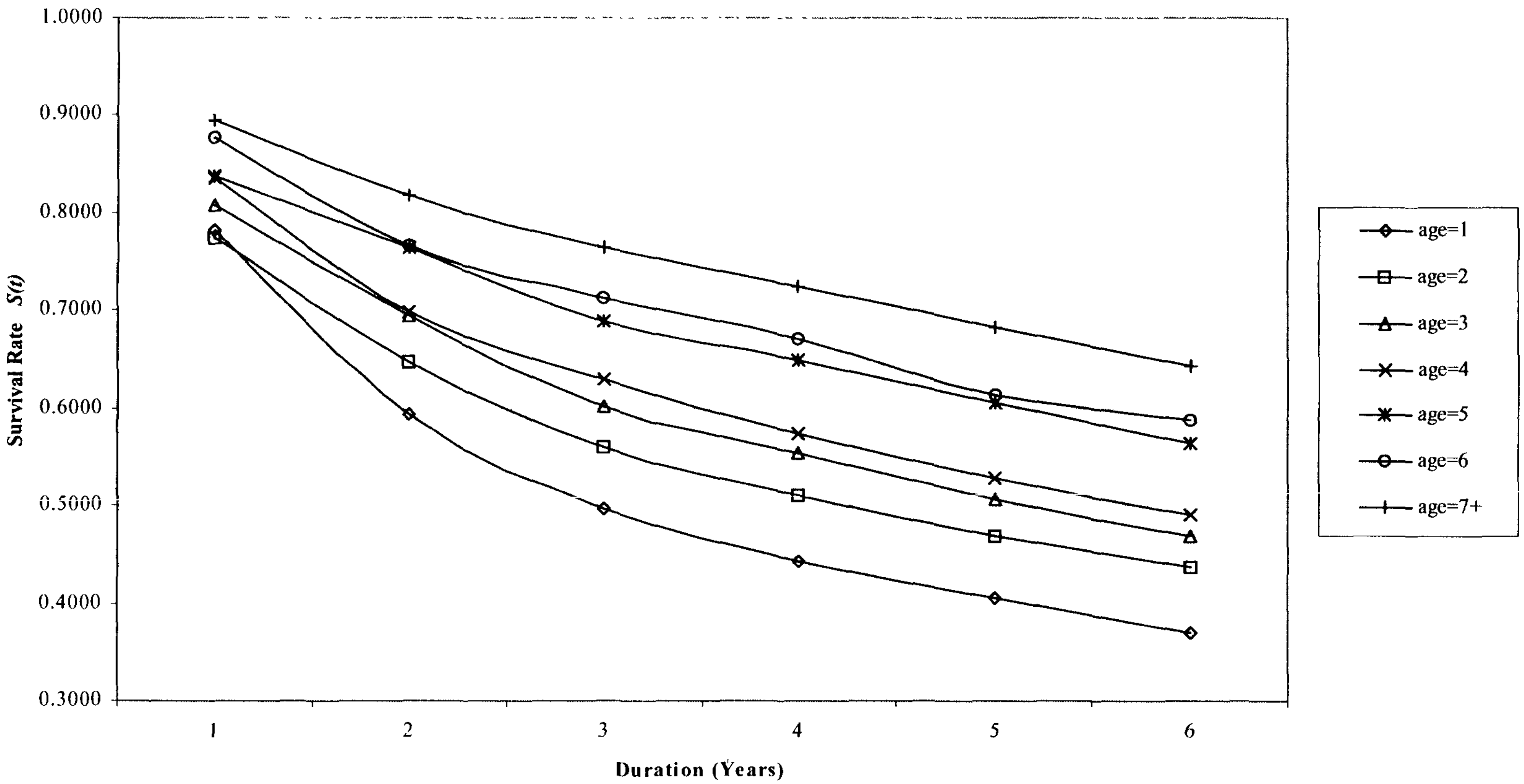
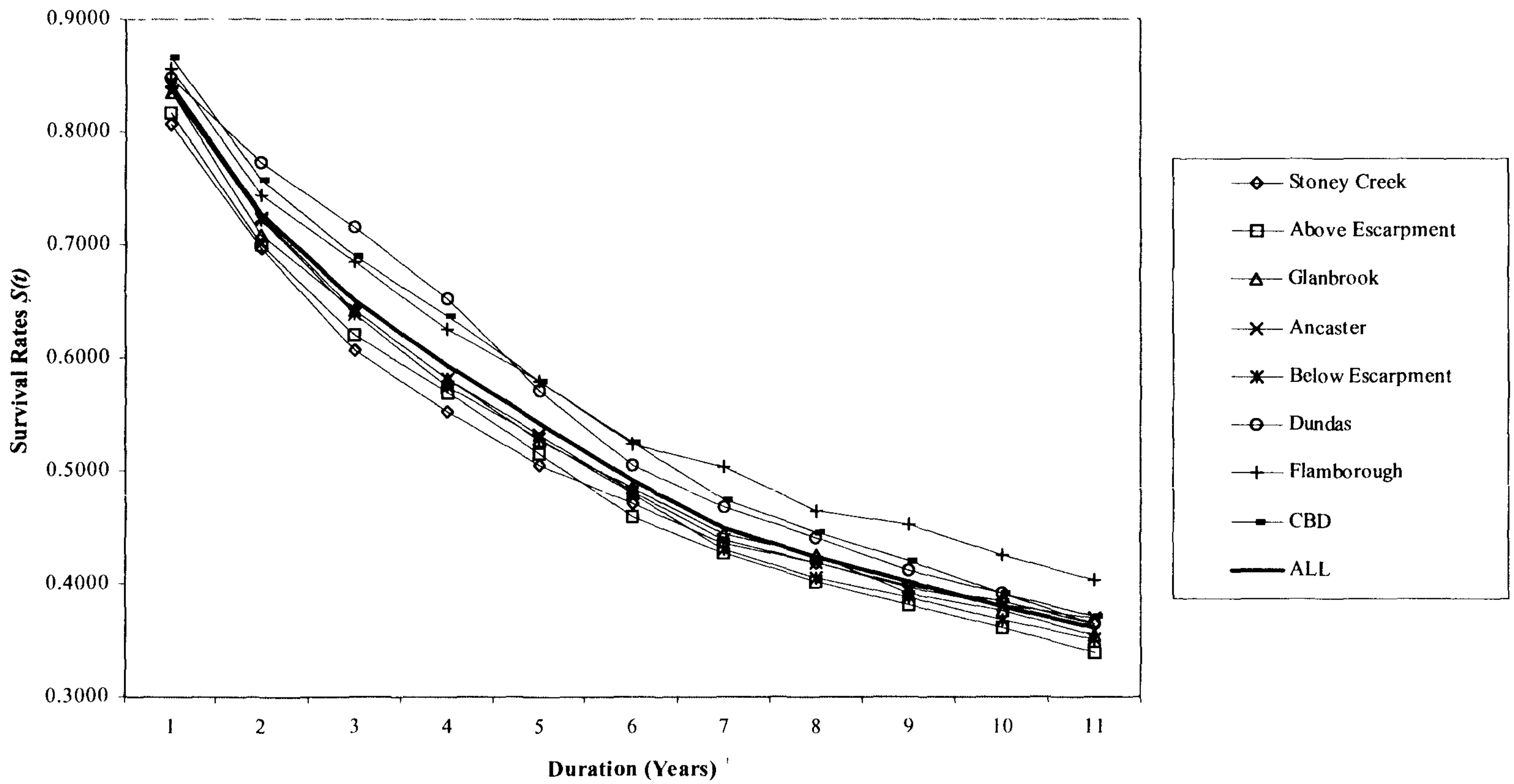


Figure 3-5: Survival rates, by geography, 1991 cohort



3.4 Survival Model

In this section, we discuss the formulation and specification of the discrete time hazard duration model that will be used to predict business failure in the City of Hamilton. The estimated model will form the basis for the survival component of the developed agent-based firmographic model. The remainder of this section first reports on the formulated model and then discusses the covariates used to specify the hazard probability function.

3.4.1 Model Formulation

We seek to estimate a model that will follow the evolution of a base year business establishment population to determine the factors that lead to failure and exit from the market. The base year population is a composite of n individual business establishments ($i = 1, 2, 3, \ldots, n$) and the establishment failure event is observed at time values that are positive and discrete ($t = 1, 2, 3, \ldots$). In our context, the base year is 1996 and the failure event is observed every year between 1996 and 2002. An observation for an individual establishment i will continue until time t_i, at which point either a failure occurs or the observation is right censored. Right censoring in our data occurs if the individual establishment is still in operation by the year 2002. It might also occur if the establishment change status from being non-incorporated to incorporated. Unfortunately, this right censoring cannot be observed or controlled and is measured as failure as explained in section 3.3.3. Based on the non-parametric analysis (Figure 3-2), this appears to have a significant effect on the larger establishments with size greater than 50 employees. Therefore, the inclusion of those observations in the hazard model might lead to misleading results. To avoid that, we decided to model the survival of small and medium (SME) establishments with size less than or equal to 50 employees.

The model will determine the hazard rate of a randomly selected business establishment i to shutdown at time t, given it was in operation between the base year and year ($T_i = t$). Following Allison (1982), the hazard rate can be expressed as the following conditional probability:

$$P_{it} = \Pr(T_i = t \mid T_i \geq t, \mathbf{x}_{it}) \tag{1}$$

The probability in equation (1) is dependent on a vector of observed explanatory variables $\mathbf{x}_{it}$ that may vary over time. In a statistical framework, each explanatory variable is associated with a parameter β that infers how the covariate contributes to the explanation of the failure events. Cox (1972) formulates the probability in equation (1) as a binomial logit model of the form:

$$P_{it} = \frac{1}{1+\exp[-(\alpha_t + \boldsymbol{\beta}'\mathbf{x}_{it})]} \tag{2}$$

where α_t is a set of constants. A consistent and efficient estimator for any of the β 's or α 's that is asymptotically normal can be achieved via the maximum likelihood estimation method. If we define δ_i as a categorical variable taking on a value of 1 if establishment i is uncensored and zero otherwise then the likelihood function for our base year population can be expressed as:

$$L = \prod_{i=1}^{n} [\Pr(T_i = t_i)]^{\delta_i} [\Pr(T_i > t_i)]^{1-\delta_i} \tag{3}$$

The estimation of equation (2) can be achieved by maximizing the log of the likelihood function in equation (3). Allison (1982) shows that the log of the likelihood function in equation (3) can be manipulated to result the log likelihood of a logistic regression model as shown in equation (4). This is done by applying elementary properties of conditional probabilities to express each of the probabilities in equation (3) as a function of the hazard rate P_{ij}. Following, if we define y_{it} as a categorical variable that is equal to 1 if establishment i fails at time t and 0 otherwise, then the log of equation (3) can be expressed as

$$\log L = \sum_{i=1}^{n} \sum_{j=1}^{t_i} y_{it} \log\left[\frac{P_{ij}}{1 - P_{ij}}\right] + \sum_{i=1}^{n} \sum_{j=1}^{t_i} \log[1 - P_{ij}] \tag{4}$$

Equation (4) implies that discrete-time hazard duration models can be estimated using the same computer routines used to estimate logistic regression models. This is achieved by treating each discrete time unit for each individual establishment as a separate observation. Thus, if an establishment failed at time k between 1996 and 2002, k different observations would be created. The dependent variable will be coded 1 for the time $t=k$ observation and zero for the other $k-1$ observations. This treatment is referred to in the literature as individual-year, which enables the estimation of equation (1) via maximizing expression (4). However, since the modeled population includes business establishments that existed prior to 1996, left truncation is present in the data. Guo (1993) notes that survival data are left truncated when the establishment has been exposed to the risk of failure for a while when it comes under observation. However, left-truncation in discrete-time hazard duration models can be accounted for following the treatment suggested by Guo (1993) and Allison (1995). Accordingly, if t_r represent the length of exposure before the observation period with $t_r \leq year\ 1996 < t_{r+1}$, where t_r and t_{r+1} are two adjacent distinct survival times, then equation (4) can be updated to

$$\log L = \sum_{i=1}^{n} \sum_{j=1}^{t_i} y_{it} \log\left[\frac{P_{ij}}{1-P_{ij}}\right] + \sum_{i=1}^{n} \sum_{j=r+1}^{t_i} \log\left[1-P_{ij}\right] \tag{5}$$

Equation (5) is still a logistic regression log-likelihood function and can be estimated using the same computer routines used to estimate equation (4). The only difference between the two is that in equation (5) the second summation in the second term sums from r+1, rather than 1 as in equation (4). The implication of this difference is reflected in the way the individual-year observations are handled for left truncated observations. If a left truncated establishment has a duration k that is equal to t_r+q, where t_r is the time of exposure to the risk before 1996 and q is the time of failure between 1996 and 2002, then we only create q individual-year observations as oppose to k individual-year observations. In other words, we drop all the observations corresponding to the t_r period and only use those corresponding to the q time period in the estimation of the model (Guo, 1993; Allison, 1995). As a result, a total of 40,667 individual-year observations were created from approximately 8,600 observations that make up the 1996 SME cohort. The explanatory variables are attributed to the *individual-year* observations based on the unit of time they belong to. The constant α_t could be restricted to a generic constant that is time invariant or as a linear, logarithmic or quadratic function of time.

3.4.2 *Model Specification*

The specification of the discrete-time hazard duration model is based on the information we gathered from the literature and exploration of the survival process in the previous sections. Using the BR data, we devised a number of explanatory variables in Table 3-3 to explain the business failure process in the city of Hamilton. We follow the same scheme highlighted in section 2 to classify our covariates into four types of generic factors that include: (1) firm specific, (2) geography specific, (3) macro-economic and (4) industry specific factors. In the firm specific factors, we use the $Age(Y)_t$ variables to test the effect of establishment age on the probability of failure. The expected sign of the age variable is positive for new and young establishments since they are more vulnerable compared to older establishments. The variable $Size96_i$ is used to examine the relation between employment size and the hazard of failure. The expected sign of this variable is negative since small establishments fail at higher rates as suggested by the literature. The $Size96_i$ squared is introduced to account for any non-linear effect that the size variable might exhibit.

Table 3-3: Description of covariates used in the failure hazard duration models

Covariate name	Covariate definition	Expected sign
Firm specific factors		
$AgeY_i$	1 if business establishment i is Y years old, 0 otherwise. Y = 1, 2, …,6 and 7+	+
$Size96_i$	Employment size of business establishment i in the year 1996	–
$Size96_i^2$	Employment size of business establishment i in year 1996 squared	+
$Growth(t)_i$	((Size(t)$_i$ – Size96$_i$) / Size96$_i$) × 100, where Size(t)$_i$ is the employment size of establishment i at the last year t of observation between 1996 and 2002	–
$Relocation(t)_i$	1 if business establishment i relocated within the city of Hamilton between 1996 – 2002, 0 otherwise	–
Geography specific factors		
$Competition^k_i$	Ratio of the total number of employees of establishments in the same k 2-digit industry of establishment i to the number of employees of establishment i. Total number of employees is measures from all business establishments of sector k within a 1500 meter radius around establishment i	+
$Agglomeration^k_i$	Total number of establishments in the same k 2-digit industry of establishment i within a 1500 meters radius around establishment I	–
Lower Hamilton	1 if business establishment i is located below the escarpment excluding the CBD, 0 otherwise	
Upper Hamilton	1 if business establishment i is located above the escarpment, 0 otherwise	
Suburbs	1 if business establishment i is located in any of the outer municipalities that includes: Dundas, Ancaster, Stoney Creek, Flambrough or Glanbrook, 0 otherwise	
Macro-economic factors		
$UEMR(t)$	Annual unemployment rate in the Hamilton region	+
$ATINC(t)$	Annual average total income (Cnd $) in the Hamilton region	–

Table 3-3 continued ...

Covariate name	Covariate definition	Expected sign
Industry specific factors		
$RASE(t)^k_i$	Ratio of the average employment size of industry *k* at the 2-digit that establishment *i* belongs to the annual employment size of establishment *i*	+
SIC1039	1 if business establishment *i* belongs to the manufacturing sector, 0 otherwise	
SIC4044	1 if business establishment *i* belongs to the construction sector, 0 otherwise	
SIC4547	1 if business establishment *i* belongs to the communication sector, 0 otherwise	
SIC5059	1 if business establishment *i* belongs to the wholesale trade sector, 0 otherwise	
SIC6069	1 if business establishment *i* belongs to the retail trade sector, 0 otherwise	
SIC7074	1 if business establishment *i* belongs to the finance insurance, 0 otherwise	
SIC7576	1 if business establishment *i* belongs to the real estate sector, 0 otherwise	
SIC77	1 if business establishment *i* belongs to the business services sector, 0 otherwise	
SIC85	1 if business establishment *i* belongs to the educational services sector, 0 otherwise	
SIC86	1 if business establishment *i* belongs to the health and social services sector, 0 otherwise	
SIC9092	1 if business establishment *i* belongs to the accommodation food and beverages services sector, 0 otherwise	

$Growth(t)_i$ is a time covariate used to examine if changes in employment size have any influence on the success of businesses. Intuitively, establishments that exhibit a growth over time reflect superiority in performance and therefore are expected to have low risk of failure. This relation is quantified with the growth variable, which compares the employment size for the last year the establishment was observed between 1996 and 2002 to that for the base year in 1996. The expected sign of the variable is negative.

Relocate$(t)_i$ is another time covariate utilized in the hazard probability to examine if a change in business location have an influence on success. The variable takes on a value of 1 if the establishment relocated at any point in time between 1996 and 2002, 0 otherwise. Relocation in our context is used as an indicator of superiority in performance. Establishments usually relocate either to expand floor space or to overcome stress. Dijk and Pellenbarg (2000) note that intrametropolitan firm location occurs when the business reacts to ongoing stress and loss in existing profit. Failure to do so will lead to a sharp decline in profit and will result in failure.

Geography specific variables include measures for local competition and agglomeration economies along with a number of dummies categorizing the geography of the city. Local competition impinged on establishment *i* is measured as the ratio of employment of establishment *i* rivals to the employment of establishment *i*. Rivals are establishments with the same 2 digit industry as establishment *i* that are located within a radius of 1500 meters from establishment *i*. This radius provided the most statistically significant results when compared with radius values of 3000, 2500, 2000, 1000 and 500 meters. According to Lerman and Liu (1984), the devised ratio implies a number of quasi competitors with the same size as establishment *i*, all of which are competing for the same local market. Since high levels of competition are a cause of failure, we expect a positive sign of this variable in the hazard probability.

The expected positive effects of agglomeration economies on the success of businesses are captured with the agglomeration variable. Previous studies for the city of Hamilton suggested a strong presence of agglomeration economies in the city (Maoh and Kanaroglou, 2004). Agglomeration in our context is measured as a count of the number of establishments within 1500 meters from business establishment *i*. The count only includes establishments with the same 2-digit industrial sector as establishment *i*. The value of 1500 meters is based on experimenting with different values as in the case of the local competition variable.

The results from the non-parametric survival analysis suggest a variability of survival by geography in the city of Hamilton. For this, a number of dummies were devised to characterize the following three geographic areas: (1) lower Hamilton, (2) upper Hamilton and (3) suburbs. These covariates other things being equal in the model, compare the failure of establishments elsewhere in the city to failure in the CBD.

The specification of the hazard probability includes two time covariate macro-economic variables that include: (1) unemployment rate *UEMR*(t) and (2) Average annual total income *ATINC*(t). *UEMR*(t) is used to capture the direct effect of economic downturn. The expected sign of the variable is positive since high levels of unemployment occur with the loss of jobs due to business failure. The variable also captures part of the decline in demand for goods and services in certain sectors of the economy such as retail trade and some services sectors. On the other hand, *ATINC*(t) is employed to account for the purchasing power and high level of demand in the city.

Higher income levels increase the demand for goods and services and thus contribute to the success of businesses enrolled in supplying these goods and services. The expected sign for this variable is negative in the hazard probability.

Industry specific factors are used to examine their contribution on the hazard of failure. Therefore, we devise a number of categorical variables that classify establishments according to the main industry they belong to. They are introduced in the model to reflect establishment heterogeneity and account for potential variation in the hazard of failure among the different industries. We also test the effect of industry size on the hazard of failure by devising the $RASE(t)^{k}{}_{i}$ time covariate. The variable compares the average employment size of the industry at the 2-digit SIC that establishment i belongs to the annual employment size of the establishment. Baldwin et al. (2000) argue that the average size of an industry can be used as a proxy for the costs of market experimentation. Here, the higher the average size of the industry, the more pronounced the input costs associated with the entry gamble. However, since part of these costs is usually sunk and irrecoverable if failure occurs, a higher average size of the industry will be positively correlated with the amount of sunk cost that will be lost in a case of a failure. Accordingly, establishments have a better chance of survival if their employment size can reach an optimal level that matches the average size of the industry they belong to. Alternatively, the hazard of failure is more pronounced if the gap between the establishment size and the average size of the industry increases over time. Thus, the expected sign of the $RASE(t)$ variable is positive.

3.5 Estimation Results

The estimation results of the hazard model are presented in Table 3-4. Five sets of models were specified and estimated for comparison purposes. The results indicate that most of the estimated parameters are significant and meet our a priori expectation with regard to the hypothesized sign. The estimated parameters of the age variables suggest that failure occurs at all ages but is more prominent among younger establishments. This is inferred from the magnitude of the different age parameters, which tend to decline as we move from *Age1* to *Age6*.

Establishment size also seems to impact the hazard probability. The negative and significant parameter of the *Size96* covariate suggests that smaller establishments are more susceptible to failure and as size increases, the propensity of failure decreases. However, this relation is not linear as indicated by the positive and significant parameter of the *Size96* squared. The hyperbolic relation between size and failure suggests that other things being equal, failure will decrease with establishment size up to a value of 20. After that point, failure will start increasing with size. This relation is best presented by the curve in Figure 3-6. The diagram shows that failure is even higher for establishments with size greater than or equal to 40 employees. We believe that this could still be a manifestation of the right censoring effect as a result of the transformation from non-incorporated (NIP) status to becoming incorporated (IP) as described earlier. Although this effect is evident for large establishments with size greater than 50 as shown in Figure 3-2, right censoring may have also occurred for medium size establishments that became

Table 3-4: Estimation results of the failure hazard duration model for all industries, 1996 – 2002

	Firm Specific	Industry Specific	Geography Specific	Macro-economy Specific	Full Model
Variable	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**
Constant	-2.1694 ***	-2.3774 ***	-1.8071 ***	7.792 ***	6.44990 ***
	(0.0359)	(0.046)	(0.0525)	(1.0452)	(1.0787)
Age1	1.1551 ***				0.99940 ***
	(0.0466)				(0.0482)
Age2	0.9083 ***				0.78510 ***
	(0.0521)				(0.0532)
Age3	0.8191 ***				0.73550 ***
	(0.0599)				(0.0607)
Age4	0.6274 ***				0.55640 ***
	(0.0717)				(0.0728)
Age5	0.3541 ***				0.31730 ***
	(0.0831)				(0.084)
Age6	0.2755 ***				0.26180 **
	(0.0807)				(0.0816)
Size96	-0.0767 ***				-0.07150 ***
	(0.0061)				(0.00669)
Size96^2	0.00189 ***				0.00176 ***
	(0.000165)				(0.000175)
Growth(*t*)	-0.00199 ***				-0.00181 ***
	(0.000155)				(0.000154)
Relocation(*t*)	-0.7126 ***		-0.6866 ***		-0.69850 ***
	(0.0473)		(0.0467)		(0.0479)
Competition		0.000032 *	5.90E-06 -		0.00001 -
		(0.000019)	(0.000017)		(0.00002)
Agglomeration		-0.00069 **	-0.00172 ***		-0.00047 -
		(0.000278)	(0.000291)		(0.000351)
Lower Hamilton			-0.1949 ***		-0.0799 -
			(0.0569)		(0.0609)
Upper Hamilton			-0.1318 ***		-0.0564 -
			(0.0600)		(0.0655)

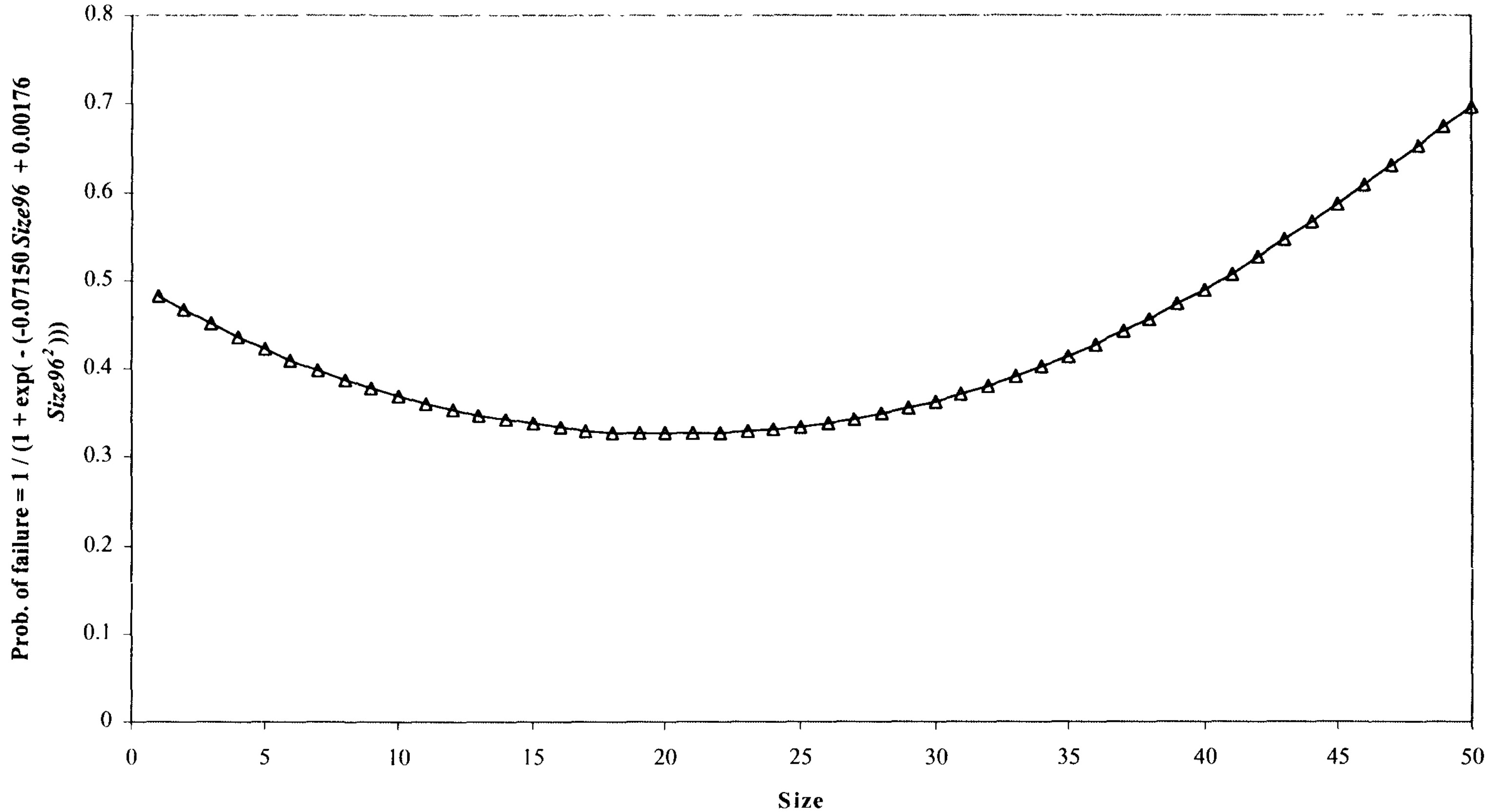
The table presents the coefficients with standard errors in parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, ** and *, respectively

Table 3- 4 continued ...

Variable	Firm Specific Coefficient	Industry Specific Coefficient	Geography Specific Coefficient	Macro-economy Specific Coefficient	Full Model Coefficient
UEMR(t)				0.1605 ***	0.1201 ***
				(0.0209)	(0.0215)
ATINC(t)				-0.00029 ***	-0.00025 ***
				(2.5E-05)	(2.5E-05)
RASE(t)		0.0192 ***			0.00573 ***
		(0.00192)			(0.00166)
Sic1039		-0.2093 **			-0.0368 -
		(0.0828)			(0.0857)
Sic4044		0.1291 **			0.1203 *
		(0.0626)			(0.065)
Sic4547		0.2871 **			0.2421 **
		(0.104)			(0.1081)
Sic5059		0.013 -			0.0599 -
		(0.0837)			(0.0864)
Sic6069		0.3072 ***			0.2389 ***
		(0.0593)			(0.0612)
Sic7074		0.7548 ***			0.571 ***
		(0.1045)			(0.1089)
Sic7576		0.1722 **			0.0656 -
		(0.0861)			(0.0884)
Sic77		0.1831 **			0.0308 -
		(0.074)			(0.0774)
Sic85		-1.5265 ***			-0.2403 -
		(0.3083)			(0.241)
Sic86		-0.5014 ***			-0.3853 ***
		(0.0774)			(0.0807)
Sic9092		0.4096 ***			0.3737 ***
		(0.072)			(0.0763)
Pseudo R2	0.0710	0.0186	0.0154	0.0309	0.1003
No. Obs.	40667	40667	40667	40667	40667
-2 Log L	24910	25944	26010	25706	24315
(%) right	66.5	58.3	55.3	53.9	69.9

The table presents the coefficients with standard errors in parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, ** and *, respectively

Figure 3-6: Relation between establishment size and the propensity of failure in Hamilton



incorporated. Another explanation for the non-linear relationship between size and failure is due to the fact that the number of establishments decreases very rapidly as establishment size increases, and the slope on *Size96* squared is therefore mainly influenced by the decreasing rate of failure from extremely small to small establishments, rather than by an actual increase in the failure rates of medium size establishments.

Unlike their cross-sectional counterparts, hazard duration models allow the analyst to examine how temporal changes in the establishment characteristics implicate the failure probability. This is possible when using time covariates, such as *Growth*(t) and *Relocation*(t). The growth parameter suggests that establishments exhibiting an increase in their employment size over time are less likely to fail. On the other hand, the parameter also implies that failure is more prominent among establishments that lost employment over time. Therefore, change in an establishment's employment size can be considered as an indicator of performance.

The relocation parameter is another performance indicator. Relocation occurs if the establishment requires additional floorspace to expand. It can also occur if the establishment is trying to overcome losses in profit due to the low attractiveness of its current location or if it is eager to increase its profit by relocating to a new site. All of these relocation actions lead to superiority in performance, associated with lower propensity of failure. Other things being equal, the relocation in our model does not suggest immunization against failure but reflects superiority in performance. An examination of the 1996 cohort shows that 34 percent of the relocating establishments have failed before 2002 after the relocation was observed. By comparison, 48 percent of non-movers have failed between 1996 and 2002.

The competition and agglomeration parameters retain the exact sign in the different models but turn out to be insignificant in the full model. However, we opt to keep them in the full model since they have the expected sign and show marginal change in their resulting standard errors. What is worth noting is the decrease in the magnitude of both parameters in the full model. We believe that the drop in the competition parameter is caused by the presence of the *UEMR*(*t*) covariate, which accounts for economic downturn in the full model. The study by Baldwin et al. (1997) shows that economic downturn has more influence on failure compared to competition. Their results also report that economic downturn and competition are related to one another in particular markets in Canada.

As for the agglomeration parameter, we believe that the drop is due to the presence of the age parameters in the full model. This was inferred from tests to include the different parameters with the agglomeration parameter to notice which ones are causing the drop. The age parameters suggest that older establishments are less susceptible to failure. Similarly, agglomeration reduces the propensity of failure. Therefore, given that older establishments are more likely to cluster together, the inclusion of the age parameters outweigh the agglomeration parameter in the full model. Industry specific variables categorizing establishments by their major industries suggest that health and social services establishments are less likely to fail. On the other hand, the results indicate that failure is more prominent among finance insurance, accommodation food and beverages, transportation and storage, retail trade and construction establishments. The results from the industry specific model also show that manufacturing and educational service establishments are less likely to fail. They also imply that failure is evident for real estate and business service establishments. Although the parameters presenting these industries turn out insignificant in the full model, they retained the expected sign. The *RASE*(t) covariate is significant and retains the expected sign. The positive parameter suggests that the size of the industry is important and will influence the propensity of failure. Therefore, establishments not able to match the average size of the industry they belong to are more likely to fail.

The parameters of the macro-economy specific covariates are significant and have the expected sign. The sign of the *UEMR*(t) parameter suggests that other thing being equal, high failure rates are associated with high unemployment rates in Hamilton. On the other hand, the sign of the *ATINC*(t) parameter suggests that an increase in the levels of income in Hamilton will lead to a higher purchasing power, decreasing the likelihood of failure.

The estimation results suggest that the full model is superior in its predictive ability, correctly predicting 70 percent of the modeled population. The model also retains the highest pseudo rho-square value. A comparison among the different models suggests that other things equal, firm and macro-economy specific factors contribute to explaining most of the variability in the full model. This is evident from the percent explained right and pseudo rho-square values in the firm specific and macro-economy specific models.

3.6 Conclusion and Direction for Future Research

The objective of this paper was to develop an econometric model with two purposes in mind: first, to investigate the failure process of small and medium size establishments in the city of Hamilton, Canada; second, to use the model as the basis for the development of a business failure module in a firmographic microsimulation system. Such a firmographic system will be developed to simulate the different firmographic processes that will drive the spatio-temporal evolution of Hamilton's business establishment population as shown in Figure 3-7. Therefore, in addition to the business failure module, the microsimulation system to be developed includes: (1) an establishment mobility module that determines if a business establishment will move within the city or leave the city over time, (2) a location choice module that determines the location of intra-urban, in-migrating and newly formed establishments in the city (Maoh and Kanaroglou, 2005), and (3) a growth-decline module that predicts the change in employment size of business establishments over time. Figure 3-8 presents a general architecture of the system, which illustrates the working mechanism of the various firmographic modules.

In developing the firmographic microsimulation system the first step will require the creation of a synthetic micro-database to replace the confidential Business Register data used in the estimation of the different econometric models. The development of the spatial firmographic system will be based on the geodatabase data model of the widely used geographic information system software environment "ArcGIS". The object-oriented nature of this geodatabase model will enable us to endow the business establishments (agents of the system) with natural behaviors, and allow any sort of relationship to be defined among them and any other geographic features (Zeiler, 1999). Visual programming via the unified modeling language (UML) will facilitate the development of such a dynamic geodatabase data model and will allow its extendibility with other future land use and transportation modules.

Once operational, the model-system will be used to evolve the synthetic establishment population over time to study the inter-play between the local economy and urban form in Hamilton. It will also be used to assess if agent-based land use models are truly superior to conventional zone-based land use models by comparing the results from the devised system to those produced by a conventional employment location model (Maoh et al., 2005)

The results in this research provide an analytical framework for studying the business establishment failure process in the urban context. The significance of the research is in two folds: (1) it demonstrated the value of using the Statistics Canada Business Register (BR) in studying the failure firmographic process (2) it provided new results on the factors influencing the failure of business establishments in Hamilton using survival analysis. The latter were well portrayed by the results achieved from the non-parametric survival analysis, which suggested a variation in the rates of survival based on the age, size, location and type of industry for the establishment. These results were further supported by the estimates of the discrete-time hazard duration model.

The model suggested that the business failure process could be explained by the establishment characteristics, the industry it belongs to, its geographic location and the macro-economic conditions in the city. Of the above factors, establishment characteristics and macro-economic conditions appear to be the most influential. Despite these results, we believe similar effort should be conducted to analyze the failure of business establishments by major economic sector. This perhaps could enhance the predictive ability of the failure module and would certainly provide more insight on how sensitive establishments of different economic sectors are to certain firm specific, geography specific, industry specific and macro-economic factors.

Lastly, our effort to model firm survival and failure succeeded in portraying the importance of the BR in devising agent-based simulation models. However, the specification suffered some limitations and was constrained by the quality of the available BR data. For instance, firm specific internal factors were dismissed from the specification of the model despite their importance in explaining the failure process as shown by Baldwin et al (1997). This suggests the need for improving the quality of the BR data in the future. Such an effort can only be realized and achieved if projection tools of the sort we are developing can become a more prominent practice in informing public policies at the local, regional or the national levels.

Figure 3-7: The evolutionary process of business establishment population over time

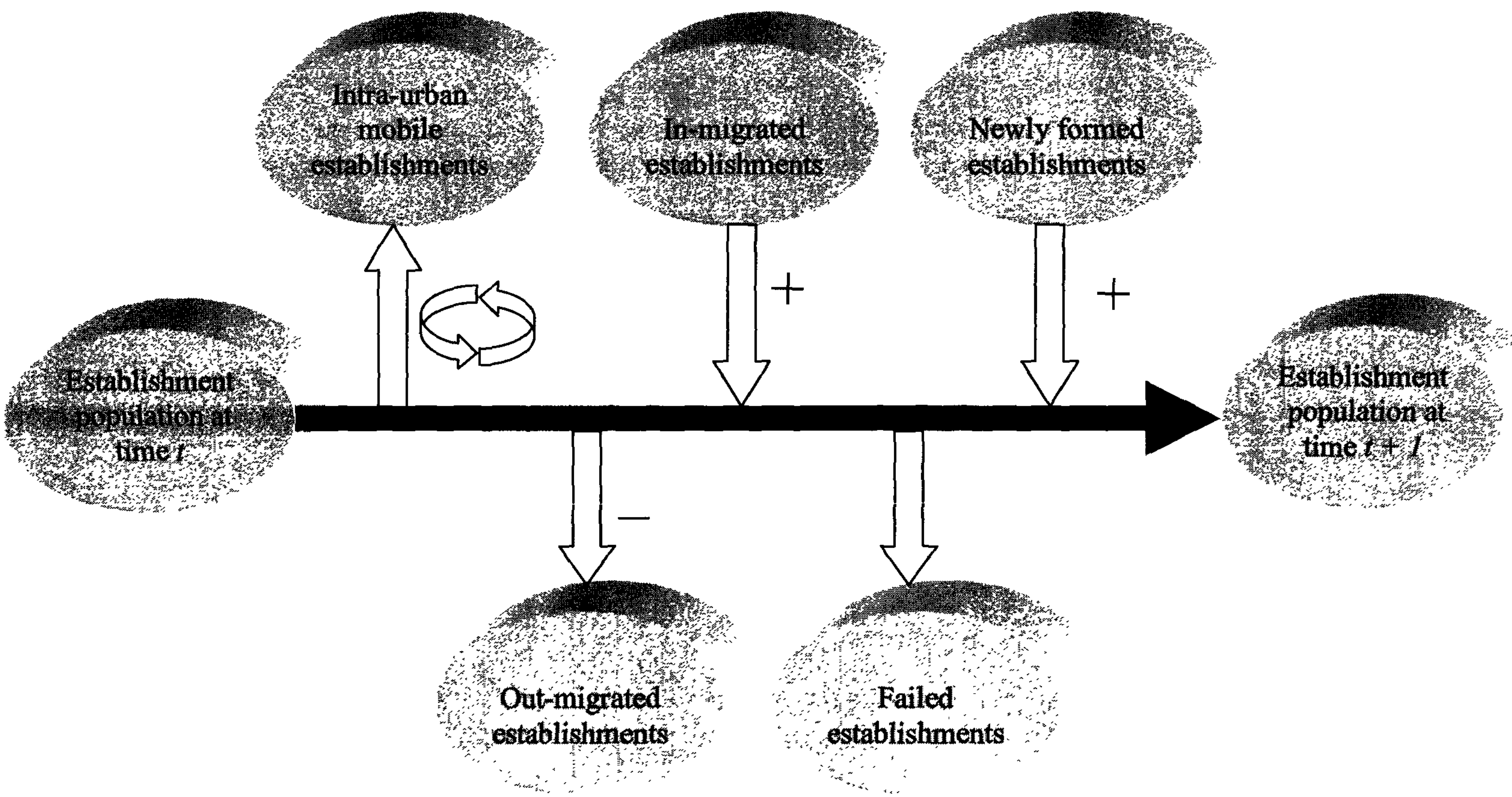
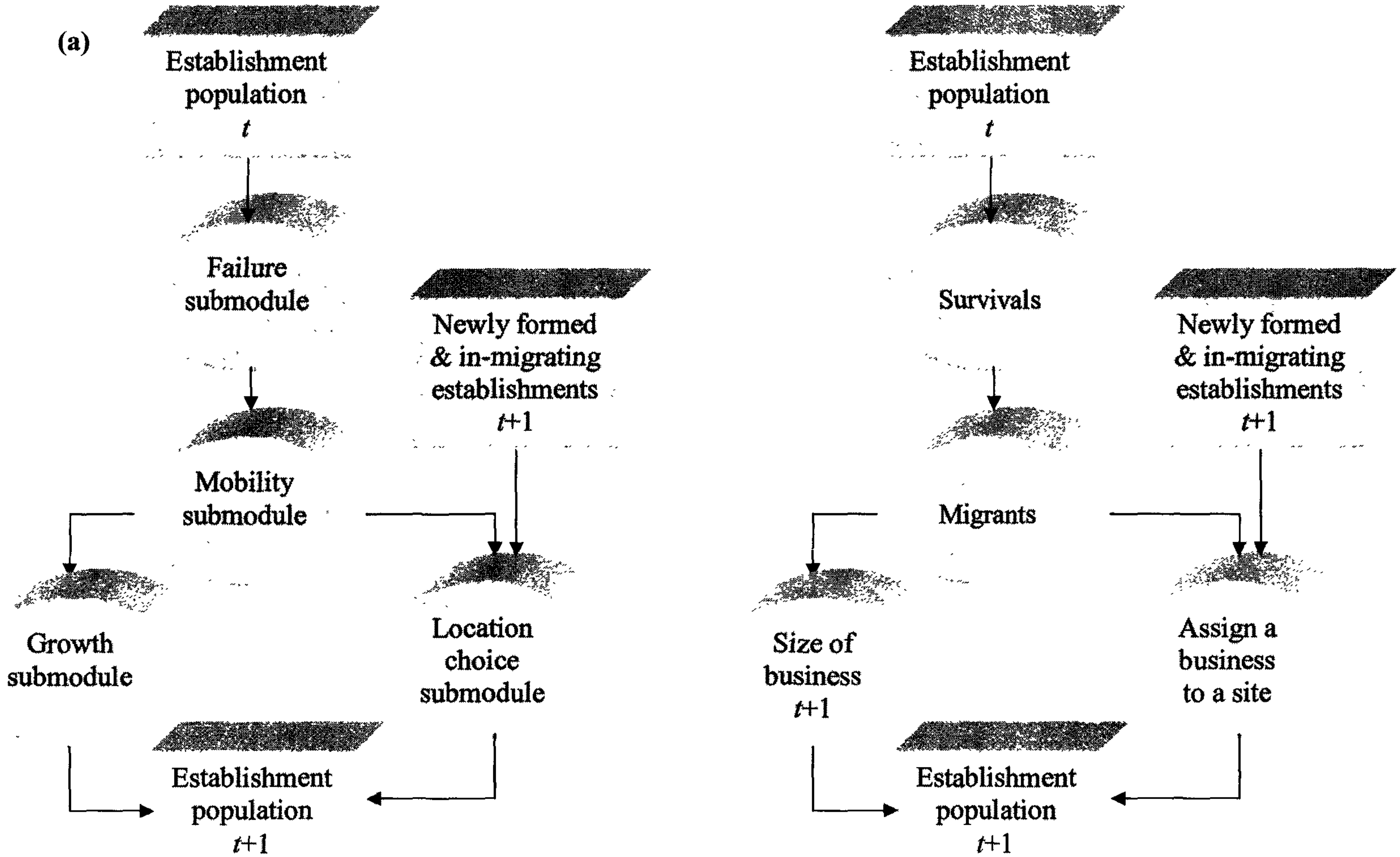


Figure 3-8: Model architecture - (a) Firmographic processes; (b) Process output



3.7 References

Agarwal, R. and D. Audretsch (2001). Does entry size matter? The impact of the life cycle and technology on firm survival. *The Journal of Industrial Economics* XLIX, 21-43.

Allison, P. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology* 13, 61-98.

Allison, P. (1995). *Survival analysis using the SAS system: a practical guide.* SAS Institute Inc.

Audretsch, D. (1991). New-firm survival and the technological regime. *Review of Economics and Statistics* 73, 441-450.

Audretsch, D. and M. Talat (1995). New firm survival: new results using a hazard function. *Review of Economics and Statistics* 77, 97-103.

Audretsch, D., E. Santarelli and V. Marco (1999). Start-up size and industrial dynamics: some evidence from Italian manufacturing. *International Journal of Industrial Organization* 17, 965-983.

Baldwin, J., T. Gray, J. Johnson and J. Proctor (1997). *Failing concerns: business bankruptcy in Canada.* Ottawa: Statistics Canada.

Baldwin, J., L. Bian, R. Dupuy and G. Gellatly (2000). *Failure rates for new canadian firms: new perspectives on entry and exit.* Ottawa: Statistics Canada.

Baldwin, J. and G. Gellaty (2004). *Innovation strategies and performance in small firms.* Cheltenham, UK: Edward Elgar.

Berglund, E. and K. Brannas (2001). Plants' entry and exit in Swedish municipalities. *The Annals of Regional Science* 35, 431-448.

Caves, R. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature* 36 (4), 1947-1982

Cox, D. (1972). Regression models and life-tables. *Journal of the royal statistical society, Series B* 34, 187-202.

Dieleman, F., M. Dijst and B. Guillaume (2002). Urban form and travel behavior: micro-level household attributes and residential context. *Urban Studies* 39, 507-527.

Dietmar, H., S. Konrad and M. Woywode (1998). Legal form, growth and exit of West Germany firms-empirical results for manufacturing, construction, trade and services industries. *The Journal of Industrial Economics* XLVI, 453-488.

Dijk, J. and P. Pellenbarg (1999). The demography of firms: progress and problems in empirical research. In *Demography of Firms: Spatial Dynamics of Firm Behaviour*, ed.Dijk, J. and Pellenbarg, P., Vol. 17, Netherlands Geographical Studies, KNAG, RUG, pp. 325-337.

Dijk, J. and P. Pellenbarg (2000). Firm relocation decisions in the Netherlands: an ordered logit approach. *Papers in Regional Science* 79, 191-219.

Dunne, P. and A. Hughes (1994). Age, size, growth and survival: UK companies in the 1980s. *Journal of Industrial Economics* 42, 115-140.

Fotopoulos, G. and H. Louri (2000). Location and survival and new entry. *Small Business Economics* 14, 311-321.

Geroski, P. (1995). What do we know about entry? *International Journal of Industrial Organization* 13, 421-440.

Harhoff, D., K. Stahl and M. Woywode (1998). Legal form, growth and exit of West Germany firms: empirical results from manufacturing, construction, trade and services industries. *The Journal of Industrial Economics* XLVI, 453-488.

Holmes, P., I. Stone and P. Braidford (2000). New firm survival: an analysis of UK firms using a hazard function. Proceedings: The Royal Economic Society 2000 Conference, University of St. Andrews, Scotland.

Hunt, J., J. Khan and J. Abraham (2003). Microsimulating firm spatial behavior. Proceedings: Eighth International Conference on Computers in Urban Planning and Urban Management "CUPUM", Sendai, Japan.

Kanaroglou, P. and D. Scott (2002). Integrated urban transportation and land-use models for policy analysis. In *Governing Cities on the Move: Functional and Management Perspectives on Transformations of Urban Infrastructure in European Agglomerations.*, ed.M. Dijst, pp. 43-75. Ashgate.

Kiefer, N. (1988). Economic duration data and hazard functions. *Journal of Economic*

Literature XXVI, 646-679.

Lerman, S. and T. Liu (1984). Microlevel econometric analysis of retail closure. In *Discrete choice models in regional science,* ed.D. Pitfield, pp. 181-201. London: Pion Limited.

Love, J. (1996). Entry and exit: a country-level analysis. *Applied Economics* 28, 441-451.

Maoh, H. and P. Kanaroglou (2004). Economic clustering and urban form: the case of Hamilton, Ontario. Working Paper No. 3, Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University, Hamilton.

Maoh, H. and P. Kanaroglou (2005b). The location of business establishments in the city of Hamilton, Canada: a micro-analytical model approach. Working Paper No. 8, Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University, Hamilton.

Maoh, H., P. Kanaroglou and R. Buliung (2005). Modeling the location of firms within an integrated transport and land use model for Hamilton, Ontario. Working Paper No. 6, Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University, Hamilton.

Mata, J. and P. Portugal (1994). Life duration of new firms. *Journal of Industrial Economics* 42, 227-246.

Mata, J., P. Portugal and P. Guimaraes (1995). The survival of new plants: start-up conditions and post-entry evolution. *International Journal of Industrial Organization* 13, 459-482.

Miller, E., D. Kriger and J. Hunt (1998). Integrated urban models for simulation of transit and landuse policies. *TCRP Web Document 9* Project H-12, 1-252.

Miller, E., J. Hunt, J. Abraham and P. Salvini (2004). Microsimulating urban systems. *Computers, environment and urban systems* 28, 9-44.

Mirron, J. (2002). Loschian spatial competition in an emerging retail industry. *Geographical Analysis* 34, 34-61.

Pinkse, J. and M. Slade (1998). Contracting in Space: an Application of Spatial Statistics to the Discrete Choice Model. *Journal of Econometrics* 85, 125-154.

Southworth, F. (1995). A technical review of urban land use-transportation models as tools for evaluating vehicle travel reduction strategies. Working Paper No. ORNL-6881, Oak Ridge National Laboratory, Tennessee.

Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke and G. Ullfarsson (2003). Microsimulation of urban development and location choices: design and implementation of UrbanSim. *Networks and spatial economics* 3, 43-67.

Wegener, M. and K. Spiekermann (1996). The potential of microsimulation for urban models. In *Microsimulation for urban and regional policy analysis,* ed.G. P. Clarke, pp. 149-163. London: Pion.

Wegener, M. (2004). Overview of land use transport models. In *Handbook of transport geography and spatial systems,* ed.D. Hensher, K. Button, K. Haynes and P. Stopher, pp. 127-146. Elsevier.

Van Wissen, L. (2000). A micro-simulation model of firms: applications of concepts of the demography of the firm. *Papers in Regional Science* 79, 111-134.

Van Wissen, L. (1997). Demography of the firm: modelling birth and death of firms using the concept of carrying capacity. In: *Population and Families in the Low Countries 1996/1997*, eds,H. Brekel and F. Deven), pp. 219-244. NIDI/CBGS, Den Haag, Brussel

Zeiler, M. (1999). *Modeling our world: the ESRI guide to geodatabase design*. Redlands: Environmental Systems Research Institute (ESRI), Inc.

CHAPTER 4: Business Establishment Mobility Behavior in Urban Areas: An Application to the City of Hamilton in Ontario, Canada

4.1 Introduction

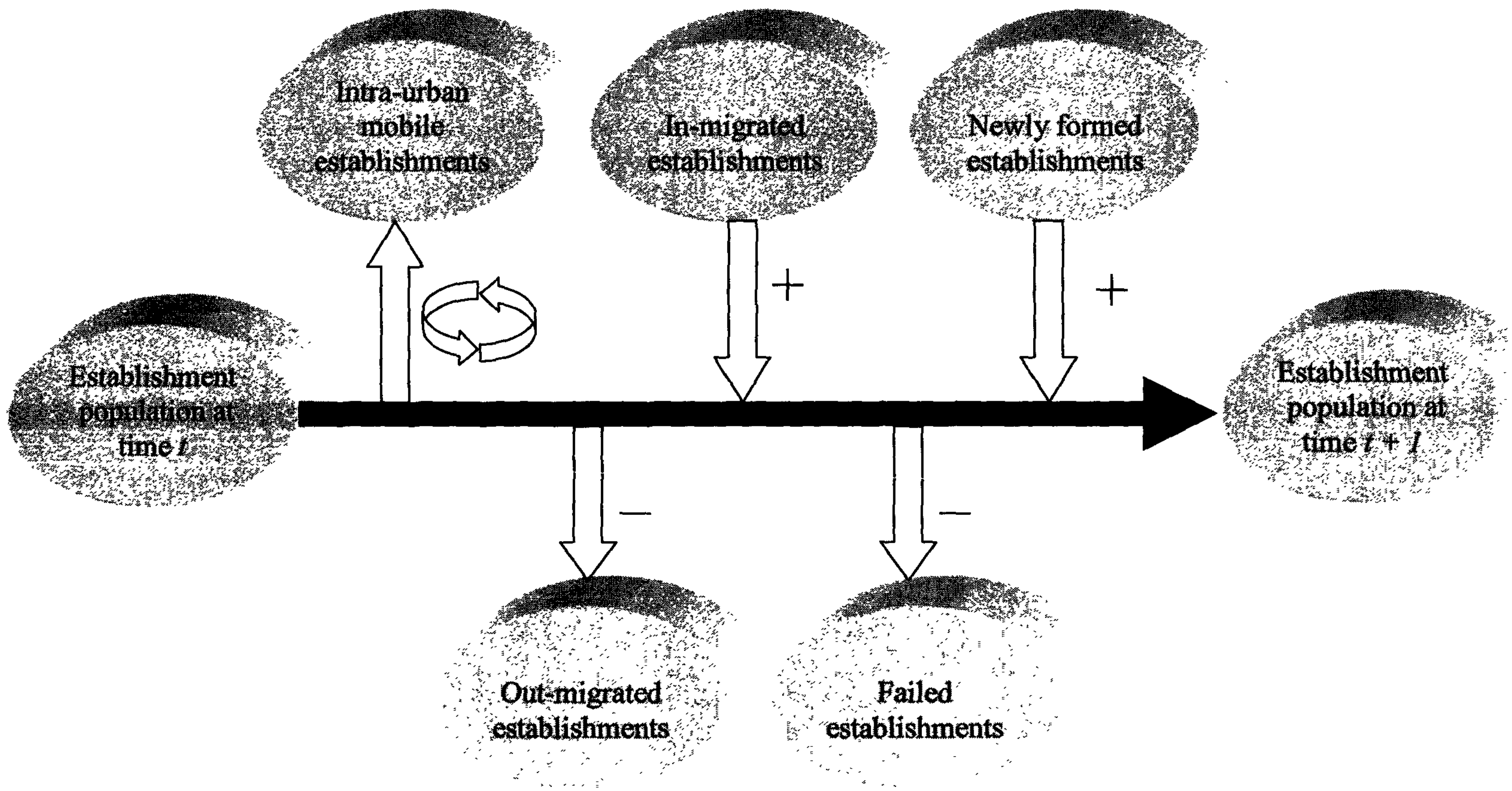
The concern about sustainable planning of cities over the past few decades has generated a considerable amount of empirical literature on the emerging urban form and its causes. A general theme drawn from this literature suggests that modern cities are becoming suburban due to decentralization of firms and people (Anderson et al., 1996). On the firm side, exodus from the traditional core of many North American cities has been fuelled by space shortage, parking problems, increasing land prices and growing congestion on city roads. Land development practices in the suburbs and the construction of beltways around cities have also aggravated decentralization and helped transform the monocentric city into a polycentric one.

Earlier efforts to assess the urban development process were based on building and using aggregated urban land use models. These simulation tools are based on zonal socio-economic data and have been used to address policies that relate to the planning process of cities. However, due to their aggregate nature, most of these models have been criticized for lacking the behavioral realism in imitating the complexities of the city development process (Miller et al., 2004). Recent thinking during the past decade and a half calls for the adoption of a microanalytical approach, whereby change in urban form is driven by the collective behavior of individual actors (persons, households, firms, developers, etc.) and their location decisions. Therefore, the total process of urban development is decomposed into subprocesses and within these the main actors (agents) and their decision behavior is identified and analyzed (Wegener and Spiekermann, 1996).

A review of the recent literature indicates that research to devise agent-based land use models is still in its early stages (Miller et al., 2004). There is a clear shortage of empirical studies that employ the micro approach to analyze the urban development process in general and the evolution of business establishments in particular. One explanation for this could be the shortage of micro data that are either confidential or too costly to collect. In the Canadian context, micro data on individuals and business establishments are collected and maintained by Statistics Canada. However, these datasets are confidential. Fortunately, Statistics Canada has a number of research programs that provide access to these micro datasets to conduct statistical analysis. Through such a program we were able to make use of data extracted from Statistics Canada Business Register (BR) to explore and study the evolution of business establishments in the City of Hamilton, Ontario, Canada. Our overall objective is to devise an agent-based microsimulation model that can be used to study and assess the interplay between the local economy and urban form in Hamilton.

To this end, we adopt the firm demographic approach illustrated in Figure 4-1 to study the evolution of the business establishment population. We investigate the causes that lead to establishment failure, formation, mobility, location choice and

Figure 4-1: The evolutionary process of business establishment population over time



growth/decline. In this paper we focus on the mobility behavior of business establishments. We specify and estimate multinomial logit models by economic sector to predict the probability of staying, moving within the city (relocating) or leaving the city (out-migrating). We attempt to fill the gap in the literature by studying the establishment mobility behavior in the City of Hamilton. The modeling framework we provide can serve as a prototype for similar studies in other cities.

Here is a brief description of the remainder of this paper. In the next section we provide a theoretical background to overview the location theories and factors influencing business establishment mobility. Section three describes the city of Hamilton and the dataset used in the analysis. It also provides a general overview of the recent mobility trends of business establishments in Hamilton. Section four discusses the formulation and specification of the discrete choice model we use to study the mobility behavior of business establishments by economic sector. Section five presents the estimation results. Finally, the last section provides a conclusion and direction for future research.

4.2 Theoretical Background

The mobility of business establishments is one of the basic firmographic processes that influence the spatial distribution of economic activities and employment in the city. It is defined as the action of leaving the current location to relocate elsewhere in or outside of the city. The analysis of establishment mobility aims to understand the underlying factors that push establishments to leave their current location. According to

Bade (1983), establishments have a strong preference to remain in situ and will only move in situations of high locational pressure caused by location deficiencies. Van Wissen (2000) notes that location deficiencies (push factors) may arise as a result of changing market orientation, space requirements, technological change, location costs, accessibility problems, local policy and labor market mismatch. Such factors are also highlighted elsewhere in the literature (See for example Brouwer et al. 2004, Dijk and Pellenbarg, 2000)

The migration process of business establishments is the outcome of two subprocesses that includes the mobility of establishments (relocation) and the location choice problem. While some studies have addressed the latter problem in the urban context (Maoh et al., 2005), little has been done to analyze establishment mobility at this geographical scale, as indicated by the absence of empirical studies on the topic. Our review to the literature also indicates a shortage of recent empirical studies to analyze establishment mobility at the regional level. According to Pellenbarg et al. (2002), firm migration research received more interest from policy makers rather than scientists. The need to solve practical planning problems or stimulate economic and employment growth was the main driving force for such an interest. Pellenbarg et al. (2002) conduct a historical review of firm relocation studies. They note that the majority of empirical studies took place during the 1970s, since firm migration was perceived as a panacea for regional development during that time. However, the idea that regions should create the conditions for innovation and creation of new economic activities that will retain firms in the region has changed this perception. Research in recent years was geared toward the issue of sustained economic growth, where migration from the firm's perspective is considered a locational strategy for removing restrictions to growth. The key point here is that firm relocation studies are rare in general and are not well addressed in the urban context in particular.

Brouwer et al. (2004) among others note that business establishment mobility has been explained via three types of factors that include: (1) firm internal factors such as the establishment size and age, (2) firm external factors such as market size and (3) location factors such as characteristics of regions. These factors are the outcome of employing the three location theories listed in Table 4-1. Dijk and Pellenbarg (2000) and Pellenbarg et al. (2002) provide a critical review of these theories. Following these reviews, we will shed light on the main themes underlining each theory in the discussion to follow.

Table 4-1: Location theories and factors influencing business mobility

Theoretical framework	Key concepts (factors)	Variables
Neo-classical theory	Market situation (Location factors)	Market size Country of location
Behavioral theory	Information/Abilities (Internal factors)	Firm size Firm age
Institutional theory	Networks (External factors)	Firm growth (positive and negative; merger; acquisition; take-over)

Source: Adapted from Brouwer et al. (2004)

The neo-classical location theory treats the firm as a profit maximizer and classifies sites into profitable and unprofitable locations. The classification is based on the difference between a least-cost surface and a revenue surface. The former can be derived from aggregating surfaces generated by (1) the transportation cost of industry inputs and outputs (2) labor cost and (3) market size. In this setting, one can define the boundaries of the spatial margins to profitability. These boundaries can then be used to model the relocation decision of firms under the assumptions of rationality and perfect information. Here, a firm moves from the current location to a new one when the first is no longer inside the spatial margins to profitability (i.e. revenue is less than cost) and the second might be profitable (i.e. revenue exceeds cost).

While the neo-classical theory sets a useful benchmark in defining the optimal behavior of firms in economic terms, it discards the internal dynamics of the firm. This limitation led to the development of a behavioral theory that is based on the more realistic notion of limited information and bounded rationality. Accordingly, the firm is treated as an agent and its optimizing behavior is replaced with one that is satisfycing. The behavioral theory treats the firm as a decision maker by focusing on the decision making process of entrepreneurs. Here, the perception of geographic configuration is what people use in their spatial decision-making. Accordingly, limited information, limited ability, perception and uncertainty all lead to a large spatial bias in relocation decision-making. Also, the theory assumes that most of the moves take place at short distances because closer locations are more familiar and easier to imagine than distant locations. Pellenbarg et al. (2002) note that push reasons leading to firm mobility can be related to both firm internal and external factors. Internal factors are related to the size and age of the firm.

The need for more floorspace to expand, given the change in size, can push the firm to relocate. Also, the age of the firm can explain the mobility behavior in cases when the current location of older firms is not attractive any more. On the other hand, external factors may include limited accessibility to site, deterioration of the building, environmental considerations, limited labor supply or high location cost.

The institutional location theory, or what is known as the "geography of enterprise", assumes that society's cultural institutions and value systems are the driving forces of any economic activity. The firm is viewed as an interacting entity that has to negotiate with various institutions (suppliers, governments, unions, etc.) to guarantee the endurance of its economic activity. Therefore, locational behavior is influenced by the outcome of the investments the firm has made through its negotiations. While the institutional location theory is best suited for large corporations who have more negotiating power, it can have an impact on the relocation of small and medium size (SME) firms. According to Pellenbarg et al. (2002), governments and real estate markets are two institutions that can impact the mobility behavior of SME firms. While the former can be more influential in the regional context through regulating subsidies, tax reduction and environmental standards, the latter plays a critical role in the urban context. The intra-urban migration behavior of SME firms is influenced by the demand for commercial property and its spatial characteristics. Therefore, the change in supply of real estate via the development of commercial and industrial floorspace, office space or industrial sites can pull firms and encourage them to relocate.

The three theories highlighted above provide a solid background for studying the mobility behavior of firms. However, we believe that the behavioral theory can be the most influential when developing agent-based firmographic models. The agent-based approach assumes that the business establishment is the decision maker and therefore its mobility behavior is more likely to be dependent on its internal characteristics. Yet, other locational and external factors may interfere and can exert an influence on the mobility decision. Therefore, external and location factors remain important determinants and should not be dismissed from the analysis.

In terms of methods, most of the recent empirical research on firm mobility utilizes one form or another of a discrete choice model to explain the mobility behavior of firms. Brouwer et al. (2004) study the relocation decision of large firms with more than 200 employees in 21 European countries using a binary logit model. Their results suggest that firm internal factors, location factors and external factors contribute to explaining the mobility behavior of firms. However, they were unsuccessful in providing a definite answer of which of the three location theories contribute the most to explaining the mobility behavior of European firms. Dijk and Pellenbarg (2000) use an ordered logit model to study firm relocation decision in the Netherlands. They use stated preference of firms with regard to their migration behavior and relate such information to a set of variables that reflects firm internal and external characterizes as well as location factors. They conclude that the decision to relocate is mainly determined by firm internal factors and to a lesser extend by site related factors. Such findings put the behavioral theory in the lead for explaining firm relocation behavior. In a similar vein, Van Wissen (2000) estimates a binary logit model to study the mobility behavior of firms in the Netherlands. The model was part of an agent-based firmographic microsimulation system developed for the whole country. His findings coincide with those of Dijk and Pellenbarg (2000) suggesting that firm internal factors such as age and size make the most significant contribution towards explaining the propensity of mobility.

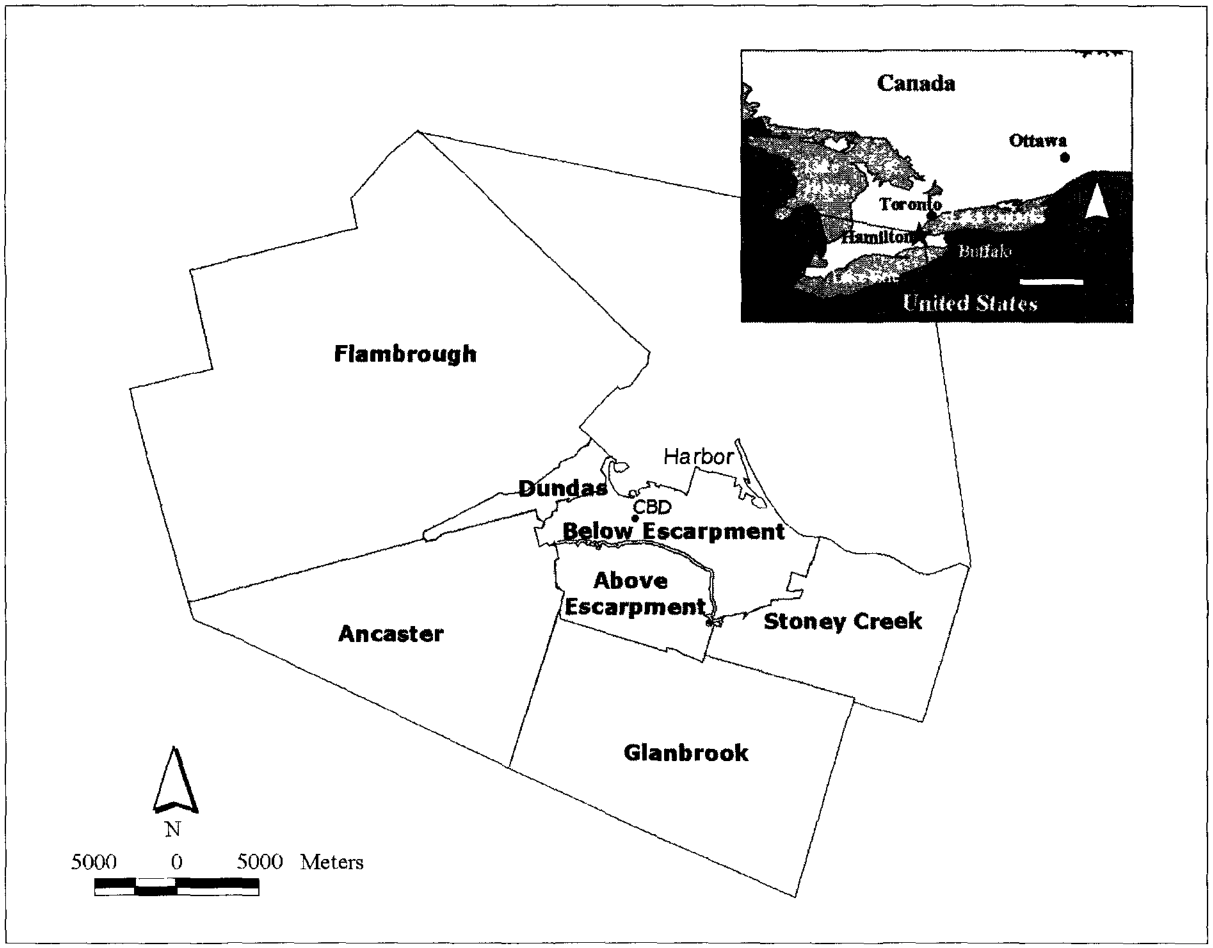
4.3 Study Area and Data

4.3.1 Study Area

The city of Hamilton in Ontario Canada is the focus of our study. The city is located approximately 60 Km southwest of Toronto and has evolved around the southwest harbor of lake Ontario as shown in Figure 4-2. The Niagara escarpment separates the city into a lower coastal area between the escarpment and the harbor to the north and a highland beyond the escarpment to the south. While early urban development took place on the lower city, the continuous population and economic growth over time led to urban development above the escarpment. Recently, the city was amalgamated with the municipalities of Ancaster, Dundas, Flamborough, Glanbrook and Stoney Creek.

The city's population is 504,513, as reported in the 2001 Canadian census. According to the Statistics Canada Business Register, the city had around 12,000 business establishments and 223,000 jobs in 2002. Hamilton has an extensive transportation network with a number of major highways and expressways that provide accessibility to

Figure 4-2: The City of Hamilton in the regional context



the suburbs and to the rest of the Greater Toronto Area (GTA). As in the case of many North American cities, urban sprawl has been one of the main features shaping the form of the city over the past decade and a half (Kanaroglou and South, 1999).

4.3.2 Data

We study the mobility behavior of small and medium (SME) size business establishments with less than 200 employees. These account for the majority (99%) of the establishment population in Hamilton and are believed to be more footloose due to the lower cost required for their infrastructure. Large establishments, on the other hand, have a more predictable mobility behavior given their small number and the high monetary value associated with their infrastructure. Therefore, the agent-based model will be developed to capture the dynamics associated with the mobility of the SME establishments. The change in location for larger establishment will be treated exogenously in the model.

The data on SME establishments is extracted from Statistics Canada Business Register (BR). The sample is constrained to self-owned non-incorporated (NIP) SME businesses since annual information is required for the estimation of the econometric models and is only available for this group of establishments in the BR. Each establishment has a list of attributes that include: (1) establishment number (2) postal code address in year *t* (3) paid worker jobs in year *t* (4) total operating revenue in year *t* (5) 1980 Standard Industrial Classification (SIC) in year *t* (6) 1986 SGC Census Division in year *t* (7) operating name in year *t* (8) street name and number in year *t* (9) city/municipality in year *t*. Applying a backward comparison enabled us to generate an age variable for each establishments in the 1996 data. Each record in the 1996 data is compared to the annual records of the BR that go back to the year 1990. The comparison is based on the establishment number (EN), which is a unique identifier generated when the establishment is created in the BR. The generated age variable took on ordered discrete values of (1, 2, …, 6 and 7+) depending on the earliest year the establishment is observed in the annual records.

General statistics from exploring the BR data reveal that NIP establishments accounted for 86% of the SME establishments in 1996. The exploration also suggests that 79% of SME establishments had employment size less than or equal to 10, 93% of which are NIP establishments. Accordingly, the NIP population is deemed appropriate for our objectives since it accounts for the majority of the SME establishments. However, the absence of the incorporated (IP) establishments from the extracted sample will undermine our knowledge of what role organizational structure has on the mobility behavior of establishments. Unlike the IP establishments, NIP establishments are simple in their organizational structure and each establishment corresponds to a unique firm in the city.

The extraction procedure of the data involves identifying the establishments that stayed at the same location, moved within the city (relocated) or left the city (out-migrated) between 1996 and 1997. First, we identified the surviving establishments between 1996 and 1997. This is achieved with the help of the establishment number (EN). The establishment is considered a survivor if its EN number existed in the two consecutive years. Next, we used the postal code address information and the Standard Geographical Classification (SGC) codes to classify a surviving establishment into *stayer*, *mover* or *out-migrant*. A *stayer* event is recorded if the establishment has the Hamilton SGC code and same postal code and coordinates in 1996 and 1997. A *mover* event is recoded if the establishment has the Hamilton SGC code but different postal code or coordinates between 1996 and 1997. Here, the mover event identifies establishments that relocated within the city. Finally, an *out-migrant* event is coded if the establishment has the Hamilton SGC code in 1996 and a different SGC in 1997.

4.3.3 Mobility Trends

Our exploration of the BR data suggest that mobility is not common among business establishments given that only 7% of the 1996 NIP establishment population relocated within Hamilton by 1997. This is compared to 2% out-migration in the same time period. An examination of these rates by economic sector suggests that communication, real estate, and accommodation food and beverage service establishments have the highest mobility rates. On the other hand, health and social service and educational service establishments have the lowest mobility rates.

We further investigate the mobility by size of establishment. For this, we consider the whole 1996 establishment population and compare it to the 2002 population since annual information on the whole population is not available as noted in the previous section. For intra-urban mobile establishments, the mean employment size is approximately 15 and the mean relocation distance is 5 km between 1996 and 2002. This implies that relocation is more prominent among smaller establishments and is happening at short distances. The results from measuring the mobility rates between eight geographic areas that includes the CBD, land below the escarpment, land above the escarpment, Ancaster, Stoney Creek, Flamborough, Ancaster, Dundas and Glanbrook suggests that more than half of the moves happened at short distances within the same geography. However, the analysis indicates that some of the intra-urban moves happened at longer distances. The maximum distance observed between the old and the new location was 34 km. This is believed to be the case as a result of decentralization and new suburban development that work on pulling establishments to the fringes of Hamilton (Maoh and Kanaroglou, 2004b).

The data also suggests that the vast majority (91%) of the out-migrating establishments moved to new locations within a radius of 100 Km around Hamilton between 1996 and 2002. The analysis also shows that 57% of the out-migrant establishments moved to close by locations in the Greater Toronto Area (GTA). These findings imply that most of inter-regional moves took place at short distances. Similar results on short distance migrations could be found else where in the literature (Pellenbarg et al., 2002).

4.4 The Mobility Model

4.4.1 Model Formulation

The model we seek to develop will be used to determine if an individual business will choose to stay at its current location, relocate to a different place within the city or will leave the city between 1996 and 1997. We use random utility theory (McFadden, 1974) to model the mobility behavior of business establishments. In this respect, the utility U_{ij} of a given alternative j for an individual establishment i can be expressed as a composite of a systematic utility V_{ij} and a random error term ε_{ij}

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

If we know that the systematic part of the utility is a linear-in-parameters function that can be measured with observed information and we make the assumption that the unobserved random errors are independent and identically distributed (IID) as Gumbel, then one can show that the probability of choosing alternative j over any of the other available alternatives can be expressed as:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{j=1}^{J} \exp(V_{ij})} \tag{2}$$

The probability in equation 2 is the well-known conditional multinomial logit (CMNL) model. The assumption of an IID structure for the random error terms in equation 1 gives rise to the independence from irrelevant alternatives (IIA) property in equation 2 (Ben-Akiva and Lerman, 1985). Under this property, business establishments are assumed to perceive the stay, relocate and leave alternatives as unique choices such that each is evaluated independently from the other. This assumption is deemed valid since intra-urban and inter-regional mobility of establishments occurs for different reasons despite the fact that both actions suggest a departure from the current location. Therefore, the IID assumption holds in our context and equation 2 can be used to model the three outcomes: {*Stay*, *relocate* and *leave*}.

It is worth mentioning that the model in equation (2) is downsized from an initial full information nested logit framework as shown in appendix 1. Within this framework, the migration behavior of the establishment is modeled as the outcome of two interlinked steps that includes: (1) the decision of mobility and (2) the decision of location choice conditional upon mobility as will be shown in the next chapter. The advantage of this approach is the gain in utilizing information about the new location through the logsum value to model the decision of mobility. Therefore, one can examine if pull factors, resembled by the characteristics of the new location, have an influence on the likelihood of mobility. However, the estimation procedure of full information maximum likelihood function for the four nested logit model structures in appendix 1 did not converge. Also, a sequential structure in which the logsum value from an estimated location choice model is calculated and specified into equation 2 as a regular covariate did not allow the estimation procedure of the model in equation 2 to converge or returned a large coefficient that is not consistent with the discrete choice theory when the logsum was specified as a the only covariate in the model.

We believe these results are due to the nature of the choice set at the lower level nest, which is based on a random sample of 9 alternatives drawn from the actual large choice set and added to the observed chosen alternative. As a result, the utilities used in the calculation of the logsum values lack the alternative specific constants normally obtained when estimating the MNL with the full choice set. Therefore, the generated logsum covariate lacks the information that is normally captured by the alternative specific constants. While there is no remedy to account for the absence of the alternative specific constants in the case of using randomly selected alternatives, a way to utilize information from the lower level nest in our case might be to use a larger set of randomly selected alternatives such as 49 or 99. This option will be addressed in future research.

4.4.2 Model Specification

The specification of the stay, relocate and leave systematic utilities of the CMNL model in equation 2 is based on the available data that were extracted from the Business Register (BR). We follow the theoretical framework discussed in section 2 to devise covariates that reflect internal characteristics of the establishments as well as location factors. Table 4-2 lists the covariates used in the specification of the stay (S), relocate (R) and leave (L) utilities of the logit formula. Internal factors include establishment size, age, growth rate and economic sectors. Location factors, on the other hand, consist of a number of categorical variables that reflect the morphology of the city. We also use local competition as a measure of location pressure. This variable is combined with a variable measuring the distance between the old location of the establishment in 1996 and its new one in 1997. Finally, we use a variable that measures the level of geographical clustering of business establishments to account for the effect of agglomeration economies.

With regard to the size of the establishment (*Size96*), we expect that small establishments can move more easily to a new location than large establishments. Dijk and Pellenbarg (2000), note that employment size can be used to capture the effect of moving cost of the establishment. They argue that the cost of moving for small establishments is expected to be much lower than for large establishments. This suggests

Table 4-2: Variables used in the specification of the logit models

Variable	Utility*	Description
Size96	R, L	Employment size of business establishment *i* in the year 1996
Size96^2	R, L	Employment size of business establishment *i* in year 1996 squared
Log(Age)	R, L	log of establishment's *i* age
Growth	R, L	employment growth rate defined as (Size97 - Size96)/ 0.5*(Size97 + Size96)
*Industry1***	S	1 if business establishment *i* belongs to SIC31, SIC33 or SIC39, 0 otherwise
Industry2	S	1 if business establishment *i* belongs to SIC401, 0 otherwise
Industry3	R	1 if business establishment *i* belongs to SIC421, SIC422, SIC423 or SIC426, 0 otherwise
Industry4	S	1 if business establishment *i* belongs to SIC456, 0 otherwise
Industry5	S	1 if business establishment *i* belongs to SIC484, 0 otherwise
Industry6	S	1 if business establishment *i* belongs to SIC52 or SIC56, 0 otherwise
Industry7	R	1 if business establishment *i* belongs to SIC63 or SIC65, 0 otherwise
Industry8	S	1 if business establishment *i* belongs to SIC70 – SIC74, 0 otherwise
Industry9	S	1 if business establishment *i* belongs to SIC75 – SIC76, 0 otherwise
Industry10	S	1 if business establishment *i* belongs to SIC77, 0 otherwise
Industry11	S	1 if business establishment *i* belongs to SIC85, 0 otherwise
Industry12	S	1 if business establishment *i* belongs to SIC90 – SIC92, 0 otherwise
Industry13	S	1 if business establishment *i* belongs to SIC96 – SIC99, 0 otherwise

* Utility: Stay (S), Relocate (R), Leave (L)

** See Table 4-3 for a definition of the SIC codes used in constructing the industry variables

Table 4-2 continued ...

Variable	Utility*	Description
Suburbs	R	1 if business establishment *i* is located in any of the outer municipalities that includes: Dundas, Ancaster, Stoney Creek, Flamborough or Glanbrook, 0 otherwise
Lower Hamilton	S	1 if business establishment *i* is located below the escarpment excluding the CBD, 0 otherwise
Upper Hamilton	S	1 if business establishment *i* is located above the escarpment, 0 otherwise
D_{od}	R	Euclidean distance between the 1996 location *o* of a relocating establishment *i* and its new 1997 location *d*
LC	R	Local competition defined as ratio of the total number of employees of establishments in the same k 2-digit industry of establishment *i* to the number of employees of establishment *i*. Total number of employees is measures from all business establishments of sector k within a 1500 meter radius around establishment *i*
Agglom	S	Total number of establishments in the same k 2-digit industry of establishment *i* within a 1500 meters radius around establishment *i*

* Utility: Stay (S), Relocate (R), Leave (L)

that as size increase the propensity of mobility will decrease. However, they also argue that this might not always be the case since small establishments tend to pay market rent, whereas medium sized and large establishments may be able to negotiate discounts for large sites or are subject to governmental incentives such as subsidies. This suggests that while large establishments may have higher relocation cost in absolute terms; this cost might be lower in relative terms. To control for this effect, we introduce the Size96 squared covariate in the model. The size variables are specified in the utility functions of the relocate and leave outcomes.

The log of the age (*Log(Age)*) variable is used to examine the hypothesis that older establishments have a higher inertia compared to young establishments. As noted by Maoh and Kanaroglou (2005a), young establishments are more susceptible to failure and face more stress compared to old establishments. Therefore, they are expected to have a higher mobility rate in an attempt to adjust to changes in the local market. However, the adjustment and the ability to penetrate the market are not commonplace among new establishments. Baldwin et al. (1997) assert that over half the new Canadian firms that failed in the first decade of their life fail within the first two years of operation. Older establishments, on the other hand, have overcame the market adjustment phase and already succeeded in penetrating the local market. As time passes, they are able to distinguish their products and services and thus gain control over the market. The variable *Log(Age)* is expected to retain a negative sign, when specified in the utility function of the relocate and leave outcomes.

Growth is introduced in the model to examine if the change in establishment size induces mobility. Dijk and Pellenbarg (2000) argue that the need to move to another location is related to the growth pattern of the firm over time. Van Wissen (2000) empirically quantifies this relation and finds that growth increases the probability of moving. A counter argument might be that growth increases the propensity of inertia. Here, growth can be a manifestation of success for small establishments that are able to grow and accommodate few employees without the need to relocate. Therefore, it is hard to determine a priori the effect of growth on mobility for SME establishments. Following Van Wissen (2000), we devise the growth variable as shown in Table 4-2 and introduce it in the utility function of the relocate and leave outcomes.

The final set of firm internal factors is presented by a number of sectoral dummies that reflect the type of industry at the 2nd and 3rd digit of the SIC codes as shown in Table 4-3. The dummies are introduced to control for the heterogeneity in the mobility behavior of establishments. They are also used to distinguish between the mobile and inert industries. Although the initial set of industry dummies included an exhaustive list based on the available SIC codes, we only present those that appeared to be significant in the final specification of the different models. Insignificant industry dummies were dropped from the final specification.

Location factors are introduced in the model to identify which parts of the city are more at risk of changing their location attractiveness. We divide the city into four

Table 4-3: 1980 Standard Industrial Classification (SIC) codes of establishments used in model specification

Code	Description	Mobility Status*
SIC31	Machinery industries (except electrical machinery)	Static
SIC33	Electrical and electronic products industries	Static
SIC 39	Other manufacturing industries	Static
SIC401	Residential building and development	Static
SIC421	Site Work industries	Dynamic
SIC422	Structural and related work industries	Dynamic
SIC423	Exterior close-in work industries	Dynamic
SIC426	Electrical work industries	Dynamic
SIC456	Truck transport industries	Static
SIC484	Postal and courier service industries	Static
SIC52	Food, beverage, drug and tobacco wholesale industries	Static
SIC56	Metal, hardware, plumbing, heating and building wholesale industries	Static
SIC63	Automotive vehicles, parts and accessories, sale and service retail industries	Dynamic
SIC65	Other retail store industries	Dynamic
SIC70 – SIC74	Finance Insurance service industries	Static
SIC75 – SIC76	Real Estate service industries	Static
SIC77	Business service industries	Static
SIC85	Education service industries	Static
SIC91 – SIC92	Accommodation food and beverage service industries	Static
SIC 96 – SIC99	Other service industries	Static

Note (*): Mobility Status is based on the results of the estimation we achieved in section 4.5

geographic areas that are categorized by the following dummies:(1) Central Business District (*CBD*), (2) land below the escarpment (*lower Hamilton*), (3) land above the escarpment (*upper Hamilton*) and (4) outer municipalities (*suburbs*). The CBD dummy, which equals 1 if the establishment is located in the CBD and 0 otherwise, is dropped from the final specification of the model since the *suburbs* variable is used instead. Both variables complement each other when specified individually in the utility function of the relocate outcome. However, the *suburbs* variable is deemed more appropriate for the specification of the model since it provides a better model fit.

Other things being equal, the variable *Agglom* is used to examine if the benefits attained from agglomeration economies increase the probability of inertia. The positive relation between agglomeration economies and the success of business establishments is well documented in the literature. Therefore, the clustering of business establishments is expected to reduce location stress and increase the propensity of staying at the same location. Our agglomeration variable is based on counting the number of establishments within 1500 meters from business establishment *i*. The count only includes establishments with the same 2-digit industrial sector as establishment *i*. We experimented with counts of establishments within different distance values that ranged between 500 and 3000 meters. Our analysis suggested that agglomeration based on 1500 meters radius produces the most significant results. Agglomeration is used as an alternative specific variable in the utility function of the stay outcome.

The composite variable (D_{od}/LC) is introduced in the model to test the hypothesis set forth by Bade (1983) regarding the relation between locational pressure and the propensity of mobility. He argued that mobility could only take place under high locational pressure that arises from locational deficiencies. In the urban context, local competition or the market power effect can be thought of as a major locational deficiency that will push the establishment to move long distance. The variable D_{od}/LC measures the additional distance of relocation for every unit increase in local competition. The variable is specified in the utility function of the relocate outcome and is expected to increase the propensity of relocation. Following Lerman and Liu (1984), local competition exerted on establishment i is defined as the ratio of employment of establishment's i rivals to the employment of establishment i. Rivals are establishments with the same 2-digit industry as establishment i that are located within a radius of 1500 meters from establishment i. The value of 1500 meters is based on experimenting with different values as in the case of the agglomeration variable.

4.5 Estimation Results

We estimated equation 2 to model establishment mobility by economic sector as shown in Table 4-4. The models were estimated using Version 3 of the NLOGIT econometric software. The significance of the alternative specific constant used in the definition of the utility function of the relocate outcome suggests that factors other than

Table 4-4: Estimation results of multinomial logit models

		Manufacturing		Construction		Communication & Transportation		Wholesale		Retail		Services	
Variable	Utility*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*
Constant	R	1.378	3.115	2.092	4.847	1.909	2.581	1.642	3.080	1.481	4.018	2.083	13.949
Size96	R, L	-0.081	-2.430	-0.572	-5.465	-0.331	-1.999	-0.332	-3.529	-0.216	-5.115	-0.041	-3.975
*Size96²****	R, L	0.343	1.709	7.250	3.939	1.342	0.378	2.874	2.462	2.207	4.061	0.162	3.848
Log(Age)	R, L	-1.488	-7.856	-1.086	-6.976	-0.943	-2.715	-0.818	-3.720	-1.188	-8.374	-0.887	-14.903
*Growth***	R, L	0.778	1.823	-0.111	-0.304	-2.146	-1.124	-0.829	-1.662	-0.431	-1.426	-0.080	-0.629
Industry1	S	2.174	3.397										
Industry2	S			1.278	3.283								
Industry3	R			0.781	2.225								
Industry4	S					1.5464	2.702						
Industry5	S					1.6774	1.704						
Industry6	S							2.028	2.878				
Industry7	R									1.143	3.411		
Industry8	S											1.936	7.364
Industry9	S											1.851	8.594
Industry10	S											1.168	8.109
Industry11	S											2.704	4.713
Industry12	S											1.070	6.682
Industry13	S											1.911	12.952

* Utility: Stay (S), Relocate (R), Leave (L)
** Growth in the manufacturing model is specified for only the relocate utility
*** Values of parameters are in thousands

Table 4-4 continued…

		Manufacturing		Construction		Communication & Transportation		Wholesale		Retail		Services	
Variable	Utility*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*	*Beta*	*t-stats*
Suburbs	R	-1.767	-3.449	-3.515	-6.663	-3.019	-3.034	-2.053	-3.347	-4.462	-7.154	-2.398	-12.362
Lower Hamilton	S	2.147	3.605	1.155	2.704	1.084	1.432	1.429	2.686	2.459	6.950	1.846	11.448
Upper Hamilton	S	1.821	3.534	1.810	4.264	4.107	2.322	3.508	3.238	2.794	6.882	2.047	11.677
D_{od}/LC***	R	0.875	4.474	18.385	7.591	20.842	4.286	3.308	4.204	44.522	8.381	1.739	7.683
Agglom	S	0.025	1.525	0.035	3.903	0.066	1.503	0.071	2.575	0.037	5.710	0.008	11.744
No. of Obs.		541		1057		254		440		1188		3990	
L(0)		-594.344		-1161.233		-279.048		-483.389		-1305.151		-4383.463	
L(β)		-140.656		-226.021		-52.902		-113.964		-296.177		-1338.486	
ρ^2		0.763		0.805		0.810		0.764		0.773		0.695	

* Utility: Stay (S), Relocate (R), Leave (L)
** Growth in the manufacturing model is specified for only the relocate utility
*** Values of parameters are in thousands

those specified in the model are important in explaining the process of establishment relocation. Such unobserved factors could be related to the cost of the location, condition of the current premises or the level of accessibility to the site. The constant could be also capturing latent information that relates to the behavior of the entrepreneur. Similar attempt to control for unobserved factors in the leave outcome did not result in a significant constant. This is believed to be the case since the number of out-migrating establishments is much lower than the number of relocating establishments. Therefore, relocating establishments account for most of the variability observed in the models.

The size parameter in the different models suggests that as the size of the establishment increases the propensity of moving decreases. On the other hand, the size-squared parameter implies a non-linear relationship between size and mobility. However, we do not believe that this suggests a higher mobility rates for larger establishments. We believe the non-linear relationship is due to the fact that the number of establishments decreases very rapidly as establishment size increases. As a result, the slope on size-squared is mainly influenced by the decrease rate of mobility from extremely small (size less than 10 employees) to small establishments (size 11 to 50), rather than by an actual increase in mobility rates of large establishments.

The models suggest that mobility decreases with age. Van Wissen (2000) and Brouwer et al. (2004) reported similar results about the effect of age on mobility. In our view, the result from the age parameter is a reflection of the life cycle of the establishment. We believe that young establishments perceive mobility as a strategy that will enable them to acquire a niche in the local market. Therefore, relocation can be thought of as part of the gambling that the establishment is willing to endure in order to penetrate the market. On the other hand, old establishments already retain market power and exhibit the skill and knowledge to adapt to market changes at a much faster pace. Therefore, old establishments are less likely to relocate or out-migrate, other things being equal.

The parameters for the variable *Growth* are significant in the manufacturing and wholesale trade models and marginally significant in the retail trade model. Growth increases the probability of relocation for manufacturing establishments in Hamilton. This is the case since growth of manufacturing establishments might be associated with an expansion in floorspace to accommodate new machinery and labor. As a result, the establishment is more likely to move in search for larger floorspace that could accommodate its new requirements. The results from the wholesale and retail trade models are the opposite of the manufacturing case. The negative parameter suggests an inverse relationship between growth and the propensity to move (relocate or out-migrate). This could be the case since these types of establishments are not heavily dependent on machinery as in the case of manufacturing. Therefore, growth could be a reflection of success at the current location. In other words, the business is more likely to remain in situ if the current location is attractive and can allow the establishment to increase profit and grow. This is particularly true for small size retail trade establishments who can afford to have one or more employees on board without the need to expand floorspace.

The parameters associated with the economic sector dummies show a variation in mobility among establishments from different industries. Our results suggest, other things being equal, that certain construction and retail trade establishments have a high tendency of relocation compared to other establishments. In the case of construction establishments, relocation is more prominent among the following industries: (1) site work industries (SIC 421), (2) structural and related work industries (SIC 422), (3) exterior close-in work industries (SIC 423) and (4) electrical work industries (SIC 426). This could be the case due to nature of these businesses, which require them to be in close proximity to construction sites. Relocation is more prominent between two retail trade industries that include: (1) automotive vehicles, parts and accessories, sale and service retail industries (SIC 63) and (2) other retail store industries (SIC 65). We believe that the dynamics of these establishments is fueled by the need to expand or to adjust to changes in market orientation. For instance, the need to be close to population may push these establishments to relocate near new residential development in the city. The estimation results also point to the fact that establishment inertia is commonplace among many types of industries. This could be found in the case of certain manufacturing, construction, transportation, communication, wholesale trade and service establishments. Table 4 lists the different industries that have static and dynamic mobility status in Hamilton as depicted from the estimated models.

The results from the location dummies suggest decentralization and suburbanization of establishments in Hamilton. The observed trends are in line with earlier research conducted by Bourne (1989), which confirmed the existence of decentralization in Hamilton among other Canadian cities. Our initial attempt to specify the *CBD* dummy in the utility function of the relocate outcome provided clear evidence on decentralization in all the modeled sectors. This is confirmed by the *Suburbs* covariate, which also suggests suburbanization of establishments in the city. While the *Suburbs* and the *CBD* covariates complement each other, the former is used in the final specification of the models because it provides a better fit. Suburbanization could also be discerned from the significance of the *Lower Hamilton* and *Upper Hamilton* covariates. The results from the location parameters confirm that mobility is more pronounced among the Central Business District (CBD) establishments. We think that these trends are due in part to the change in the real estate market and the way land development practices have been emerging in the city. The high costs of rent and parking in the CBD combined with low demand for goods and services due to low-income residents will continue to push establishments out from the core of the city. On the other hand, new developed land and the availability of free parking in the suburban areas will pull the relocating establishments towards to the edge of the city.

With regard to the *Agglom* variable we find that industrial clustering increases the propensity of inertia. This relation is less evident for manufacturing, communication and transportation establishments although the agglomeration parameter for these sectors retains the correct sign. The results also imply that the effect of agglomeration economies is more significant in the case of retail and service establishments. These results meet our a priori expectation and highlight the market share effect that continues to manifest itself in the city. As for location pressure, the estimated models accept the hypothesis that the increase in local competition (location pressure) will push the establishment to move long distance. This result is in accordance with the neo-classical theory and imply that the change in the boundaries of the spatial margins to profitability due to local competition will increase the propensity of moving.

4.6 Conclusion

This paper investigates the underlining factors that drive the mobility behavior of business establishments in the city of Hamilton in Ontario, Canada. We specify and estimate multinomial logit models to study the mobility process of small and medium size establishments using micro data for 1996 and 1997. These data were extracted from Statistics Canada Business Register. The results from the logit models will be used as the basis for the development of a business mobility module in a firmographic microsimulation system. Such system will be used to simulate the interplay between the local economy and the urban form of Hamilton.

Our analysis suggests that firm internal factors such as age, size, growth and economic sector are very important determinants of mobility. While these factors do not translate into policy, the age and growth variables provide an insight on the relation between the life cycle of the establishment and its willingness to stay or move. On the other hand, the results from the models allow us to draw general conclusions about the relation between public policy and establishment mobility. For instance, the models succeed in emphasizing the decentralization and suburbanization trends in the city. These trends will have a direct impact on the emergence of urban form. The location factors clearly suggest a decline in the attractiveness of sites in the core of Hamilton. As a result, business establishments are moving towards the suburbs due to a number of factors that pull establishments from the core (Maoh and Kanaroglou, 2005b). We also find that geographical clustering based on the agglomeration variable discourage mobility. As a result, zoning and the way land uses are distributed in the city may impact the decision of staying or moving. Location stress as depicted by the local competition variable also appears to influence the mobility behavior of establishments. The variable suggests that the growth of rival establishments increases location pressure and will push the establishment to relocate.

Finally, we conclude that mobility factors can be linked to the neo-classical, behavioral and institutional theories highlighted in section 2. However, future research is still needed to thoroughly scrutinize the relation between public policy and establishment mobility behavior in the urban context if data becomes available. Future research should include some policy related factors that we believe can influence the mobility decision of establishments. These include variables that account for tax difference over space and between municipalities and changes in rent values over space and time. Also, the design, collection and analysis of stated preference information on firm relocation as in Dijk and Pellenbarg (2000) will prove to be helpful in understanding the role of policy in the context of establishment mobility. Meanwhile, the BR remains the most reliable resource to conduct firm demographic research, albeit more effort is needed to enrich the information of the BR records.

4.7 References

Anderson, William, Pavlos Kanaroglou and Eric Miller. 1996. "Urban form, energy and the environment: A review of issues, evidence and Policy," *Urban Studies*, 33, 7-35.

Bade, Franz-Josef. 1983. "Locational behavior and the mobility of firms in West Germany," *Urban Studies*, 20, 279-297.

Baldwin, John, Tara Gray, Joanne Johnson and Jody Proctor 1997. *Failing concerns: business bankruptcy in Canada*. Ottawa: Statistics Canada.

Ben-Akiva, Moshe and Steven R. Lerman 1985. *Discrete choice analysis: theory and application to travel demand*. Cambridge: The MIT Press.

Bourne, Larry. 1989. "Are new urban forms emerging? Empirical tests for Canadian urban areas," *The Canadian Geographer*, 33, 312-328.

Brouwer, Aleid E., Ilaria Mariotti and Jos N. van Ommeren. 2004. "The firm relocation decision: an empirical investigation," *The annals of regional science*, 38, 335-347.

Dijk, Van and P Pellenbarg. 2000. "Firm relocation decision in the Netherlands: an ordered logit approach," *Papers in Regional Science*, 79, 191-219.

Kanaroglou, Pavlos and Robert South. 1999. "Can urban form affect transportation energy use and emissions? An analysis of potential growth patterns for the Hamilton Census Metropolitan Area," *Energy Study Reviews*, 9, 22-40.

Lerman, S and T Liu 1984. "Microlevel econometric analysis of retail closure," in D. Pitfield (eds.) *Discrete choice models in regional science*, London: Pion Limited, pp. 181-201.

Maoh, Hanna and Pavlos Kanaroglou. 2004. "Economic clustering and urban form: the case of Hamilton, Ontario," Working Paper No. 3,Center for Spatial Analysis, School of Geography and Geology, McMaster University.

Maoh, Hanna, Pavlos Kanaroglou and Ronald Buliung. 2005. "Modeling the Location of Firms within an Integrated Transport and Land-use Model for Hamilton, Ontario," Working Paper No. 6,Center for Spatial Analysis, School of Geography and Geology, McMaster University.

Maoh, Hanna and Pavlos Kanaroglou. 2005a. "A discrete-time hazard duration model of survival for small and medium business establishments with an application to Hamilton, Ontario," *Unpublished manuscript*, Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University,

Maoh, Hanna and Pavlos Kanaroglou. 2005b. "The Location of Business Establishments in the City of Hamilton, Canada: A Micro-Analytical Model Approach," *Unpublished manuscript*, Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University,

McFadden, Daniel 1974. "Conditional logit analysis of qualitative choice behavior," in P. Zarembka (eds.) *Frontiers in Econometrics*, New York: Academic Press, pp. 105-142.

Miller, Eric, John Hunt, John Abraham and Paul Salvini. 2004. "Microsimulating urban systems," *Computers, environment and urban systems*, 28, 9-44.

Wegener, Micheal and Klaus Spiekermann 1996. "The potential of microsimulation for urban models," in G. P. Clarke (eds.) *Microsimulation for urban and regional policy analysis*, London: Pion, pp. 149-163.

Van Wissen, Leo 2000. "A micro-simulation model of firms: applications of concepts of the demography of the firm," *Papers in Regional Science*, 79, 111-134.

4.8 Appendix

Figure 4-3: Nested structures of the business establishment mobility problem

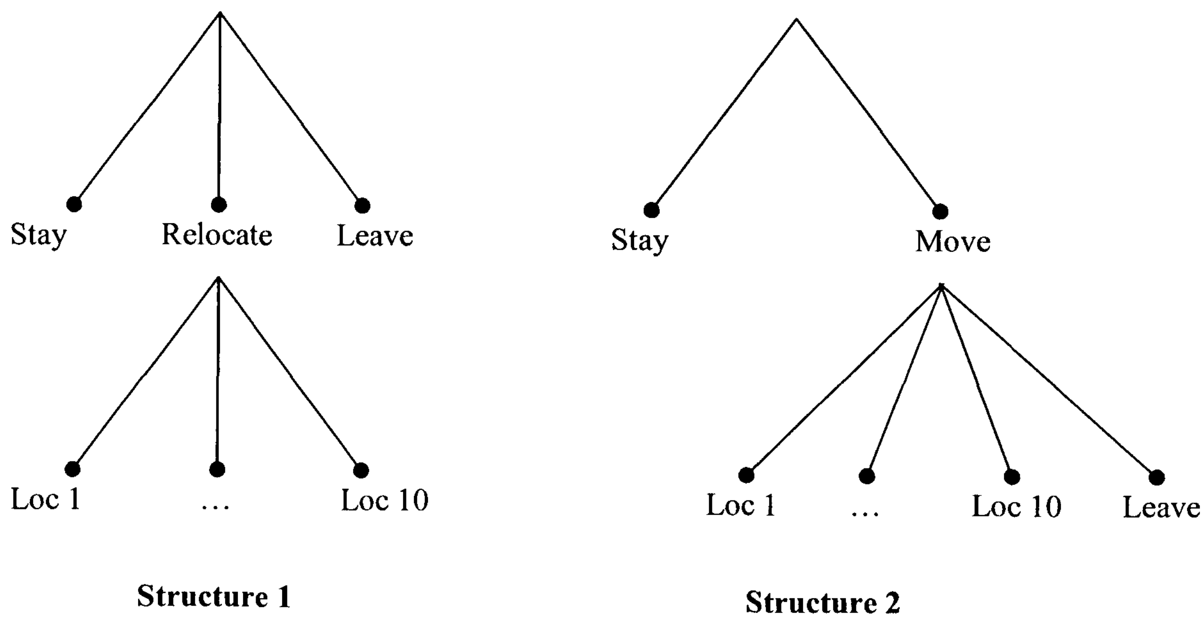
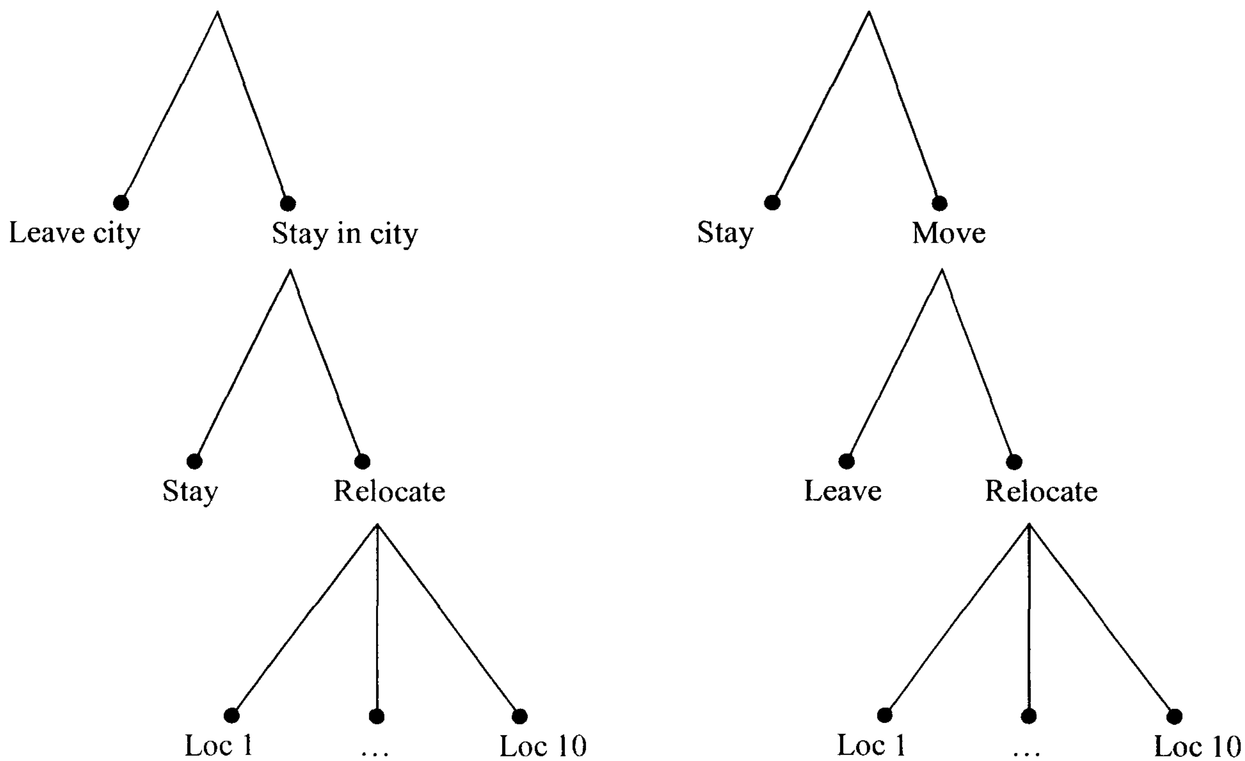


Figure 4-3 continued …



Structure 3 **Structure 4**

CHAPTER 5: The Location of Business Establishments in the City of Hamilton, Canada: A Micro-Analytical Model Approach

5.1 Introduction

Over the past three decades, a growing body of research has focused on studying sustainable cities by assessing the inter-relation between urban land use and transportation systems. Computer simulation programs, known as Integrated Urban Models (IUMs), were developed for many cities around the globe to address the issue (Southworth, 1995; Miller et al. 1998; Kanaroglou and Scott, 2002; Timmermans, 2003; Wegener, 2004). IUMs are planning tools used to evaluate the long-term repercussions of proposed land use and transport network changes. These include the differential estimation of land development, location decision of households and firms and the trips generated between the different parts of the city. These trips are then used to estimate congestion and associated tailpipe emissions in all the links of the urban transport network. The results achieved from simulating the urban system could enable decision makers to carry on, rectify or abandon certain aspects of their infrastructure projects.

Research to develop and improve the performance of IUMs received more attention in recent years in response to the environmental concerns raised by the different governments on the green house gas emissions problem (Dieleman et al., 2002). The research community has recognized the drawbacks of existing IUMs that make them ineffective for long-term policy planning. These include: (1) the aggregate nature of models with traffic zone or census tract as the unit of analysis; (2) their static nature and lack of behavioral realism due to large time lags and (3) the weak response to long-term planning policies due to their low level of policy sensitivity (Miller et al., 2004).

The proposed concept to rectify these drawbacks calls for adopting a micro-analytical approach to handle complex urban processes. Within such a framework, urban development is driven by the micro behavior of agents (such as persons, households, business establishments, developers) that interact in the city. Accordingly, a micro-analytical theory of urban development has to decompose the processes of urban change into subprocesses to identify the main agents and their decision behavior (Wegener and Spiekermann, 1996). The dynamic behavior within each process and for each individual agent is then explicitly simulated over both time and space to generate aggregate system behavior (Miller et al., 2004).

Research initiatives are underway to improve the land use submodel of IMULATE; an operational IUMs that has been developed for the City of Hamilton since 1994. IMULATE was developed along the theoretical guidelines of most existing IUMs and has been used to study urban problems in the city of Hamilton (Anderson et al., 1996; Scott et al., 1997; Kanaroglou and South, 1999). Towards developing a micro-analytic model, in this paper we focus on modeling the changing location of firms in the city of Hamilton. Adopting the concept of firm demography we study processes that relate to the establishment of new firms and the growth, decline, failure, or migration of existing firms. Firm demography, a bottom-up approach, is ideal for the development of agent-based models. It makes use of individual businesses as the units of analysis to explain the dynamic changes that take place within the population of business establishments in a geographic area over time. Miller et al. (2004) note that attempts to develop agent-based land use and transportation models have been rare to date. Furthermore, we note the lack of studies that have used firm demography for projections in order to assess land use planning policies. With the exception of the work by Waddell et al. (2003), most of the existing land use models are based on aggregate employment location models that predict the change in employment at the zonal level.

This paper reports on the research efforts we undertook to develop a location choice model for an agent-based firm demographic model for Hamilton. We estimate a number of econometric models to explain the location behavior of single business establishments with less than 200 employees in the city of Hamilton. Information about these business establishments was extracted from the Statistics Canada Business Register. The modeling framework divides the city into grid cells of 200 by 200 meters and predicts the annual location choice of each establishment at this fine spatial resolution. Land use, population census and business establishment data were geo-coded to grid cells of the developed land to create the explanatory variables of the econometric models.

Following Lee (1982) and Martinez (1992), each establishment is assumed to face the decision of selecting a location that will maximize its profit. Accordingly, discrete choice theory is applied and a number of multinomial logit models by economic sector are specified and estimated for the periods 1996 - 1997 and 2001 – 2002. The estimation of the models for the second period is to evaluate if the explained location behavior has changed over time. This is a neglected aspect in the literature of econometric models that are used as the basis for simulation models probably due to the lack of temporal data. This paper offers a novel framework for utilizing different sources of datasets to study the location behavior of business establishments at the micro level. It also provides new empirical evidence of the factors influencing the location behavior of individual business establishments in the city of Hamilton.

The remainder of this paper is organized as follows. Section two provides theoretical background on firm location behavior in the urban context and highlights the most recent research that has addressed this issue. Section three describes the city of Hamilton, which forms the geographical domain of our study. It also highlights the different data sources used in the empirical research. Section four discusses the methods of analysis used to formulate and specify the location choice models while section five provides a discussion about the estimation results of these models. The last section provides a conclusion to our study.

5.2 Intrametropolitan Firm Location Factors

Employment location modeling has made remarkable advancements since the work of Alonso (1964) who emphasized the role of bid-rent theory in explaining the location of residential land use activities in the city. In general terms, bid-rent theory treats land use activities as consumers of land, thus the demand for land varies with price subject to budget constraints. This implies that land use activities maximize their location utility by trading off transportation cost for land consumption. In the case of a monocentric city, land use activities out-bid each other for a central location. This is particularly true for commercial and industrial activities where firms out-bid each other for a central site in order to maximize their utility by means of maximizing their profit (Mirron, 1982).

While the theory of location choice behavior was propounded in the early 1960s, empirical models were not developed until the early 1980s. The theoretical advances achieved by McFadden (1975, 1978) in discrete choice models and the success in applying them to residential location problems (Ellickson, 1981) has paved the road for empirical research on industrial firm location. Lee (1982) was a pioneer in applying discrete choice theory to model the intraurban employment location in the City of Bogota. Using a multinomial logit model, Lee estimates the probability that a firm of a particular type will occupy a given site in the urban area. His logit model is formulated under the assumption that firms out-bid each other for a particular site, which is assigned to the highest bidder. Later efforts diverted from the concept of bid-rent theory and modeled the location decision of firms as a process driven by profit maximization (Hansen, 1987; Shukla and Waddell, 1991). Within this framework, urban space is partitioned into a set of discrete zones to form the location choice set. The firm is then modeled as making the choice of the zone that will maximize its profit. Martinez (1992) notes that the idea of profit maximization via the maximum utility framework was perceived as a practical alternative of bid-rent theory to model firm location in contemporary cities.

However, the work of Martinez (1992) changed this perception by providing a theoretical basis for using discrete choice theory in the firm location choice problem. He showed that at equilibrium, a logit probability, like the one formulated by Lee (1982) under the assumptions of bid-rent theory, is the same as the logit probability that results under the assumption of profit maximization. The conclusion is that bid-rent and profit maximization theories are equivalent and complement each other. A review of the recent firm location literature suggests a number of studies that applied utility theory to model the location decision of firms. The utility in these studies has been usually specified as a composite function of both the characteristics of the chosen site and the attributes of the firm.

Among the site characteristics that are used to explain the location of firms, the importance of the Central Business District (CBD) is well documented (see for example Vahaly, 1976; Clapp 1980; Erickson and Wasylenko, 1981; Lee, 1982; Ihlanfeldt and Raper, 1990; Shukla and Waddell, 1991; Waddell and Shukla, 1993; Coffey et al., 1996; Yarish, 1998; Wu, 1999). The effect of the CBD in the urban context has been used to test the centralization or decentralization of firms in urban areas. While some firms favor locating centrally, others tend to locate on land away from the CBD.

There is clear evidence in the literature on the direct relationship between land use and transportation systems (Hanson and Giuliano, 2004). The construction of highways and expressways over the past 6 decades has reshaped the form of many cities and led to decentralization and suburbanization. Simmons et al. (1998) point out that decentralization of the commercial market is accompanied by decentralization of commercial activities along arterials such as major roads, highways and expressways. The effect of highway and expressway proximity on firm location has been tested in previous studies. Coffey et al. (1996) argue that accessibility by automobile to firms in inner and outer Montreal suburbs represents an important pull locational factor that influenced the location of high services firms. Erickson and Wasylenko (1981), Vahaley (1976), Shukla and Waddell (1991), Waddell and Shukla (1993), Ihlanfeldt and Raper (1990) and Yarish (1998) report similar results when testing the influence of highways and expressways on the location of firms in metropolitan areas.

While trends of decentralization continue to shape metropolitan areas (Bourne 1991; Hanson and Giuliano, 2004; Wegener 1986; Lee 1985; Simmons et al., 1998 and Anderson et al., 1996), some land use features other than the CBD appear to influence the location of firms in cities. The study conducted by Jones et al. (1998) on the commercial structure in the Dallas Forth Worth CMSA suggested that commercial firms tend to cluster in box stores of 1000 - 5000 sq. m around regional malls. Simmons et al. (1998) observed the same trends for the commercial structure of metro Toronto. The empirical results achieved by Clapp (1980), Ihlanfeldt and Raper (1990) and Yarish (1998) highlighted the influence of regional malls on a firm's location decision. Moreover, the studies of Yarish (1998) and Shukla and Waddell (1991) indicated that airport proximity influence the location decision of firms.

Firms, other things being equal, seek to locate on lower price land. While such a factor has been dismissed by some studies, others highlighted the importance of land prices in the location decision of firms. Hansen (1987) found that land prices do not add any explanatory power to his model of industrial location choice. Coffey et al. (1996) report that land prices rank as the first factor affecting the location of firms in the CBD of Montreal, and as the third factor at the CMA level. Erickson and Wasylenko (1980) used the CBD proximity as a surrogate to capture the effect of land prices on the location decision of firms in the municipalities that form the Milwaukee region. Yarish (1998) and Clapp (1980) used rental rate in their models to test the effect of land prices on the intrametropolitan office location in Metro Toronto and Los Angeles, respectively. Clapp's findings overlap with what is suggested by Erickson and Wasylenko (1980), regarding the relation between land prices and the CBD. They suggest that buildings farther away from the CBD commanded less rent, reflecting the dropping land prices away from the Centre.

Agglomeration economies are at the heart of the urban processes and their use in firm location modeling is essential. The effect of agglomeration economies has been tested by many researchers (see for example Erickson and Wasylenko, 1980; Ihlanfeldt and Raper, 1990; Shukla and Waddell, 1991; Waddell and Shukla, 1993; Yarish, 1998; Wu, 1999; Waddell et al., 2000; Frenkel, 2001). All studies confirmed the importance of agglomeration economies in firm location and the benefits firms achieve from clustering and localizing in space. The urban economic literature refers to two types of agglomeration economies that influence the locational decision of firms in urban areas. These are localization economies and inter-industry linkages. Localization economies represent positive externalities from locating in close proximity to firms of the same industry and may be due to shared labor pool, site amenities or the benefits gained from knowledge spill over. Inter-industry linkages refer to agglomeration economies associated with locating at a site that has greater access to business in strategically related but different industries (Maoh and Kanaroglou, 2004).

Land use zoning is also a factor that influences firm location. Wu (1999) shows that foreign direct investment (FDI) firms are more likely to locate in the Economic and Technology development zone (ETDZ). The prestige level of a given zone can also be a determining factor in firm location. Wu (1999) found that FDI firms are attracted to areas that are in close proximity to the prestigious Garden Hotel or the Guangzhou international convention center. Similarly, Frenkel (2001) shows that for high technology firms prestige of a region ranks fifth among 14 location factors. Other location factors used in the literature include population size and income. Shukla and Waddell (1991) found that all types of firms show a tendency towards locating in populated areas. Jones et al. (1998), Ihlanfeldt and Raper (1990), Hansen (1987), Yarish (1998) and Shukla and Waddell (1991) found that population income and the spatial variation in population income affect where firms locate in a city.

Industry and firm characteristics such as firm size, age and organizational structure influence the location decision. Waddell and Shukla (1993) employed such variables to study the location decision of manufacturing firms in Dallas-Fort Worth area. Their findings suggest that large and medium firms tend to locate in some subcenter as opposed to the CBD. Frenkel (2001) used characteristics like firm size, age, date of establishment and type (multiplants or otherwise) to test how hi-tech firms behave when choosing a site. His findings suggest that the centre is attractive to small plants with high technological ability. Also, 58% of the multiplant firms are located outside the metropolitan area, of these, 83% are new plants established after 1980. The findings also show that 69% of the multiplant firms, located within the metropolitan area, are older plants established before 1980.

5.3 Study Area And Data

5.3.1 Study Area

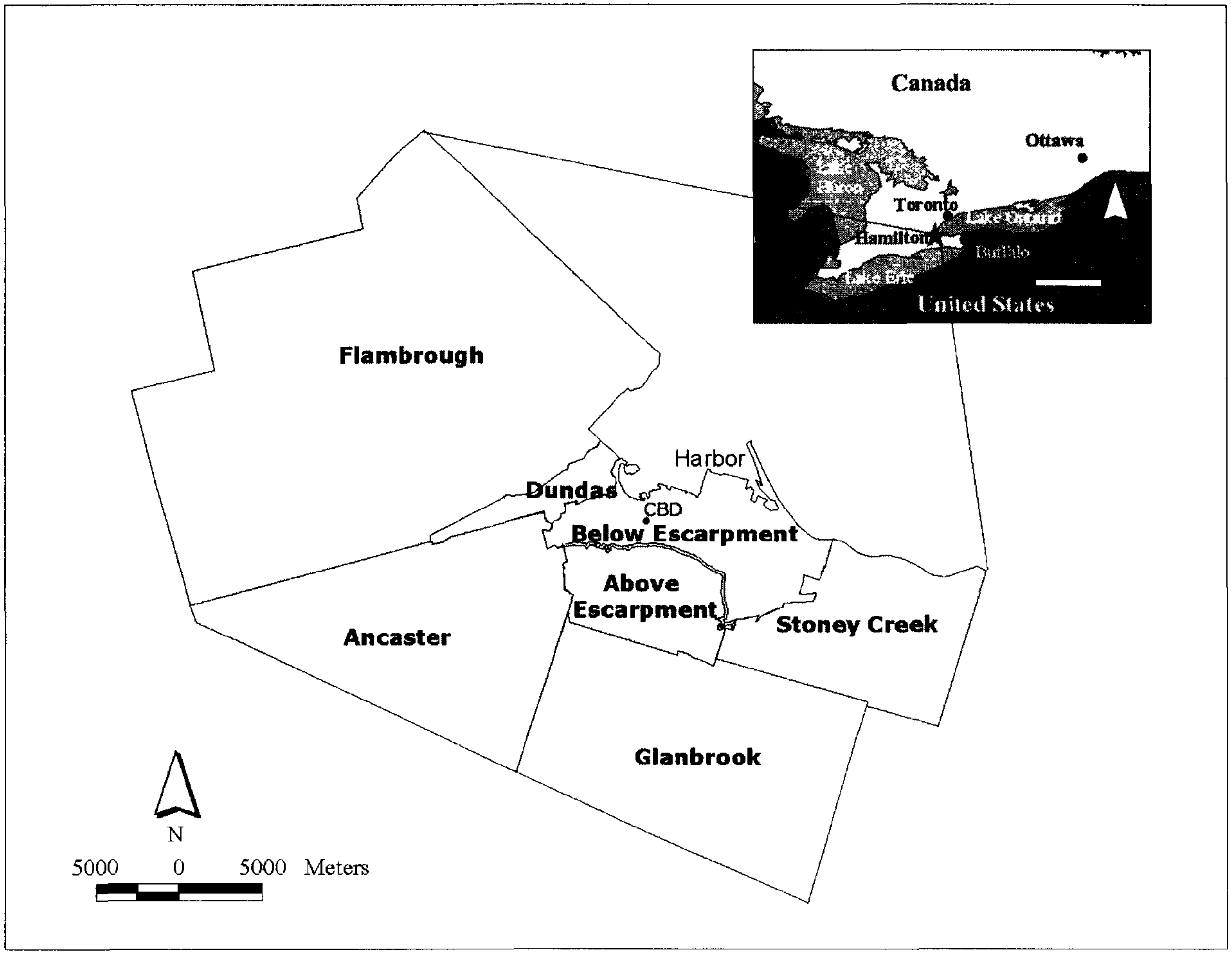
We specify and estimate location choice models for the city of Hamilton in Ontario Canada. The city is located south west of the Greater Toronto. The topography is shaped by the Niagara escarpment, which separates the city's geography into a lower coastal area between the escarpment and the harbor to the North and a highland beyond the escarpment to the south. While early urban development took place on the lower city, the continuous population and economic growth over time led to urban development above the escarpment. Recently, the city expanded its boundary and annexed the municipalities of Ancaster, Dundas, Flamborough, Glanbrook and Stoney Creek as shown in Figure 3-1.

The total population in Hamilton has increased from 465,153 in the year 1996 to 504,513 in 2001 (Statistics Canada, 1996 and 2001). Like most North American cities, Hamilton has an extensive transportation network with a number of highways and expressways that provide accessibility to the suburbs. Urban sprawl is one of the features shaping the form of the city (Kanaroglou and South, 1999). In recent years, continuous land development-projects in the suburbs have encouraged residential, commercial and industrial sprawl and have led to a clear decentralization of employment and population (Maoh and Kanaroglou, 2004). As a result, location of commercial and industrial business establishments have been influenced by the various development regimes implemented in the city. Despite being one of the major manufacturing cities in Canada, Hamilton's retail trade and service sectors have noticeably grown over the past two decades. Many of these retailers and service firms emerged in the various regional malls in the city's core and suburbs.

5.3.2 Data Sources

The empirical analysis is based on three sources of data. These are: (1) Statistics Canada Business Register (BR), (2) Statistics Canada Census data and (3) Desktop Mapping Technologies Incorporation (DMTI) GIS data. The BR was used to extract annual information on individual business establishments for the years 1996 –

Figure 5-1: The city of Hamilton in the regional context



1997 and 2001 – 2002. The extracted data included small and medium size (SME) establishments with less than 200 employees. Larger establishments with more than 199 employees were excluded from the analysis since they have a small number and their location behavior is predictable given the high monetary value associated with their infrastructure. SME establishments, on the other hand, account for the majority (99%) of business establishment population in Hamilton. They are also more footloose and tend to relocate more frequently due to the lower cost required for their infrastructure. Consequently, the devised agent-based model will be developed to capture the dynamics of the location choice behavior driven by SME establishments whereas the location choice behavior of large establishments will be handled exogenously.

The extracted SME data was constrained to the self-owned non-incorporated (NIP) establishments. This restriction was imposed since annual information is required and is only available for this group of business establishments in the BR. General statistics from exploring the BR data in Table 5-1 indicates that NIP establishments accounted for 86% and 84% of the SME establishments in the years 1996 and 2002, respectively. The exploration also suggests that 79% and 77% of SME establishments in 1996 and 2002 are small with employment size less than or equal to 10. Accordingly, the NIP population is deemed appropriate for our objectives since it accounts for the majority of the SME establishments as shown in Table 5-1. However, the absence of the incorporated (IP) establishments from the extracted sample will undermine our understanding of what role organizational structure has on the location choice behavior of establishments. Unlike the IP establishments, NIP establishments are simple in their organizational structure and each establishment corresponds to a unique firm in the city.

Table 5-1: Distribution of IP and NIP business establishments in Hamilton by size for the years 1996 and 2002

	1996			2002		
Size	**IP (%)**	**NIP (%)**	**Total (%)**	**IP (%)**	**NIP (%)**	**Total (%)**
1 – 10	7	93	100	9	91	100
11 – 20	28	72	100	27	73	100
21 – 50	50	50	100	43	57	100
51 – 100	71	29	100	62	38	100
101 – 199	80	20	100	80	20	100
1 – 199	**14**	**86**	**100**	**16**	**84**	**100**

Source: Special tabulation from the Business Register, Statistics Canada (2004)

The extraction procedure of the firm data from the BR involved identifying firms that relocated between 1996–1997 and 2001–2002, respectively. It also entailed identifying newly formed and in-migrating firms in 1997 and 2002. The process of identifying the relocating firms was achieved in two steps. The first step involved determining which firms continued to exist in Hamilton between 1996–1997 and 2001–2002, respectively. This was achieved with the help of the establishment number (EN) and the Standard Geographic Code (SGC) representing the city of Hamilton. The EN is a unique identifier that is generated when the establishment is created in the BR. The firm is considered a continuer if its EN number existed in the two consecutive years in the BR with the Hamilton SGC. The second step was to compare the postal code addresses and coordinates for the continuer firms. A relocation event is coded if the firm has a different postal code address and/or coordinates between the two consecutive years.

Newly established and in-migrating firms were also identified using the EN number and the Hamilton SGC. A firm in the 1997 or 2002 population was considered new if its EN did not exist in the 1996 or 2001 Canada's establishment list, respectively. On the other hand, a firm was considered an in-migrant in 1997 or 2002 if its EN existed in the 1996 or 2001 establishment list with a SGC different from the one for Hamilton.

Census information was obtained from Statistics Canada census population surveys for the years 1996 and 2001. The information for 1996 and 2001 was obtained at the enumeration area level and dissemination area level, respectively. The information included total population, newly developed dwellings in different time periods, total number of households and average household income in 1996 and 2001. Land use information, on the other hand, was obtained from the 1999 land use cover extracted from the (DMTI) GIS data for the city of Hamilton. We also made use of the DMTI road network to identify major roads, highways and expressways in the city.

5.4 Location Choice Modeling

5.4.1 Model Formulation

The location choice models we seek to formulate will be used to simulate the probability of choosing a particular developed site in the city as a place for the firm to be established. To account for the heterogeneity in firm population we develop different models by economic sector. For each model, relocating, newly established and in-migrating firms are grouped into one pool. Each firm is assumed to be searching for an urban location that will maximize its profit. Therefore, firms will out-bid each other for a particular site, which is assigned in the end to the highest bidder. Following the work of Lee (1982) and Martinez (1992), the process of bidding for a location in the city could be modeled using the concept of utility maximization.

While our first intention was to use the delineation of developed land parcels to present the set of location alternatives, the idea was dismissed due to the nature of geographical coordinates used with the BR data. These coordinates are based on the postal code addresses, which have a one-to-many relation with the land parcels. It was, therefore, impossible to geo-code a firm to the land parcel it belongs to. We opted for dividing the developed land available to firms into grid cells of 200 by 200 meters. This spatial resolution is deemed sufficient for an agent-based model (Waddell *et al*, 2003). The result is a total of 2635 and 2855 grid cells that form the sets of location alternatives for the periods 1996-1997 and 2001-2002, respectively. The developed land was determined by intersecting the grid cells with the postal code addresses of all the establishments that existed in Hamilton between 1990 to 1997 and 1990 to 2002 for the two periods of study, respectively. Land use, population census and firm data were geo-coded to the generated grid cells and each firm in the searching pool was assigned to one location represented by a grid cell.

Assuming that the grid cells form the set of location alternatives, the probability of selecting grid cell i as the new location of a firm can be estimated using a multinomial logit model:

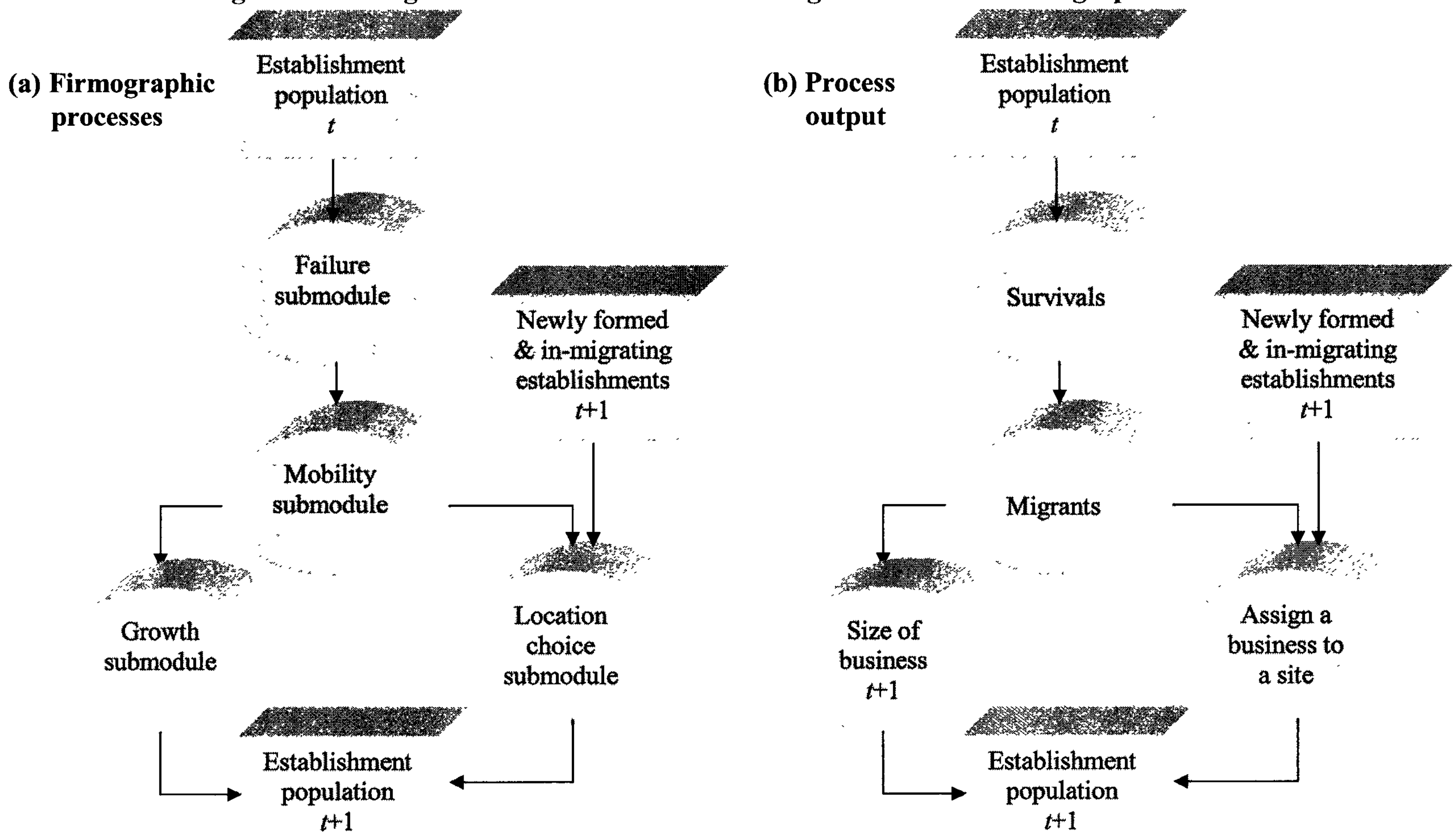
$$P_n(i) = \frac{\exp(V_{ni})}{\sum_{j=1}^{J_n} \exp(V_{nj})}$$

Here V_{ni} is a linear-in-parameter function that depends on covariates characterizing firm n and the attributes of site i. Following McFadden (1978), the estimation of the logit model is based on a random sample of alternatives that is drawn from the choice set of grid cells. We use a uniform distribution to randomly sample a set of nine grid cells in addition to the chosen cell. The location choice problem is therefore modeled among 10 alternative locations. While the method overcomes the problem of dealing with a large choice set, it does not allow us to obtain estimates for the alternative specific constants in the logit model. However, the method produces consistent coefficient estimates that can be used to understand the location behavior of firms in the urban context.

5.4.2 Simulating Firm Location Choice

The formulated location choice model will be an integral part of the developed agent-based firmographic model for the city of Hamilton shown in Figure 5-2. For each simulation year, the search pool is created from grouping the relocating with the newly established and in-migrating firms. While newly established and in-migrating totals are determined exogenously, relocating firms are identified from processing the base year

Figure 5-2: The general structure of Hamilton's agent-based firm demographic model



establishment population via the *failure* and *mobility* submodels as shown in Figure 5-2. For each firm in the search pool, a sample of ten locations is randomly selected from the full set of alternative locations. The location choice probability is then calculated for the ten locations using the logit model that corresponds to the industry in which the firm belongs. Upon calculating the location choice probabilities, Monte Carlo simulation is used to select the specific alternative location to be assigned to the firm.

5.4.3 Model Specification

The specification of the models involved using information from the BR, the population census and GIS layers to create variables at the grid cell level. GIS tools were employed to create a number of spatial measures to reflect the variables listed in Table 5-2. These were devised based on the information we gathered from the intrametropolitan firm location literature and the exploration of economic clustering and urban form in the city of Hamilton (Maoh and Kanaroglou, 2004). Distance to CBD is used to examine how centrality influences the location of firms in the city of Hamilton. The variable was calculated as the Euclidean distance between the core and the centroid of each grid cell. Proximity to highways and main roads is used to test what influence land accessibility has on the location behaviour of firms. The variable was based on intersecting grid cells with highways and main roads in the city. Developed land within 1500 meters from regional malls is specified since it is an attraction to certain types of firms.

TABLE 5-2: List of covariates used in the specification of the location choice models

Variable*	Definition
Location Specific Factors	
CBDPRO	Euclidian distance between the centroid of grid cell i and the centroid of the CBD
HWYMRPRO	1 if grid cell i intersects with a main road or a highway, 0 otherwise
MALLPRO	1 if grid cell i is within 1500 metes from a regional mall, 0 otherwise
AGGLOM1	Number of manufacturing business establishments within 500 meters from the center of grid cell i
AGGLOM2	Number of construction business establishments within 500 meters from the center of grid cell i
AGGLOM3	Number of communication business establishments within 500 meters from the center of grid cell i
AGGLOM4	Number of wholesale trade business establishments within 500 meters from the center of grid cell i
AGGLOM5	Number of retail trade business establishments within 500 meters from the center of grid cell i
AGGLOM6	Number of services business establishments within 500 meters from the center of grid cell i
MOUNTAIN	1 if grid cell i is located above the escarpment, 0 otherwise
SUBURBS	1 if grid cell i is located in the suburban municipalities of the city of Hamilton, 0 otherwise
*NEWDEVLOP***	Density of residential dwellings built in grid cell i between 1991 and 1996
*OLDDEVLOP***	Density of residential dwellings built in grid cell i between 1981 and 1990
POPDENS	1996 population density in grid cell i
HHLDINCDENS	Density of average household income in grid cell i
HHLDDENS	Density of number of households in grid cell i
AVGDWELLVAL	Weighted average dwelling value in grid cell i

* Values of covariates were based on the 1996 and 2001 census data for the 1996 – 1997 and 2001 – 2002 models, respectively.

** New development in the 2001 – 2002 models was based on the density of newly developed dwellings between 1996 and 2001. On the other hand, old development was based on the density of newly developed dwelling between 1981 and 1995.

Table 5-2 continued ...

Variable	Definition
Location Specific Factors	
LANDUSE1	Percentage of open space land use in grid cell i
LANDUSE2	Percentage of resource and industrial land use in grid cell i
LANDUSE3	Percentage of park and recreational land use in grid cell i
LANDUSE4	Percentage of residential land use in grid cell i
LANDUSE5	Percentage of commercial land use in grid cell i
LANDUSE6	Percentage of governmental land use in grid cell i
Firm Specific factors* **	
NEWBORN	1 if the firm is newly established in the year 1997, 0 otherwise
INDUSTRY1	1 if the firm belongs to SIC 28, 0 otherwise
INDUSTRY2	1 if the firm belongs to SIC 28 or SIC 39, 0 otherwise
INDUSTRY3	1 if the firm belongs to SIC 426, 0 otherwise
INDUSTRY4	1 if the firm belongs to SIC 401, 421, 422 or 427, 0 otherwise
INDUSTRY5	1 if the firm belongs to SIC 424, 0 otherwise
INDUSTRY6	1 if the firm belongs to SIC 45 – SIC 47, 0 otherwise
INDUSTRY7	1 if the firm belongs to SIC 48 – SIC 49, 0 otherwise
INDUSTRY8	1 if the firm belongs to SIC 52, 0 otherwise
INDUSTRY9	1 if the firm belongs to SIC 59, 0 otherwise
INDUSTRY10	1 if the firm belongs to SIC 60, 61 or 65, 0 otherwise
INDUSTRY11	1 if the firm belongs to SIC 61, 62 or 63, 0 otherwise
INDUSTRY12	1 if the firm belongs to SIC 64 or 65, 0 otherwise
INDUSTRY13	1 if the firm belongs to SIC 601, 633, 635 or 641, 0 otherwise

*** See Table 5-3 for details about the corresponding industry of each SIC number

Table 5-2 continued ...

Variable	Definition
INDUSTRY14	1 if the firm belongs to SIC 70 – SIC 74, SIC 77, SIC 91 – SIC 92 or SIC 96 – SIC 99, 0 otherwise
INDUSTRY15	1 if the firm belongs to SIC 96 – SIC 99, 0 otherwise
INDUSTRY16	1 if the firm belongs to SIC 75 – SIC 76, SIC 77, SIC91 – SIC92 or SIC 96 – SIC 99, 0 otherwise
INDUSTRY17	1 if the firm belongs to SIC 86, 0 otherwise
INDUSTRY18	1 if the firm belongs to SIC 91 – SIC 92, 0 otherwise

Table 5-3: 1980 Standard Industrial Classification (SIC) codes of establishments used in model specification

SIC code	Industry description
SIC 28	Manufacturing: Printing, publishing and allied; industries
SIC 39	Other manufacturing industries
SIC 401	Construction: Residential building and development industries
SIC 421	Construction: Site work industries
SIC 422	Construction: Structural and related work industries
SIC 424	Construction: Plumbing, heating and air conditioning industries
SIC 426	Construction: Electrical work industries
SIC 427	Construction: Interior and finishing work industries
SIC 45 – SIC 47	Transportation and storage industries
SIC 48 – SIC 49	Communication and other industries
SIC 52	Wholesale: Food, beverage, drug and tobacco industries
SIC 59	Wholesale: Other product industries
SIC 60	Retail: Food, beverage and drug industries
SIC 601	Retail: Food stores
SIC 61	Retail: Shoe, apparel, fabric and yarn industries
SIC 62	Retail: Household furniture, appliances and furnishings industries
SIC 63	Retail: Automotive vehicles, parts and accessories, sales and services
SIC 633	Retail: Gasoline service stations
SIC 635	Retail: Motor vehicle repair shops
SIC 64	Retail: General retail merchandizing industries
SIC 641	Retail: General merchandize stores
SIC 65	Retail: Other retail store industries
SIC 70 – SIC 74	Services: Finance insurance industries
SIC 75 – SIC 76	Services: Real estate industries
SIC 77	Services: Business services industries
SIC 86	Services: Health and social services industries
SIC 91 – SIC 92	Services: Accommodation food and beverages services industries
SIC 96 – SIC 99	Services: Other services industries

Agglomeration economies are accounted for using a spatial measure that counts the number of firms of a particular sector within a 500 meters radius from the center of a given grid cell. The value of 500 meters radius was selected after experimentation with values at 1500, 1000 and 500 meters. The chosen radius provided more plausible and statistically significant results. Land use in grid cells is used as a control for the type of land development. It is used to examine if land specialization due to zoning practices influences the location behavior in the city of Hamilton. Percentage of land use in a given grid cell was calculated from intersecting the grid cell layer with the land use GIS cover. The procedure provided the land use classification of each grid cell and the associated land use area in the grid cell.

Due to the aggregate nature of the census population, a disaggregation was required to obtain information at the grid cell level. Following Bracken and Martin (1989) we use cross-area interpolation. The method makes use of kernel estimation to transform a spatial population distribution from census centroid data into a continuous surface of population density. This surface is then superimposed on a grid to make it compatible with other GIS data. Using this method we generated a number of variables at the grid cell level that reflect the population density, household density, average household income density, average dwelling value density and the intensity of developed dwellings in the time periods 1981 – 1990, 1991 – 1996, 1991 – 1995 and 1996 – 2001. Finally, we derived a number of categorical variables as shown in Tables 5-2 and 5-3 to represent the industry where a firm belongs. These were interacted with other variables to account for firm heterogeneity and to examine if location factors vary among the different firms.

5.5 Estimation Results

Estimation results for logit models of six economic sectors in periods 1996–1997 and 2001–2002 are presented in Tables 5-4 and 5-5, respectively. The models were estimated using Version 3 of the NLOGIT econometric software. The figures in the tables are the estimated coefficients of the specified covariates with their associated *t-statistics* in parentheses. Specification of models in Table 5-5 was based on estimated models for the 1996–1997. The 2001–2002 models are estimated to examine if differences in location preferences is likely to emerge over time. Covariates in the 2001–2002 models were adapted using the 2001 census and 2001–2002 BR data. Since the coefficients are based on random sampling of alternatives, alternative specific constants were not specified in the models. The fit for all the estimated models was reasonable with rho-squared values ranging between 0.10845 and 0.21632. In the remainder of this section, we first explain the location behavior of firms in Hamilton on the basis of the results obtained for the 1996–1997 model. We then emphasize the similarities in the location behavior of firms for the two time periods.

The positive and significant parameters of the *CBDPRO* variable indicate that firms in general prefer to locate away from the Central Business District (CBD). This

Table 5-4: Parameter estimates of location choice models, 1996 – 1997

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
CBDPRO	0.000110	0.000097	0.000092	0.000061		0.000055
	(2.804)	(4.717)	(3.243)	(2.326)		(4.800)
HWYMRPRO	0.663567	-0.362410	0.580206	0.498985	0.376421	0.332224
	(3.355)	(-2.516)	(2.508)	(2.342)	(3.168)	(4.642)
MALLPRO		-0.646879	-0.959108			
		(-2.072)	(-1.782)			
AGGLOM1	0.047804					
	(6.447)					
AGGLOM2		0.080863				
		(8.405)				
AGGLOM3			0.234618			
			(4.971)			
AGGLOM4				0.094018		
				(6.703)		
AGGLOM5					0.022177	
					(6.829)	
AGGLOM6						0.005285
						(13.799)
MOUNTAIN		0.783139				
		(4.458)				
SUBURBS				0.618772	0.426167	
				(2.423)	(2.294)	
NEWDEVLOP		0.004611		0.002791	0.003345	
		(3.410)		(1.838)	(3.189)	

Table 5-4 continued …

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
OLDDEVLOP		-0.001936 (-2.573)			-0.000629 (-2.734)	
POPDENS	-0.000258 (-2.970)					
HHLDINCDENS					0.000002 (1.698)	0.000001 (2.538)
AVGDWELLVAL		-0.000002 (-2.017)				
HHLDDENS		0.000801 (2.722)				
LANDUSE1		1.559011 (4.300)	0.984063 (2.479)			0.661705 (5.049)
LANDUSE2	0.750878 (2.167)		2.136029 (2.952)			
LANDUSE3						-0.771938 (-1.916)
LANDUSE4					-0.559876 (-2.719)	
LANDUSE5					1.152924 (2.946)	0.628379 (2.891)
LANDUSE6						0.398759 (1.712)
NEWBORN x CBDPRO	-0.000058 (-1.318)					

Table 5-4 continued ...

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
INDUSTRY1 x CBDPRO	-0.000110 (-1.385)					
INDUSTRY2 x POPDENS	0.000339 (2.552)					
INDUSTRY3 x HWYMRPRO		0.806584 (1.740)				
INDUSTRY3 x MOUNTAIN		-0.891513 (-1.443)				
INDUSTRY4 x LANDUSE1		-0.918098 (-2.185)				
INDUSTRY5 x LANDUSE1		1.753567 (2.133)				
INDUSTRY6 x LANDUSE2			-2.205140 (-2.577)			
INDUSTRY7 x CBDPRO			-0.000136 (-2.291)			
INDUSTRY8 x CBDPRO				-0.000159 (-2.051)		
INDUSTRY9 x HWYMRPRO				-0.763197 (-1.889)		
INDUSTRY10 x MALLPRO					0.650485 (2.890)	
INDUSTRY11 x HHLDINCDENS					-0.000002 (-1.535)	

Table 5-4 continued …

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
INDUSTRY12 x LANDUSE4					0.643616	
					(1.888)	
INDUSTRY13 x SUBURBS					-0.425008	
					(-1.752)	
INDUSTRY14 x CBDPRO						-0.000028
						(-2.027)
INDUSTRY15 x MOUNTAIN						0.286058
						(2.021)
INDUSTRY16 x NEWDEVLOP						0.001207
						(1.902)
INDUSTRY17 x NEWBORN x HWYMRPRO						0.882275
						(3.552)
INDUSTRY18 x NEWBORN x HWYMRPRO						0.546541
						(2.732)
No. of Observations	150	321	99	153	396	1318
L(0)	-345.388	-739.130	-227.956	-352.296	-911.824	-3034.807
L(B)	-270.674	-618.704	-189.137	-313.941	-812.938	-2632.406
Rho^2	0.21632	0.16293	0.17029	0.10884	0.10845	0.13260
Adj. Rho^2	0.21165	0.15885	0.16278	0.10429	0.10544	0.13164
No. of iterations to converge	6	6	5	5	6	9

The table presents the estimated coefficients with their *t-stats* in parentheses

Table 5-5: Parameter estimates of location choice models, 2001 – 2002

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
CBDPRO	0.000015	0.000053	0.000033	0.000051		0.000026
	(0.264)	(2.698)	(1.200)	(1.582)		(1.922)
HWYMRPRO	0.426131	0.281376	0.214104	0.442450	0.383930	0.279315
	(1.641)	(2.117)	(0.943)	(1.896)	(2.798)	(3.271)
MALLPRO		0.101787	0.153578			
		(0.285)	(0.279)			
AGGLOM1	0.077710					
	(4.981)					
AGGLOM2		0.102579				
		(10.465)				
AGGLOM3			0.308413			
			(5.776)			
AGGLOM4				0.129769		
				(5.505)		
AGGLOM5					0.049132	
					(8.889)	
AGGLOM6						0.008592
						(13.689)
MOUNTAIN		0.006991				
		(0.040)				
SUBURBS				0.033516	0.389474	
				(0.110)	(1.845)	
NEWDEVLOP		0.003908		-0.001604	0.001438	
		(2.783)		(-0.674)	(1.706)	

Table 5-5 continued ...

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
OLDDEVLOP		-0.001346			-0.000925	
		(-1.795)			(-3.287)	
POPDENS	-0.000233					
	(-1.962)					
HHLDINCDENS					0.000002	0.000000
					(1.844)	(-0.348)
AVGDWELLVAL		-0.000001				
		(-1.405)				
HHLDDENS		0.000281				
		(2.259)				
LANDUSE1		0.153947	0.580904			0.835142
		(0.395)	(1.419)			(5.466)
LANDUSE2	0.946233		-0.236922			
	(1.860)		(-0.223)			
LANDUSE3						-0.158355
						(-0.384)
LANDUSE4					-0.567610	
					(-2.192)	
LANDUSE5					0.870484	0.845052
					(2.149)	(2.716)
LANDUSE6						0.355804
						(1.231)
NEWBORN x CBDPRO	0.000050					
	(0.900)					

Table 5-5 continued …

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
INDUSTRY1 x CBDPRO	0.000021					
	(0.318)					
INDUSTRY2 x POPDENS	0.000193					
	(1.076)					
INDUSTRY3 x HWYMRPRO		-0.548186				
		(-1.060)				
INDUSTRY3 x MOUNTAIN		0.090806				
		(0.155)				
INDUSTRY4 x LANDUSE1		0.874414				
		(2.067)				
INDUSTRY5 x LANDUSE1		1.427339				
		(1.973)				
INDUSTRY6 x LANDUSE2			-0.377331			
			(-0.321)			
INDUSTRY7 x CBDPRO			0.000037			
			(0.654)			
INDUSTRY8 x CBDPRO				0.000057		
				(0.913)		
INDUSTRY9 x HWYMRPRO				-0.046973		
				(-0.092)		
INDUSTRY10 x MALLPRO					0.704633	
					(2.018)	
INDUSTRY11 x HHLDINCDENS					-0.000001	
					(-0.863)	

Table 5-5 continued ...

Covariate	Manufacturing	Construction	Communication & Transportation	Wholesale Trade	Retail Trade	Services
INDUSTRY12 x LANDUSE4					-0.100075	
					(-0.255)	
INDUSTRY13 x SUBURBS					-0.494691	
					(-1.720)	
INDUSTRY14 x CBDPRO						0.000021
						(1.290)
INDUSTRY15 x MOUNTAIN						0.207736
						(1.175)
INDUSTRY16 x NEWDEVLOP						0.000036
						(0.071)
INDUSTRY17 x NEWBORN x HWYMRPRO						0.191259
						(0.651)
INDUSTRY18 x NEWBORN x HWYMRPRO						0.865340
						(3.457)
No. of Observations	83	332	106	118	303	907
L(0)	-191.1146	-764.4583	-244.074	-271.705	-697.6833	-2088.4447
L(B)	-144.8538	-651.4804	-217.7875	-250.4326	-600.3443	-1806.131
Rho^2	0.24206	0.14779	0.1077	0.07829	0.13952	0.13518
Adj. Rho^2	0.23385	0.14378	0.10015	0.07218	0.13571	0.1338
No. of iterations to converge	6	6	6	5	6	7

The table presents the estimated coefficients with their *t-stats* in parentheses coincides with previous findings we achieved on the decentralization trends in the city of Hamilton (Maoh and Kanaroglou, 2004). While decentralization seems to be a prominent phenomenon, the interaction terms we specified in some of the models provide a more elaborate picture on the central effect we seek to capture with the *CBDPRO* covariate. For instance, the *INDUSTRY1* x *CBDPRO* interaction term suggests that firms in the printing, publishing and allied industries (SIC28) favor central locations. The same could be said about firms that belong to the communication and other utilities industries (SIC 48 – SIC 49) as suggested by the sign of the *INDUSTRY 7* x *CBDPRO* interaction term.

The estimated parameters of the interaction term *INDUSTRY8* x *CBDPRO* suggest that wholesale firms specializing in sales of food, beverage, drug and tobacco have a preference for central locations. Results from the Services model suggest that finance insurance, business services, accommodation food and beverages, and other services prefer to locate in the CBD, as indicated from the estimated parameter of the interaction term *INDUSTRY14* x *CBDPRO.*

The use of the interaction term *NEWBORN* x *CBDPRO* suggests that new manufacturing firms gravitate towards the center of the city. These results support the *plant incubation hypothesis* about the location of newly established manufacturing firms. In this hypothesis, it is argued that central locations attract a disproportionate number of new firms, since the core of the city hosts many support services and external economies that can strengthen immature firms (Barf, 1987). The results from the manufacturing model with regard to the *CBDPRO* covariate suggest that manufacturing firms in

Hamilton locate in the core in early stages of their life but will relocate away from the CBD as they mature and grow.

Land accessibility in the models is measured by the *HWYMRPRO* covariate. The estimated parameters of this variable suggest that accessibility plays a role in the location decision of the firms. The magnitude of the estimated parameters suggests that manufacturing, communication, transportation and wholesale trade firms are strongly affiliated with land in close proximity to main roads and highways. The same could be said about new firms that belong to health and social services, and accommodation, food and beverages industries. These newly established firms seem to favor accessible land, as depicted from the magnitude of the estimated parameters associated with the two interaction terms *INDUSTRY17* x *NEWBORN* x *HWYMRPRO* and *INDUSTRY18* x *NEWBORN x HWYMRPRO*. On the other hand, construction firms avoid places in close proximity to main roads and highways except for construction firms specializing in electrical work (SIC 426). Also, wholesale trade firms specializing in other product industries (SIC 59) do not locate on land in close proximity to main roads and highways.

Land in close proximity to regional malls seems to attract retail trade firms that specialize in selling food, beverage and drugs (SIC 60), shoe, apparel, fabric and yarn (SIC 61), and general retailing stores (other retail store SIC 65). Moreover, the estimated parameter of the *MALLPRO* covariate suggests that construction, and communication and transportation firms tend to avoid locating on land in close proximity to regional malls.

The results from all models indicate that agglomeration economies exert a positive influence on the location choice of firms. All the agglomeration variables specified are highly significant and meet our a priori expectation with regard to their sign. Results suggest that other things being equal, firms of the same type appreciate the benefits of agglomeration economies, leading to economic clustering in the city of Hamilton.

The *MOUNTAIN* and *SUBURBS* covariates were specified to capture urban sprawl in Hamilton. The estimated parameters of these covariates suggest that construction and other service (SIC 96 – SIC 99) firms favor land above the escarpment. The results also suggest that wholesale trade and retail trade firms are becoming suburban, as depicted by the sign and magnitude of the *SUBURBS* variable. While the general trend from the construction and retail trade models suggest that most such firms are becoming suburban in their location behavior some notable exceptions can be observed. The interaction term *INDUSTRY3* x *MOUNTAIN* suggests that construction establishments specialized in electrical work gravitate away from land above the escarpment. Moreover, the interaction term *INDUSTRY13* x *SUBURBS* indicate that food stores (SIC 601), gasoline service stations (SIC 633), motor vehicle repair shops (SIC 635) and general merchandize stores (SIC 641) are still central and tend to avoid locating on land in the suburbs.

The variable *NEWDEVLOP* reflects the intensity of recent residential development that took place in 1991–1996. On the other hand, *OLDDEVLOP* reflects the intensity of previous residential development that took place in 1981–1990. The specification of these two variables in the different models suggests that construction, wholesale trade and retail trade firms favor land in close proximity to new residential development. These results complement our findings on the suburbanization trends, since most of the recent residential development is happening in the suburbs. We also believe that new residential development is a catalyst for commercial development. Firms in retail and wholesale trade gravitate towards newly formed residential nodes to benefit from proximity to these populated areas. Since construction firms are the major contributors to residential development, their proximity to the construction sites is commonplace. The estimation results also suggest that certain types of service firms favor land in close proximity to areas with new residential development. The interaction term *INDUSTRY16* x *NEWDEVLOP* suggests that real estate, businesses, accommodation food and beverages, and other services gravitate towards areas with new residential development. All of these services are population oriented and as in the case of retail and wholesale trade will benefit from being in close proximity to newly populated areas. Our results also indicate that construction and retail trade firms avoid areas with old residential development, given that these are more saturated compared to newly developed areas.

The *POPDENS* variable is a measure of population density at a given location. The parameter associated with this variable suggests that manufacturing firms tend to avoid densely populated areas. However, the interaction term *INDUSTRY2* x *POPDENS* suggests that print, publishing and allied industries (SIC 28), and other light manufacturing industries (SIC 39) will locate in densely populated areas. Such results are expected since the central part of the city is densely populated and leading to a preference for firms in SIC 28 for a central location, as noted earlier. Light manufacturing industries, non hazardous to the general population and operating in small stores, are more likely to locate in populated areas.

Results related to the density of household income suggest that retail trade and services favor richer areas in the city. However, the interaction term *INDUSTRY 11* x *HHLDINCDENS* suggests that, other things being equal, shoe, apparel, fabric and yarn retailers (SIC 61), household furniture, appliances and furnishing retailers (SIC 62) and automotive vehicles, parts and accessories, sales and service (SIC 63) do not prefer high income areas when they make their location decisions. The density of average dwelling value was used as a surrogate for land prices; however, the variable was only significant in the construction model. The negative estimated parameter suggests that construction firms tend to locate on inexpensive land in the city of Hamilton. Household density was also specified to account for the effects of urban sprawl on the location decision of firms. Here, we attempted to examine if certain firms will be attracted to locations with low household densities. However, the parameter was insignificant in all models except for the construction model. The positive parameter coupled with information about the spatial distribution of household density, suggests that construction firms prefer locating in areas with medium densities. These are mainly areas above the escarpment.

The land use variables were specified to control for land use zoning within a given grid cell. The estimated coefficients of these variables provide interesting results on the nature of land allocated to the different types of firms in Hamilton. The results suggest that construction, and communication and transportation firms tend to locate on open space land. However, this tendency is more prominent among construction firms specialized in plumbing, heating, and air conditioning and mechanical work (SIC 424). On the other hand, construction firms specializing in residential building and development (SIC 401), site work (SIC 421), structural and related work (SIC 422), interior and finishing work (SIC 427) do not favor open space land, as depicted from the negative sign associated with the interaction term *INDUSTRY4* x *LANDUSE1*. As expected, manufacturing firms locate on resource and industrial land use. The same could be said about communication firms. The communication and transportation model, however, indicates that transportation firms are not affiliated with resource and industrial land uses, as indicated by the negative sign of the interaction term *INDUSTRY6* x *LANDUSE 2*.

Retail trade firms show a strong affiliation with commercial land use, as suggested by the *LANDUSE5* parameter. Also the interaction term *INDUSTRY12* x *LANDUSE4* suggests that population oriented retail firms, such as general merchandize stores (SIC 64) and other retail stores (SIC 65), show affiliation with grid cells characterized by residential land use. Other retailers show no affiliation with areas that are mainly residential, as suggested by the negative sign of the *LANDUSE4* parameter in the retail trade model. Service firms show strong affiliation with commercial, governmental and open space land uses and no affiliation with park and recreational land uses.

Next, we examine the consistency in location choice behavior over time by comparing the estimated coefficients from Tables 5-4 and 5-5. The results from Table 5-5 indicate that most of the specified covariates remain significant over time. However, to examine if there is a significant difference between the coefficients for any given covariate over time, we calculate a 1-degree-of-freedom Wald chi-square statistic for testing the null hypothesis that the (1996-1997) estimated coefficient equals the (2001-2002) estimated coefficient. The formula for this statistic is

$$\chi^2 = \frac{(\beta_{t+1} - \beta_t)^2}{[s.e.(\beta_{t+1})]^2 + [s.e.(\beta_t)]^2}$$

where β_{t+1} is the coefficient for the 2001-2002 period, β_t is the coefficient for the (1996-1997) period and s.e(.) is the estimated standard error (Allison, 1999). Table 5-6 lists the Wald chi square statistic for the different coefficients. The results suggest an overall consistency in location choice behavior over time given that only few coefficients appear to be significantly different. However, except for the agglomeration parameters in the retail trade and service sector models, all the other significantly different parameters were not statistically significant or had the opposite sign in the (2001-2002) location choice models. This is an interesting finding that sheds light on some of the location choice factors that may vary in their importance over time. More interestingly, the agglomeration coefficients in the retail trade and service sector models were positive and significant in both time periods but appear to be significantly different as reported in table 5-6. We believe this is the case because of the change in the clustering pattern of those firms over time. We show in Maoh and Kanaroglou (2004) that firms from the retail trade and service sector have decentralized over time. As a result, the intensity of the clustering pattern has declined in the core and increased in a number of nodes over the city. If this is a continuous trend that will emerge over time, then locations that were less attractive in (1996-1997) will become more attractive in (2001-2002) due to the emerging geographic clustering. Therefore, agglomeration effects will exert more influence on the location choice decision in (2001-2002). That is, the new clusters are more likely to attract a disproportionate number of firms. This can also be discerned from the increase in the magnitude of the agglomeration parameters in the (2001-2002) models.

Table 5-6: Wald chi-square for difference in parameter for the (1996-1997) and (2001-2002) models

Covariate	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6
CBDPRO	1.89	2.39	2.23	0.06		2.68
HWYMRPRO	0.53	10.79*	1.28	0.03	0.00	0.23
MALLPRO		2.49	2.09			
AGGLOM1	3.00					
AGGLOM2		2.50				
AGGLOM3			1.07			
AGGLOM4				1.70		
AGGLOM5					17.68*	
AGGLOM6						20.23*
MOUNTAIN		9.81*				
SUBURBS				2.17	0.02	
NEWDEVLOP		0.13		2.42	2.01	
OLDDEVLOP		0.31			0.66	
POPDENS	0.03					
HHLDINCDENS					0.00	6.44*
AVGDWELLVAL		0.67				
HHLDDENS		2.65				
LANDUSE1		6.97*	0.50			0.74
LANDUSE2	0.10		3.41			
LANDUSE3						1.13
LANDUSE4					0.00	
LANDUSE5					0.25	0.33

Key: Sector 1 = Manufacturing, Sector 2 = Construction, Sector 3 = Communication & Transportation, Sector 4 = Wholesale Trade, Sector 5 = Retail Trade and Sector 6 = Services
* Significantly different at the 0.05 level

Table 5-6 continued ...

Covariate	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6
LANDUSE6						0.01
NEWBORN x CBDPRO	2.32					
INDUSTRY1 x CBDPRO	1.61					
INDUSTRY2 x POPDENS	0.43					
INDUSTRY3 x HWYMRPRO		3.81				
INDUSTRY3 x MOUNTAIN		1.33				
INDUSTRY4 x LANDUSE1		9.04*				
INDUSTRY5 x LANDUSE1		0.09				
INDUSTRY6 x LANDUSE2			1.58			
INDUSTRY7 x CBDPRO			4.45*			
INDUSTRY8 x CBDPRO				4.71*		
INDUSTRY9 x HWYMRPRO				1.21		
INDUSTRY10 x MALLPRO					0.02	
INDUSTRY11 x HHLDINCDENS					0.33	
INDUSTRY12 x LANDUSE4					2.05	
INDUSTRY13 x SUBURBS					0.03	
INDUSTRY14 x CBDPRO						5.27*
INDUSTRY15 x MOUNTAIN						0.12
INDUSTRY16 x NEWDEVLOP						2.08
INDUSTRY17 x NEWBORN x HWYMRPRO						3.23
INDUSTRY18 x NEWBORN x HWYMRPRO						0.99

Key: Sector 1 = Manufacturing, Sector 2 = Construction, Sector 3 = Communication & Transportation, Sector 4 = Wholesale Trade, Sector 5 = Retail Trade and Sector 6 = Services
* Significantly different at the 0.05 level

5.6 Conclusion

This paper presented a micro-analytical firm location model for the city of Hamilton. The model will be used within an agent-based firm demographic model to simulate the life trajectory of individual establishments over time and space. The research presented here offers a new way to approach the firm location problem at the micro level. Using spatial analysis and GIS, it was possible to develop critical variables that enabled us to extend previous firm location modeling techniques and adapt them for the use in microsimulation models. This was possible due to the existence of micro firm data that we extracted from Statistics Canada Business Register (BR). This research also highlights the potential of employing the BR data in the development of urban firm demographic agent-based models. Discrete choice models were applied to investigate the location choice behavior of individual relocating, new and in-migrating establishments.

The results from the estimated econometric models reinforced previous findings obtained by the authors on the decentralization and urban sprawl phenomena in the city of Hamilton (Maoh and Kanaroglou, 2004). This is particularly true for wholesale and retail trade firms which exhibit clear trends of suburbanization. We also found that accessibility to main roads and highways matter to all kinds of firms. Not surprisingly, our agglomeration economies variables reflect the high tendency of industrial clustering in the city. The analysis indicates that agglomeration economies are influential in the location choice decision of firms. The results also suggest that new residential development is a catalyst for commercial development and attracts construction, wholesale and retail firms towards these newly populated areas. The results emphasize the dependence of retail trade and service firms on population. This was reflected with the household income and residential land use variables in the different models. Finally, the introduction of interaction terms to account for firm heterogeneity proved to be very effective in revealing some of the latent behavior that cannot be captured by the generic variables in the different models. For instance, we show that while retail trade firms tend to avoid selecting locations that are predominantly residential, general retail merchandizing (SIC 64) and other retail stores (SIC 65) show high affiliation with residential land use. In other cases, while we were able to show the importance of main road and highway proximity, we were also able to emphasize how newly established health and social service firms, and accommodation food and beverages firms react towards location in proximity to main roads and highways.

Temporal comparison of model performance is important in the development of simulation models. The availability of BR data for different years enabled us to examine if drastic changes are likely to occur in the location behavior over time. This was accomplished by estimating the same 1996–1997 models for the 2001–2002 period. The estimation results suggest a consistency in location behavior over time. In a nutshell, the estimated models were successful in portraying the micro location choice behavior of firms in the city of Hamilton. We also demonstrated that such models could be useful for the development of a firmographic agent-based model for Hamilton.

5.7 References

Allison, Paul 1999. "Comparing Logit and Probit Coefficients Across Groups," *Sociological Methods and Research*, 28 (2), 186-208

Alonso, William 1964. *Location and Land use: Towards a General theory of Land Rent.* Cambridge, Mass: Harvard University Press.

Anderson, William, Pavlos Kanaroglou and Eric Miller. 1996. "Urban form, energy and the environment: A review of issues, evidence and Policy," *Urban Studies*, 33, 7-35.

Anderson, William, Pavlos Kanaroglou, Eric Miller and Ronald Buliung. 1996. "Simulating automobile emissions in an integrated urban model," *Transportation Research Record*, 1520, 71-80.

Barff, Richard. 1987 "Industrial clustering and the organization of production: A point pattern analysis of manufacturing in Cincinnati, Ohio," *Annals of the Association of American Geographers,* 77, 89-103.

Bourne, Larry 1991. "Addressing the Canadian city: contemporary perspectives, trends and issues," in T. Bunting and P. Filion (eds.) *Canadian Cities in Transition*, Toronto: Oxford University Press, pp. 25-44.

Bracken, I and D Martin. 1989. "The generation of spatial population distribution from census centroid data," *Environment and Planning A*, 21, 537-543.

Clapp, John. 1980. "The intrametropolitan location of office activities," *Journal of Regional Science*, 20, 387-399.

Coffey, William, Rejean Drolet and Mario Polese. 1996. "The intrametropolitan location of high order services: patterns, factors and mobility in Montreal," *Papers in Regional Science*, 75, 293-323.

Dieleman, Frans, Martin Dijst and Guillaume Burghouwt. 2002. "Urban form and travel behavior: micro-level household attributes and residential context," *Urban Studies*, 39, 507-527.

Ellickson, Bryan. 1981. "An alternative test of the hedonic theory of housing market," *Journal of Urban Economics*, 9, 56-80.

Erickson, Rodney and Michael Wasylenko. 1980. "Firm relocation and site selection in suburban municipalities," *Journal of Urban Economics*, 8, 69-85.

Frenkel, Amnon. 2001. "Why high technology firms choose to locate in or near metropolitan areas," *Urban Studies*, 38, 1083--1102.

Hansen, Eric. 1987. "Industrial location choice in Sao Paulo, Brazil: a nested logit model," *Regional Science and Urban Economics*, 17, 89-108.

Hanson, Susan and Genevieve Giuliano 2004. *The geography of urban transportation*. New York: The Guilford Press.

Ihlanfeldt, Keith and Michael Raper. 1990. "The intrametropolitan location of new office firms," *Land Economics*, 66, 182-198.

Jones, Ken, Shizue Kamikihara and Jim Simmons. 1998. "Dallas-Fort Worth: commercial structure and change," *Progress in Planning*, 50, 273-289.

Kanaroglou, Pavlos and Darren Scott 2002. "Integrated urban transportation and land-use models for policy analysis," in M. Dijst (eds.) *Governing Cities on the Move: Functional and Management Perspectives on Transformations of Urban Infrastructure in European Agglomerations.*, Ashgate, pp. 43-75.

Kanaroglou, Pavlos and Robert South. 1999. "Can urban form affect transportation energy use and emissions? An analysis of potential growth patterns for the Hamilton Census Metropolitan Area," *Energy Study Reviews*, 9, 22-40.

Lee, Kye. 1982. "A model of intraurban employment location: an application to Bogota, Colombia," *Journal of Urban Economics*, 12, 263-279.

Lee, Kye. 1985. "Decentralization trends of employment location and spatial policies in LDC cities," *Urban Studies*, 22, 151-162.

Maoh, Hanna and Pavlos Kanaroglou. 2004. "Economic clustering and urban form: the case of Hamilton, Ontario," Working Paper No. 3, Center for Spatial Analysis, School of Geography and Geology, McMaster University.

Martinez, Francisco. 1992. "The bid-choice land use model: an integrated economic framework," *Environment and Planning A*, 24, 871-885.

McFadden, Daniel. 1975. "Aggregate travel demand forecasting from disaggregate behavioral models," *Transportation Research Record*, 534, 24-37.

McFadden, Daniel 1978. "Modeling the choice of residential location," in A. Karlqvist, L. Lundqvist, F. Snickars and J. Wiebull (eds.) *Spatial Interaction Theory and Planning Models*, Amsterdam: North Holland Amsterdam, pp. 75-96.

Miller, Eric, John Hunt, John Abraham and Paul Salvini. 2004. "Microsimulating urban systems," *Computers, environment and urban systems*, 28, 9-44.

Miller, Eric, David Kriger and John Hunt. 1998. "Integrated urban models for simulation of transit and landuse policies," *TCRP Web Document 9*, Project H-12, 1-252.

Mirron, John 1982. "Economic equilibrium in urban land use," in L. S. Bourne (eds.) *Internal Structure of the City*, New York Oxford: Oxford University Press, pp. 124-136.

Scott, Darren, Pavlos Kanaroglou and William Anderson. 1997. "Impacts of commuting efficiency on congestion and emissions: a case of the Hamilton CMA, Canada," *Transportation Research Part D*, 2, 245-257.

Shukla, Vibhooti and Paul Waddell. 1991. "Firm location and land use in discrete urban space: A study of the spatial structure of Dallas-Fort Worth," *Regional Science and Urban Economics*, 21, 225-253.

Simmons, Jim and Maurice Yeates. 1998. "Toronto: commercial structure and change," *Progress in Planning*, 50, 253-272.

Southworth, Frank. 1995. "A technical review of urban land use-transportation models as tools for evaluating vehicle travel reduction strategies," ORNL-6881,The Office of Environmental Analysis and Sustainable Development, U. S. Department of Energy.

Timmermans, Harry. 2003. "The Saga of Integrated Land Use-Transport Modeling: How Many More Dreams Before We Wake Up?," paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne.

Vahaly, John. 1976. "The location of service and office activities in Nashville-Davidson county, 1970," *Land Economics*, 52, 479-492.

Waddell, P, A Borning, M Noth, N Freier, M Becke and G Ullfarsson. 2003. "Microsimulation of urban development and location choices: design and implementation of UrbanSim," *Networks and spatial economics*, 3, 43-67.

Waddell, Paul and Vibhooti Shukla. 1993. "Manufacturing location in a polycentric urban area: a study in the composition and attractivness of employment subcenters." *Urban Geography*, 14, 277-296.

Wegener, Michael. 1986. "Transportation network equilibrium and regional deconcentration," *Environment and Planning A*, 19, 437-456.

Wegener, Michael 2004. "Overview of land use transport models," in D. Hensher, K. Button, K. Haynes and P. Stopher (eds.) *Handbook of transport geography and spatial systems*, Amsterdam: Elsevier, pp. 127-146.

Wegener, Micheal and Klaus Spiekermann 1996. "The potential of microsimulation for urban models," in G. P. Clarke (eds.) *Microsimulation for urban and regional policy analysis*, London: Pion, pp. 149-163.

Wu, Fulong. 1999. "Intrametropolitan FDI firm location in Guangzhou, China: a poisson and negative binomial analysis," *The Annals of Regional Science*, 33, 535-555.

Yarish, Michael. 1998. "The Intrametropolitan Location of New Office Firms in Metropolitan Toronto," M.A dissertation, School of Geography and Geology, McMaster University.

CHAPTER 6: Conclusions

Recently, the agent-based approach has become the favorable method by land use and transportation modelers. The reason for this is that it is based on a micro-analytical framework that simulates the dynamic change in urban form as an outcome of the behavior pertaining to individual agents (households, business establishments, developers, etc.) in the city. It is believed that the adoption of this approach can lead to models that are truly behavioral and policy sensitive, capable of promoting a wide range of sustainability practices in urban areas (Miller et al., 2004). Therefore, agent-based models attempt to rectify the pitfalls inherited in the conventional zone-based integrated urban models (IUMs). However, research to develop operational agent-based IUMs is still in its early stages.

This dissertation contributes significantly to this new paradigm by focusing on the demographic behavior of small and medium size establishments in the city of Hamilton in Ontario, Canada. The main objective, as discussed in Chapter 1, is to set the basis for the development of a comprehensive agent-based firmographic Decision Support Tool (DST). Accordingly, the research described in this dissertation applies principles from the firmographic literature to devise the DST. This represents a pioneering effort towards using firmography as the basis for developing a projection tool to assess planning policies in the urban context. Similar efforts were undertaken by van Wissen (2000) to develop a projection tool at the regional level. To this end, the proposed modeling framework serves as a prototype for developing operational agent-based firmographic models at the urban level. The theoretical foundation of the framework is influenced by the existing empirical firmographic literature. Therefore, the framework simulates the evolution of the establishment population as an outcome of three main firmographic processes that relate to the failure, mobility and location choice of individual establishments.

Chapters 3 to 5 report on the research activities that are conducted to model these firmographic processes. Specifically, Chapter 3 explores the survival and failure trends of businesses using non-parametric survival analysis. It also provides parameter estimates for a survival model that is developed to explain the establishment failure process. Chapter 4 is concerned with the mobility behavior of establishments. It presents a mobility model that is estimated to explain the driving forces that lead the establishment to abandon its existing location and move to new site elsewhere within or outside the city. Finally, Chapter 5 is dedicated to study the location choice behavior of business establishments. It describes the formulation of a location choice model estimated to explain the underlying factors that drive the establishment to select a certain developed location as a destination. The findings of Chapters 3 to 5 provide significant information about the factors driving the firmographic behavior of small and medium size establishments in the urban context. The contributions of these findings to the literature are discussed in the following section.

6.1 Contributions of the Dissertation

6.1.1 Contribution to the Literature

The research presented in this dissertation makes important methodological and theoretical contributions to the literature. From the methodological perspective, the dissertation provides the most comprehensive available account of an agent-based firmographic system in the urban context. It also achieves the goal of getting the system towards becoming operational by providing the required parameters for its formulation. The research activities in Chapters 3 to 5 illustrate how data from Statistics Canada Business Register (BR) can be used to study firm demographic processes. This is extremely important since such data have never been used before to model firmographic processes in the Canadian context.

The discrete-time hazard duration model in Chapter 3 emphasizes the importance of the temporal dimension in studying the establishment failure process. It provides an advantage over the widely used cross-sectional models since it allows incorporating temporal information to assess the underlying phenomena. Moreover, Chapter 5 makes another methodological contribution in applying spatial data analysis methods to derive critical land use information that can be used to study the establishment location choice problem at the micro level.

From the theoretical perspective, the findings of Chapters 3 to 5 have a number of contributions. The results in Chapter 3 assert that firm-specific factors are the most influential in determining the fate of an establishment. Macro-economic conditions are ranked second and appear to have a significant role in explaining part of the variability in the failure process. The findings also emphasize the variation in failure rates among the different industries. Surprisingly, agglomeration economies and local competition exert the least effect on the propensity of failure in the presence of other factors. However, this is not the case in the mobility model. The results in Chapter 4 suggest that local competition is an influential stress factor that can push the establishment to move long distances. On the other hand, agglomeration economies increase the propensity of inertia. The results also reinforce earlier findings regarding the effect of an establishment's age and size on the propensity to move. More interestingly, growth conveys different effects in the mobility model. On the one hand, it appears as a push factor in the manufacturing model that increases the propensity of mobility. On the other hand, it appears as a performance factor in the wholesale and retail trade models that decreases the likelihood to move. Overall, the models in Chapters 3 and 4 provide novel work that aims to fill the existing gap in the literature about modeling firmographic processes that have not been well addressed in the urban context.

Chapter 5 makes a direct contribution to the urban economic literature. Unlike most of the existing firm location studies, the estimated models account for the heterogeneity in location choice behavior using interaction terms that combine the characteristics of the individual establishments and the characteristics of the locations. Such efforts reveal some interesting aspects of the location choice behavior. For instance, the results from the manufacturing model support the *plant incubation hypothesis* about the location of newly established manufacturing firms. They also show that while retail trade firms tend to avoid selecting locations that are predominantly residential, general retail merchandizing and general retail stores show high affiliation with residential land use. In other cases, the models were able to emphasize how newly established health and social service firms, and accommodation food and beverages firms react towards location in proximity to main roads and highways. The validity of the estimated models for 1996 was examined by estimating the same models with data for 2001. The estimation results suggest a consistency in location behavior over time, which provides confidence in utilizing such models for simulation purposes. Overall, the models in Chapter 5 provide a new way to approaching the old firm location choice problem.

6.1.2 Implications for Planning and Policy

The results of this dissertation have some implications for city planning and economic policy. As noted earlier, the survival of business establishments, other things being equal, is influenced by the macro-economic conditions in the city. Accordingly, the creation of new job opportunities can decrease unemployment rates and increase the demand for services and goods in the city. This may suggest the need to design policies that can help attracting new businesses in order to create more job opportunities. Moreover, the average size of the industry exerts an influence on the propensity of establishment survival. The findings suggest that small establishments in industries that are dominated by large enterprises may not have the chance to survive. Therefore, policy might be geared towards promoting an environment within which small establishments in certain industries can compete effectively.

From the planning perspective, the results succeed in emphasizing the decentralization and suburbanization trends in the city. These trends will have a direct impact on the emergence of urban form. The location factors in both the mobility and location choice models clearly suggest a decline in the attractiveness of sites in the core of Hamilton. As a result, business establishments are moving towards the suburbs due to a number of pull factors. The results also point to the fact that geographical clustering based on the agglomeration variables discourages mobility and attracts businesses that are searching for new locations. As a result, zoning and the way land uses are planned in the city may impact on the mobility and location choice decisions. As a conclusion, city planning should be geared towards policies that will lead to a reduction in commercial decentralization and suburbanization by creating the proper atmosphere that can sustain businesses in the central parts of the city and attract new business to these areas. Here, the use of well designed IUMs becomes essential in informing sustainable planning policies. The work of this dissertation is a major contribution towards achieving such goals.

6.2 Directions for Future Research

The research discussed in this dissertation expands our understanding of firm behavior and provides the initial and most critical step towards developing a firmographic agent-based microsimulation model. The empirical results are based on a rich and extensive dataset that enabled us to study the firmographic behavior of individual business establishments over space and time. The work was successful in portraying the most important factors that contribute to explaining the failure, mobility and location choice behavior of small and medium size establishments in the urban context. However, more work is needed in the near future to explore some of the questions that emerged during the course of this research. More specifically, some work is required to examine if the low rate of survival among large non-incorporated establishments in Chapter 3 is entirely due to right-censoring in the data. In the same vein, the proposed hypothesis that large non-incorporated establishments have lower survival rates due to high levels of competition from their large incorporated counterparts needs further investigation. Another issue is to examine if firm internal factors such as managerial experience and financial management impact the probability of failure. This, however, will depend on the availability of data that are not currently available in the BR.

Some work is also required to establish the link between the mobility models and the location choice models presented in Chapters 4 and 5. An avenue for future research is to run nested logit models after increasing the number of the randomly selected alternatives at the lower level nest. Currently, the generated logsum variable from the observed chosen alternative and the 9 randomly selected alternatives did not enable us to examine how pull factors influence the probability of moving as noted in Chapter 4. Therefore, the purpose of this future research activity will be to examine if the size of the random sample can remedy this limitation. Also, if data become available, one should examine if commercial rent values affect the propensity of moving in the mobility models. Another issue that was not accounted for due to the lack of data is the affect of floorspace cost (price or rent) and quantities on establishment location choice behavior. Therefore, future research should try to account for this limitation in the location choice models when data becomes available.

The successful implementation of the model as an operational system opens the door to several avenues for future research. Estimating a growth/decline model to predict the change in establishment size will complement the framework presented in Chapter 2. This is an important aspect since growth is a factor in the failure and mobility submodules. Another pressing issue is the generation of a synthetic list of individual business establishments to replace the confidential list that was extracted from BR to estimate the different statistical models. Such list will be used as the basis for running simulations. Here, the focus should be on achieving the best possible accuracy when generating the synthetic population since the predictive accuracy of the simulations will depend on this list.

Another worthwhile research endeavor is the implementation of the system within an object-oriented Geographic Information System (GIS) environment. Such approach will facilitate the storage, maintenance and manipulation of the micro data during simulations. It will also add GIS capabilities that can facilitate the design of scenarios and the display of spatial output. The utilization of the ArcObjects technology of the ArcGIS platform can leverage the development of the system since it affords developing dynamic geodatabase data models. The latter is instrumental for the development of spatial microsimulation systems. An advantage of using the ArcGIS technology is also due to the ability of designing the geodatabase with Unified Modeling Language (UML) diagrams. This can save immensely on the amount of time needed to write programming scripts and will facilitate the expandability of the system in the future.

Finally, other research activities include (1) devising a real estate development module to be used with the developed location choice module. The purpose of such module will be to simulate the addition of new floor space either through the construction of new development, intensification (add more to the same) or conversion (redevelopment) of existing development, and (2) building a region-based models to predict the number of newly formed establishments at the city level and determine the mobility trends between regions. Information from these models can be used to replace the exogenous list of newly formed and in-migrating establishments used by the location choice model.

6.3 References

Miller, E., J. Hunt, J. Abraham and P. Salvini (2004). Microsimulating urban systems *Computers, environment and urban systems,* 28, 9-44.

Van Wissen, L. (2000). A micro-simulation model of firms: applications of concepts of the demography of the firm *Papers in Regional Science,* 79, 111-134.