2013

Modelling Non-residential Real Estate Prices and Land Use Development in Windsor with Potential Impacts from the Windsor-Essex Parkway

Kevin Gingerich
University of Windsor

Modelling Non-residential Real Estate Prices and Land Use Development in Windsor with Potential Impacts from the Windsor-Essex Parkway

Written by
Kevin Gingerich

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

Windsor, Ontario, Canada
2013

Modelling Non-residential Real Estate Prices and Land Use Development in Windsor with Potential Impacts from the Windsor-Essex Parkway

Written by
Kevin Gingerich

APPROVED BY:

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

____
Dr. C. Lee, Department Reader
Department of Civil and Environmental Engineering

____
Dr. W. Anderson, Outside Program Reader
Department of Political Science

____
Mr. J. Tofflemire, Special Member
Department of Civil and Environmental Engineering

____
Dr. S. Cheng, Chair of Defense
Department of Civil and Environmental Engineering

January 11, 2013

DECLARATION OF CO-AUTHORSHIP / PREVIOUS PUBLICATION

I. Co-Authorship Declaration

I hereby declare that this thesis incorporates material that is result of joint research, as follows: This thesis incorporates the outcome of a submitted research paper for publication considerations under the supervision of Dr. H. Maoh and Dr. W. Anderson. Segments of this research are located throughout the thesis. In all cases, the key ideas, primary contributions, data analysis and interpretation, and modelling efforts were performed by the author, and the contribution of co-authors was primarily conducted through the provision of supervision and final editing.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Declaration of Previous Publication

This thesis includes one original paper that has been previously submitted for two conference presentations and publication in a peer reviewed journal as follows:

Gingerich, K., Maoh, H. & Anderson, W. 2012, "Location and transportation effects on non-residential real estate: Price regressions in Windsor, Ontario"
- Submitted for publication in the Transportation Research Record
- Presented at the 59th Annual NARSC Conference, Ottawa, Ontario
- Presented at the TRB 92nd Annual Meeting, Washington, D.C.

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

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

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

ABSTRACT

A study of non-residential land use in the Windsor, Ontario CMA was undertaken to examine possible local implications from construction of the Windsor-Essex Parkway. Two distinct model types were employed. The first consisted of price regressions for industrial, vacant, commercial, office, retail, restaurant, and plaza properties. The second set studied the discrete choice of land use types within commercial and industrial zoning. The commercial logit model had four alternatives: office, retail, restaurant, and other. The industrial logit model had three alternatives: warehouse, factory, and other. The results obtained from these models provide a useful account of interacting land use processes that can inform future transportation and land use policies. Moreover, the empirical analysis is particularly valuable given the larger amount of research into residential land use compared to non-residential. Finally, the models may be useful in the future as part of a more complex integrated urban model.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Hanna Maoh for his supervision, guidance, and encouragement during the two years completing this thesis. In addition to the assistance throughout completion of work for this thesis, the extra time exerted to help me prepare and submit journal articles, presentations, and scholarships was particularly appreciated. I would also like to thank Dr. William Anderson for his time spent assisting in the direction of the project and allocating the resources required for me to complete this work. I would also like to thank Dr. Chris Lee for his time spent as the department reader for this project. In addition, I have had the pleasure of taking numerous undergraduate and graduate classes with him that introduced me into the field of transportation. I also appreciate the time he has taken to write several recommendations for me over the years and co-supervising my capstone project during my final undergraduate year. I would also like to thank John Tofflemire for also taking the time to review and guide my work on this project with his knowledge and expertise, especially given that as someone outside of the university his involvement was outside any work obligations. Finally, thanks to Dr. Shaohong Cheng for taking the time to chair my final defense meeting.

Additionally there are numerous people and companies that have made this work possible through contributions of data. Thanks to the Windsor Essex Real Estate Board for providing real estate data used for the price regressions. Specifically, the assistance of Ms. Sandra Bistany and Ms. Monica Tuck for accommodating our data requirements and allowing me as much time needed in their computer lab collecting data. Also thanks to the City of Windsor for providing permit data for new development used by the logit models. Again, several individuals were very helpful in obtaining this data including Chris Aspila, Darlene Craig, and Dave Dean. In addition I would like to thank Marcel Fontin from the University of Toronto for providing fabric parcel data for the Windsor area.

I would also like to thank Serena, Yingqi, and Shakil whom I have had the pleasure of sharing an office with over the past two years and Jason for giving me guidance as someone completing his thesis ahead of me. Finally, I would like to thank my family for their love, continual support and leaving me a key for a place to sleep whenever needed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Overview

Increasingly, Canada is seeing a trend that has been ongoing within most modern post-industrial nations as the shifting general population relocates away from traditional rural locations towards focal points within major cities. This urbanization brings with it several challenges that become ever more important as the trend continues. One key to solving these challenges lies with effectively integrating and understanding the processes within and between the transportation system and land use. Due to the integrated relationship between the two, it is expected that a significant infrastructure project such as the Windsor-Essex Parkway (WEP) connecting the Highway 401 in Ontario to the proposed Detroit River International Crossing (Figure 1-1) will have a noticeable impact to accessibility and land use.

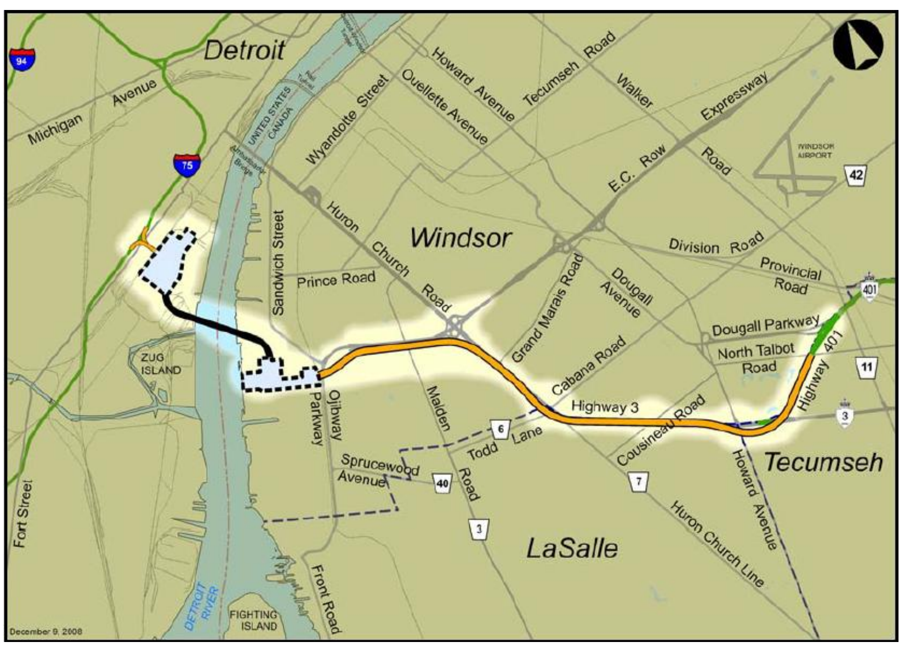

Figure 1-1: Proposed layout for the WEP and DRIC
(Source: URS Canada, 2008)

Nearby to the potential DRIC site and the nearest similar border crossing is the Ambassador Bridge. This privately owned crossing features the single largest volume of truck traffic between Canada and the U.S. *(Taylor et al., 2004)*. This integral gateway between the two countries not only sees large volumes of daily commuters but is also particularly appealing for the logistics industry as the Great Lakes area boasts a high density of manufacturing industries that rely on just-in-time (JIT) delivery *(Austin et al., 2008)*.

On a broader scale, the Canadian economy is highly dependent upon foreign trade with 45% of its GDP attributed to said trade *(Trading Economics, 2011)*. An increasing proportion of these transactions are based in the resource sector, but the manufacturing/automotive industry in the Great Lakes region is still a predominant trade industry *(Andrea and Smith, 2002)*. While there may be both pros and cons to the construction of a new crossing, an infrastructure project of this magnitude will have a significant impact not only on a national/international scale, but also on a local scale affecting land use in the nearby area. For instance, the Eix Transversal highway in Catalonia, Spain, improved both the residential and commercial industries in the close vicinity in comparison to all of Catalonia *(Obregón-Biosca & Junyent-Comas, 2011)*. The 11 km parkway that is currently being built will not only have an impact based on accessibility but also include some 300 acres of green space and 20 kilometers of recreational trails *(Windsor Essex Mobility Group & MTO, 2011)* that will impact local land use.

To forecast and monitor this change, both transportation and land use as well as guiding policies are required as shown in Figure 1-2. Land use can be further broken down into different key agents that are involved at a local level – residents, firms and developers along with the land prices that are affected by all three agents. To this end, the primary purpose of this project is to focus on firms and the potential impact of the WEP infrastructure project on accessibility and land use change for non-residential properties.

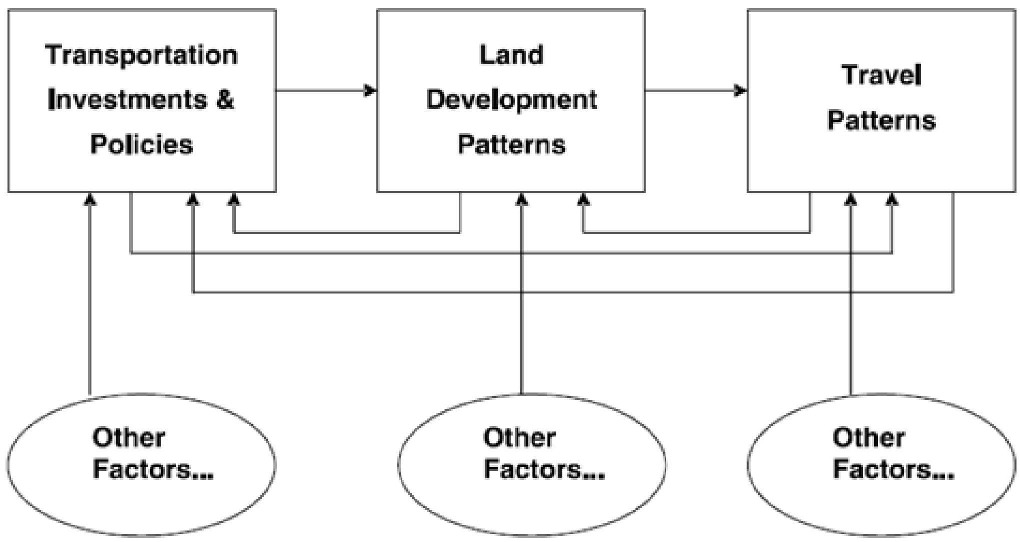

Figure 1-2: Interactions between transportation and land use
(Source: Handy, 2005)

1.2 Research Goals

During the preliminary stages of this project, several general over-arching goals were created for guidance, though the original goals were altered over time due to data complications. A summary of the primary goals can be described as:

1) Determine the effects of various spatially based phenomena for the Windsor CMA using price regressions and discrete choice models of non-residential land use
2) Investigate potential impacts on local land use caused by the development of the Windsor-Essex Parkway

1.3 Thesis Outline

This report follows a logical progression that should be familiar to most readers. Chapter Two begins with a discussion of previous literature on land use and transportation that provides a foundation for the models and their parameters. The following chapter outlines primary and secondary sources of data utilized by the models while Chapter Four details the methodological approaches taken. Chapter Five provides the results of those models along with discussions on their relevance toward possible effects from construction of the Windsor-Essex Parkway. Finally, Chapter Six provides conclusions on results, limitations, and future considerations before concluding with references and additional information provided in the appendices.

2 LITERATURE REVIEW

2.1 The Role of Land Use on Development

Before creating the models themselves, a literature review was undertaken of both the broad subjects of land use/transportation as well as more focused aspects important for this project. This was done to inform and provide a theoretical basis for the variables that were included within the models. The first part of this literature review focuses on aspects related to land use.

2.1.1 Land Use and Urban Form

To better understand the processes that drive the spatial configuration of modern cities, it is important to understand several terms. *Anderson et al. (1996)* broke down land use into three distinct topics: urban form, urban interaction, and urban structure.

Urban form relates directly to the spatial patterns that are developed: compact, radial, and sprawl are several common examples of urban form. Cities with only one central business district (CBD) are considered monocentric while those with multiple CBDs demonstrate polycentric patterns *(Anderson et al., 1996)*. Many other types of urban forms exist, along with various combinations of them. Even the self-iterating fractal patterns often utilized in mathematics have been found to exist within the urban form *(Anas et al., 1998*). The association with fractals lends to the notion that the pattern of the city is repeated at various scales within smaller segments of the city. In terms of modelling, *Hu and Lo (2007)* utilized a fractal dimension to determine the optimal resolution of cell sizes to include in their logit model.

One of the longstanding gaps in literature is that the patterns themselves are often characterized simply by observation instead of discrete, quantifiable terms. One such example of this is the urban sprawl pattern that often dominates the characterization of modern North American culture. Instead, *Galster et al. (2001)* proposed defining urban sprawl in discrete terms that could compare the level of sprawl in various cities quantitatively instead of by simple observation. They describe the term based on eight distinct characteristics that can each be calculated, with lower values indicating higher levels of urban sprawl.

Interestingly, Los Angeles, a city often considered the typical case of urban sprawl was found to have a relatively low level of sprawl based on the calculations of *Galster et al. (2001)*. *Anas et al. (1998)* also make this claim when noting that *Batty and Longley (1994)* found LA to have a high fractal dimension of 1.93 (the maximum possible value is 2) which indicates "a relative absence of fine-structure irregularities in development patterns" *(Anas et al., 1998).* This is contrary to what one would expect a city with large amounts of urban sprawl to exhibit. This conclusion leads one to believe that the actual patterns that develop may not coincide with commonly held notions. While the increased accessibility from a highway project is expected to affect development in certain ways, modelling the land use processes for Windsor may provide additional insight not found through a simple examination of the urban landscape.

Urban interaction is the term used to describe the flow of objects between various points within the spatial area. These objects can be representative of various types including people, goods, and information *(Anderson et al., 1996)*. Demand for the flow of people and goods are supported by the transportation network within the city. This forms a relationship between land use and transportation that will be discussed in more detail later. The urban spatial structure is a term developed by *Bourne (1984; via Anderson et al., 1996)* that encompasses both the aforementioned terms of urban form and urban interaction, as well as a set of principles that drive the relationship between those terms. According to *Anderson et al. (1996)*, this emphasizes that the flows within a city are not driven solely by the urban form. Land use itself is often considered a very dynamic system that possesses a temporally lagged nature. Therefore the current spatial structure of a city is the product of changes that have occurred in the past. For instance, *Woudsma et al. (2008)* found that accessibility to major access points most significantly affected the land use in Calgary, Alberta after a five to ten year period.

Anas et al. (1998) describes the history of the spatial structure for most cities in North America from the origins to suburbanization and urban sprawl to a final tendency in modern times towards sub-centers or "edge cities" that consist of commercial office space with other land uses integrated within. This polycentric pattern moves closer towards what *Anderson et al. (1996)* call a compact multinucleated form that describes a relatively energy efficient pattern of urban form compared to urban sprawl.

2.1.2 Land Developers' Behaviour

When studying the impact of an infrastructure project on non-residential parcels, it is important to remember that prior to a firm locating there, the land must be developed. Thus, developers are an integral component in that they develop the land for suitable use. *Maruani and Amit-Cohen (2011)* contend that the development planning process consists of three unique agents – the developer, the planning system that enforces regulations, and the public who can alter the demand for various types of development. Of the three groups, developers are considered the most influential as denoted by the characteristics of activism, leadership, and dominance in the planning process shown in Table 2-1 *(Maruani and Amit-Cohen, 2011)*.

Table 2-1: Characteristics of Developers

	Leadership	**Activism**	**Dominance**	**Overall**
High	Developers	Developers	Developers	Developers
Medium	Government	Government	Government	Land Owner / Builders
Low	Public	Public	Public	
Sources	*Maruani and Amit-Cohen (2011)*			*Buttimer et al. (2008)*

A review of the literature on developers finds that one of the most important aspects about them is the large amount of financial risk inherent in their field *(Buttimer et al. 2008; Maruani and Amit-Cohen, 2011)*. *Buttimer et al. (2008)* divides the development of a home into three distinct groups of agents: the land owners; land developer; and the home builder. The developer normally carries the most risk of the group because of the time needed to purchase the land and comply with regulations in addition to the amount of change in the market price that can fluctuate within that time. To counter this, developers often employ risk-mitigating measures such as presale options for homebuilders *(Buttimer, 2008)*. Due to the high risk nature of work for developers, this risk is often mitigated by maintaining large volumes of work to spread the risk. This leads to a small number of developers generally active within a community. While the modelling presented herein uses a scale at the property parcel level, individual developers can have an influence on the results found creating a possible bias in the development type choice of a given parcel of land.

2.1.3 Firm Location and Agglomeration

When looking at the land use patterns of a city, it can be seen that the location of firms will be a significant factor as one of the primary agents involved in the shaping the urban landscape. Where or if a firm decides to locate is often a complex mix of many variables. *Van Dijk and Pellenbarg (2000)* organized these factors into three specific groups: internal firm factors such as the size and age of the firm; external firm factors such as government policies and regional market demands; and spatial factors such as accessibility and agglomeration.

A look at the important stages of development for a firm can help determine important internal characteristics. *De Bok (2009)* performed a microscopic model of the location decision of firms by modelling their four distinct lifecycle phases: firm formation, growth, migration, and dissolution. A general flow chart for the decision of a firm to relocate is given in Figure 2-1. *De Bok* found that the notion of a "breeding ground scenario" often occurs where firms originate in a diverse area then later move to a more specialized area. Similarly, *Maoh and Kanaroglou (2009)* found that new manufacturing firms favoured locating near the CBD whereas older firms gravitated towards the opposite. These two theories are highly related since the CBD tends to be a highly diverse area with many types of firms. In general, several papers conclude that the internal characteristics of a firm were more significant than other factors (For example: *de Bok, 2009; Van Dijk and Pellenbarg, 2000*).

Agglomeration is often considered among the spatial factors for firm location (for example – *Woudsma et al., 2008; de Bok, 2009*). *McCann and Shefer (2003)* discuss that there are three idealized types of clustering: pure agglomeration; industrial complex; and

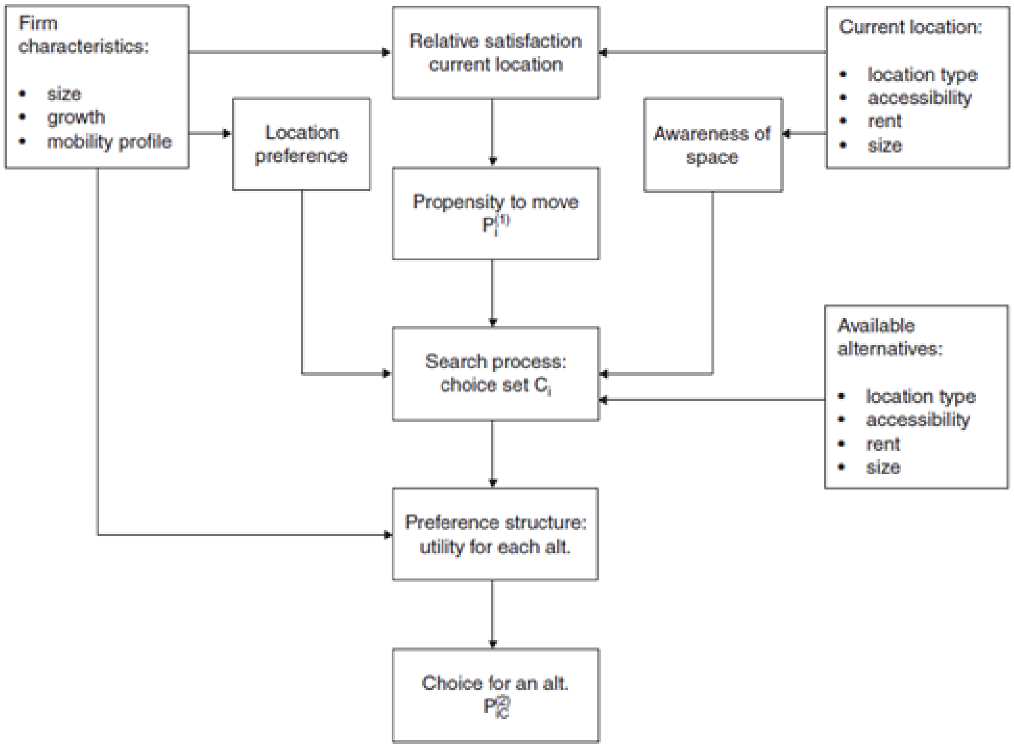

Figure 2-1: Modelling Firm Location
(Source: de Bok and Sanders, 2005)

social network. The pure agglomeration model is based on the traditional notion of agglomeration developed by *Alfred Marshall (1890, 1920; via - Van der Panne, 2004; McCann and Shefer, 2003)* where firms in close proximity to each other gain beneficial externalities such as a more skilled labour pool, information spillovers, and specialized suppliers. This model represents individual firms as atoms or points and does not consider the relationships between the firms as the other two models do.

The agglomeration model developed by *Marshall* focusing on the benefits of specialization is also often contrasted with an opposing view associated with *Jane Jacobs (1969; via van der Panne, 2004*). *Jacobs* believed that the clustering of firms gain externalities from close access to different, diverse types of industries instead of the knowledge spill over of firms in a similar industry. *De Bok and Van Oort (2011)* measure both of these agglomeration models through production specialization (PS) and production diversification (PD) indices that function as indicators of Marshall and Jacobs agglomerations, respectively. Their results found that the PS index was more significant compared to the PD index. The only sector found to have a statistical significance for the PD index was in the transportation industry (*De Bok and Van Oort, 2011)*. This exception seems applicable given that the transportation industry is reliant on other industries to generate their demand. In general, however, the literature is often undetermined in the debate on whether Marshall externalities or Jacobs externalities are more prevalent (*Van der Panne, 2004; de Bok and Van Oort, 2011).* As is often the case, the real impacts of agglomeration are likely to be a mix of the two instead of one or the other.

2.2 The Role of Transportation on Land Development

In light of the definitions for land use, the transportation system can be thought of as the links that facilitate urban interaction (the flow of people and goods). While both public and private modes of transportation are available, the implementation of high speed rail and public transit in North America has mostly been left by the wayside, especially in Canada. For instance, the City of Windsor in Ontario moves roughly 3% of all travelers by public transit (*City of Windsor, 1999*). In fact, North America has seen a shrinking quantity of railroad corridors with the United States containing only 272,000 kilometers in modern times compared to 416,000 kilometers in 1920 (*Garrison and Levinson, 2006* via *Xie and Levinson, 2008*). Because of this decreasing trend for railroads, a larger proportion of freight travels by road. This increases the impact and importance that an infrastructure project such as the WEP and DRIC can have on accessibility for cross-border trade.

Highway improvements in particular can have a significant impact on land use and location patterns that can vary depending on the context of the situation. For example, *Funderburg et al. (2010)* studied three different scenarios for new highway improvements and found different results for each. This underscores the importance of assessing each scenario individually.

2.2.1 Logistics and Warehousing

Because of the location of Windsor inside a large trade corridor, it is logical to assume that the logistics and warehousing industries exert a prominent role in local land use. Several recent events highlighted in the media also support the importance of this industry. First, the City of Windsor is in the planning process to create a 60 hectare air-cargo hub (*Windsor Star, 2011a*). This hub could act as an international intermodal facility since it is planned to include American CBP staff and the ability to provide pre-clearance for flights directly into the U.S. Secondly, the former site of a Chrysler minivan plant will contain a warehousing hub, including a 755,000 square foot warehouse to be leased to Chrysler (*Windsor Star, 2011b*). The close proximity of the warehouse to the Chrysler Assembly Plant emphasizes the need for short transfer times of parts in a just-in-time process. Utilizing warehouses close to a large trade corridor at the border is presented by *Capineri and Leinbach (2006)* as one way of making cross border interactions more efficient.

On a broad scale, modern literature suggests that logistics and warehousing industry has changed in recent years. *Hesse (2004)* states that as globalization occurred, warehousing and logistics became more centralized. This culminates with an increased preference for large distribution centres (DC's) over traditional warehouses where the focus is shifted toward the efficient flow of goods instead of efficient storage (*Hesse, 2004*).

Even more recently, *Torbianelli (2009)* suggests that a combination of the older decentralized warehouses and newer centralized DC's is becoming more popular. This hybrid combination uses both primary DC's along with secondary warehouses to provide more flexible solutions. In particular, it becomes easier to utilize intermodal methods of transportation since the final delivery to the end use must usually be delivered by truck *(Torbianelli, 2009).*

Aside from the change in how logistics systems operate, the financing of the property itself is handled differently. In the past, firms requiring extra logistics or warehousing capacity would simply buy the property required. This has largely been replaced however, by specialized developers speculating and developing the land. As seen in previous sections, this is indicative of the large influence developers have on land use. Subsequently, firms that are interested simply lease the space instead of outright buying it. This provides a strong rate of return on the investment by developers while also providing firms with required space and the flexibility to relocate if the need arises *(Hesse, 2004)*.

2.2.2 Transportation and Land Use Interactions

It is clear that a transportation network has a larger effect then simply providing a convenient transfer from one point to another. As Sir Rod Eddington described it:

> "A good transport network is important in sustaining economic success in modern economies. The transport network secures connectivity between different parts of a country, as well as to the rest of the world: linking people to jobs; delivering products to markets; underpinning supply chains and logistics; and supporting domestic and international trade" (Eddington, 2006)

This economic change can be partially attributed not only to the efficiency with which goods and people are moved, but also the change in land use that occurs. While transportation and land use can be thought of as two distinct systems, they are largely interconnected with one another. For instance, dating back to the mid twentieth century, *Stopher and Mehburg (1975)* stated that the urban transportation planning process consisted of seven consecutive steps: Inventory, land-use forecast, trip generation, trip distribution, modal split, network assignment, and evaluation. This process is now commonly seen with the middle four steps comprising the Urban Transportation Modelling System (UTMS). According to *Stopher and Mehburg (1975)*, land use was a component of the planning process for many years but a feedback process between the two systems was only starting to emerge in the 1970's (*Stopher and Mehburg, 1975*). Currently, models that include this feedback have become more common. For instance, two such models that feature a large degree of interaction between the two systems include "IMULATE" calibrated for the Hamilton, Ontario CMA (*Kang et al. 200*9) and the open source model "UrbanSim" utilized for a number of cities in North America (*Waddell, 2002; Noth et al., 2003*).

2.2.3 Accessibility

This change in land use that occurs due to a transportation network is largely attributed to the change in accessibility. For example, firm agglomeration mentioned earlier will frequently occur along important transportation links due to the increased access that these roads provide. Various accessibility measures are used by researchers to help formulate better models for land use and transportation (For instance - *van Dijk and Pellenbarg, 2000; de Bok, 2009; Mataloni, 2011; Song et al, 2011; Woudsma et al, 2008; Straatemeier, 2008*). The significance of the accessibility measure for various topics, such as the choice of firm location can vary. For instance, while *Mataloni (2011)* found that the amount of road infrastructure was a significant factor, *van Dijk and Pellenbarg (2000)* and *de Bok (2009)* found that accessibility was not nearly as significant compared other firm specific factors.

Batty (2009) explains that the earlier versions of accessibility of an area were generally calculated proportional to the size of the opportunities available and inversely proportional to the distance to that point. This can be seen in a much earlier paper produced by *Hansen (1959)* that describes this accessibility as a potential for opportunities that measures the "intensity of the probability of interaction". *Batty (2009)* contends that in more recent times, there are many different measures of accessibility and believes that a unified theory is required to bring these measures together; much like the definitions for urban sprawl mentioned earlier.

Geurs and Ritsema van Eck (2003) group various accessibility measures into three categories: infrastructure measures that indicate traffic mobility such as travel speed and congestion; activity based measures that calculate accessibility to various activities such as the definition already provided by *Hansen*; and utility based accessibility that measures the benefits of accessibility to the users. The activity based measure of accessibility can be different based on the type of activity chosen. For instance, *Kumar and Kockelman (2008)* include both a factor for job accessibility as well as population accessibility. It should be noted that the level of job accessibility can vary significantly if job competition, education level, etc. are included *(Geurs and Ritsema van Eck, 2003; Geurs et al, 2006)*. It is based upon this background that a number of parameters are included in the models for this project. For instance, the potential accessibility attribute is used to determine accessibility to the residential population with a negative exponential distance decay function.

2.3 Types of Operational Modelling

The literature reviewed indicates several types of current modelling that have been recently performed. The integrated urban model UrbanSim has a component for real estate development as one portion of the entire model. The real estate sub-module is described by *Waddell (2002)* as a bottom-up process where the developer for each location decides whether to develop said location and what type of development to perform. Many studies also use a bottom-up process by utilizing Agent Based Models (ABM) to simulate land development. For instance, *Kieser and Marceau (2009)* use an ABM model to study changes in land use by using agents specified as land developers, citizens, and planners. Similarly, *Magliocca et al. (2011)* also use an ABM model with agents represented as developers, farmers/landowners, and consumers. The developer agent for their model purchases and develops land to maximize profit based on the success of past developments in the area. The models presented later in this thesis validate this argument – several parameters measuring the success of nearby real estate proved to be statistically significant.

A further example of ABM modelling is performed by *Ligmann-Zielinska and Jankowski (2010)* with a top-down process simulating planning and zoning regulations to accompany the bottom-up ABM model. The developer agents in their model purchase and develop land based on their preference for various factors including accessibility, land value and nearby natural amenities. *Waddell and Ulfarsson (2003)* believe that a bottom-up model for land development can follow two scenarios – “use looking for a site” and “site looking for a use”. In the first scenario following a destination choice framework, a developer has a specific project/use in mind and looks for the best site. In the latter scenario the landowner looks to sell property for the use that will give the most profits. Based on the work of *Martinez (1992)* it was found that the two scenarios are complementary, reaching the same conclusions regardless of approach.

Hu and Lo (2007) and *Kamyab et al. (2010)*, however, believe that using a discrete choice model such as a logistic regression can provide more optimal results over the ABM models. *Hu and Lo* assert that a logistic regression can encompass more demographic and econometric factors than cellular automata and ABM models when modelling land use change in addition to providing measures of statistical significance.

2.3.1 Integrated Urban Models and Land Prices

The analysis conducted in this thesis will be useful in the future by providing components required to create more complex models. Many contemporary models are comprehensive and include both land use and transportation components to simulate the relationship between the two. In recent times, integrated urban models are becoming increasingly popular. While many large cities have been utilizing these models, smaller cities have also begun to create them. For instance, the City of Montgomery, Alabama, with a population of roughly 200,000 people in the year 2000 is similar in size to the City of Windsor. Montgomery has started calibrating their own model based on the Production, Exchange, Consumption Allocation System (PECAS) model (*Clay, 2010*).

According to *Iacono et al. (2008)*, integrated urban models were considered to begin with the Lowry model in the 1960s, named after its developer *Ira Lowry*. The group of models derived from the Lowry model, were labeled as spatial interaction/gravity based models by *Iacono et al (2008)*. These models were based around the gravity theory adapted from physics as well as the spatial interaction that is produced as a result of demand for basic/non-basic industries. In general, these early models did not include explicit land development. A brief diagram on the history of integrated urban models is given in Figure 2-2.

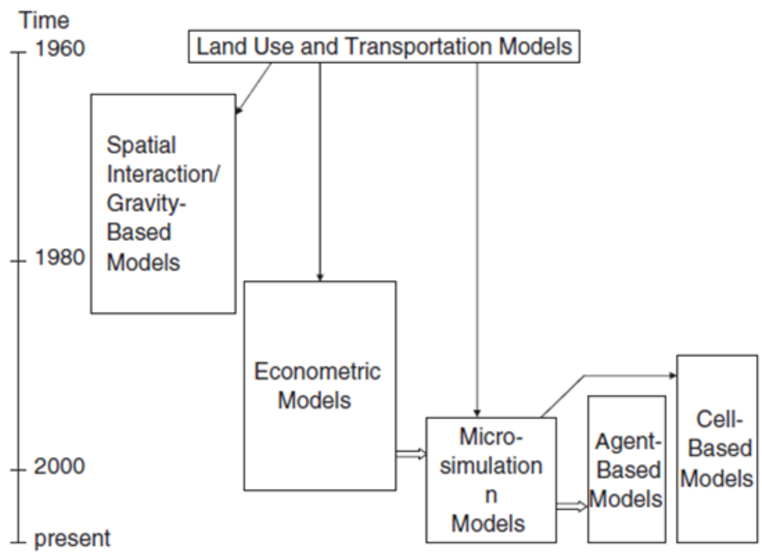

Figure 2-2: Chronological Timeline of Integrated Urban Models

(Source: Iacono et al., 2008)

A more accurate representation of the land development market was attributed to the second generation of integrated urban models named “econometric models” by *Iacono et al. (2008)*. Many of these econometric models contain sub-modules that allocate land development based on profit maximization for land developers. Later models in the 1990s created more advanced land development models. For instance, the MUSSA model developed for Santiago, Chile incorporated bid-rent theory into the market for land development *(Martinez, 1996)*.

In the UrbanSim model, the real estate development module utilizes a multinomial discrete choice logit model with one model for each initial development type. The procedure for creating the model begins by discovering the historical “events” of a cell developing or changing its type of development. A sample of cells with no events are also selected (the same size as the number with events) to properly account for bias within those that were chosen (*Waddell and Ulfarsson, 2003*). Using the outcome of logit model, the type of development is often chosen through a Monte Carlo simulation (similar to *Maoh and Kanaroglou, 2009*). The results are then added to a “development template” that determines the degree of change that occurs (*Waddell and Ulfarsson, 2003*).

Land prices have become an integral component of most IUMs. In UrbanSim, a hedonic regression analysis is used for land price model where the price of each individual lot is broken down into sub-prices for individual characteristics (*Waddell and Ulfarsson, 2003*). The market prices are represented by both a mean price level (average price determined by supply and demand, interest rates, the economy, and population characteristics) and relative prices based on the characteristics of the land around each cell. While the price function itself is not dynamic, the change in variables used is. The land price model is performed after all other models are completed to calculate the final land prices for the chosen time period (*Waddell and Ulfarsson, 2003*).

2.4 Factors Affecting Land Use Changes

A brief look at some of the factors in literature that were found to influence land use is given here. While many papers studied firm locations, this is assumed to exhibit similar relationships to those found in land development. For many of the models that focused on firm migration, internal firm characteristics were often more significant than outside characteristics. Some of the variables used in academic papers cover significantly large areas (more than one metropolitan area) and therefore contain factors that would not be viable when modelling a smaller area. For instance, the model presented by *Cheng and Stough (2006)* includes labour cost differences between cities. Many factors were therefore filtered to ensure they could have some implications for land use in one city. Table 2-2 provides a summary of the types of variables found to be significant in academic literature but is not comprehensive due to time constraints.

2.4.1 Impacts of Nearby Neighbourhoods

Similar to the idea of cellular automata, the state of neighbourhoods surrounding an area can have an influence on the nearby zone. One important recent change in literature is that researchers are now preferring to use methods that include autocorrelation between zones including simultaneous auto-regressive (SAR) models and spatially correlated logit models (*Maoh and Kanaroglou, 2007; Nguyen and Sano, 2010; Woudsma et al., 2008; Sener et al., 2011*).

Table 2-2: Factors in Literature Influencing Land Development

	Parameter Description	Residential	Commercial Industrial
Parcel Characteristics	Perimeter to Area Ratio	+	
	Slope	-	
	Flood Plains / Wetlands	-	
	Percentage of parcel containing roads	+	
	Rent		-
Impact of Neighbouring Zones	Existing Res. Land Use	+	+
	Higher Res. Density	+	
	Existing Civic Land Use		
	Existing Com. Land Use	+	+ / -
	Existing Undeveloped Land	+	
	Existing Developed Open Spaces	+	
	Existing Mixed Use (Entropy)	-	-
	No. Of Physically Active Recreation Centres	+	
	Avg. Household Incomes		+ / -
	Nearby Parking Prices		+
Accessibility	Job Accessibility		+
	Population Accessibility		-
	Highway Accessibility	-	+
	Total Commute Time	-	
	Distance to CBD		+ / -
	Distance to Malls		+ / -
	Transit in Both Home and Work Zones	+	

+/- denote positive and negative relationships

Zhou and Kockelman (2008) found that the impact of surrounding neighbourhoods up to 2 miles away had a significant effect on development. For residential development, it was found that existing residential land use and higher densities increased the likelihood of further residential development (*Waddell and Ulfarsson, 2003; Zhou and Kockelman, 2008)* as well as increased commercial and industrial development *(Waddell and Ulfarsson, 2003).* Furthermore, specific types of firms such as retail firms may gravitate towards areas of new residential development *(Maoh and Kanaroglou, 2009)* to capture a market that would be considered relatively open. In fact, of all the types of land use, *Zhou and Kockelman (2008)* found that only the civic land use had a negative impact on the further nearby development of its own type. The authors attribute this outlier to the equitable nature of public services that attempt to spread apart to service an optimal quantity of people.

Nearby existing commercial land use was found to increase residential development but its effect on further commercial development was mixed. *Zhou and Kockelman (2008)* found that it encouraged the development of commercial use while *Waddell and Ulfarsson (2003)* found that the opposite occurred. *Zhou and Kockelman (2008)* also looked at undeveloped land and found that its close proximity had a significantly positive effect on residential development. This indicates a strong demand for housing located further away from highly developed areas and is indicative of the urban sprawl common in many urban areas today. Developed open spaces also had an impact, increasing the value of land nearby as expected (*Waddell and Ulfarsson, 2003*). The amount of mixed use in an area had a negative impact, shown through an entropy factor by *Zhou and Kockelman (2008)* and *Sener et al. (2011)* to be negatively impeding the development of both commercial and residential land use. For the choice of residential location, nearby access to local amenities such as physical recreation centres have a positive influence (*Sener et al., 2011*). The decision for commercial firms to locate in an alternative can also depend on the general income of the area. Some industries may prefer areas with a higher average income while other industries tend to prefer those with less (*Maoh and Kanaroglou, 2009*).

2.4.2 Characteristics of the Parcel

The physical characteristics of the parcel in question were also studied through various factors. *Zhou and Kockelman (2008)* found that a greater ratio of perimeter to area had a positive correlation on residential development. Slope was also a factor, negatively affecting the development of land (*Zhou and Kockelman, 2008*) and reducing the overall value (*Waddell and Ulfarsson, 2003*). According to *Waddell and Ulfarsson*, several other attributes also reduce the land value including flood plains/wetlands. Perhaps the most important characteristic of the parcel itself is the price. For instance, *Hunt (1997)* found that as the rent increased, the probability of a commercial firm choosing that location significantly decreased.

2.4.3 Accessibility Parameters

It should first be noted that accessibility here is either a distance/travel time between locations, or a measure of potential accessibility to all locations. Accessibility as an economic measure is not considered in order to limit the scope of the project. Several different types of accessibility were found to impact development. *Kumar and Kockelman (2008)* look at both job accessibility and population accessibility. Job accessibility is significantly positive indicating firms tend to locate in close proximity to other jobs (firms). Population accessibility, in contrast to the job accessibility, is significantly negative indicating that firms tend to avoid areas with high populations. *Kumar and Kockelman (2008)* also find that firms are more likely to locate closer to the CBD. *Maoh and Kanaroglou (2009)* found most retail firms tended to locate closer towards malls while construction and transportation industries tended to avoid them.

Extensive literature has been documented on the general accessibility of a transportation system. For instance, *de Bok (2009)* finds that proximity to transportation infrastructure is a significant factor in the performance of firms from various industries. Similarly, *Woudsma et al. (2008)* found that accessibility to various access points in a city as well as congestion were significant determinants for logistics firms. For highways, areas with a close proximity to this infrastructure exhibited increased amounts of commercial and industrial firms (*Kumar and Kockelman, 2008; Hunt, 1997*) while also increasing the value of land by up to 9% (*Waddell and Ulfarsson, 2003*). Conversely, areas near highways decreased the amount of residential development (*Sener et al., 2011*). *Sener et al.* also looked at commute times and transit. Results of the analysis found that a higher commute time had a negative influence on the choice of residential location. On the other hand, access to transit in both work and home zones had a positive influence, indicating that all things being equal, residents will choose areas with easy access to public transportation.

2.4.4 Divided Parcels

One final consideration for factors affecting the development process is whether the land parcels in question will be divided to increase profits. *Zhou and Kockelman (2008)* look at this issue extensively with a model specifically addressing this issue. They found that factors affecting development were different for their two separate models of non-subdivided and subdivided parcels. This indicates that the developments of parcels that are divided and undivided are not the same but instead follow a slightly different process. This issue of divided parcels should be kept in mind when looking at the results from the case study performed in Chapter 5.

3 DATA

The data utilized for this project consisted of two primary datasets in addition to numerous sources of secondary data. The primary data is used as the dependent variables within the models while the secondary data are used to provide independent parameters as input variables.

3.1 Study Area

The study area encompasses the Windsor CMA and nearby outlying areas located in Southwestern Ontario, Canada. The region shown in Figure 3-1 is bounded by the Detroit River to the west/northwest, Lake St. Clair to the north, Lake Erie to the south and the Town of Tilbury to the east. For the two primary datasets, the real estate data covered this region in its entirety, while the permit data for new development only pertained to the City of Windsor.

The area is known primarily (and historically developed around) the automotive manufacturing industry due to its proximity to Detroit, Michigan. This area has developed into a strong trade corridor, especially after the introduction of the now defunct Canada-United States Automotive Products Agreement (Auto Pact) allowing tariff free trade across the border. In recent times, the area has suffered economically. For instance, there has been a decrease in per capita income in Detroit since 2003 *(Harpel, 2011)*. More recent recessions may have exasperated the toll on Detroit's economy as it has supposedly moved close to bankruptcy and a state takeover *(Windsor Star, 2012a)*. Though Windsor has significantly fewer residents and is separated politically by the border and physically by the Detroit River, close spatial proximity and high levels of local trade has created a symbiotic relationship between the two cities. It is within this economic context that forms the backdrop for the property transactions employed here.

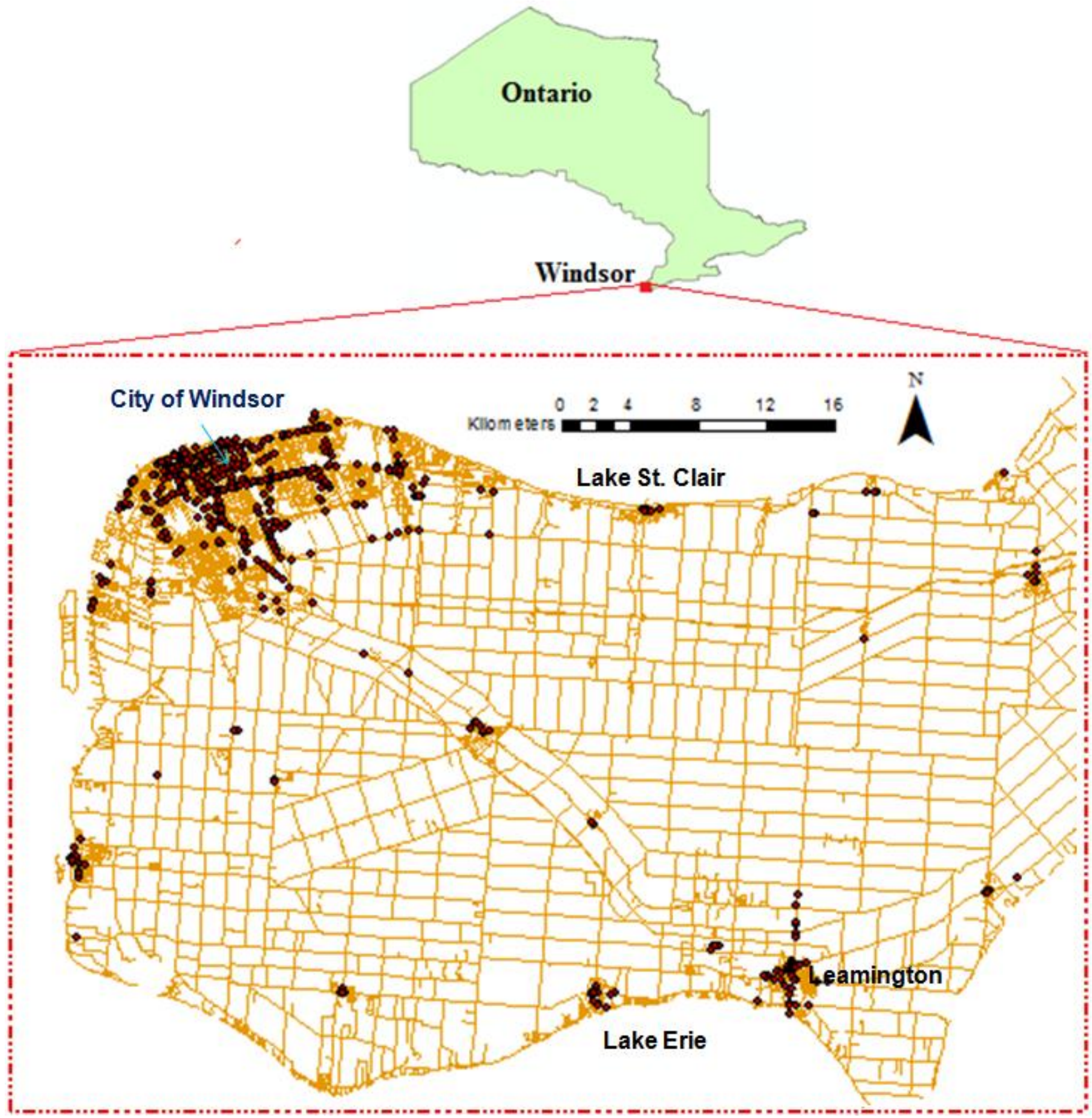

Figure 3-1: Study Area for Price Regression Models

3.2 Real Estate Prices

3.2.1 Overview

One of the primary data sources obtained for this project was non-residential property listing records from the Windsor-Essex Real Estate Board (REB). This data was collected for the 1991 – 2011 period as earlier records were not available. The search for listings was performed based on entry date – the last day that someone entered or edited the listing. Since the closing date of the listing is not given (unless it was sold), it is assumed that the listing was deactivated from the market during the same month/year of the entry date (the last date the record was edited). This assumption was found to not hold perfectly due to extreme circumstances. However, given the use of this dataset with only properties that were sold and their corresponding sale price, this is not a concern due to the inclusion of a sale date. A sample of the data obtained and corresponding descriptions are found in Table 3-1 and Table 3-2, respectively.

Table 3-1: Sample Real Estate Board data

MLS Number	Status	District	Sub District	Address	Age	List Price	List Date	Sale Price	Sale Date	Type	Lot size
0400644	Leased	00 - Windsor, Lasalle, Tecumseh	7 - South Central	[redacted] RHODES # [redacted]	1-10	$8.00*	20-JAN-2004	$38,316	11-MAR-2010	Office	3193 SQ FT
0613795i	Expired	00 - Windsor, Lasalle, Tecumseh	3 - Central Windsor/Downtown	[redacted] OUELLE-TTE	OL	$3000**	09-NOV-2006	-	-	Office, Retail	0.499 ACRES
0706992i	Cancell-ed	00 - Windsor, Lasalle, Tecumseh	5 - Tecumseh	[redacted] TECUMSEH		$159**	04-JUN-2007	-	-	Vacant Land	53 X 180 SQ FT
0710359i	Sold	00 - Windsor, Lasalle, Tecumseh	7 - South Central	[redacted] TECUMSEH E	40	$399**	17-AUG-2007	$382**	29-JAN-2010	Office	91 X IRREG

* Real estate prices for leasing given as dollars per square foot per year
** Prices are shown as actual value x 10^{-3}

Table 3-2: REB data attributes

Listing Attributes	Description
MLS Number	A unique identifier for each listing
Status	The *current* status of a listing based on 4 statuses (sold, leased, expired, or cancelled)
City, District, Sub-District and Address	Identifies the location of the property
Age	Age of the building at the time of listing (often not included or broad in description, i.e. 50-100 years old)
List Price and List Date	The price in dollars or dollars per square foot if the listings are for sale or lease, respectively
Sale Price and Sale Date	The total price sold (if for sale) or total price for one year (if for lease) and the date sold
Type	Industry type; many records also included sub-types
Lot Size	Size of the lot listed in ft^2 unless listed otherwise

3.2.2 Geocoding

To connect the listings spatially, the dataset was uploaded into ArcGIS and geocoded. To do this, address locators are required to determine the physical location of each listing. Due to the limited information of the address locators and the original data, several different locators were used to geocode as many listings as possible. The first address locator was created using the DMTI road network. The results from this exercise indicated that roughly two thirds of the full dataset were successfully geocoded using the locator as seen in Figure 3-2. To increase the accuracy and efficiency of the DMTI address locator, areas East of Tilbury were excluded. This ensures that erroneous matches are not made with similar street names in other cities.

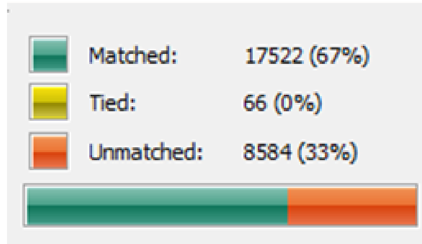

Figure 3-2: Geocoding results using DMTI address locator

Some roads within the DMTI road shapefile lacked address number information, explaining why one third of the records were not placed. To further increase the proportion of records that are spatially located, a second address locator was used. This locator is provided in ArcGIS and contains information for all of North America using projected (x,y) coordinates. Use of this second locator added 3,827 valid data points, resulting in 21,415 geocoded data points out of a possible 26,172. This gave an 82% overall success rate for spatially locating the data.

3.2.3 Initial Observations

Before investigating the data further, some initial observations regarding the data could be seen. Table 3-3 shows the quantity of records by year listed while Figure 3-3 shows the effect of geocoding on the data. Figure 3-4 shows the percentage of listings sold and leased. The average number of listings that were sold corresponded to roughly 10 % of the total available listings.

The most significant discrepancy in the data is the large decrease in available listings from 1997 to 1998 causing an increase in the percentage of listings sold in Figure 3-4. This is the result of a significantly large number of listings having a final entry date in 1997. Because this entry date was assumed to correspond to the final year a listing was available (a listing would normally be last edited sometime near the date it was taken off the listings, except for unusual circumstances). In this case, it is believed that a change in the computer system used to store the listings occurred[1]. It is more likely that a gradual trend occurred instead of the larger variation shown.

Table 3-3: REB count based on year listed

Year	Before Geocoding Process			After Geocoding Process		
	Total Listings	Listings for Sale	Listings for Lease	Total Listings	Listings for Sale	Listings for Lease
***Before 1991**	632	528	104	459	356	103
1991	1313	1119	194	1012	832	180
1992	1366	1130	236	1081	862	219
1993	1297	1084	213	1003	813	190
1994	1291	1060	231	1021	810	211
1995	1181	976	205	928	741	187
1996	1262	1006	256	1005	774	231
1997	1276	972	304	975	705	270
1998	1059	844	215	879	683	196
1999	1114	873	241	937	716	221
2000	1072	844	228	915	708	207
2001	1092	868	224	882	692	190
2002	946	725	221	807	617	190
2003	1007	753	254	874	651	223
2004	1151	880	271	993	748	245
2005	1324	947	377	1142	805	337
2006	1375	1033	342	1168	864	304
2007	1285	951	334	1079	796	283
2008	1241	890	351	1045	739	306
2009	1275	905	370	1062	746	316
2010	1326	916	410	1089	751	338
2011	1285	875	410	1059	714	345
AVG	1216	936	280	998	751	247
SUM	26170	20179	5991	21415	16123	5292

*Not a complete record for the period before 1991

[1] Based on correspondence with the Windsor-Essex Real Estate Board

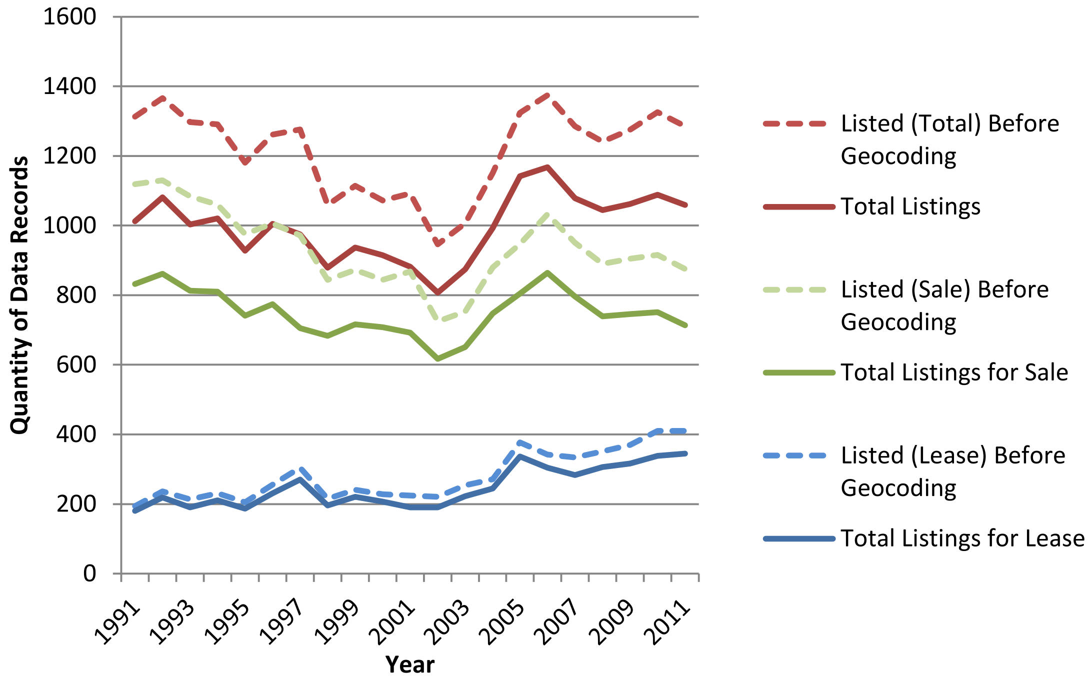

Figure 3-3: REB records listed by year

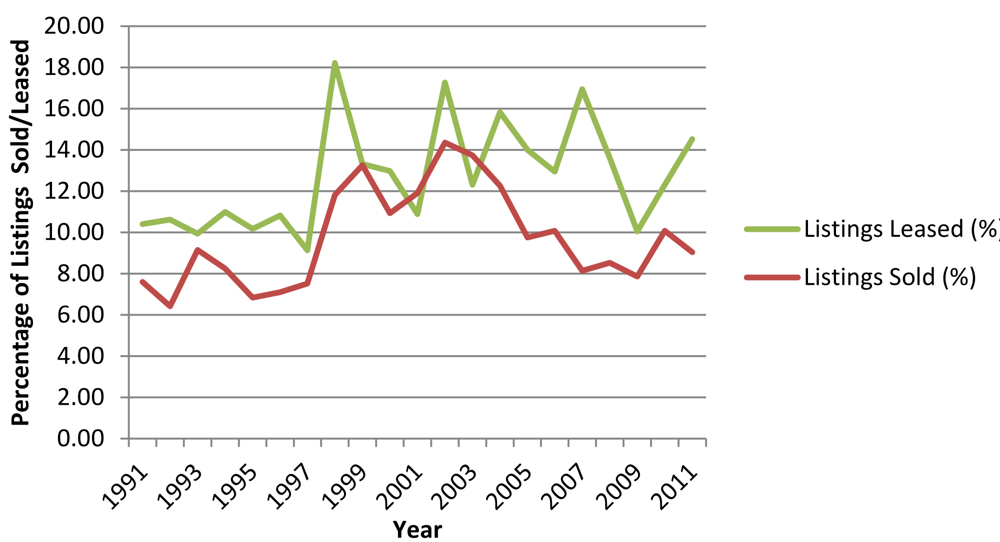

Figure 3-4: Percentage of listings sold and leased

3.2.4 Industry Types

To categorize the industries of given properties, the data records were sorted into industry types. For example, the 1997-2001 time period had 33 unique industry names as seen in Table 3-4. These industries were placed into two distinct groups – once based on their general industry type (commercial, industrial, etc.), and again based on more specific sub-types within the commercial industry (retail, office, food and plaza) to account for heterogeneity. Kernel density maps denoting the areas of firm clusters based on industry types are shown in Appendix A.

The counts of listings for each group are shown in Figure 3-5. These counts are based on the number of properties within their respective industries that sold. This was done due to the risk of listed properties being placed on the market at unreasonable prices. Using sold properties allowed for actual real estate value to be captured. In addition, properties that were put up for lease were disregarded to avoid expected differences in the valuation and utility for those looking to buy or rent. Moreover, the lease / rent prices are contract prices. Due to provisions within the contracts themselves, the effective prices can be difficult to determine *(Colwell et al., 1998)*. Some listings were also disregarded due to insufficient information available.

Table 3-4: REB data types

Listings sorted into general and specific industries (1997-2001)

Original Given Industry	General Type	Detailed Type
Restaurant / Foods	Commercial	Food
Plaza	Commercial	Commercial Centre
Dry Cleaning / Laundry, Laundry	Commercial	Services
Office(S)	Commercial	Offices
Office(S), Retail	Commercial	C_Mixed
Shopping Centre	Commercial	Commercial Centre
Retail	Commercial	Retail
Dry Goods / Fashion	Commercial	Retail
Beauty / Hair	Commercial	Services
Hotel	Commercial	Accomodation
Furniture / Household Furn.	Commercial	Retail
Florist / Gifts, Gifts	Commercial	Retail
Other Retail	Commercial	Retail
Variety Store	Commercial	Retail
Entertainment	Commercial	Services
Bar / Hotel	Commercial	Accomodation
Grocery / Mini Mart	Commercial	Retail
Other Services	Commercial	Services
Hardware / Decor	Commercial	Retail
Motel	Commercial	Accomodation
Automotive / Aircraft	Commercial	Retail
Sports / Recreation, Recreation	Commercial	Services
Daycare / Children	Commercial	Services
Pets	Commercial	Services
Electronic	Commercial	Retail
Industrial	Industrial	N/A
Manufacturing / Wholesale	Industrial	N/A
Institutional	Institutional	N/A
Remarks	Unknown	N/A
Contract Maint	Unknown	N/A
Vacant Land	Vacant	N/A
Vacant Land, Vacant Land, Vacant Land	Vacant	N/A
Warehouse	Warehousing	N/A

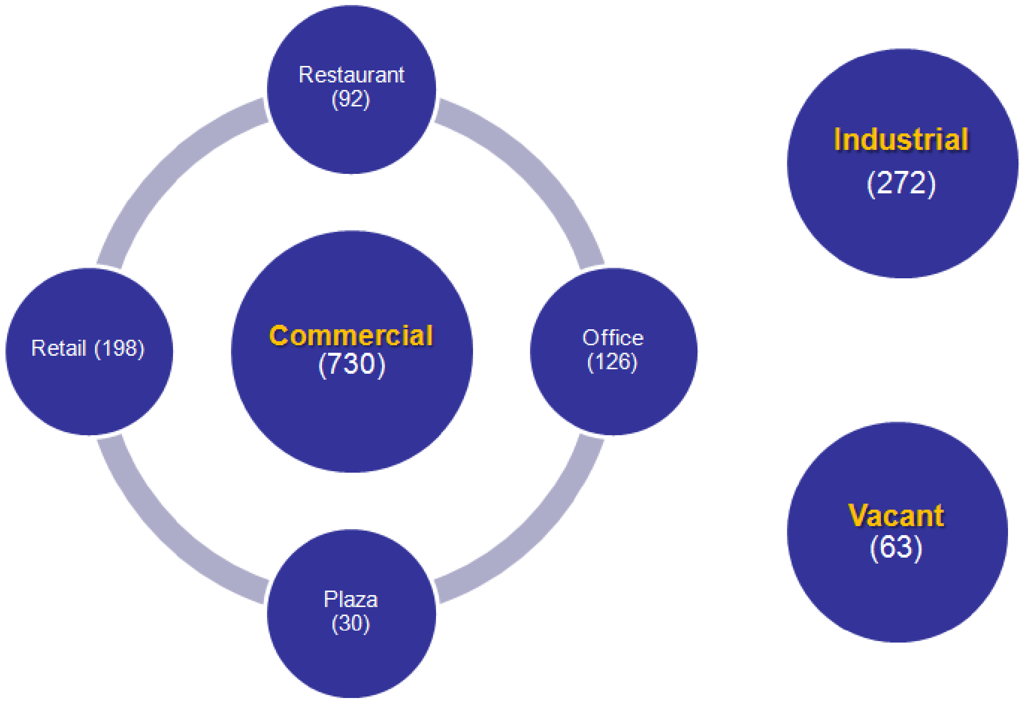

Figure 3-5: Quantity of Listings by Industry Type

3.3 Development Permits

Data on new development in the Windsor area were obtained from the City of Windsor in the form of records for development permits. They contained the following attributes for each individual construction permit record:

- Date Issued
- Land Use Type
    - Residential / Res. Accessory
    - Institutional
    - Industrial
    - Commercial
- Sub-Type (ex. Retail)
- Address
- Property Roll Number
- Frontage Area
- New/Addition
- Work Area
- Construction Cost
- Detailed Description

While several of the attributes were not used due to a number of missing values, primary interest with this dataset is with the type of land use and corresponding spatial location. Geocoding the spatial location for each permit was performed comparatively different from the real estate data. In order to determine the spatial locations of these permits, parcel roll numbers were used to join the parcel shapefile and permit database (both obtained from the City of Windsor). Within the permit dataset, there were 235 duplicate roll numbers indicating the presence of multiple permits for individual parcels. Since the roll numbers were used as the link between the two datasets, a standard join of the tables within ArcGIS would result in the loss of these duplicates due to the required one-to-one relationship for the join. To overcome this, the two files were placed within a geodatabase to create a one-to-many join between city parcels permits.

While the parcel shapefile provided by the City of Windsor was recently updated, not all property roll numbers were listed resulting in the loss some permit data. This resulted in the total permit count dropping from 3057 down to 2880. Moreover, the total count that would be used for modelling was further reduced by including only permits issued for new construction as shown in Table 3-5.

Table 3-5: Permit Data by Land Use Type

Land Use Type	New	Addition	Total
Residential Dwelling	1835	118	1953
Accessory Structures	420	31	451
Commercial	119	56	175
Industrial	74	40	114
Institutional	116	62	178
Other	8	1	9
Total	**2572**	**308**	**2880**

The dataset consisted of permits inclusive of the years 2005-2011. Moreover, these permits were only available within the city limits of Windsor, excluding outlying suburbs and rural areas. While permits for other areas in the region could potentially be obtained from corresponding municipalities, this was not pursued due to time constraints. Additionally, while the quantity of permits for Windsor is already lower than what are preferred, outlying municipalities would contain even fewer permits presenting a greater probability of bias due to small data counts.

3.4 Secondary Data

Aside from the primary datasets outlined above, several sources of secondary data were also used to provide input variables in the models. Table 3-6 summarizes many of these data sources. For instance, digital elevation maps (DEMs) with 30m x 30m raster cells obtained from Desktop Mapping Technologies Inc. (DMTI) were used for elevations. A road network available from DMTI was also utilized.

Table 3-6: Summary of Secondary Data Sources

Data	Source
Digital Elevation Model	DMTI
Slope	DMTI
Heritage Sites, Police Stations, etc.	City of Windsor
Transit Bus Routes	City of Windsor
Basic Parcel Data	City of Windsor
Fabric Parcel Data	University of Toronto
Census Data	U. of T CHASS Analyzer
Windsor Streets	DMTI
Network Assignment Data	COMMUTE
Distance to CBD, mall, etc.	ArcGIS
Windsor CMA Coast Line	Statistics Canada

City of Windsor data taken from their website was also used. Among these files, the spatial location of public transit lines proved to be a significant factor in several models. Census data was also included providing demographic variables between different census tracts in the area. Data for this model came from the Computing in the Humanities and Social Sciences (CHASS) census analyzer from the University of Toronto. Examples of demographic data provided includes: the number of people travelling to work by type of mode; age demographics; number of private dwellings; immigration numbers; employment numbers including a breakdown by industry type; and average income / education levels. Due to their aggregated nature, some of the variables in the census data exhibited correlated behaviour with other census attributes and spatial variables. Therefore while several demographic attributes could be found to be statistically significant on their own in the model, they were often capturing the effects of other variables (and each other) and are therefore not included within the final models to avoid multicollinearity. Finally, several variables were created in GIS. For example, the rail lines were drawn in ArcGIS due to unacceptable location deviations in the rail shapefile created by DMTI.

4 METHODOLOGY

As was demonstrated throughout the literature review, complex relationships exist between/within land use and transportation. A simplified view of these relationships is shown in Figure 4-1 to illustrate the importance of the two modelling sections performed within the scope of this thesis. Items demarked with a green outline indicate specific areas covered while those shown in red denote additional items that were not addressed due to a limited scope and schedule.

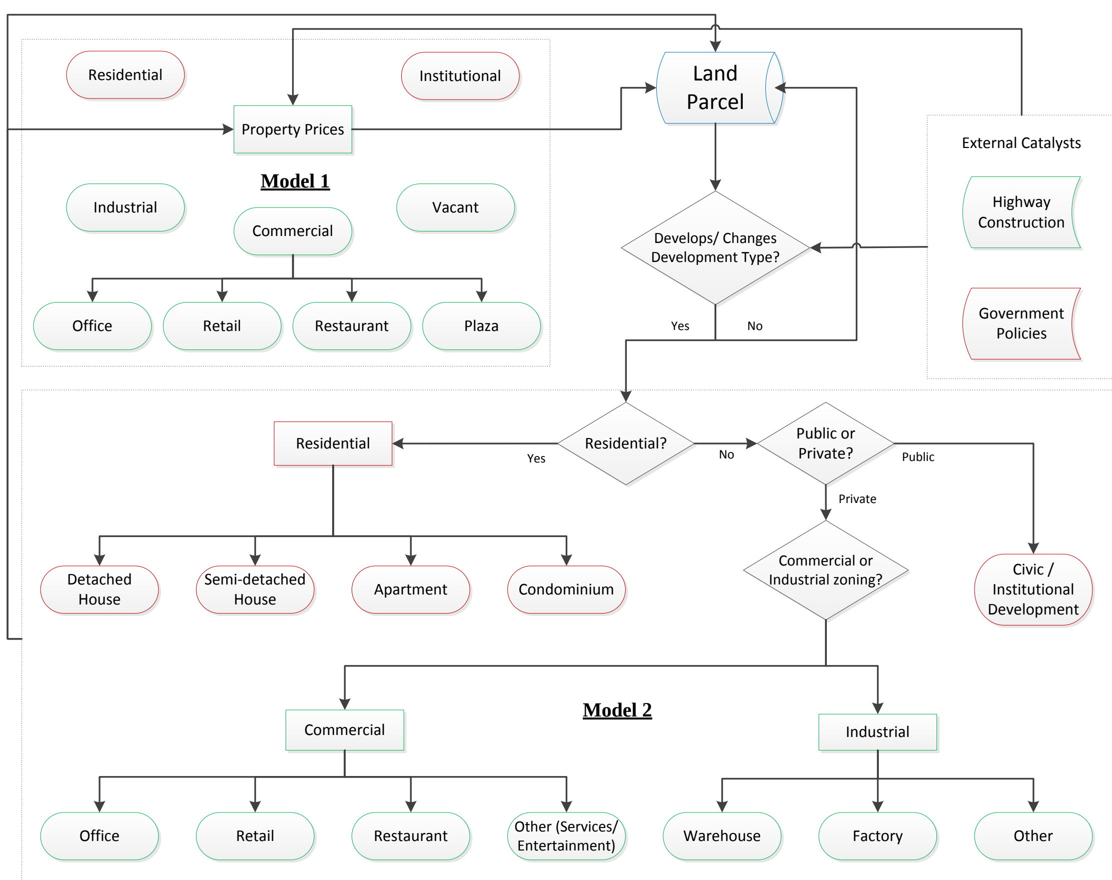

Figure 4-1: Methodology Flow Chart

Two model types were performed corresponding to the primary datasets given in Chapter 3. The first set of models consisted of OLS regressions analyzing real estate prices for seven groups of non-residential properties. The second modelling type was composed of separate logit models for commercial and industrial development with the subtype of development providing the alternative choices for the decision maker. The two models are closely linked together with the development choice in Model 2 affecting the property price of the land parcel in Model 1. The development type and price will lead to further evaluations of land to determine if a change in development type will occur, producing a looping effect. External factors will also carry influence on the price and development of the land parcel. Specifically, this thesis seeks to understand implications from the external influence of new highway transportation infrastructure.

4.1 Model 1: Price Regressions

4.1.1 Background

Paramount to the understanding of land use in a region are the prices that represent the physical utility a particular lot is worth to potential buyers. The perspective of utility depends on the type of property sought after. For instance, residential home buyers will consider a mix of financial and non-financial benefits while firms will tend to focus almost exclusively on potential profit when purchasing a property.

Price regressions are useful for the understanding of factors that influence these prices. For example, they provide a measure of the benefits obtained from transportation improvements *(Du and Mulley, 2006; Chalermpong, 2007).* Additionally they can be used as a component within modern integrated urban models. While the increased prevalence of geographic information systems (GIS) has led to a large influx of price regression models, they tend to be focused on residential prices due to the larger quantity of data available. In addition, residential data often benefit from individual features such as the number of bedrooms/bathrooms, allowing for hedonic regressions to determine individual prices for pieces of the structure. Examples of studies on residential prices include: *Du and Mulley (2006)*; *Chalermpong (2007)*; *Kockelman (1997)*; *Iacono and Levinson (2011)*; *Martinez and Viegas (2009)*; and *Srour et al. (2002)*.

By contrast, transactions for non-residential real estate properties occur less often, especially outside of the core city areas *(Tu et al., 2004)*. Moreover, the data does not usually gain the benefit of detailed structural attributes that can be used in hedonic regressions. Despite these obstacles, commercial properties have a large impact on urban areas and should therefore be duly considered. Examples of papers on non-residential price models include: *Tu et al. (2004)*; *Montero-Lorenzo et al. (2009)*; *Colwell et al. (1998)*; *Füss et al. (2012)*; and *Dunse et al. (2005)*.

In terms of modelling, OLS regressions are still frequently used *(Srour et al., 2002; Dunse et al., 2005; Ten Seithoff and Kockelman, 2002; Ozus, 2009)* but have started to give way to more advanced methods. The latter include spatial regressions *(Chalermpong, 2007; Martinez and Viegas, 2009; Páez et al., 2001)* and a focus on local effects *(Du and Mulley, 2006; Páez et al., 2001; Hannonen, 2008)*. Due to the scarcity of commercial price regressions, this thesis employs a number of regression models of real estate market prices in the Windsor, Ontario metropolitan area.

4.1.2 Price Normalization

To account for different land use types within the dataset, it was important to model them explicitly. One option is to use a single model that incorporates land use type dummies as control variables within the model *(Ten Seithoff and Kockelman, 2002)*. Instead, separate models were specified for seven categories as shown in Figure 3-5. This allows each type to be fully scrutinized by all specified regressors and limits heterogeneity within the models.

While sale prices were used as the base determinant for the dependant variable, several measures were employed to normalize these prices across all listings. First, the data spanned two decades from 1991 to 2011. A common measure of inflation in Canada, the consumer price index (CPI), was used to adjust all prices to 2011 values. Secondly, listing prices were also adjusted to reflect the price per unit area. Based on the data available, developed square footage was either completely unavailable for a listing or found to be somewhat inconsistent. The lot size parameter was used in its place. Unfortunately, using the lot size is not as desirable compared to the actual square footage since it is expected that prices are based primarily on developed land. An inherent drawback to using the lot size is the possibility for properties to develop vertically (e.g. offices) in addition to utilizing the available horizontal space. This could have an influence on errors occurring in the models. For instance, positive outlier results could occur due to buildings with a developed height larger than other properties in the dataset. The negative repercussions, however, are partially mitigated due to the separation based on land use type.

4.1.3 Dependent Parameters

Covariates for our econometric analysis were based almost exclusively on spatial variables since structural attributes of the properties themselves were not available (as is common for non-residential properties). A number of categorical (dummy) variables were devised by conferring a value of 1 when the listing was within a specified buffer distance from a certain land feature, otherwise a value of 0 was assigned. Since the most significant distance can vary depending on the variable in question *(Maoh et al., 2012)*, a sensitivity analysis with multiple buffer values was employed (i.e. 200 m, 400 m, 600 m). The buffer area providing the most significant results was then selected for the final model. This leads to the range of buffer sizes displayed in Table 4-1. While a continuous variable representing time to the CBD was included, others were created as dummy variables to avoid capturing the effects from unknown latent variables as the distance increases. In addition, the use of dummy variables provides useful information on the size of direct effect that each feature exerts on prices. Finally, dummy variables generally had a greater significance over alternative continuous variables. Over 100 variables were initially considered, but many were either insignificant or correlated with each other. For instance, demographic census tract data was initially included but appeared to be highly correlated with the location variables.

4.1.4 Model Specifications

To create the price models, a basic OLS regression was first used. This regression took on two forms with both a linear regression as well as a log transformation of the dependent variable as follows:

$$P = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \epsilon \quad \text{(Commercial model only)} \qquad (4.1)$$

$$\mathrm{Ln}(P) = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \epsilon \quad \text{(All other models)} \tag{4.2}$$

Where P is the price per ft^2 in 2011 dollars, β_0 is the constant (intercept), X_i are the set of independent spatially-oriented parameters along with their beta coefficients (β_i), and ϵ is an error value for unknown variables and influences centered on a mean of 0. The majority of models used equation *4.2* but the commercial model was more significant following equation *4.1*. Although the logarithmic transformation in equation 4.2 results in several negative values for the dependent variable, Ln(*P*), this did not affect the results from the regressions since the price P for these data points return to positive when transformed back using the exponential function.

In light of an analysis of our data indicating the presence of spatial autocorrelation in land prices, spatial regression models were also introduced to tease out any spatial effects. A spatial lag model, known as Simultaneous Autoregressive (SAR), was used. The model takes on the following form *(Anselin et al., 2006)*:

$$P = X\beta + \rho W(P - X\beta) + \epsilon \tag{4.3}$$

Where ρ represents the spatial lag coefficient, *W* is a matrix for the weights used, and the remaining variables are similar to those used in equations *4.1/4.2*.

Finally, to account for the bias resulting from outlier observations, dummy variables (Positive and Negative) were introduced in the models. These variables were created based on analyzing the residuals obtained from the original regression models. They were then included within the regressions in a similar manner as other explanatory variables. We contend that the observed outliers pertain to properties that differ in terms

Table 4-1: Explanatory Variables of Land Price Regressions

Variable	Description	Measured as:
Lot Size	Size of the exterior lot	ft^2
CBD Time	Time to reach the center of CBD from each property	Minutes
Ln(CBD Time)	Logarithmic transformation of CBD Time	Minutes
Potential Accessibility	Accessibility of sold property i to residential population R in census tract j. t_{ij} is the travel time (minutes) between i and j	$\sum_j R_j * e^{-0.07 t_{ij}}$
$CBD_{200\ M}$	Properties within 200 meter buffer of the CBD	1 (true) or 0
$CBD_{400\ M}$	Properties within 400 meter buffer of the CBD	1 or 0
$Rail_{200\ M}$	Properties within 200 meter buffer of a rail line	1 or 0
$Rail_{400\ M}$	Properties within 400 meter buffer of a rail line	1 or 0
$Transit_{200\ M}$	Properties within 200 meter buffer of a transit line	1 or 0
$Ramp_{800\ M}$	Properties within 800 meter buffer of a highway ramp	1 or 0
$Coast_{400\ M}$	Properties within 400 meter buffer of the coast	1 or 0
$Auto_{600\ M}$	Properties within 600 meter buffer of the 3 largest automotive plants in the area	1 or 0
Year Sold	The specific year a property was sold (e. g. 1999)	Year
Sandwich	Properties located in Sandwich Town	1 or 0
Leamington	Properties located in Leamington	1 or 0
Locations 1-6	Properties located in specified zone – see Figure 5-1	1 or 0
Positive / Negative	Outlier properties with prices largely deviating from the norm	1 or 0

of their internal structure when compared to their neighboring counterparts. Typically, we would have concluded that the positive or negative outliers were solely influenced by geographic location if they exhibited a clustered pattern. However, since they were scattered with no apparent spatial pattern, the nature of their internal characteristics caused their prices to deviate from the norm. Positive outliers (i.e. properties with significantly high prices) suggest that, other things being equal, the property must have better desirable features (e.g. flooring, internal architecture, larger developed space compared to lot size, availability of infrastructure such as sewage, etc.) when compared to neighbouring properties of the same size. On the other hand, negative outliers (i.e. properties with significantly low prices) suggest that, other things being equal, the property must lack the general desirable features found in most neighbouring properties of the same size and type.

4.2 Model 2: Land Development Type Choice

4.2.1 Background

Academic literature on land use often follows one or a combination of several categories of agents that are a part of and intrinsically effect the composition of urban form for a city. Specifically, this predominantly consists of the developers that prepare the land, and the residents and firms that move to occupy those lands. The results of their combined activity is an important outcome when noting that there are numerous facets of society that are influenced by them including but not limited to economic *(Páez, 2009)* and environmental impacts *(Anderson et al. 1996)*.

Therefore in an effort to better understand and adapt to the changing urban landscape, it is important to study characteristics of the various agents. While access to demographic and household data has led to a large number of studies on residential development, the number of papers studying non-residential development is much more limited. It is with this in mind that Model 2 studies the choice of land use for new construction of commercial and industrial land. More specifically, an analysis of the explicit choice of development type *within* commercial and industrial zoning is studied through the use of logit models. In order to achieve this, permit data on new construction from the City of Windsor is employed to create the models.

4.2.2 Discrete Choice Theory – Logit Models

Of the papers reviewed and discussed, land development models were either based on logit models, agent based modelling (ABM), cellular automata (CA) or a combination thereof. Typically the CA option reveals an apparent flaw in that it does not provide a strong statistical correlation with the underlying factors (*Hu and Lo, 2007; Kamyab et al., 2010*). Rather, the CA models reveal only potential outcomes based on the rules and parameters set for the simulation. Therefore a logit model is more appropriate in this case as it lends itself to not only predicting future outcomes but also studying the significance of the variables utilized.

While logit models are now fairly common, the use of a logistic curve dates back as early as the 18^{th} century. At that time, the curve was used primarily to model biological growth such as the human population (*Cramer, 2003*). *Cramer* notes that in the 1970's, the logit model was connected to discrete choice theory by Nobel Laureate *Daniel McFadden*, giving it a firm theoretical background and applying it to modelling the choice of destinations (*Cramer, 2003; McFadden, 1974*). Since that time, the logit model has gained popularity and is used frequently in the transportation field. As noted in the literature review, one of the more recent groups of integrated urban models are econometric models that incorporate the logit model into the modelling processes.

All three primary land use agents may be modelled using the logit regression – residential location (*Sener et al., 2011*), firm location (*Nguyen and Sano, 2010*), and land development (*Waddell and Ulfarsson, 2003*). The basic logit models at their core consist of two terms that result in the utility of an alternative for a decision maker. The first part of the equation is a deterministic term representing the observable utility for the decision maker. The second is a random error term of unobservable characteristics. In mathematical terms, the utility for an alternative can be written in its simplest form as:

$$U_{it} = V_{it} + \epsilon_{it} \tag{4.4}$$

Where U_{it} is the total utility of alternative i for individual t, V_{it} is the observable deterministic component, and ϵ_{it} is the random error component. This utility represents the value attributed to each alternative but is not necessarily explicitly based on physical finances. However, for non-residential properties where financial achievements are often the primary goal, the utility will be mainly representative of potential profit. For example, consider if the alternative with the highest utility for a specific parcel within commercial zoning was for retail development. This would indicate that for the owner / developer of the parcel, developing it for retail use will be the most profitable alternative and therefore the most likely to be chosen.

The error term is assumed to follow a pattern that is independently and identically distributed (IID). Generally this choice of pattern is between a normal distribution and a logistic distribution, the former being a probit model and the latter being a logit model. Though the two patterns are similar as shown in Figure 4-2, the logistic distribution function is mathematically simpler.

4.2.2.1 Logit Methodology

Several recent papers have divided groups to create more homogeneous clusters allowing the modeller to study differences between them. *Maoh and Kanaroglou (2007)* studied the degree of clustering on 13 different types of firms and performed a simultaneous autoregressive (SAR) model with a spatial lag parameter. *Song et al. (2011)*

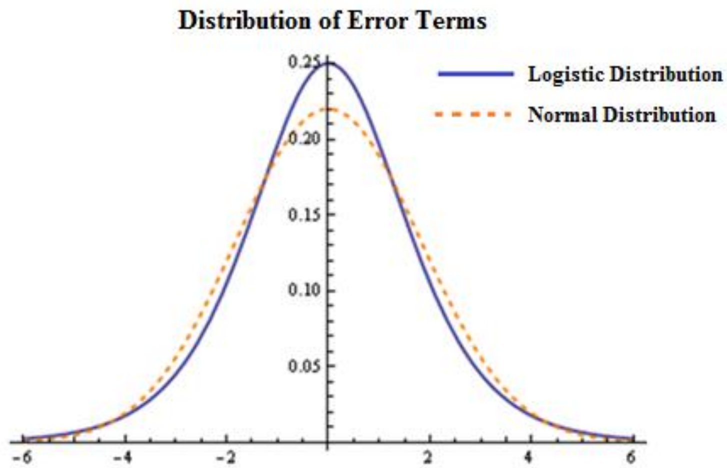

Figure 4-2: Distribution of Error Terms
(Source: Adapted from http://www.johndcook.com)

studied the impacts of accessibility for the locations of industrial agglomeration in Seoul using destination choice logit models of 12 groups of industries. This type of methodology with separate models for industry groups provides valuable information on significant spatial and demographic factors that can influence them. They can also provide information on differences arising from these varying groups. However, many these models provide these differences between industry preferences implicitly through comparisons. The models performed here provide a different methodology framework that views these differences explicitly. These models are based on the assumption that general zoning for a municipality is usually pre-determined or guided by master plans. The primary implication for these models then is that they explore the likely types of development that will be in demand within this pre-determined zoning.

Several main categories of logit models are used in modern times including: binomial / multinomial (*De Bok, 2009*), ordered (*Van Dijk and Pellenbarg, 2000*), nested

(Mataloni, 2011), and mixed *(Nguyen and Sano, 2010)*. All the types listed are variants of the multinomial logit model that is characterized by the equation:

$$P_{it} = \frac{\exp(V_{it})}{\sum_j \exp(V_{jt})} \qquad (4.5)$$

Where P_{it} is the probability of decision maker, *t*, choosing alternative *i* among all *j* alternatives and V_{it} is the observable utility.

The choice of logit model can be based on several factors. First, it is important that the logistic regression can effectively model the land development in the region appropriately. For example, two land development scenarios as stated by *Waddell and Ulfarsson (2003)* are the "use looking for a site" (destination choice) and "site looking for a use" as seen in Figure 4-3 and Figure 4-4, respectively.

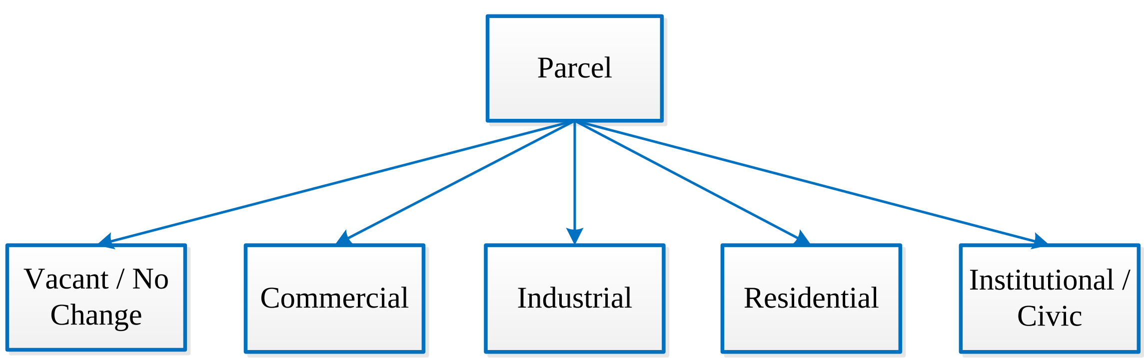

Figure 4-3: Site looking for a use

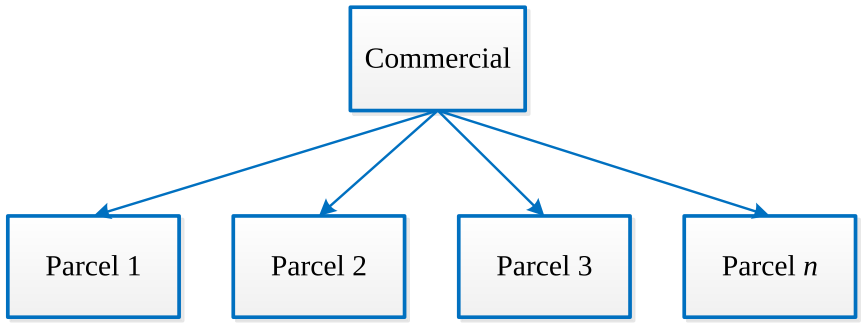

Figure 4-4: Use looking for a site

Both the destination choice model and its reverse could be chosen. However, an increased complexity would occur when choosing the destination choice methodology since the number of parcels available as alternatives could be very large. The site looking for a use scenario is simpler to implement in this case because the alternative list (the sub-types of development) is known and limited to a smaller set of choices. For this reason the "site looking for a use" methodology was subsequently chosen. The alternatives for the models are conditional on commercial or industrial zoning and their data counts are shown in Table 4-2. As it stands, the categories chosen give an optimal compromise between the level of detail and homogeneity in the groups and the size of the alternative groups.

Table 4-2: Logit Model Alternatives and Data Counts

Alternative no.	Commercial Model		Alternative no.	Industrial Model	
1	Office	35	1	Warehouse	20
2	Retail	47	2	Factory	18
3	Restaurant	18	3	Other	36
4	Other	19			
	SUM	**119**		**SUM**	**74**

4.2.2.2 Independence of Irrelevant Alternatives

Even with the general methodologies for the models known, further questions exist about the type of logit model. Specifically, the more basic model type (i.e. multinomial) must not violate the independence of irrelevant alternatives (IIA) property (*Mataloni, 2011*; *Cheng and Stough, 2006)*. This property is "Axiom 3" proposed in the paper by *McFadden (1974)* and assumes that all alternatives are completely independent of one another. Consider an example with 3 modes of travel – car, train and bus. Using a multinomial logit model, a change in train use would incur equal amounts of change in car and bus use. Realistically, however, the two modes of public transportation are not independent of each other and the change for bus use would be expected to far exceed that for cars. *Cheng and Stough (2006)* and *Mataloni (2011)* both test for the IIA property using a method called the Hausman-McFadden test. The test works by removing one (or a group) of alternatives from the model. In *Mataloni*'s case, the IIA property was violated and as such a nested logit model was used in lieu of the multinomial logit model. Typically, one may resort to the nested logit model when the IIA property is violated. This relaxes the assumption of independence between some of the alternatives.

4.2.2.3 Multinomial Logit

The first model presented here was the multinomial model using equations *4.4* and *4.5*. The decision maker in this model pertains to each individual parcel where development occurred (representing the decisions of the owner of the parcel). Here, four alternatives for the commercial properties and three alternatives for the industrial properties were used to form the choice sets, as shown in Figure 4-5 and Figure 4-6.

Qualitatively, this represents a choice for the owner to decide which type of development will be established on the parcel assuming that the more general type of commercial and industrial land use are already predetermined.

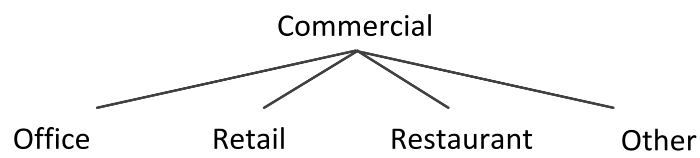

Figure 4-5: Multinomial Commercial Structure

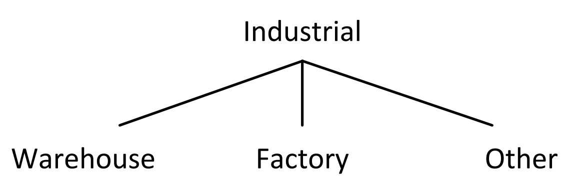

Figure 4-6: Multinomial Industrial Structure

4.2.2.4 Nested Logit

In addition to the multinomial logit, a nested logit with two levels was also created. The nested structure allows for multiple alternatives that share similarities to be grouped together. Careful selection of the structure helps account for the independence of irrelevant alternatives (IIA) assumption imposed when creating the model. Numerous configurations of the nested structure were selected and modelled for both the commercial and industrial cases, but was only significant in the commercial case. A sample of these configurations is given in Figure 4-7. The probability of the top nest for a given decision maker and alternative can be given as:

$$P_{jt} = \frac{\exp(V_{jt} + \delta_j IV_{jt})}{\sum_k \exp(V_{kt} + \delta_k IV_{kt})} \quad (4.6)$$

Where j is an alternative in the top nest, V_{jt} is the observable utility of alternative, δ_j is the inclusive scale parameter between 0 and 1 denoting the magnitude of effect for the inclusive parameter IV_{jt}. This inclusive parameter gives the total observable utility of all the alternatives in the lower nest that belong to alternative j in the upper nest. This can be shown as:

$$IV_{jt} = \ln[\textstyle\sum_i \exp(V_{it|j})] \quad (4.7)$$

Where $V_{it|j}$ is the observable utility of alternative i in the lower nest as a subset of alternative j. The probability for alternative i in the second nested level is similar to a standard multinomial logit:

$$P_{it} = \frac{\exp(V_{it|j})}{\sum_k \exp(V_{kt|j})} \quad (4.8)$$

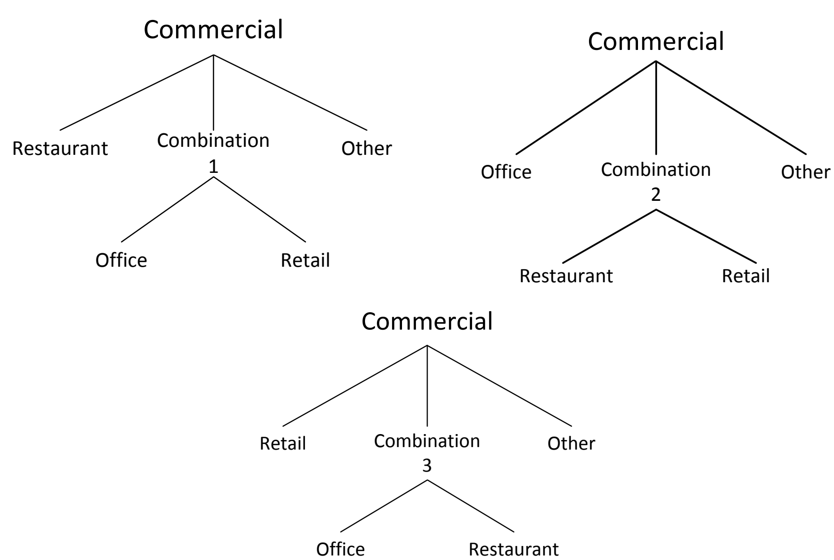

Figure 4-7: Sample Configurations of Commercial Nested Logits

4.2.2.5 Multinomial Logit with Spatial Effects

When creating any spatially oriented model, one key consideration is to be aware of spatial autocorrelation within the data. In recent times, several studies have proposed that there is a correlation between decision makers. More specifically for logit models, the alternatives chosen nearby will have an impact on the decision maker. Presumably, this impact will be a positive influence further increasing the likelihood of the decision maker choosing the same alternative leading to an inertia effect for particular choices. *Bhat and Guo (2004)* proposed the use of a Spatially Correlated Logit (SCL) model wherein the destination choice for a residential home is correlated with other lots that are contiguous. *Sener et al. (2011)* further expanded on this with a Generalized Spatially Correlated (GSCL) model that includes non-contiguous lots by incorporating a decay function to decrease their impact with increasing distance. Furthermore, *Mohammadian et al. (2008)* observed homebuilders as the decision maker choosing from several different housing alternatives to be built on a lot. Again, the parameter for correlation from the alternative choices of other decision makers had a significant positive impact. While the studies done by *Bhat and Guo (2004), Sener et al. (2011)* and *Mohammadian et al. (2008)* provide evidence of recent research into correlation among decision makers, their scope is limited to residential developments. Subsequently, this thesis looks to address this issue for commercial and industrial development to determine if similar conditions exist.

To this end, another model created was similar to the multinomial logit but included a spatial variable, S_i, associated with parameter lambda (λ). This parameter characterizes the extent to which the choice of the decision maker i is impacted by other nearby decision makers. Quantitatively, S_i is the sum of all rival decision makers j who chose the same alternative as the current decision maker. The values are weighted based on a negative exponential distance decay function. The S_i covariate is added to the model and behaves in the same manner as all other independent input variables, and can be defined as:

$$S_i = \sum_{j=1}^{n} X_{ij} \exp^{(-0.00037 d_{ij})} ; \quad X_{ii} = 0 \qquad (4.9)$$

Where S_i is a covariate based on a negative exponential distance decay function with X_{ij} = 1 when parcels i and j choose the same alternative (0 otherwise), and d_{ij} is the Euclidean distance between them in meters. The distance decay parameter associated with d_{ij} was set to a constant value of 0.00037. The latter was determined through trial and error in the models and proved to be the most significant. Based on this curve with the rate of decay decreasing with increased distance, the weight is decreased to 50% at a distance of roughly 1850 meters, as can be seen in Figure 4-8. A nested logit model with spatial effects was also created for the commercial group by combining both a nested structure and the spatial covariate, S_i. It should be noted that the distance decay parameter could be estimated empirically. However, this causes the logit model to rely on non-linear systematic utilities. The NLogit 4.0 software used in this thesis to estimate the models can only handle linear-in-parameter utilities.

While the mixed logit was modelled for this thesis, the results indicated that the model could not improve upon earlier results. For instance, set distributions (normal, logarithmic, uniform, triangular) were attempted to capture variations in the beta coefficients but proved to be insignificant and failed to increase the overall ρ^2 results.

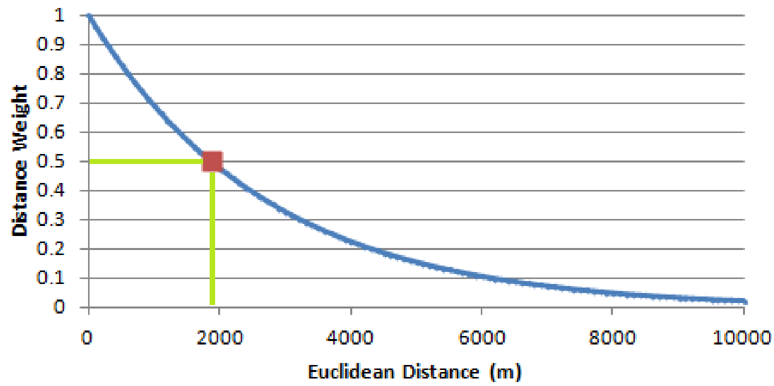

Figure 4-8: Distance Decay Function for S_i

4.2.3 Parameters

Several major types of attributes were utilized for the development type choice models as alternative specific variables. The first represents spatial parameters quantified by either the distance/time from the decision maker`s property to the spatial feature or a dummy variable where the property is denoted as one when within a specified buffer zone of the spatial feature and zero otherwise. This includes the Rhodes and Tunnel covariates representing areas found to exhibit a high propensity for certain land use development types. In addition, real estate prices from the Windsor Essex Real Estate Board and demographic census data from Statistics Canada were included. For these models, the 2006 census was used due to the 2005-2011 time period of the development permits. All parameters included in the final models are given in Table 4-3. Similar to the covariates included in the price regression models, a sensitivity analysis using multiple buffer values was performed (i.e. 200 m, 400 m, 600 m) to determine the area of effect that best captures the utility from various spatial locations.

Table 4-3: Significant Covariates in the Logit Models

Variable	Description
CBD Time	Time from center of CBD to property (Minutes)
$Transit_{200m}$	Dummy variable for 200 m buffer around transit lines
ECR_{400m}	Dummy variable for EC Row Expressway 400 m buffer
$Rail_{600m}$	Dummy variable for properties within 600 m buffer of rail lines
Tecumseh Rd_{200m}	Dummy variable for properties within 200 m buffer of Tecumseh Rd.
AP Ratio	Area to Perimeter ratio of lot geometry
Rhodes	Dummy control variable for Rhodes Drive
Tunnel	Dummy control variable for area south of international tunnel crossing
Com. $Sold_{200m}$	Absolute quantity of commercial properties sold within 200 m buffer of property (1996-2005)
Ind. Sold $Prop_{400m}$	Proportion of industrial properties sold over listed within 400 m buffer (1996-2005)
Median	Median personal income by census tract (2006)
Com. Price	Real estate price per square foot (all commercial properties sold minus office) extracted from kriging interpolation results
Ind. Price	Real estate price per square foot (all industrial properties) extracted from kriging interpolation results
$OC_{96\text{-}00}$	Proportion of occupied dwellings built between 1996-2000 over all occupied dwellings by census tract (2006)
S_i	Spatial correlation parameter for the influence of other decision makers choosing the same alternative
δ	Inclusive parameter for the interaction between tiers in the nested models

4.2.4 Kriging Surface Interpolations

Kriging interpolations were used with the assistance of the Geostatistical Analyst in ArcGIS to create the price surfaces used as independent variables for Model 2. While there are several options for surface interpolations using ArcGIS such as inverse distance weighted (IDW) and local polynomial, the kriging method was chosen for several reasons. The first is the general shapes created by kriging. For example, consider recent real estate prices based on active MLS listings[2]. Using this data, surface maps shown in

[2] Online listing data gathered and geo-referenced by Kunal Gulati in 2011

Figure 4-9 were interpolated using two methods: IDW and kriging. The three-dimensional views allow for a better visualization of the surface properties. The IDW surface on the left shows a very smooth surface but with several large peaks, located at points with high outliers. The peaks found in the IDW method are due to the surface exhibiting the exact price at the location of each known point. For a subject such as real estate prices, this surface will be unrealistic. The kriging surface on the right shows a rougher surface but does not have pronounced peaks. Another reason why kriging was used is because it is a more statistically oriented method relying on analysis of the data.

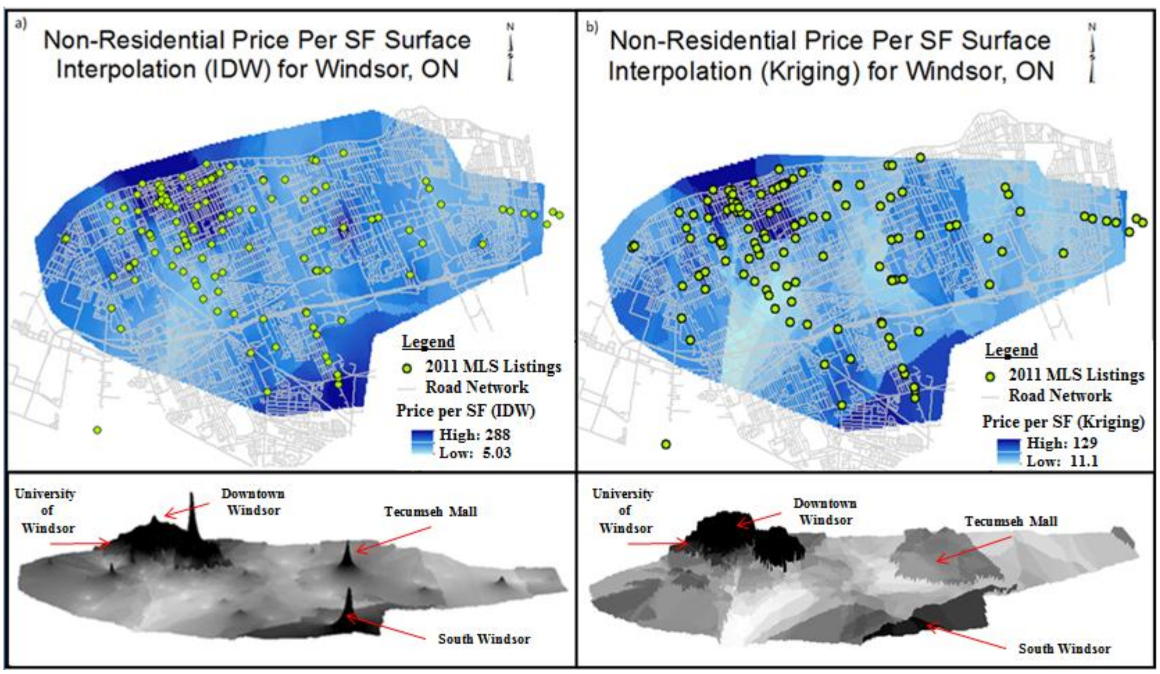

Figure 4-9: IDW vs. Kriging Surface Interpolations

The kriging function assumes that the data is normally distributed. In order to accommodate this distribution type, it was necessary to explore the data and transform it. For instance, a histogram for commercial properties is shown in Appendix B to be positively skewed. Moreover, taking the natural logarithm creates a distribution closer to normality, but a negative skew is created instead. As an alternative, Box-Cox transformations were able to better demonstrate a distribution close to normal. The Box-Cox transformation is a power relationship with the exception being a basic logarithm when its parameter, λ, is equal to 0 as shown below:

$$y_i' = \frac{y_i^{\lambda} - 1}{\lambda} \; ; \lambda \neq 0 \qquad (4.14)$$

$$y_i' = ln(y_i) \; ; \lambda = 0 \qquad (4.15)$$

Where λ is the Box-Cox parameter, y_i, is the original variable being transformed (property prices in this case), and y_i' is the final transformed value. In the case of commercial values, the parameter providing the greatest fit was 0.31 but this value varied slightly depending on the land use category ranging from 0.08 to 0.49. To ensure that the distribution is close to normal, QQ plots were also generated in ArcGIS with both original and transformed data. The y-axis contains the values associated with specific quartiles in the dataset while the x-axis gives the value based on a normal distribution for the same quartile. Data that follows a normal distribution will show points following a linear relationship while non-normal distribution will show deviations from the line as shown in Figure 4-10.

In ArcGIS, the given transformation is first applied to the dataset before calculations are performed, then transformed back to the surface at the end of the process for ease of use. Another possible step for the kriging interpolation is to remove global trends from the data to better account for local variations. Similar to the transformation to a normal dataset, the removed trend is later reapplied to the final surface automatically.

Finally, kriging relies heavily on spatial autocorrelation. In order to quantify this in the data, the software uses semi-variogram plots with distance and semi-variance between properties on the x-axis and y-axis, respectively. The trend within these semi-variograms provides a necessary function used to calculate the weights of each nearby property when creating the surface. To optimize this function for kriging, ArcGIS also includes an automatic tool that was used to calibrate the data to provide the best overall surface fit. Given the complexities involved in creating a properly specified interpolation surface, background information on the capabilities possessed by ArcGIS for kriging was obtained through tutorials prepared by ESRI as a guide (ESRI Inc., 2010).

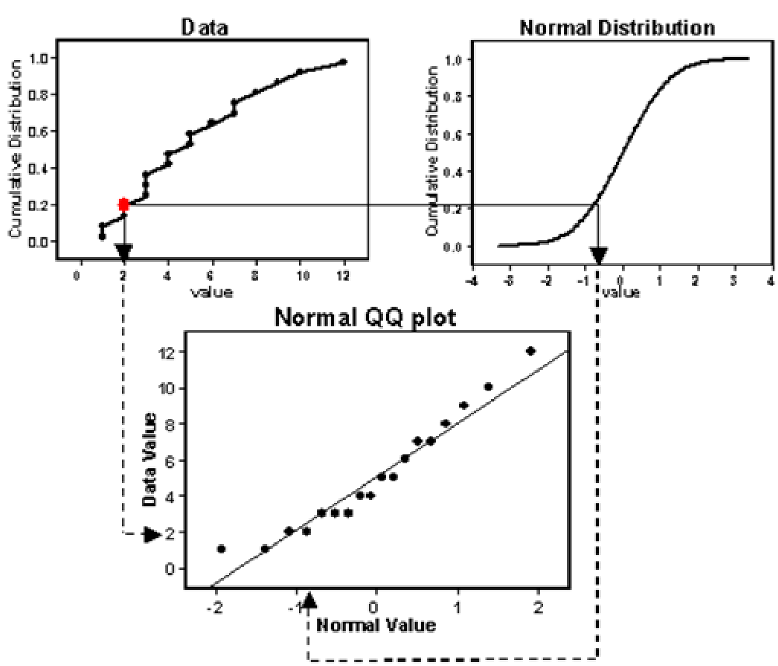

Figure 4-10: Normal QQ Plots
(Source: ESRI Inc., 2010)

4.3 Software Programs

Creation of the models presented here was made by possible through the use of a variety of computer software programs, all of which had valid student/educational licenses. First, Microsoft Office was used extensively through the use of Office for typed documents, Excel for data manipulation, and finally PowerPoint and Visio for presentation purposes.

Due to the highly spatial nature of the data used, ArcGIS was used for both basic functions such as distance buffers and also more complex purposes such as the creation of road networks using the Network Analyst and kriging surface interpolations with the Geospatial Analyst. Exploring the spatial data was also performed with GeoDa. More specifically GeoDa was used to study possible spatial autocorrelation for data used in Model 1 and to perform the spatial lag regression as a possible alternative to the linear OLS regression models.

Finally, the logit specifications in Model 2 were created using the commercial software Limdep. In addition, Limdep was also used to create several statistical measures for the regressions in Model 1 and adjust the t-statistics due to the presence of heteroscedasticity.

5 RESULTS AND DISCUSSION

5.1 Model 1: Price Regression Analysis

5.1.1 Original Results

Initial regressions were first conducted between the dependent price variable and each independent variable separately. This allowed for a determination of variables that were more likely to be significant though other variables would still be tested throughout the model building process. Results of these individual regressions are shown in Appendix C.

An analysis of the original regressions based off equations *4.1* and *4.2* revealed R^2 coefficients varying among the models considerably from 0.18 to 0.69. To improve on these results, SAR models using equation *4.3* were attempted. Unfortunately among all land use types the results of the spatial models did not increase the fit or significance. Moreover, the spatial lag parameter, ρ, for equation *4.3* was not significant.

Adjusting for extreme outliers within the data proved highly beneficial. The dummy variables for the outliers used to do this were found to be significant when applied to values whose residuals were greater than one standard deviation from the mean. The lowest R^2 in the final models was in the commercial group (which contains a large degree of heterogeneity within the group) with 0.73 while both the plaza and vacant models were greater than 0.90. Final results are shown in Table 5-1. The original R^2 values are given Table 5-2. Additionally, the areas bounded by the location dummy variables are given in Figure 5-1.

Table 5-1: Regression Results

Variables	Commercial	Retail	Office	Food	Plaza	Industrial	Vacant
Intercept	41.549 (21.94)	3.795 (86.68)	-43.835 (-2.37)	3.700 (49.40)	-0.532 (-2.37)	1.558 (14.46)	2.125 (12.71)
Lot Size*	-15.673 (-4.85)	-1.050 (-9.16)	-1.714 (-5.87)	-1.622 (-6.58)	-	-0.307 (-18.91)	-0.185 (-6.10)
CBD Time	-	-0.014 (-6.17)	-	-0.011 (-3.32)	-	-	-
Ln(CBD Time)	-3.108 (-3.79)	-	-	-	-	-	-0.227 (-4.89)
Pot. Accessibility*	-	-	-	-	2.567 (14.76)	0.690 (8.40)	-
CBD $_{200\ M}$	35.798 (10.89)	-	-	-	-	-	-
CBD $_{400\ M}$	-	-	0.255 (1.80)	-	-	-	-
Rail $_{200\ M}$	-	-	-	-0.480 (-3.36)	-	0.099 (1.86)	-
Rail $_{400\ M}$	-8.248 (-5.65)	-	-	-	-	-	-
Transit $_{200\ M}$	-	-	-	-	-	-	0.626 (5.27)
Ramp $_{800\ M}$	-	-	-	-	-	-	-0.684 (-4.38)
Coast $_{400\ M}$	13.571 (4.81)	-	-1.057 (-4.31)	-	-	-	-
Auto $_{600\ M}$	-	-	-	-	-	0.574 (4.32)	-
Year Sold	-	-	0.024 (2.56)	-	-	-	-
Sandwich	-	-0.389 (-2.69)	-	-	-	-	-1.025 (-5.30)
Leamington	8.562 (2.76)	0.533 (4.32)	-	-	2.230 (8.48)	1.611 (11.04)	1.240 (5.00)
Location 1	30.311 (7.76)	-	-	-	-	-	-
Location 2	-38.794 (-9.97)	-	-	-	-	-	-
Location 3	-23.675 (-3.38)	-	-	-	-	-	-
Location 4	-	0.936 (11.60)	-	-	-	-	-
Location 5	-	-	-	-	-	-2.352 (-8.79)	-
Location 6	-	-	-	-	0.687 (6.20)	-	-
Positive	64.201 (17.19)	0.960 (14.18)	1.398 (6.04)	1.121 (12.64)	0.630 (5.52)	1.073 (18.62)	0.953 (8.42)
Negative	-39.621 (-22.09)	-1.205 (-14.16)	-2.479 (-8.24)	-1.929 (-5.87)	-0.690 (-5.89)	-1.235(-12.96)	-1.163 (-4.94)
Dependent Variable	P	Ln(P)	Ln(P)	Ln(P)	Ln(P)	Ln(P)	Ln(P)
Observations	730	198	126	92	30	272	63
R^2	0.727	0.818	0.786	0.772	0.932	0.802	0.925
Adjusted R^2	0.723	0.811	0.775	0.759	0.918	0.796	0.914
F	174.24	121.83	72.93	58.21	66.3	132.82	83.06

Values shown in the following format: β (t-stat); * Parameters are x 10^{-5}

Figure 5-1: Location Control Variables

Table 5-2: Original Regression Fit

Land Use	Original R^2
Commercial	0.180
Retail	0.268
Office	0.234
Food	0.293
Plaza	0.661
Industrial	0.347
Vacant	0.689

5.1.2 Analysis of Final Results

The results suggest a variation in land prices by geographic location. Like other Canadian cities, certain parts of Windsor have developed into attractive areas while others lost their potential for sustaining businesses and clients over time. For instance, properties in locations 1, 4 and 6 (as shown in Figure 5-1) enjoy higher prices, other things being equal. These locations demark key commercial corridors in the city (e.g. Ottawa street, Dougall Ave. and Ouellette Street in Downtown). On the other hand, properties in the northwestern area of Windsor have lower land prices. This area houses a lower income population, which in turn has a negative impact on retail. Also, vacant land in this area tends to have lower prices when compared to vacant land elsewhere in the region.

5.1.2.1 Proximity to CBD

Several variables that are related to the CBD were also significant regressors in all of the models as demonstrated in Figure 5-2. The direct time to reach the CBD proved to be negatively significant in the majority of models, offices being the sole exception (no significance). The sign was as expected due to the way this variable was measured – properties located further away from the CBD had larger values as expressed in minutes. Thus, a negative trend indicates a loss of value for properties located further away. This is not a new finding and has a strong foundation based on historical bid rent theory documented for several centuries *(Shieh, 2003).*

Another significant variable tied closely to the CBD is potential accessibility. Measured in terms of potential to the residential population, it is shown as positively significant for two models; plaza and industrial listings. Originally, however, the variable was significant for the majority of models, but was later removed from most due to high correlations with the CBD time variables. Dummy variables for properties in close proximity to the CBD were also created to control for higher prices within that part of the region for the commercial and office models. Moreover, the retail model also had a nearby area that was located close to the CBD.

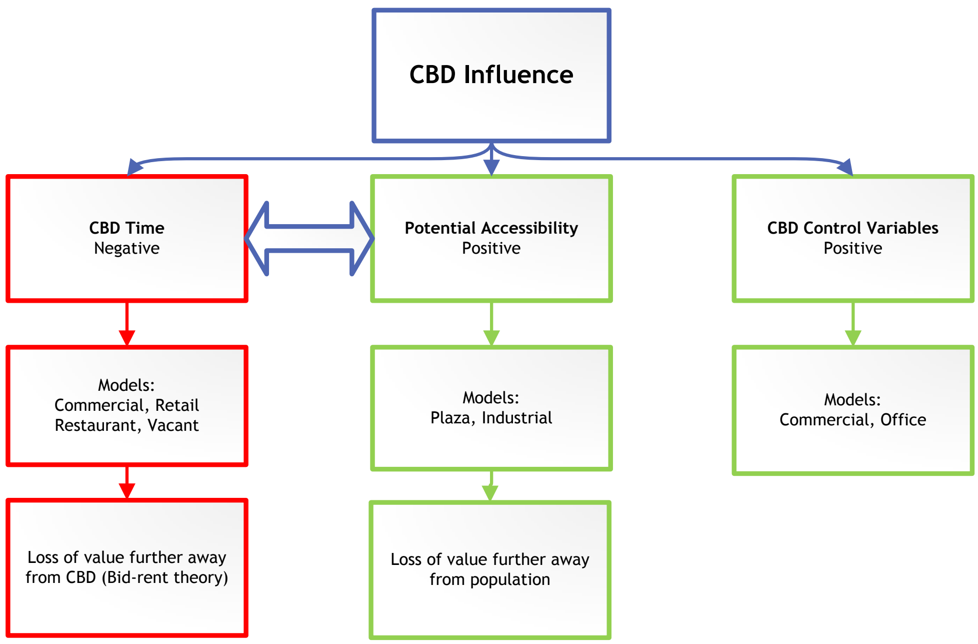

Figure 5-2: Price Regression Results – CBD

5.1.2.2 Transportation Infrastructure

In light of the levels of interaction between land use and transportation, many location factors involve transportation infrastructure. Previous studies that have addressed this issue have mixed results *(e.g. Du and Mulley, 2006)*. In our models, proximity to railway lines showed mixed results with industrial and commercial listings increasing and decreasing the price, respectively. This is not surprising when noting that industrial properties are more likely to gain utility from access to nearby rail compared to their commercial counterparts. Proximity to public transit was also found to be significant but only for vacant lands. This could be a result of the potential accessibility for customers/workers that this public infrastructure attracts. The effect that transit has may be lower than in other cities though, as only an estimated 3% of the Windsor population makes use of public transit as of 1999 *(City of Windsor, 1999)*. Variables representing direct proximity to highways, however, were generally not shown to be significant except for a negative correlation with vacant lands. By contrast, *Dunse et al. (2005)* found that proximity to a highway junction was a heavily significant and positive impact, even compared to the influence of the CBD. The result found here could be a product of the scale used for these models. A larger macroscopic model that envelopes a larger area such as a province or state may be able to better describe the effect of highways.

Several indirect measures of transportation were also found to be positively associated with land prices. As mentioned, both the travel time to CBD as well as potential accessibility showed a positive impact on prices closer to the CBD itself. Underpinning these variables is the transportation network that accommodates accessibility into the central areas of the city. Therefore improvements in the mobility of vehicles through transportation investments would allow for even greater access to the CBD, increasing its impact further.

5.1.2.3 Miscellaneous Results

Another variable that stands out in the results is the lot size. Similar to the results found by *Ten Siethoff and Kockelman (2002)*, this attribute was negative in six out of the seven models. Several reasons could explain this. First, a larger lot can cost more money

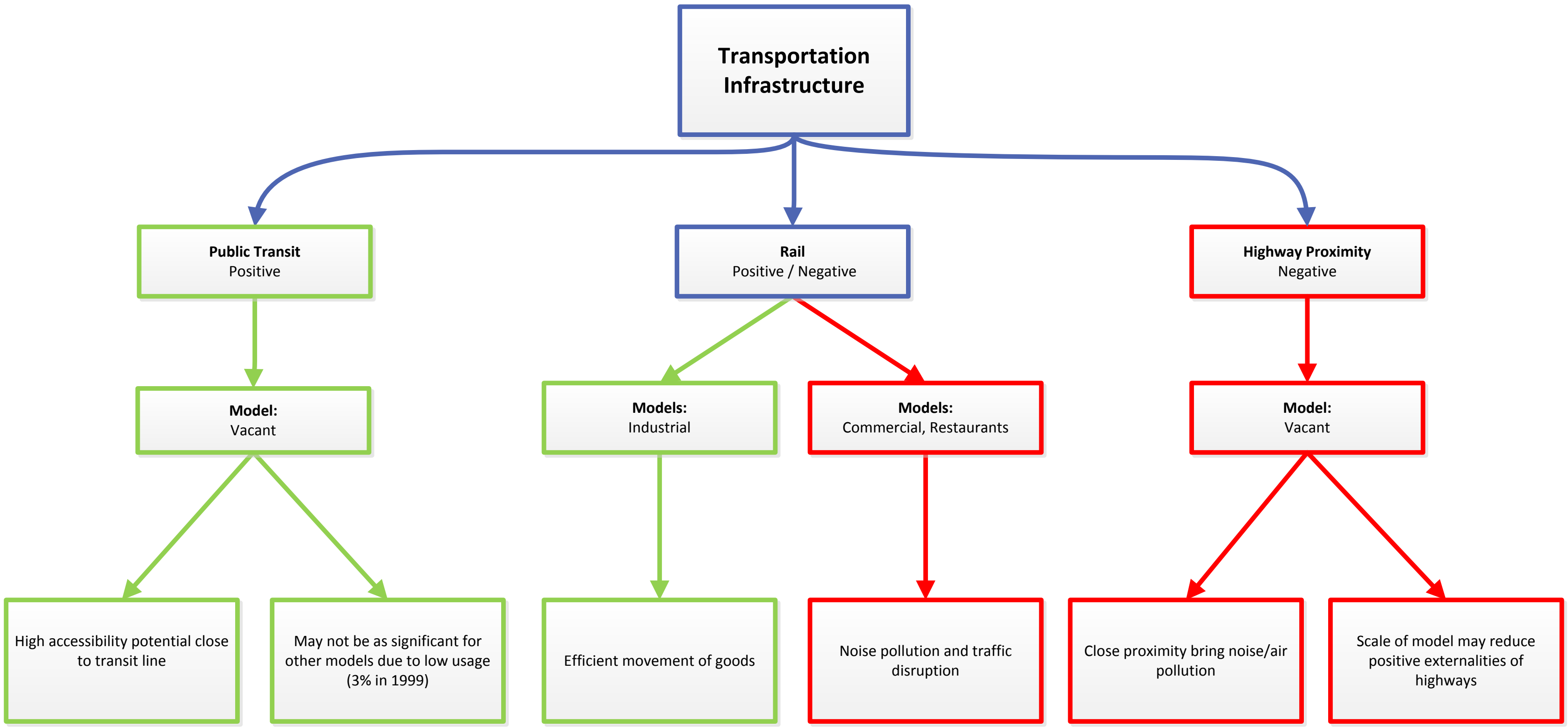

Figure 5-3: Price Regression Results - Transportation

to the buyer in several respects: a greater cost for connecting to public utilities as well as paying for those utilities, in addition to increased maintenance needs. A second possible explanation is that a larger lot will receive a discount per unit area because of its larger size. In essence this would be a bulk discount as seen when buying large quantities of other consumer goods. The lot size variable also tended to be one of the most significant contributors to the models' overall fit when analyzing the variables individually. The only exception to this was the plaza model. This could be due to the importance that plazas place on their size. While most types saw a per unit discount for larger lots, commercial plazas can increase the attractiveness and visibility to potential consumers.

It was also found that industrial properties within a 600 meter buffer around the three major automotive plants in Windsor was a positive influence on prices. This parameter reveals the influence of agglomeration due to the automotive sector. Finally, proximity to the coast was also found to be significant. While this resulted in an increase of prices for all commercial models grouped together, it was found to have a negative influence specifically on office properties. This could be the result of the lack of a need for office locations to attract retail customers.

5.1.3 Spatial Autocorrelation

Preliminary investigation of the data included exploring spatial autocorrelation in prices through the Moran's I (*MI*) statistic, which is specified as:

$$MI = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(P_i-\bar{P})(P_j-\bar{P})}{\left(\sum_{i=1}^{n}(P_i-\bar{P})^2\right)\left(\sum\sum_{i\neq j} w_{ij}\right)} \qquad (5.1)$$

Where $\bar{P}$ is the mean value for n individual dependent prices, and w_{ij} is an element in matrix W having a value 1 if property i is adjacent (neighbour) to property j, 0 otherwise.

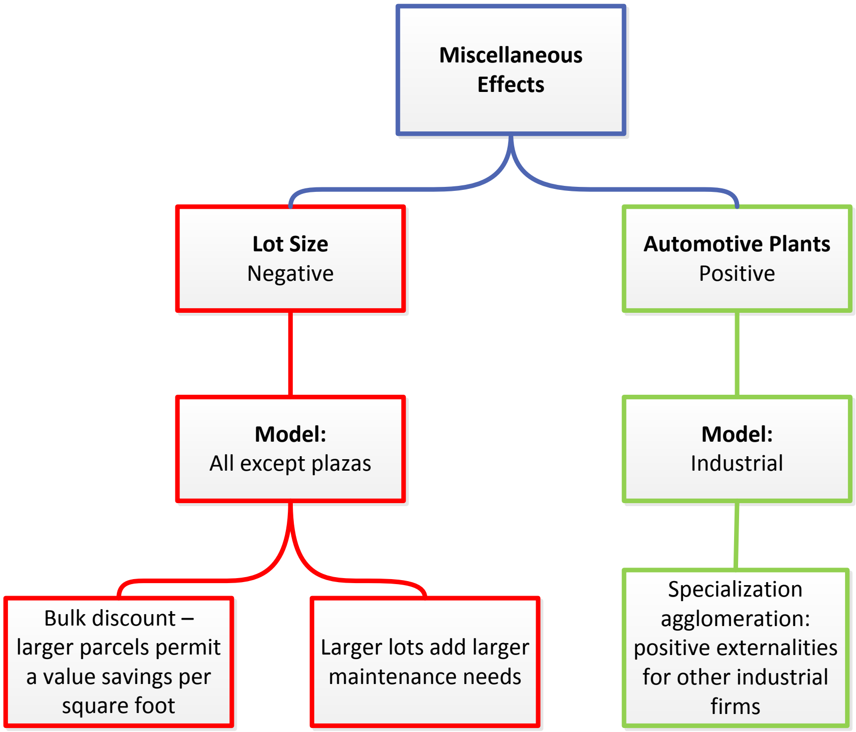

Figure 5-4: Price Regression Results – Miscellaneous

The initial *MI* tested positive for five of the models; an indication of positive spatial autocorrelation between land prices for nearby listings. For example, *Anselin*'s Moran scatter plot is shown in Figure 5-5 with the numerator of equation *5.1* along the y-axis and the denominator along the x-axis. The slope of the scatter plot gives a *MI* of 0.33 for vacant land (all values are shown in Table 5-3). To test the significance, a number of permutations (999 in this case) representing random draws are calculated using Monte Carlo simulations *(Anselin et al., 2006)* to create an empirical distribution. The addition of the actual *MI* is then included in the distribution and the subsequent probability is measured. The results from these permutations were found to be highly significant with a p-value of 0.015 for the vacant case. Restaurant and plaza models were the only ones to

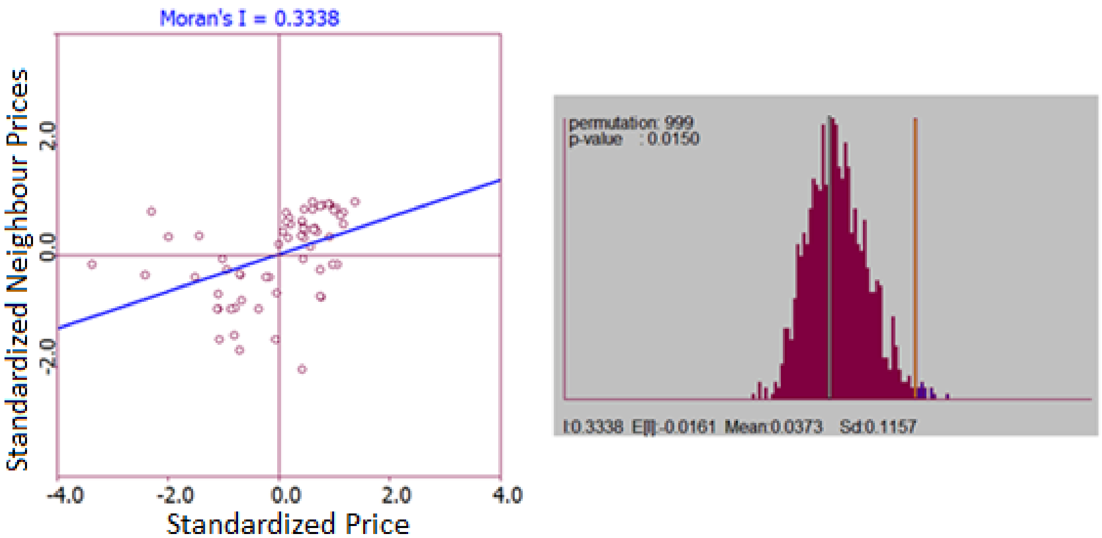

Figure 5-5: Initial Moran`s I results for vacant land using GeoDa

not indicate statistically significant positive spatial autocorrelation. The *MI* statistic was also calculated on the error term following the conclusion of the regression models and given as *MI* (after regression) in Table 5-3. In this case the *MI* was only statistically significant for two models, commercial and office. The difference indicates that the independent variables used for modelling were able to account for some of the spatial autocorrelation found in the initial data.

To further verify the presence of spatial autocorrelation in the error term, a second measure known as the Durbin-Watson (*DW*) test was also prepared. In this case, the only two models that show significant spatial autocorrelation are the commercial and industrial models. Comparing these results with the *MI* (after regression) values, it can be seen that only the commercial model shows significant signs of positive autocorrelation in both tests. Industrial and office models show signs of positive autocorrelation in one test each. The others indicate no presence of spatial autocorrelation.

5.1.4 Homoscedasticity

An assumption made when performing the regressions is that the variance of errors remains constant throughout all independent variable data points - known as homoscedasticity. If this assumption does not hold, t-stat values inferring significance in the models can be erroneous. A test that can indicate the presence of heteroscedasticity in the data is the *Jarque-Bera* (*JB*) test. While not a direct measure, the *JB* test measures if the data follows a normal distribution by using skewness and kurtosis. The *JB* statistic can be given by:

$$JB = n\left(\frac{S^2}{6} + \frac{EK^2}{24}\right) \tag{5.2}$$

Where n represents the number of observations, S and EK are the skewness and excess kurtosis attributes of the distribution, respectively. They can be written as:

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}} \tag{5.3}$$

$$EK = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{2}} - 3 \tag{5.4}$$

Where x_i variables represent the data points for each covariate and $\bar{x}$ is the mean.

In reality, the equations above are also adjusted for small sample sizes. For this measure, the null hypothesis, H_0, is true if the distribution of data is normal. Five out of the seven models show a non-normal distribution in the data. A more direct measure for the presence of heteroscedasticity would also be useful. The *Breusch-Pagan* (*BP*) Test is suited for this role, specifically testing for this phenomenon. The results of this test ended up similar to the *JB* test, though retail and vacant models were both significant in only one of the two.

To remedy the violation, a heteroscedasticity consistent covariance matrix was used to standardize the variance in the errors and provide a better representation of the t-values. Based on previous findings using this matrix *(Long and Ervin, 2000)*, sample sizes less than 250 should use a variant of the matrix named HC3 and was subsequently used on all models except commercial and industrial.

5.1.5 Multicollinearity

Another assumption embedded in the linear regression model is that there is no multicollinearity. That is, the regressors used to determine the dependant variable should ideally be completely independent of each other. While a significant amount of multicollinearity will not affect the overall fit, the significance of individual variables becomes doubtful *(Paul, 2008)*.

In order to be aware of any multicollinearity, several techniques were used. During the compiling of models, the procedure involved adding individually significant variables one at a time. Careful scrutiny of the effect that an added variable would have on others was conducted. Any significant change in the beta coefficient or significance of the regressors was a strong indication for the presence of correlations between them. For example, consider the two variables CBD time and potential (residential) accessibility. The former is a measure in minutes of the time from the CBD to the property. The potential accessibility as a measure of the residential population generally coincides with the former variable since denser populations tend to live closer to the CBD. It could be clearly seen while building the models that the two variables were highly correlated with one another. To account for this, the two would be modelled together and one variable would typically be more dominant (measured in terms of statistical significance). Thus, the dominant variable would be chosen. Four out of the seven models favoured the CBD time (or the natural logarithm of the CBD time) while two favoured the potential accessibility measure.

A second, more quantitative method is the multicollinearity condition number (κ). Since no multicollinearity would indicate that two variables are orthogonal, the κ variable measures the amount of correlation within the entire model by taking the ratio of the greatest eigenvalue ($\lambda_{Largest}$) over the smallest ($\lambda_{smallest}$). In mathematical terms this can be easily expressed as:

$$\kappa = \frac{\lambda_{Largest}}{\lambda_{smallest}} \tag{5.5}$$

While the threshold for significance with this value varies based on source, a κ of less than 100 indicates a relatively small amount of multicollinearity. Values between 100 and 1000 are considered to have a moderate amount while values greater than 1000 are considered to be very significant *(Paul, 2008)*.

The κ values for all seven models are shown in Table 5-3. Six models held values of 16 or less which shows a very minimal amount of correlation between the independent variables. The office model, on the other hand, exhibited a κ of 941. While still in the moderate range, it is close to being in the highly significant category. A further exploration of multicollinearity in the regression model for office listings was conducted with a correlation matrix on the covariates. The $Coast_{400M}$ variable was found to exhibit some correlation with the Positive and Negative variables introduced to account for the lack of hedonic property attributes suggesting a large degree of variation among properties along the coast.

Table 5-3: Statistical Measures

Variables	Null Hypothesis (H_0)	Comm.	Retail	Office	Food	Plaza	Industrial	Vacant
Multicollinearity Condition Number	No Multicollinearity at 0	7.211	4.163	941.1	3.320	16.07	10.98	9.42
Jarque-Bera Test	Normal Distribution	1032 (0.00)	6.204 (0.04)	35.53 (0.00)	97.19 (0.00)	0.518 (0.77)	6.378 (0.04)	0.739 (0.69)
Breusch-Pagan Test	Homoskedasticity	331.1 (0.00)	7.949 (0.34)	29.89 (0.00)	50.76 (0.00)	2.639 (0.76)	17.11 (0.03)	17.25 (0.03)
Moran's *I* (Before Regression)	No Spatial Autocorrelation	0.413 (0.00)	0.313 (0.00)	0.240 (0.02)	-0.051 (0.21)	0.250 (0.09)	0.229 (0.00)	0.334 (0.02)
Moran's *I* (After Regression)	No Spatial Autocorrelation	0.123 (0.00)	-0.016 (0.98)	0.142 (0.02)	-0.058 (0.63)	0.099 (0.14)	0.036 (0.25)	0.273 (0.12)
Durbin-Watson Test	No Spatial Autocorrelation at 2	1.744 (*)	1.804 (†)	2.198 (†)	2.259 (†)	2.379 (†)	1.317 (*)	1.994 (†)

Values shown in the following format: β (t-stat)

* / † = statistically / not statistically significant based on Durbin-Watson charts for p = 0.05

Table 5-4: Durbin-Watson Test Results

Variables	Commercial	Retail	Office	Food	Plaza	Industrial	Vacant
Durbin-Watson Test	1.744	1.804	2.198	1.932	2.379	1.317	1.994
Negative DW	N/A	N/A	1.802	N/A	1.621	N/A	N/A
Observations	730	198	126	92	30	272	63
Regressors (no intercept)	11	7	6	5	5	8	8
Upper Bound	1.933	1.832	1.803	1.776	1.606	1.92	1.894
Lower Bound	1.907	1.637	1.55	1.542	0.877	1.894	1.298
Autocorrelation	Positive	None	None	None	None	Positive	None

5.1.6 Validation

In order to check the validity of the model itself, two variations of methodology were used. For larger models (commercial, retail, industrial), a sample of data was removed from the set, similar to *Case et al. (2004)*. For this project, a 10% random sample was removed and the models were then recalibrated on the remaining 90% random sample. Next, the prices of the properties pertaining to the 10% sample were estimated and compared to the observed prices. For the remaining models, the datasets were considered too small for an unbiased random sample. In such cases, removal of each listing individually was performed in lieu of the 10% sample as done by *Montero-Lorenzo et al. (2009)*. As shown in Appendix D, the linear trend line in each case reveals slopes between 0.95 and 1.13 where 1 would be considered the most optimal (direct linear relationship between observed and predicted values). In light of the inherent randomness and unaccounted effects in real estate prices, these results indicate well behaved models.

It is worth mentioning that in all cases, the outlier dummy variables were assumed to be known and were included in the calculation of the estimated property prices. From a practical point of view, the estimated models can be used in a predictive sense if any given property is classified into one of three types: average, inferior or superior. The classification would be based on the internal structure of the property. For instance, a property with regular internal characteristics would be considered average. By comparison, a property with exceptional internal characteristics would be considered superior while a property with poor internal characteristics would be considered inferior.

This situation can be characterized by the following three regression equations in which $\beta_i X_i$ pertain to the set of spatial variables in Table 5-1, excluding the positive and negative dummies:

Average property price model: $\mathrm{Ln}(P) = \beta_0 + \sum_{i=1}^{n} \beta_i X_i$ *(5.6)*

Superior property price model: $\mathrm{Ln}(P) = (\beta_0 + \beta_{positive}) + \sum_{i=1}^{n} \beta_i X_i$ *(5.7)*

Inferior property price model: $\mathrm{Ln}(P) = (\beta_0 - \beta_{negative}) + \sum_{i=1}^{n} \beta_i X_i$ *(5.8)*

In the average property price model, the constant β_0 captures the effect of missing variables that cannot be easily measured. For example, factors such as negotiations between the buyer and seller will affect the price but this action cannot be quantified. By comparison, the superior property price model will have a further adjustment to the constant β_0 through the effect captured by the $\beta_{positive}$ parameter as described earlier. Consequently, the predicted price for those properties will be higher than their average counterparts, other things being equal. The same could be said about the inferior property price model where the constant is deflated by the $\beta_{negative}$ parameter.

It should be mentioned that in several of the validation charts it can be seen that the price dips into negative values for a few observations. For most models this is the result of a log transformed dependant price variable. However, this also happens for the commercial model, where the dependant variable is the actual price per unit area. For this model, one outlier value was observed close to zero. Consequently, the regression predicted this point as slightly negative. This does not represent the market realistically and if the model is used to estimate prices in the future a minimum value would need to be set to avoid negative predictions.

5.2 Model 2: Land Development Type Choice

5.2.1 Overall Results

Results from the logit models are given in Table 5-5. Intercept values (constants) were not found to be significant or enhance the commercial and industrial models. This indicates that the utility functions were able to capture the propensity of each alternative with the specified covariates already included in the models. Therefore the constants were removed from the final results. While the mixed logit was created for the commercial and industrial models, no significance was found in the distributions for beta variables. The commercial dataset found the original multinomial logit model to have a ρ^2 of 0.22. Moreover, the nested model was found to significantly increase the model fit with a ρ^2 of 0.31. In this case, the nested structure followed the framework shown in the top left of Figure 4-7 with office and retail properties located within the same branch and both the restaurant and other alternatives located within their own branches. Moreover, the inclusive scale parameter was also found to be significant indicating that the tiered structure is a suitable configuration. This reveals that office and retail developments behave similarly to some extent, while restaurants are independent despite sharing with retail development the need to sell finished goods.

Including the spatial parameter within the commercial model also increased the fit of ρ^2 over the multinomial equivalent to a value of 0.24. The variable was found to be positively significant for the three groups (office, retail, and restaurant) excluding the other category. This indicates that the decisions of nearby individuals have a positive influence. This coincides with similar results found in models for residential land use *(Maoh et al 2012; Mohammadian, 2008)* and other areas such as public school boundary consolidation *(Parent and Brasington, 2012)*. Finally, the nested and spatial models were combined and created the most significant model with a ρ^2 of 0.33. Due to the non-linear nature of the logit model, the ρ^2, known as the pseudo-R^2, is not equal to the R^2 seen in the earlier regression models. However, the relationship is empirically known (*Hensher et al., 2005*) and illustrated in Figure 5-6. Therefore the ρ^2 value of 0.33 is comparable to 0.65 – 0.70 for a linear R^2.

Table 5-5 also shows the results for the industrial case. The multinomial logit model here is shown to have a ρ^2 of 0.30. While the spatial parameter was not found to be statistically significant to 90% within the model, the adjusted ρ^2 increased indicating that the model is still an improvement over the MNL logit. The nested logit model saw a decrease in ρ^2 to 0.27 in addition to an insignificant inclusive scale parameter greater than one indicating a poorly structured model. With only three total alternatives, the lack of improvement is not a surprising result.

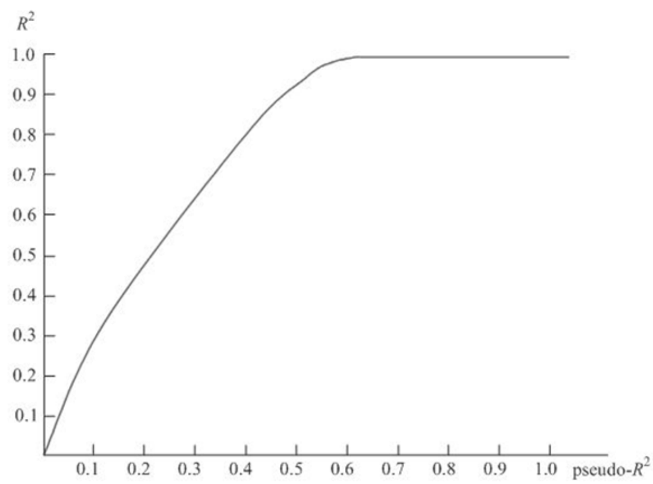

Figure 5-6: Relationship Between R^2 and Pseudo R^2
(Source: Hensher et al., 2005)

Table 5-5: Logit Model Results

Variables	Utility	MNL	MNL Spatial	Nested	Nested Spatial
Commercial Models					
$Transit_{200m}$	OF	3.63 (4.71)	3.79 (4.76)	3.91 (3.52)	4.08 (4.29)
AP Ratio	OF	-0.05 (-1.98)	-0.06 (-2.09)	-0.06 (-1.93)	-0.06 (-2.10)
Rhodes	OF	2.26 (1.93)	2.33 (2.00)	2.31 (1.91)	2.35 (1.95)
Median*	OF,RT	-4.83 (-1.97)	-5.66 (-2.26)	-4.65 (-1.70)	-5.14 (-1.77)
CBD Time	RT	0.28 (4.12)	0.28 (4.08)	0.30 (3.26)	0.30 (3.75)
Com. $Sold_{200m}$	RT	0.14 (3.51)	0.15 (3.61)	0.15 (3.01)	0.16 (3.33)
Tunnel	RS	4.31 (4.04)	4.36 (4.06)	4.12 (3.56)	4.15 (3.77)
Tecumseh Rd_{200m}	RS	1.55 (2.57)	1.43 (2.38)	1.47 (2.33)	1.37 (2.28)
Com. Price	OT	0.02 (2.36)	0.03 (3.18)	0.02 (2.17)	0.03 (3.20)
λ	OF,RT,RS	-	0.54 (2.12)	-	0.60 (2.17)
δ **	OF/RT	-	-	0.87 (2.58)	0.86 (4.42)
Naive ρ^2		0.2200	0.2406	0.3143	0.3332
Adjusted ρ^2		0.2000	0.2187	0.2946	0.3120
119 Observations					

Industrial Models

Variables	Utility	MNL	MNL Spatial	Nested	Nested Spatial
$Transit_{200m}$	W	1.20 (2.03)	1.32 (2.18)	0.90 (1.62)	0.96 (1.53)
ECR_{400m}	W	1.25 (1.90)	1.32 (2.00)	0.96 (1.61)	1.03 (1.63)
Ind. Price	F	0.22 (3.96)	0.21 (3.70)	0.18 (3.05)	0.17 (2.76)
$Rail_{600m}$	F	-3.26 (-3.90)	-3.40 (-3.98)	-2.88 (-3.32)	-3.03 (-3.34)
Ind. Sold $Prop_{400m}$	F	4.07 (2.64)	4.03 (2.61)	3.55 (2.41)	3.53 (2.36)
$OD_{96\text{-}00}$	F	-13.9 (-1.78)	-15.4 (-1.84)	-13.5 (-2.00)	-14.7 (-2.00)
Ind. Price	O	0.14 (3.30)	0.15 (3.35)	0.17 (2.92)	0.17 (2.97)
λ	F	-	0.35 (1.49)	-	0.31 (1.38)
δ **	W/F	-	-	1.44 (2.36)	1.41 (2.21)
Naive ρ^2		0.2959	0.3095	0.2672	0.2804
Adjusted ρ^2		0.2609	0.2701	0.2253	0.2338
74 Observations					

Values shown in the following format: β (Wald)
Utilities: OF – Office; RT – Retail; RS – Restaurant; OT – Other (commercial); W – Warehouse; F – Factory; O – Other (industrial)
* Parameters are x 10^{-5}
** Inclusive scale parameter, δ, set to 1.00 for branches with only one alternative

5.2.2 Time to the CBD

Within the commercial model, the retail group found the distance to the CBD to be a statistically significant positive influence on its utility for retail lots. This indicates that within the City of Windsor, the further a property is from the CBD, the greater the utility for retail firms compared to other commercial businesses. This can be attributed to a longstanding decentralization tendency of the population towards the suburbs associated with urban sprawl. Similar results were found by *Maoh and Kanaroglou (2007)* showing retail as the most dominant industry causing an exodus away from the CBD, though several other types of development observed this to some degree as well. Similarly, *Waddell and Ulfarsson (2003)* found that commercial development tends to occur more prominently in areas with nearby residential development.

These findings combined indicate that residents are moving outwards into the suburbs while retail businesses are following behind in attempt to stay in proximity to their customers. Moreover, the price regression models indicated that retail prices (along with others) had a statistically significant decrease corresponding with further distances from the CBD. Therefore the strong affinity for suburban growth of retail stores is likely a mix of lower prices in addition to following the customer base as development expands in the suburbs.

5.2.3 Transportation

5.2.3.1 $Transit_{200m}$

Transit was found to have an impact on both the commercial and industrial logit models. For commercial construction, a close proximity to transit lines in the city increased the desirability for office buildings, everything else being equal. Reasons for this could be the result of increased utility for office employees to be able to commute to work using public transit. This may also be the result of differing organization goals between firm types. Offices generally do not sell products directly to customers so the focus is on retaining employees for better productivity. On the other hand, retail and restaurants provide goods directly to consumers and are less likely to be concerned with employee retention. Therefore these goods oriented industries are focused less on appeasing workers and more on attracting customers. This results in retail development moving towards suburbs (as seen in the CBD Time variable) where transit is less established.

Transit was also found to exert a significant influence on the development of warehouses. However, similar to other spatial variables here, this is measured based on a buffer area and does not take into account the level of service provided by transit. For warehousing, the significance of close proximity to transit is likely to be based more on the spatial location of the transit line and less on the actual benefits it provides. One issue that could influence this finding is space requirements. Storage space is going to be at a higher demand in highly dense areas where availability of space is at a premium. This would correlate with areas along bus routes that are typically denser (and therefore more viable for the placement of transit routes). When moving outwards from the city into suburbs and further into rural areas, accessibility to space becomes much more prevalent thus decreasing the potential profits of dedicated storage. On the other hand, this represents a decrease in utility when looking at factory development and other industrial types. Bus lines tend to develop in dense areas whereas factories tend to develop with larger spacing for manufacturing requirements and to avoid heavy environmental impacts and regulations.

5.2.3.2 ECR_{400m}

For warehouse development, proximity to EC Row Expressway was also found to increase their utility. This indicates the importance of accessibility for warehouses to provide a suitable service storing and supplying goods efficiently. Typical results from other studies show an increased utility for the majority of industries when close to highways (e.g. *Song et al. 2011*). The results here indicate that warehouses impose a much greater affinity for close proximity to highway access compared to other industrial development. This result may also be indicative of the impact of land prices on warehouses. Previous price regression results found that land prices near highway access ramps decreased significantly for vacant land. This decreased price near highways may also attract warehouses who will presumably seek cheaper land due to the larger requirements for warehouse space.

5.2.3.3 $Rail_{600m}$

Another transportation system variable that held a strong influence on the utility of factories was proximity to rail lines. *Dieleman (2004)* found that nearby rail terminals increase the utility for industrial firms. While industrial zones may be located near rail, factories found a negative influence from rail compared to warehouses and the other industrial category. While these results seem counter-intuitive at first, there are several explanations for why this result might have occurred. It should be noted again that similar to transit, the buffer area is measured with respect to all rail lines and therefore does not take into account actual levels of service. Therefore the significance of this proximity to the rail variable could be the result of other latent variables near the physical location of rail lines in addition to the rail line itself. A possible latent variable could be the level of noise pollution that would be associated with busy rail lines. That can prove to be a deterrent factor for certain types of factories that require high levels of attentiveness from their employees. Another latent variable could be that certain factories would want to locate far from busy rail lines to avoid the risk of derailment of trains carrying hazardous material. Another explanation for the negative significance for factories in close proximity to rail is that modern factories typically receive materials and goods via trucks.

This can also be viewed as a positive influence on the warehouse and other development alternatives. Warehouses are often used as logistic hubs in Windsor, transferring goods from rail to trucks. Intuitively, warehouses likely prefer to locate in areas that are highly accessible by rail for a city such as Windsor where the bulk of transportation activities associated with warehousing are dedicated towards moving goods across the border. The positive influence on the other development category could be influenced by the inclusion of properties owned by rail companies as well as the occurrence of quarries/mines shipping their goods via rail.

5.2.4 Geometry

In this analysis, the only variable describing physical properties of the developed parcel was the area to perimeter ratio derived from lot dimensions. Based on the work of *Zhou and Kockelman (2008)*, it was found that the area to perimeter ratio was negative for residential development. On the other hand, they found it to be positive for commercial and office uses. In our case, it was found to be negative for offices, indicating that developers prefer a smaller area to perimeter ratio, seeking to maximize their perimeter. Differences between the results found here and the study by *Zhou and Kockelman* could be the result of the inclusion of residential properties in the latter. Discrepancies could also occur due to variances between preferences in Austin, Texas and Windsor, Ontario.

A lower ratio represents a smaller area in proportion to perimeter. As shown in Figure 5-7, this lower ratio corresponds to shapes following a square pattern. Conversely, a larger AP Ratio variable will be seen for lots that consist of a more rectangular or irregular geometric pattern. Therefore the results of this model represent a preference for office development to try to maximize their perimeter, corresponding with a shape closer in dimension to a square. This finding could be due to an affinity for offices to develop vertically, capturing more square footage compared to one story buildings where other types are less likely to have that advantage. As another example, consider a 2x2 lot and a 4x4 lot. Both are square lots, but due to the areas, larger AP Ratio values will result from the larger lot with ratios of 0.5 and 1, respectively. Again, this points towards offices

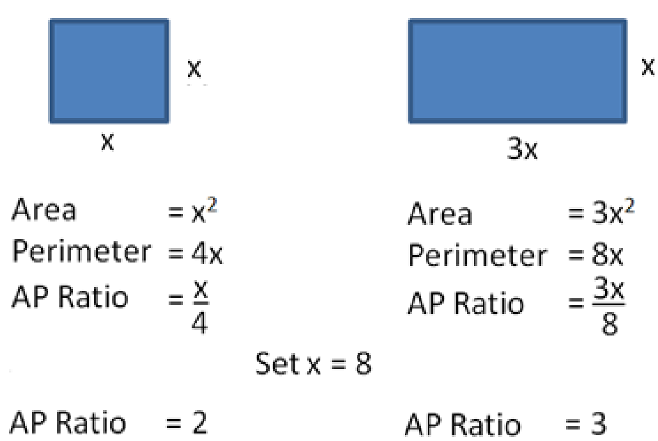

Figure 5-7: AP Ratio Example

preferring lots with smaller areas that could be the result of the models inability to capture vertical development.

5.2.5 Real Estate Influences

5.2.5.1 Rhodes, Tunnel, Com. Sold$_{200m}$, and Ind. Sold Prop$_{400m}$

The Rhodes and Tunnel parameters are both location control variables whose spatial boundaries were created based on the location of areas showing high utilities for specific alternatives that could not be accounted for through other parameters. In this model, Rhodes Drive represents a popular area for office development. Also, a zone near the international tunnel crossing between Canada and the US included a high representation of developing restaurants. With a large $34 million investment in the tunnel plaza (*Windsor Star, 2012b*), this trend near the tunnel crossing may continue to drive restaurants to the area. These variables are a necessary inclusion to control for the effects of inertia in local real estate that would bias the results were they not included.

In addition, two other attributes related to the real estate market were significant. Com. Sold$_{200m}$ and Ind. Sold Prop.$_{400m}$ showed a positive relationship with the number of nearby properties sold in the 10 years prior to the building permits (1996-2005). Com. Sold$_{200m}$ indicates the influence of commercial properties sold on the choice of retail development while the Ind. Sold Prop.$_{400m}$ shows a positive relationship between the proportion of industrial properties sold and the choice to develop land for factories.

5.2.5.2 OD$_{96\text{-}00}$

Another variable revealing the influence of other properties is OD$_{96\text{-}00}$ indicating the proportion of occupied dwellings built between 1996 and 2000 compared to the total by census tract. Due to the negative significance this holds on factory development, this result suggests that nearby residential development deters factories (or vice-versa that factories deter residential development from occurring due to smog and noise pollution).

This seems counterintuitive to the results found by *Waddell and Ulfarsson (2003)* that residential development increases all three general groups - residential, commercial, and industrial. However, the industrial properties were not split into sub-types in the literature whereas the model created here found results between factories and other industrial properties.

5.2.5.3 Tecumseh Rd_{200m}

One area that showed significance in the commercial model is Tecumseh Rd. - one of the main roadways traversing east-west in the city. A dummy variable using a 200 meter buffer along this road found a positive correlation with new restaurant development. While sensitivity analysis was performed on various buffer distances, the 200 meter buffer variable proved to be the most significant. The significance of this buffer variable indicates that the developed properties have a strong preference for locating along this busy roadway. While the road itself may increase accessibility for vehicles traveling in the east-west direction, this is also attributed towards the inertia effect that encourages the development of restaurants along this commercial strip of Windsor.

5.2.6 Prices

5.2.6.1 Median

The median attribute taken from 2006 census tract data showed a negative association for office and retail alternatives. Therefore office and retail development tend to avoid high population income locations. However, causation may occur the other way around, exhibiting the tendency for upper class residents to prefer living in suburbs away from high intensity job zones.

5.2.6.2 Com. Price and Ind. Price

An interesting similarity between the commercial and industrial logit models is that the variables for real estate prices were positively significant in the ‘other’ category for both. In addition, factories in the industrial model also had a propensity for developing in areas with higher real estate prices. For the commercial model, the kriging variable utilized prices from all commercial real estate sold except for offices. In the industrial model, this variable was based on all industrial properties sold.

The ‘other’ alternative for commercial zoning is made up of service and entertainment industries (e.g. hair salons and movie theatres, respectively). This indicates that, all things being equal, these types of commercial land uses will be more prominent in areas with higher commercial real estate prices. In the industrial model, the other permits consist of utility companies, mines, and those permits where the industry was not identified. In many cases, these types of industries may not have a choice in their location. For instance, quarries need to be placed where the intended resource is located.

5.2.7 Estimated Parameter Elasticities

An important analysis of the logit models presented here is the elasticities of the parameters. These values calculate the importance of each parameter by determining how much change exists on the alternatives for a 1% change in the covariate. Based on this statistical attribute applied to the commercial logit model as shown in Table 5-6, those variables that carry an elasticity greater than 0.8 include (from greatest to least) transit, CBD Time, AP ratio, median (for offices), and real estate price per square foot. The total direct elasticities are larger than 0.1 in all cases except for the Rhodes control variable where the impact is lower due to the parameter's small area of effect.

Based on the elasticities for the industrial model in Table 5-7, the more influential covariates include rail proximity and real estate price per square foot. Both contain a direct influence on the factory alternative. Variables among the lower end of elasticities include proximity to EC Row Expressway, proximity to properties sold and newly developed occupied residences with all direct elasticities above 0.15.

Table 5-6: Model 2 Commercial Elasticity Results

Variable	Alternative	Branch	Choice	Total	Variable	Alternative	Branch	Choice	Total
Transit $_{200m}$	**Office**	**0.482**	**1.918**	**2.400**	Spatial (S_i)	Office	-0.167		-0.167
	Retail	0.482	-1.751	-1.269		Retail	-0.167		-0.167
	Restaurant	-1.029		-1.029		**Restaurant**	**0.624**		**0.624**
	Other	-1.029		-1.029		Other	-0.167		-0.167
AP Ratio	**Office**	**-0.111**	**-0.974**	**-1.085**	CBD Time	Office	0.379	-1.621	-1.242
	Retail	-0.111	0.405	0.294		**Retail**	**0.379**	**0.865**	**1.244**
	Restaurant	0.238		0.238		Restaurant	-1.019		-1.019
	Other	0.238		0.238		Other	-1.019		-1.019
Rhodes	**Office**	**0.007**	**0.013**	**0.020**	Com. Sold $_{200m}$	Office	0.141	-0.391	-0.250
	Retail	0.007	-0.086	-0.079		**Retail**	**0.141**	**0.273**	**0.141**
	Restaurant	-0.067		-0.067		Restaurant	-0.196		-0.196
	Other	-0.067		-0.067		Other	-0.196		-0.196
Median	**Office**	**-0.145**	**-0.811**	**-0.956**	Tunnel	Office	-0.174		-0.174
	Retail	-0.145	0.550	0.405		Retail	-0.174		-0.174
	Restaurant	0.330		0.330		**Restaurant**	**0.105**		**0.105**
	Other	0.330		0.330		Other	-0.174		-0.174
Median	Office	-0.212	0.811	0.599	Tecumseh	Office	**-0.057**		**-0.057**

	Retail	**-0.212**	**-0.550**	**-0.762**	Rd $_{200m}$	Retail	-0.057		-0.057
	Restaurant	0.488		0.488		**Restaurant**	**0.172**		**0.172**
	Other	0.488		0.488		Other	-0.057		-0.057
Spatial (S_i)	**Office**	**0.070**	**0.511**	**0.581**	Com. Price	Office	-0.179		-0.179
	Retail	0.070	-0.280	-0.211		Retail	-0.179		-0.179
	Restaurant	-0.172		-0.172		Restaurant	-0.179		-0.179
	Other	-0.172		-0.172		**Other**	**0.882**		**0.882**
Spatial (S_i)	Office	0.121	-0.511	-0.390					
	Retail	**0.121**	**0.280**	**0.402**					
	Restaurant	-0.320		-0.320					
	Other	-0.320		-0.320					

Values presented in **bolded** typeface represent direct elasticity effects

Table 5-7: Model 2 Industrial Elasticity Results

Variable	Warehouse	Factory	Other
Transit $_{200\ m}$	**0.293**	-0.195	-0.195
ECR $_{400m}$	**0.152**	-0.135	-0.135
Rail $_{600m}$	0.220	**-1.895**	0.220
Ind. Sold Prop $_{400m}$	-0.175	**0.313**	-0.175
OD$_{96\text{-}00}$	0.118	**-0.541**	0.118
Ind. Price	-0.741	**1.869**	-0.741
Ind. Price	-0.894	-0.894	**0.852**

Values presented in **bolded** typeface represent direct elasticity effects

5.3 Case Study – Lasalle Ontario

To apply the development type choice models, the Town of Lasalle was selected due to its immediate proximity to the Windsor-Essex Parkway currently under construction. The roughly 66 km^2 area was used to examine the effect of the new parkway on non-residential land development. While the models shown previously were initially created based on individual parcel polygons and properties, the application for Lasalle instead incorporated the models on a uniform 100 x 100 meter cell grid. Each of these grid cells approximates the utility of a land parcel on the basis of the centroid of the grid cell. The principal reason for diverging from the parcel oriented process used to create the original models was to avoid complications due to divided lots and the geometry based parameter, AP ratio. This would be necessary since possible new development might be in underdeveloped areas that require parcel division as performed by the land developer.

To avoid creating bias on the model equations, the average AP Ratio was included for all centroids. Therefore this parameter would have no bearing between differences across the grid. Concurrently, attributes associated with real estate prices were also set constant. This was done due to the scarcity of real estate properties available in the area bounded around Lasalle. If the same kriging methodology for determining price surfaces in the City of Windsor was used for the Town of Lasalle, the variation in prices becomes highly insignificant and misleading. Therefore the few points available within Lasalle were averaged and the constant values were set for all grid centroids in their respective commercial and industrial models.

5.3.1 Industrial Land Use

The parameter associated with EC Row Expressway variable in the industrial logit model holds significant interest for this thesis. To determine possible impacts on warehousing due to the addition of the Windsor-Essex Parkway (WEP), we made the assumption that the WEP and the EC Row Expressway will hold the same influence on the type of industrial land use development in the vicinity of these major transportation corridors. Consequently, the buffer area around the WEP was used to calculate an ECR_{400m} equivalent variable for the locations in the vicinity of the Windsor-Essex Parkway. It is important to note that the assumption of equivalency between the EC Row and the WEP corridors is rather a strong one. The former expressway was created to facilitate movement across the city while the latter parkway is being developed to increase the efficiency of goods (and to some extent people) movement between Windsor and Detroit. However, both highways facilitate greater accessibility in general and as such are likely to impact warehousing development in the region. If such hypothesis holds, then the assumption of equivalency is a valid one and worth considering.

The first scenario pertained to the status quo where there is no new highway constructed. The second looks at predicted probabilities assuming that the WEP has been constructed by extending the ECR parameter to include a similar buffer zone around the WEP. In comparison, the second case is similar to the first in all respects except for a change to the ECR_{400m} buffer extending to include the WEP. As can be seen in Figure 5-8, the two maps showing probabilities for the warehouse development built within industrial zones show a noticeable difference in the absolute probability. The first shows some areas in the northeast corner having a maximum probability around 54% due largely to the presence of transit routes from Windsor ending nearby. Meanwhile, the probability jumps to a maximum of 80% in the WEP scenario. In its current state, the area within this high probability zone is primarily residential before the WEP is built. Since this model is based on the premise of industrial zoning, this would negate any potential results. However, it is not unreasonable to assume that while the original model showed a 400 meter buffer as the most significant, the effect of an increased propensity for warehousing could likely spread further outwards due to increased accessibility. Furthermore, many of the residential buildings that were occupied near the WEP would have been bought and demolished prior to the parkway's construction. A possibility exists that certain areas where buildings were demolished could be rezoned in the future.

Looking at the other variables and their impact in the Lasalle area, proximity to rail will likely impact the west side of Lasalle going north-south. The increasing probability

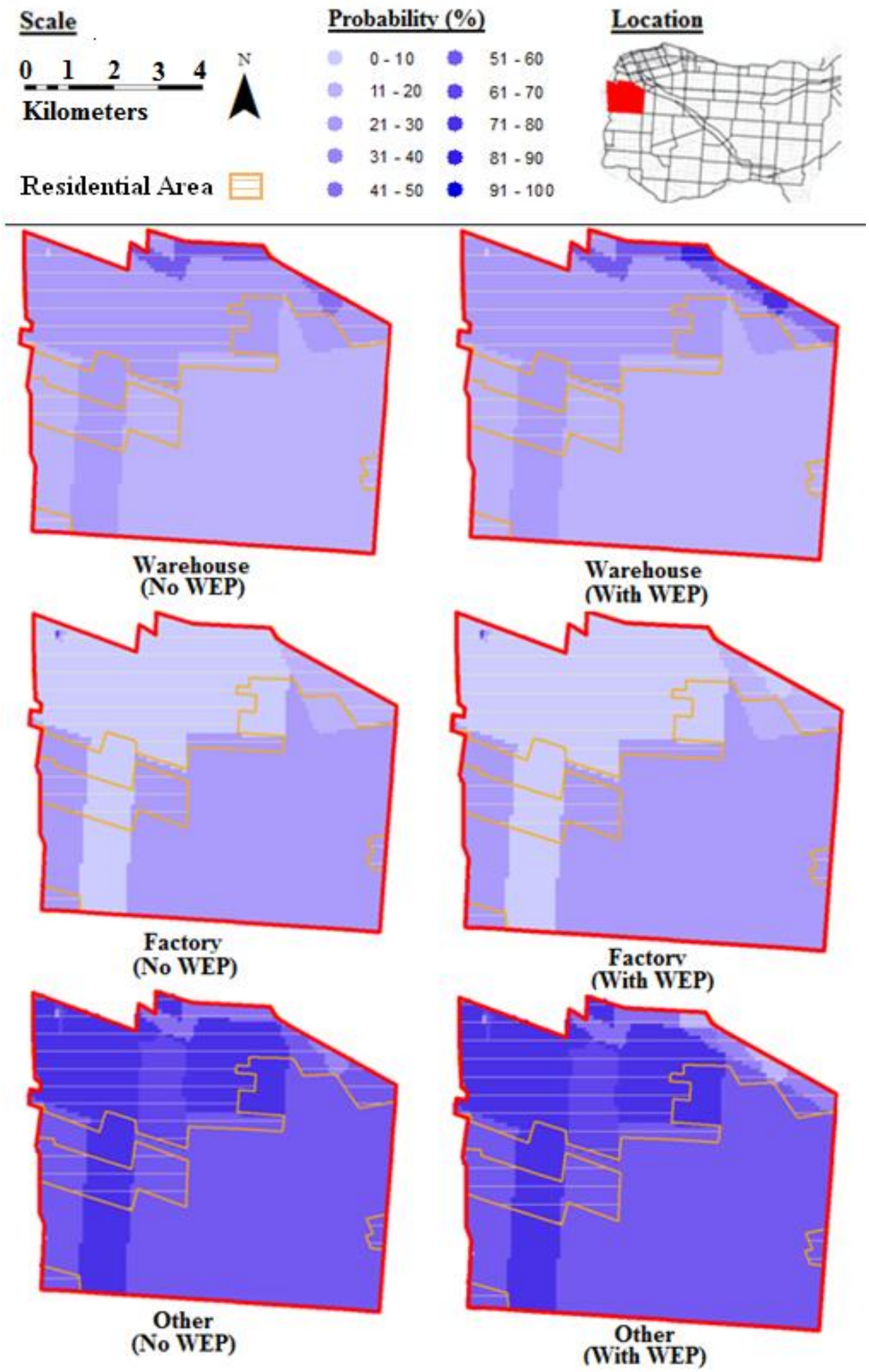

Figure 5-8: Lasalle Industrial Land Use Probability Maps

of warehousing in that area is due to the direct decrease in the choice of developing factories. Transit also has an influence as it will likely increase the potential attractiveness for warehouses, as shown in Figure 5-8.

The probability for factories had four significant variables that were used to explore the propensity of factory development in Lasalle. Proximity to rail decreases the potential for factory development. Also, as the proportion of new occupied residential dwellings increases at a given census tract, the potential for developing factories at a location within that tract tends to decrease. Conversely, one of the two positive influences for factories was close proximity to areas with high real estate inertia (measured as a higher proportion of industrial buildings sold) that accounts for the very small area of variation in northwest Lasalle, increasing the propensity significantly from 10% to 68%. The other positive parameter for factories was the price of industrial real estate that was set constant and did not influence differences across the surface.

The ‘other’ industrial development alternative only had the price of industrial real estate producing variance in the probability maps. Due to this alternative possessing the largest proportion in the dataset, overall percentages are generally the highest here. The surface map shows high percentages in the northwestern area of Lasalle and along the rail corridor with probabilities reaching as high was 73%.

Final maps of the Lasalle region based on the industrial alternatives with the highest utility are shown in Figure 5-9 with the status quo scenario on top and impacts from constructing the WEP on the bottom. Note that this only reflects the alternative with the highest utility and is not representative of Monte Carlo predictions using the utility values as weighted probabilities in a random draw. Based on these surfaces, it can be seen that the northeast area of Lasalle has a noticeable difference between the two scenarios. The simulations suggest that the area will potentially be attractive for warehousing development, all other things being equal. The other category is the alternative with the highest probability without the WEP but warehousing development replaces it as the highest alternative in the WEP scenario.

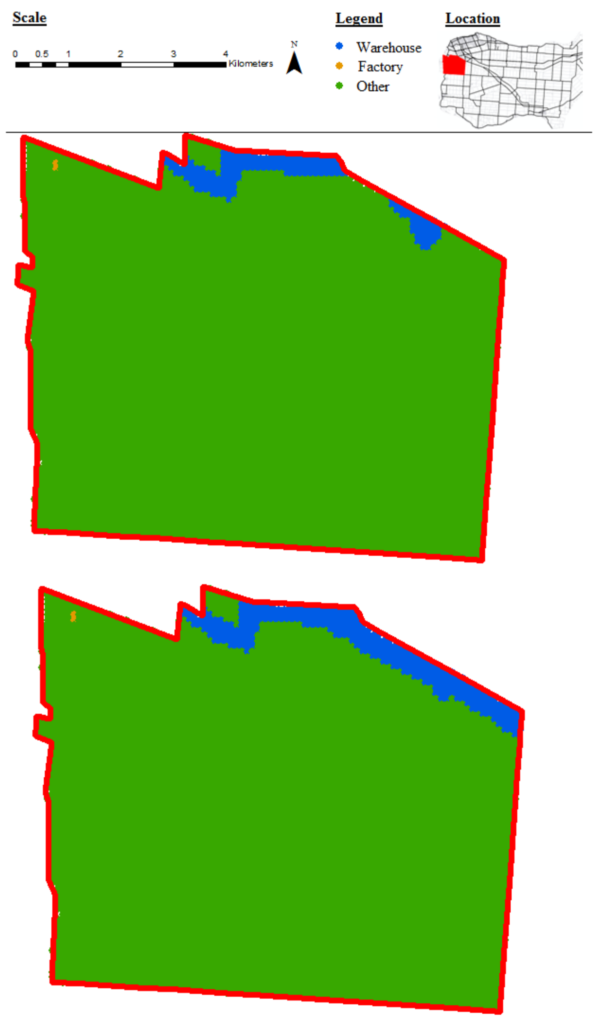

Figure 5-9: Lasalle Industrial Land Use Projections

Top Diagram – Scenario 1: No Windsor-Essex Parkway
Bottom Diagram – Scenario 2: Windsor-Essex Parkway included using ECR covariate as a proxy

5.3.2 Commercial Land Use

In a similar manner to the industrial model, the statistically significant real estate price attribute was set at a constant value based on the small number of commercial properties that were sold in Lasalle. In addition, the spatial variable was unusable here due to the need for data on the choices of nearby decision makers. Therefore in lieu of the nested spatial model, the nested model with no spatial parameter was used. Both models are similar with the same signs, variables and a minor difference in beta coefficients.

Determining the utility for office development resulted in direct effects only due to the median income within census tracts and proximity to transit. The AP Ratio was set constant with no variation across the map and the Rhodes parameter is a location control variable situated within the City of Windsor.

Originally, the retail alternative utilized the parameter measuring time required to reach the CBD. Due to the location of Lasalle outside the boundaries of the original Windsor logit models, this resulted in an unrealistic increase in utility. This created an imbalance in the model that left retail as the highest alternative in any location in Lasalle. To remove this imbalance, the spatial location to which time is measured was relocated from the Windsor CBD to a commercial area on the west side of Lasalle. The results of this parameter can be seen in Figure 5-10 with the probability for retail increasing towards the east end of Lasalle. In addition, proximity to real estate hotspots measured by sold commercial properties increased the probability for retail. The utilities for the restaurant and other alternatives both have no direct influence on variations across the probability surfaces.

Looking at the surface maps of probabilities in Figure 5-10, the office development alternative shows a marginal probability covering the surface with one exception where proximity to transit increased the probability to 30%. Similarly, the restaurant probability surface has a maximum of 30% located on the west side of Lasalle. Restaurant and office development, however, are predominantly overshadowed by retail and other development with retail stronger on the western side and other development significant on the eastern side.

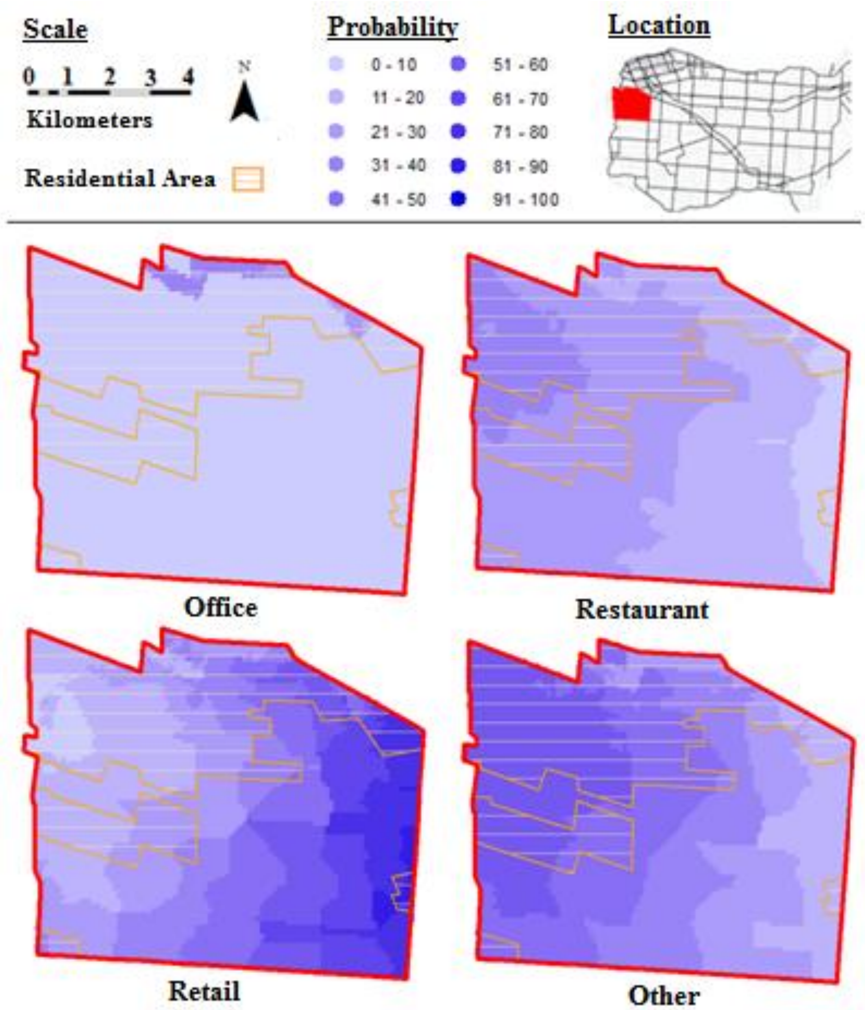

Figure 5-10: Lasalle Commercial Land Use Probability Maps

Figure 5-11 shows the surface representing alternatives with the highest probabilities. The surface is primarily composed of other development on the western half and retail development on the eastern half. Due to the influence of transit, office development becomes the most likely development to occur in some areas in northern Lasalle. Since these results are based on a direct transfer of the logit models developed for the City of Windsor to the Town of Lasalle, it is expected that other influences not captured in the models will have some impact on the validity of these scenarios. However, the applications performed here enhance knowledge in the area with regards to possible influences on development choice from various spatial phenomena.

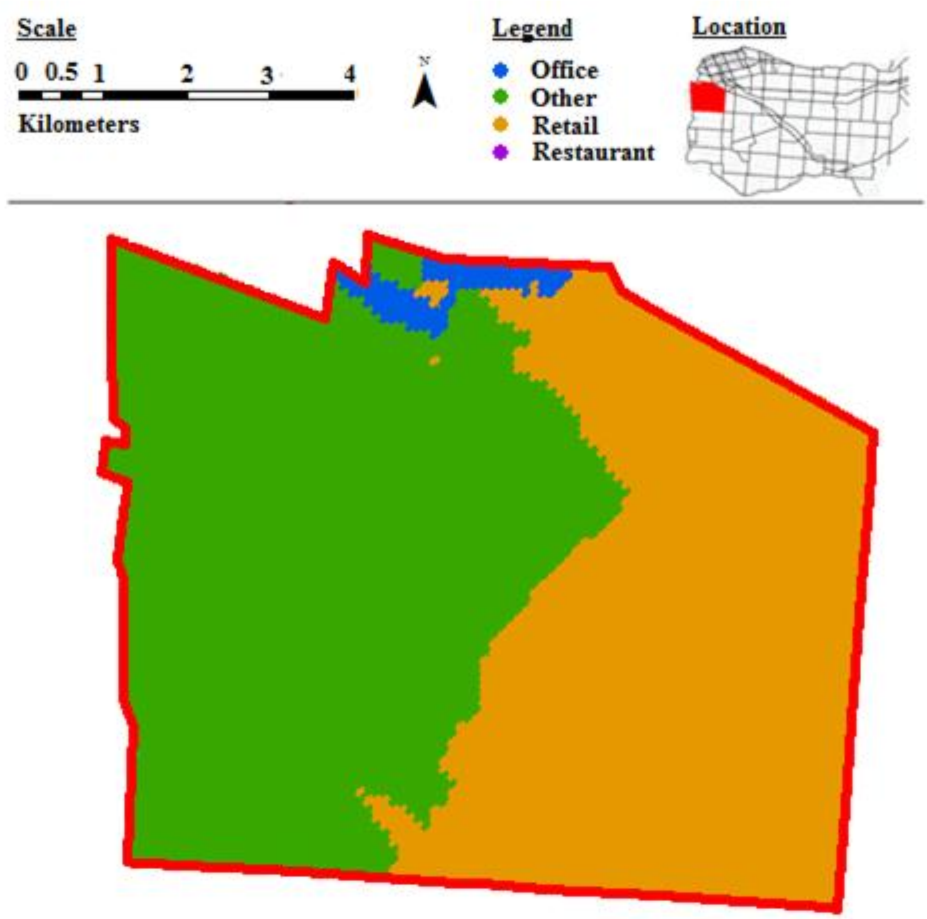

Figure 5-11: Lasalle Commercial Land Use Projections

5.3.3 Possible Price Regression Implications

An analysis of the results found in the price regression models also provides valuable feedback into the effect that a new highway may have on the local area. A direct effect on real estate prices can be found with the $Ramp_{800m}$ variable that was calibrated in the model for vacant land with a beta coefficient of -0.684. Since the dependent price variable was log transformed, an estimate of the price loss per square foot on vacant land within 800 meters of a highway ramp can be given by $e^{0.684}$ equaling $1.98/ft^2 (2011 currency). As a comparison, the average price of vacant lots throughout the Windsor region was equal to $7.43/ft^2. Therefore, a vacant lot within the 800 meter vicinity of a highway ramp is estimated to reduce in value by 27%. However, due to the large amount of landscaping and parks included with the construction of the WEP this effect may not be as pronounced.

Additionally, indirect effects can be seen through other variables found to be significant. The potential accessibility is based on proximity to the residential population which will increase due to extra capacity generated from a new highway. In the regression models, an increase in accessibility was found to increase the prices of both industrial and plaza listings. Furthermore, several variables indicated a negative relationship between the time required to travel from the listing to the CBD and the sale price of the listing. Similar to potential accessibility, a new highway would increase the accessibility of nearby areas, decreasing the time needed to travel to the CBD. This in turn could result in an increase in prices for commercial, retail, restaurant, and vacant listings.

6 CONCLUSIONS

6.1 Summary of Methods

The broad purpose of this thesis was to analyze and model non-residential land use in the Windsor region in order to explore the potential impact that new highway infrastructure such as the Windsor-Essex Parkway could have for the local area. To this end, two non-residential land use processes were analyzed and modelled. The first process was concerned with determining the factors that influence the sale price of non-residential properties in the Windsor-Essex region. On the other hand, the second process was focused on explaining the factors that give rise to a specific type of non-residential land development in the City of Windsor.

For the first process, a set of seven price regression models were specified and estimated based on the price (in 2011 Canadian currency) of several types of non-residential real estate properties provided by the Windsor-Essex Real Estate Board. The seven types included: commercial, office, retail, restaurant, plaza, industrial, and vacant. While the commercial model used the direct price as the dependent variable, the six others utilized a logarithmic transformation due to increased fit for the data.

The second process was handled via two discrete choice models that considered the development choice type that a developer will make when constructing a land parcel for commercial or industrial uses, respectively. The choices were deduced from new development data provided by the City of Windsor in the form of new development permits. Alternatives in the commercial model were broken down into four qualitative categories – office, retail, restaurants, and other. The industrial model, on the other hand, had three alternative types – warehouse, factory, and other.

6.2 Summary of Results

6.2.1 Model 1: Price Regressions

While residential models generally benefit from hedonic attributes describing the structure, it can be seen from the models shown here that non-residential data can still be modelled through the input of only outside (location) variables. However, a consequence of the exclusion of internal lot characteristics was that control parameters for outlier variables had a significant influence on the overall fit. The models improved their R^2 from a range of 0.18-0.69 in the original models to 0.73-0.93. Despite the addition of covariates to adjust for the effect of major outliers, the significance of the other variables remained similar to models without this adjustment. Dissecting the data into homogenous groups was seen to be an important tool in properly analyzing land prices. The commercial model itself had the lowest R^2 value of 0.73 while the sub-types ranged from 0.78 to 0.93.

As it stands, the estimated models were able to predict reasonable land prices for non-residential properties over the urban landscape when including a basic categorization of the quality of the building. This is again due to the ability of these correction variables to capture internal site characteristics that were missing from the modelled data. Therefore while the location and transportation effects have been seen to impact land prices, some measures of additional information regarding the hedonic characteristics of the property are also significant. The models also validate bid rent theory through variables measuring the time to CBD. In addition, among the significant variables were several indicating that proximity to transportation has an impact on property prices. While these proximity based variables were significant in the models, they could be partially influenced by other latent variables coinciding with the spatial locations of the transportation variables.

The regressions can be a viable source of information for local city planners and officials in determining areas in the Windsor region where demand and prices are particularly high and low. For example, the location specific variables could be scrutinized further to determine why certain areas are either flourishing with high prices or slumping with low prices. The latter is of particular interest from a city planning perspective. Moreover, since some of the models themselves are relegated to sub-types of industry, these models may be useful for policies and plans that are targeting these specific groups.

With the uneven distribution of academic research in favour of residential prices over commercial/industrial, the results and relationships developed here will be useful for future studies. The differences in results between models also validate that varying industry types react differently to location and transportation phenomena and should be modelled separately.

6.2.2 Model 2: Land Use Development Type Choice

The multinomial and nested logit models provided substantial insight into the non-residential land development process in the City of Windsor. The attempts to identify randomized parameters via the mixed logit model estimation did not produce any significant differences from the conventional multinomial and nested logit models. However, accounting for spatial effects in the multinomial and nested logit models improved the estimation results.

For the commercial logit model, the nested spatial structure was found to be the most significant with a ρ^2 of 0.33. The nested configuration that gave the greatest significance was grouping office and retail land development type alternatives together in the lower nest while restaurants and other alternatives are left as single degenerate branches. This result implies that while retail and restaurants both share a common goal of creating a profit through the sale of goods, their development pattern differs over space. The significance of the spatial parameter in our models is in line with the previous residential land development studies. These studies found a positive effect of nearby alternative choices for residential properties. Similarly, the effects of nearby alternative choices for commercial properties have a positive impact on the choice type of the developed parcel. In this case the spatial parameter was significant for all three defined alternatives (excluding the other group).

By contrast, the industrial model did not find significant improvement using nested or mixed logit models. However, the spatial effects were able to provide a modest improvement to ρ^2 though the spatial parameter was not significant at the 90% confidence level. The lack of improvement in the nested model is not a surprise given that there are only three alternative groups to model. The small sample size may have also impacted the ability of more complex models to adequately fit the data. Despite the nested logit model not providing significant improvements, the ρ^2 of 0.31 for the industrial multinomial model with spatial effects is closer in comparison to the commercial nested spatial model than the multinomial counterpart.

The results and significance of the models provide a strong affirmation of the recent trends to model land use categories in homogenous groups within generalized firm categories. Furthermore, the results also reiterated the importance of several types of variables including transportation, prices, and market inertia. The commonly mentioned land use topic of urban sprawl could also be seen for the Windsor area through a propensity for retail development to occur closer to the perimeter of the city compared to other commercial land uses. This trend could be troublesome for environmental concerns associated with urban sprawl (*Anderson et al., 1996*; *Su, 2012*).

6.2.3 Transportation Policy Implications

The results obtained from the models presented in this thesis can be used to inform land use policy and transportation planning decisions, particularly in the Windsor region. For example, the parameters used in the development type choice models may be used to entice specific industry types such as increased transit coverage creating a greater demand for office development. However, should detailed data become available, further research into many of the proximity based covariates should be performed to determine the extent to which the infrastructure or other underlying land use phenomena are contributing to the significance seen in the models. Based on the findings of the industrial logit model for development type choice, it was found that transportation has a strong influence on the location preferences of the warehousing industry compared to other industrial development, particularly towards close proximity to transit and highways.

Based on this information, policies aimed at attracting warehousing in these areas could further enhance the appeal of these sites for potential development.

The simulation case study performed on the Town of Lasalle reinforced the implications of the statistical modelling results on land use development. Although the simulation experiments were insightful, the reader should exercise caution when deducing the likely impacts that the new WEP will have on Lasalle's land use development. This is because the simulations are based on specific assumptions that might or might not hold in practice. However, as in any modelling exercise simulations can provide an intelligent guess about the potential impacts that a specific infrastructure project might have on the region's transportation and land use systems. Based on the conducted simulations, the addition of new highway infrastructure was found to increase the probability of warehousing development by 25-30%. Of course, this is triggered by the assumption that the Windsor-Essex Parkway and EC Row Expressway will have a similar effect on land development. This difference significantly altered the surface map showing the most probable development sub-type resulting in warehousing surpassing other industrial development as the most likely to occur. However, since the surrounding area appears to be predominantly residential, it should be noted that this development of warehousing would be conditional on the pursuance of industrial zoning.

Additionally, land prices of non-residential properties could also be influenced by the presence of the Windsor-Essex Parkway, as discerned by the price regression models. For instance, the industrial price regressions found that properties within an 800 meter buffer of a highway ramp decreases the value by an average of 27% though this may be mitigated for the WEP with the inclusion of attractive features such as trails. Conversely, the construction of new transportation infrastructure will increase potential accessibility to the residential population and reduce the time required to reach the CBD which may in turn increase the prices of commercial, retail, restaurant and vacant properties based on results from the price regression models.

6.3 Limitations and Directions for Future Research

6.3.1 Small datasets

One of the largest caveats for this study is the relatively small sample sizes when categorizing land uses by sub-type. The logit models for commercial and industrial models contained 119 and 74 development permits, respectively. This could have been the reason for the mixed logit's inability to increase ρ^2. More specifically, the mixed logit model failed to find significance in the standard deviation of various distributions of beta coefficients. As such, the mixed model collapsed to a conventional multinomial logit model with the static mean representing the beta coefficient of the specified variables. Similarly, the price regression models also faced low record counts in some of the land use groups. For instance, three of the seven commercial groups had less than one hundred records with the commercial plaza sub-type only containing thirty observations. It should be noted that smaller samples were somehow unavoidable for a smaller sized city such as Windsor. However, the results are still insightful and can be useful in future research targeting cities of various population sizes. Furthermore, the smaller sized city is beneficial because fewer hidden latent variables are likely to arise from unknown interactions.

6.3.2 Lot Size Versus Property Size

While the records used for real estate prices included a field for the size of property footprint on the land, this attribute was inconsistent and had many records that left the field blank. For example, multiple listings of the same property sometimes gave errors based on the location of the decimal place (10,000 and 100,000). Because of this, the lot size variable was used in its place to standardize the price among listings. While this will lead to an inability to capture variations in building footprint compared to lot size, the effect was partially mitigated due to the categorization of the groups used.

6.3.3 Spatial Autocorrelation

When it came to accounting for spatial effects in the price regressions, the *Moran's I* and *Durbin-Watson* statistical tests for spatial autocorrelation in the error term found a positive association for commercial (both tests), office (*Moran's I*) and industrial (*Durbin-Watson*) property listings. However, the spatial lag models were no different than the ordinary least square regression models since the spatial lag parameters were insignificant. This indicates that spatial autocorrelation in the data was captured using properly specified covariates that were introduced in the models so that the spatial lag parameter is no longer needed.

Other forms of spatial modelling may prove to be viable options to remove the remaining spatial dependency and increase the validity of the models. The modelling exercise suggests that the region has several distinct areas that could only be accounted for through the inclusion of location indicator variables such as Sandwich and Leamington. A regression method that separates the entire space into several groups such as a switching regression *(Páez et al., 2001)* may prove useful for future modelling efforts.

The commercial development model also found the presence of spatial correlation through the introduction of the spatial lag variable, λ, quantifying the impact on a decision maker of those nearby who chose the same alternative. While this was found to be a positive influence on the retail, office, and restaurant alternatives in the commercial development model, the industrial model did not find any significance here. Further research into this area would be useful for confirmation of these findings.

6.3.4 Type of Analysis

While useful, the case study highlighting the possible impact of a highway on land use in the Windsor area is based on cross-sectional studies of land use in recent years. An ex-ante/ex-post analysis would have provided a strong comparison to determine possible changes in land use over time*(Iacono and Levinson, 2011).* For instance, studying land use before and after construction of the EC Row Expressway moving traffic east-west within Windsor could have provided a direct analysis of the effect from a highway. This would be advantageous given the temporal nature land use allowing the correlation to infer the direction of the relationship. However, this road was predominantly built in the 1970s and would have required data before this time. The real estate data and development permits that were available date back to 1991 and 2005, respectively.

Even with an ex-ante / ex-post comparison, the impact from EC Row Expressway (ECR) would not be the same as the Windsor-Essex Parkway (WEP) due to the unique purpose of the highway. ECR was built specifically to move traffic locally, whereas the main purpose of the WEP is primarily meant to facilitate trade between Ontario and

Michigan in a more efficient manner. In the future, an analysis of the changes to land use in the region due to the construction of the WEP will provide a useful comparison with the results found in this thesis.

6.3.5 Interactions Between Models

The methodology flow chart given in Figure 4-1 describes some of the interactions between land use processes. These interactions can be seen within the results of the modelling performed in this thesis. For example, the commercial and industrial development type choice models found the price of their respective real estate markets to have an influence on the choice of development sub-type. Furthermore, the various price regression models developed here demonstrate that a change in development can lead to changes in real estate pricing. Therefore a complex system exists that can be better calibrated through careful adherence to these interactions.

For example, a simulation model could be developed where the real estate prices and choice of development are determined iteratively until some form of convergence is achieved. Beyond that, including additional modules representing other land use processes and external catalysts would further increase the capabilities of the model. In combination with other models for the Windsor region, the information provided in this thesis provides a strong foundation for the creation of these complex integrated models in the future.

REFERENCES

Anas, A., Arnott, R. & Small, K.A. 1998, "Urban spatial structure", *Journal of Economic Literature,* vol. 36, no. 3, pp. 1426-1464.

Anderson, W.P., Kanaroglou, P.S. & Miller, E.J. 1996, "Urban form, energy and the environment: a review of issues, evidence and policy", *Urban Studies,* vol. 33, no. 1, pp. 7-35.

Andrea, D. & Smith, B. 2002, "The Canada – U.S. border: An automotive case study", *Altarum Institute: Center for Automotive Research*, Ann Arbor, MI.

Anselin, L., Syabri, J. & Kho, Y. 2006, "Geoda: an introduction to spatial analysis", *Geographical Analysis*, vol. 38, pp. 5–22.

Austin, J., Dezenski, E. & Affolter-Caine, B. 2008, "The vital connection: reclaiming Great Lakes economic leadership in the bi-national US-Canadian region", *The Brookings Institution*, Washington, D.C.

Batty, M. & Longley, P. 1994, *Fractal Cities,* San Diego: Academic Press Inc.

Batty, M. 2009, "Accessibility: In search of a unified theory", *Environment and Planning B: Planning and Design,* vol. 36, no. 2, pp. 191-194.

Bhat, C.R. & Guo, J. 2004, "A mixed spatially correlated logit model: Formulation and application to residential choice modeling", *Transportation Research Part B*, vol. 38, no. 2, pp. 147-168.

Buttimer Jr., R.J., Clark, S.P. & Ott, S.H. 2008, "Land development: Risk, return and risk management", *Journal of Real Estate Finance and Economics*, vol. 36, no. 1, pp. 81-102.

Capineri, C. & Leinbach, T.R. 2006, “Freight transport, seamlessness, and competitive advantage in the global economy”, *European Journal of Transportation and Research*, vol. 6, no. 1, pp. 23-38.

Case, B., Clapp, J., Dubin, R., & Rodriguez, M. 2004, “Modeling spatial and temporal house price patterns: A comparison of four models” *Journal of Real Estate Finance and Economics,* vol. 29, no. 2, pp. 167-191.

Chalermpong, S. 2007, “Rail transit and residential land use in developing countries: A hedonic study of residential property prices in Bangkok, Thailand” *Transportation Research Board 2007 Annual Meeting.* Washington, D.C.

Cheng, S. & Stough, R. 2006, “Location decisions of Japanese new manufacturing plants in China: A discrete-choice analysis”, *The Annals of Regional Science*, vol. 40, no. 2, pp. 369-387.

City of Windsor. 1999, “Windsor Area Long Range Transportation Study” [online] Available at: <www.citywindsor.ca/002931.asp>

Clay, M.J. 2010, "Developing an integrated land-use/transportation model for small to medium-sized cities: Case study of Montgomery, Alabama", *Transportation Planning and Technology,* vol. 33, no. 8, pp. 679-693.

Colwell, P.F., Munneke H.J. & Trefzger, J.W. 1998, “Chicago's office market: Price indices, location and time”, *Real Estate Economics,* vol. 26, no. 1, pp. 83-106.

Cramer, J.S. 2003, “The origins and development of the logit model”, *University of Amsterdam and Tinbergen Institute,* [online] Available at: <http://www.cup.cam.ac.uk/resources/0521815886/1208_default.pdf>

De Bok, M. & Sanders, F., 2005, “Firm relocation and accessibility of locations”, *Transportation Research Record: Journal of the Transportation Research Board,* no. 1902, pp. 35-43.

De Bok, M. 2009, “Estimation and validation of a microscopic model for spatial economic effects of transport infrastructure”, *Transportation Research Part A,* vol. 43, no. 1, pp. 44-59.

De Bok, M., & van Oort, F. 2011, “Agglomeration economies, accessibility, and the spatial choice behaviour of relocating firms”, *Journal of Transport and Land Use,* vol. 4, no. 1, pp. 5-24.

Dieleman, F. & Wegener, M. 2004, "Compact city and urban sprawl", *Built Environment,* vol. 30, no. 4, pp. 308-323.

Du, H. & Mulley, C. 2006, “Relationship between transport accessibility and land value: Local model approach with geographically weighted regression” *Transportation Research Record*, *Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., no. 1977, pp. 197-205.

Dunse, N., Jones, C., Brown, J. & Fraser, W.D. 2005, “The spatial pattern of industrial rents and the role of distance”, *Journal of Property Investment and Finance,* vol. 23, no. 4, pp. 329-341.

Eddington, R. 2006, “Transport’s role in sustaining the UK’s productivity and competitiveness: the case for action”, *London: Department of Transport* [online] Available at: <http://www.dft.gov.uk/about/strategy/eddingtonstudy>

ESRI Inc., 2010, "Geostatistical Analyst Tutorial" [online] Available at: <http://help.arcgis.com/en/arcgisdesktop/10.0/pdf/geostatistical-analyst-tutorial.pdf>

Funderburg, R.G., Nixon, H., Boarnet, M.G. & Ferguson, G. 2010, "New highways and land use change: Results from a quasi-experimental research design", *Transportation Research Part A: Policy and Practice*, vol. 44, no. 2, pp. 76-98.

Füss, R., Stein, M. & Zietz, J. 2012, “A Regime-Switching Approach to Modeling Rental Prices of U.K. Real Estate Sectors”, *Real Estate Economics*, vol. 40, no. 2, pp. 317-350.

Galster, G., Hanson, R., Ratcliffe, M.R., Wolman, H., Coleman, S. & Freihage, J. 2001, "Wrestling sprawl to the ground: Defining and measuring an elusive concept", *Housing Policy Debate*, vol. 12, no. 4, pp. 681-717.

Geurs, K.T. & Ritsema van Eck, J.R. 2003, "Evaluation of accessibility impacts of land-use scenarios: The implications of job competition, land-use, and infrastructure developments for the Netherlands", *Environment and Planning B: Planning and Design*, vol. 30, no. 1, pp. 69-87.

Geurs, K., van Wee, B., & Rietveld, P. 2006, “Accessibility appraisal of integrated land-use – transport strategies: methodology and case study for the Netherlands Ranstad area”, *Environment and Planning B: Planning and Design*, vol. 33, no. 5, pp. 639-660.

Handy, S. 2005, "Smart growth and the transportation-land use connection: What does the research tell us?", *International Regional Science Review*, vol. 28, no. 2, pp. 146-167

Hannonen, M. 2008, “Predicting urban land prices: A comparison of four approaches”, *International Journal of Strategic Property Management*, vol. 12, pp. 217-236.

Hansen, W. 1959, “How Accessibility Shapes Land Use”, *Journal of the American Institute of Planners*, vol. 25, no. 2, pp. 73-76.

Harpel, E.D. 2011, “Services in the Detroit regional economy: Implications for a city in transition”, *Service Industries Journal,* vol. 31, no. 4, pp. 597-612.

Hensher, D., Rose, J. & Greene, W. 2005, *Applied Choice Analysis,* New York: Cambridge University Press.

Hesse, M. 2004, "Land for logistics: Locational dynamics, real estate markets and political regulation of regional distribution complexes", *Tijdschrift voor Economische en Sociale Geografie,* vol. 95, no. 2, pp. 162-173.

Hu, Z. & Lo, C.P. 2007, "Modeling urban growth in Atlanta using logistic regression", *Computers, Environment and Urban Systems,* vol. 31, no. 6, pp. 667-688.

Hunt, J.D. 1997, "Stated preference examination of location choice behavior of small retail firms", *Transportation Research Record: Journal of the Transportation Research Board,* no. 1606, pp. 25-32.

Iacono, M., Levinson, D. & El-Geneidy, A. 2008, "Models of transportation and land use change: A guide to the territory", *Journal of Planning Literature,* vol. 22, no. 4, pp. 323-340.

Iacono, M. & Levinson, D. 2011, “Location, regional accessibility, and price effects: Evidence from home sales in Hennepin County, Minnesota”, *Transportation Research Record, Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., no. 2245, pp. 87-94.

Kamyab, H.R., Salman, M.A., Hossini, S.M. & Gholamalifard, M. 2010, "A knowledge-based approach to urban growth modeling in Gorgan city using logistic regression", *Journal of Environmental Studies*, vol. 36, no. 54, pp. 89-96.

Kang, H., Scott, D., Kanaroglou, P.S. & Maoh, H. 2009, "An exploration of issues related to the study of generated traffic and other impacts arising from highway improvements", *Environment and Planning B: Planning and Design,* vol. 36, no. 1, pp. 67-85.

Kieser, M. & Marceau, D. 2009, "Multi-agent modeling to understand the impact of stakeholders' decisions in a residential land development project in Southern Alberta", *Proceedings of the 42nd Annual Hawaii International Conference on System Sciences, HICSS.*

Kockelman, K.M. 1997, “Effects of location elements on home purchase prices and rents in San Francisco Bay area”, *Transportation Research Record, Journal of the Transportation Research Board,* Transportation Research Board of the National Academies, Washington, D.C., no. 1606, pp. 40-50.

Kohn, C. & Brodin, M.H. 2008, "Centralized distribution systems and the environment: how increased transport work can decrease the environmental impact of logistics", *International Journal of Logistics: Research and Applications,* vol. 11, no. 3, pp. 229-245.

Kumar, S. & Kockelman, K. 2008, “Tracking size, location, and interactions of businesses: Microsimulation of firm behavior in Austin, Texas”, *Transportation Research Record: Journal of the Transportation Research Board,* Transportation Research Board of the National Academies, Washington, D.C., no. 2077, pp. 113-121.

Ligmann-Zielinska, A. & Jankowski, P. 2010, "Exploring normative scenarios of land use development decisions with an agent-based simulation laboratory", *Computers, Environment and Urban Systems,* vol. 34, no. 5, pp. 409-423.

Long, J.S. & Ervin, L.H. 2000, “Using heteroscedasticity consistent standard errors in the linear regression model”, *The American Statistician*, vol. 54, no. 3, 2000, pp. 217-224.

Magliocca, N., McConnell, V. Walls, M. & Safirova, E. 2011 "Zoning on the urban fringe: Results from a new approach to modeling land and housing markets", *Regional Science and Urban Economics*, vol. 42, no. 1-2, pp. 198-210.

Maoh, H. & Kanaroglou, P. 2007, “Geographic clustering of firms of urban form: a multivariate analysis”, *Journal of Geographical systems*, vol. 9, no. 1, pp. 29-52.

Maoh, H. & Kanaroglou, P. 2009, “Intrametropolitan location of business establishments: Microanalytical model for Hamilton, Ontario, Canada”, *Transportation Research Record*, *Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., no. 2133, pp. 33-45.

Maoh, H., P. Kanaroglou, & Koronios, M. 2012, “Developed Housing Types at the Land Parcel Level: A Micro-Analytical Modeling Approach”, Submitted to *Journal of Geographical Systems*.

Maruani, T. & Amit-Cohen, I. 2011, "Characteristics of developers and their relations to open space conservation", *Land Use Policy*, vol. 28, no. 4, pp. 887-897.

Martinez, F.J. 1996, “MUSSA: land use model for Santiago city”, *Transportation Research Record*, *Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., no. 1552, pp. 126-134.

Martinez, F.J. 1992, “The bid-choice land-use model: an integrated economic framework”, *Environment & Planning A*, vol. 24, no. 6, pp. 871-885.

Martínez, L.M. & Viegas, J.M. 2009, “Effects of transportation accessibility on residential property values: Hedonic price model in the Lisbon, Portugal, metropolitan area" *Transportation Research Record, Journal of the Transportation Research Board,* Transportation Research Board of the National Academies, Washington, D.C., no. 2115, pp. 127-137.

Mataloni, R. 2011, "The structure of location choice for new U.S. manufacturing investments in Asia-Pacific", *Journal of World Business,* vol. 46, no. 2, pp. 154-165.

McCann, P. & Shefer, D. 2003, "Location, agglomeration and infrastructure", *Papers in Regional Science*, vol. 83, no. 1, pp. 177-196.

McFadden, D. 1974, "Conditional logit analysis of qualitative choice behaviour", *Frontiers in Econometrics,* ed. By P. Zarembka, New York: Academic Press, pp. 105-142.

McKinnon, A. 2009, "The present and future land requirements of logistical activities", *Land Use Policy*, vol. 26, pp. 293-301.

Mohammadian, A., Haider, M. & Kanaroglou, P. 2008, "Homebuilders choice behaviour analysis", *Canadian Journal of Regional Science*, vol. 31, no. 2, pp. 225-240.

Montero-Lorenzo, J.M., Larraz-Iribas, B. & Páez, A. 2009, "Estimating commercial property prices: an application of cokriging with housing prices", *Journal of Geographical System*s, vol. 11, pp. 407-425.

Nguyen, Y. & Sano, K. 2010, "Location choice model for logistic firms with consideration of spatial effects", *Transportation Research Record*, *Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., no. 2168, pp. 17-23.

Noth, M., Borning, A. & Waddell, P. 2003, "An extensible, modular architecture for simulating urban development, transportation, and environmental impacts", *Computers, Environment and Urban Systems*, vol. 27, no. 2, pp. 181-203.

Obregón-Biosca, S. & Junyent-Comas, R. 2011, "Socioeconomic impact of the roads: Case study of the "Eix Transversal" in Catalonia, Spain", *Journal of Urban Planning and Development*, vol. 137, no. 2, pp. 159-170.

Ozus, E. 2009, “Determinants of office rents in the Istanbul metropolitan area”, *European Planning Studies*, vol. 17, no.4, pp. 621-633.

Páez, A., Uchida, T. & Miyamoto, K. 2001, “Spatial association and heterogeneity issues in land price models”, *Urban Studies*, vol. 38, no. 9, pp. 1493-1508.

Páez, A. 2009, "Spatial analysis of economic systems and land use change", *Papers in Regional Science*, vol. 88, no. 2, pp. 251-258.

Paul, R.K. 2008, “Multicollinearity – Causes, effects, and remedies”, *Indian Agricultural Statistics Research Institute*, [online] Available at:
<http://www.iasri.res.in/seminar/AS-299/ebooks/2005-2006/Msc/trim2>

Parent, O. & Brasington, D. 2012 “Using a spatial bivariate probit model to analyze public school consolidation”, *2012 NARSC Conference Presentation*, Ottawa, ON.

Sener, I.N., Pendyala, R.M. & Bhat, C.R. 2011, "Accommodating spatial correlation across choice alternatives in discrete choice models: An application to modeling residential location choice behavior", *Journal of Transport Geography*, vol. 19, no. 2, pp. 294-303.

Shieh, Y. 2003, “An early use of bid rent functions”, *Urban Studies*, vol. 40, no. 4, 2003, pp. 791-795.

Song, Y., Lee, K., Anderson, W. & Lakshmanan, T. 2011, “Industrial agglomeration and transport accessibility in metropolitan Seoul”, *Journal of Geographical Systems*, pp. 1-20.

Sridhar, K.S. & Wan, G. 2010, "Firm location choice in cities: Evidence from China, India, and Brazil", *China Economic Review*, vol. 21, pp. 113-122.

Srour, I.M., Kockelman, K.M. & Dunn, T.P. 2002, “Accessibility indices: Connection to residential land prices and location choices”, *Transportation Research Record, Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., no. 1805, pp. 25-34.

Stopher, P.R. & Mehburg, A.H. (1975). *Urban Transportation Modeling and Planning.* Lexington, MA: D.C. Heath and Company.

Straatemeier, T. 2008, "How to plan for regional accessibility?", *Transport Policy,* vol. 15, no. 2, pp. 127-137.

Su, Y. 2012, “Modelling the future of land use, transportation and the environment in the Windsor CMA, Ontario, Canada” *M.A.Sc Thesis*, University of Windsor, Windsor, Ontario.

Taylor, J., Robideaux, D. & Jackson, G. 2004, “U.S. – Canada transportation and logistics: Border impacts and costs, causes, and possible solutions”, *Transportation journal*, vol. 43, no. 4, pp. 5-21.

Ten Siethoff, B. & Kockelman, K.M. 2002, “Property values and highway expansion timing, size, location, and use effects”, *Transportation Research Record, Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., no. 1812, pp. 191-200.

Torbianelli, V.A. 2009, "Road congestion generated by Distribution Centres in European port regions: Risks and opportunities of the development of hybrid logistics", *Pomorstvo*, vol. 23, no. 1, pp. 21-39.

Trading Economics, 2011, “Canada’s GDP growth rate”, [online] Available at: < http://www.tradingeconomics.com/canada/gdp-growth>

Transportation Research Board, 2011, “Discrete Dependant Variable Models”, [online] Available at: <http://onlinepubs.trb.org/onlinepubs/nchrp/cd-22/v2chapter5.html>

Tu, Y., Yu, S. & Sun, H. 2004, “Transaction-based office price indexes: A spatiotemporal modeling approach”, *Real Estate Economics*, vol. 32, no. 2, 2004, pp. 297-328.

URS Canada Inc. 2008, “Detroit River International Crossing Environmental Assessment Study”, [online] Available at: <http://www.partnershipborderstudy.com>

Van Der Panne, G. 2004, "Agglomeration externalities: Marshall versus Jacobs", *Journal of Evolutionary Economics*, vol. 14, no. 5, pp. 593-604.

Van Dijk, J. & Pellenbarg, P. 2000, “Firm relocation decisions in The Netherlands: An ordered logit approach”, *Papers in Regional Science*, vol. 79, pp. 191-219.

Waddell, P. 2002, "UrbanSim: Modeling urban development for land use, transportation, and environmental planning", *Journal of the American Planning Association*, vol. 68, no. 3, pp. 297-343.

Waddell, P. & Ulfarsson, G.F. 2003, “Dynamic simulation of real estate development and land prices within an integrated land use and transportation model system”, *Transportation Research Board, 2003 Annual Meeting*, Transportation Research Board of the National Academies, Washington, D.C.

Windsor Essex Mobility Group & Ontario Ministry of Transportation, 2011, "Public Information Open House 1", [online] Available at: <http://www.weparkway.ca/pdfs/PIOH-presentation_2011-03-09.pdf>

Windsor Star, 2011a, "Air Cargo Hub Could Benefit, Mayor Says", [online] Available at: <http://www.windsorstar.com/news/cargo+could+benefit+mayor+says/5828333/story.html>

Windsor Star, 2011b, "Chrysler warehouse to bolster Windsor's tax base", [online] Available at: <http://www.windsorstar.com/Chrysler+warehouse+bolster+Windsor+base/5562355/story.html>

Windsor Star, 2012a, "This year's fireworks show may be the last", [online] Available at: <www.windsorstar.com/business/This+year+fireworks+show+last/6767017/story.html>

Windsor Star, 2012b, "Shovels to hit tunnel plaza", [online] Available at: <http://www.windsorstar.com/Shovels+tunnel+plaza/7305920/story.html>

Woudsma, C., Jensen, J., Kanaroglou, P. & Maoh, H. 2008. "Logistics land use and the city: A spatial-temporal modeling approach", *Transportation Research Part E,* vol. 44, no. 2, pp. 277-297.

Xie, F. & Levinson, D. 2008, "The weakest link: The decline of the surface transportation network", *Transportation Research Part E: Logistics and Transportation Review,* vol. 44, no. 1, pp. 100-113.

Zhou, B. & Kockelman, K.M. 2008, "Neighborhood impacts on land use change: A multinomial logit model of spatial relationships", *Annals of Regional Science,* vol. 42, no. 2, pp. 321-340.

APPENDICES

Appendix A: Kernel Density Maps

To understand the areas of spatial locations for records listed and those that are sold, kernel density maps were utilized. The kernel density can be given by the following equation based on a quartic kernel and with no edge correction:

$$\boldsymbol{\lambda_\tau(s) = \sum_{h_i \leq \tau} \frac{3}{\pi\tau^2}\left(1 - \frac{h_i^2}{\tau^2}\right)^2}$$

Where $\lambda_\tau(s)$ are the values of density at point *s* for threshold τ (the maximum radius, 1 km was used for this exercise), and h_i is the distance between point *s* and each observed point.

A brief look at the density maps indicates several points. First, the maps for all listings show the CBD as the most popular location for listings as expected. The density decreases, but is still significant, following south down Ouellette Ave. as well as just east of the downtown core. A higher density is also noticeable along the majority of the shore contiguous to the Detroit River as well as following Tecumseh Rd., particularly at the Tecumseh mall location near the Forest Glade residential area. Several areas south of EC Row Expressway are also noticeably dense.

Figure A-1: Kernel Density Maps 1997-2001

Appendix B: Kriging Surface Interpolations

As was previously stated, several steps were required to complete the necessary surface interpolations for prices through the use of kriging. First, histograms and QQ plots were utilized to determine a mathematical transformation that would resemble a normal distribution. Second, plots showing the general trends in prices were examined to determine the most appropriate trend to remove before performing the interpolations. Finally, semi-variograms were used determine the rate of decay of influence between points. This is then used to determine the weights that are applied to neighbouring data points. Listed below are some of the charts and graphs that were used to assist in understanding the optimal parameters during the kriging process.

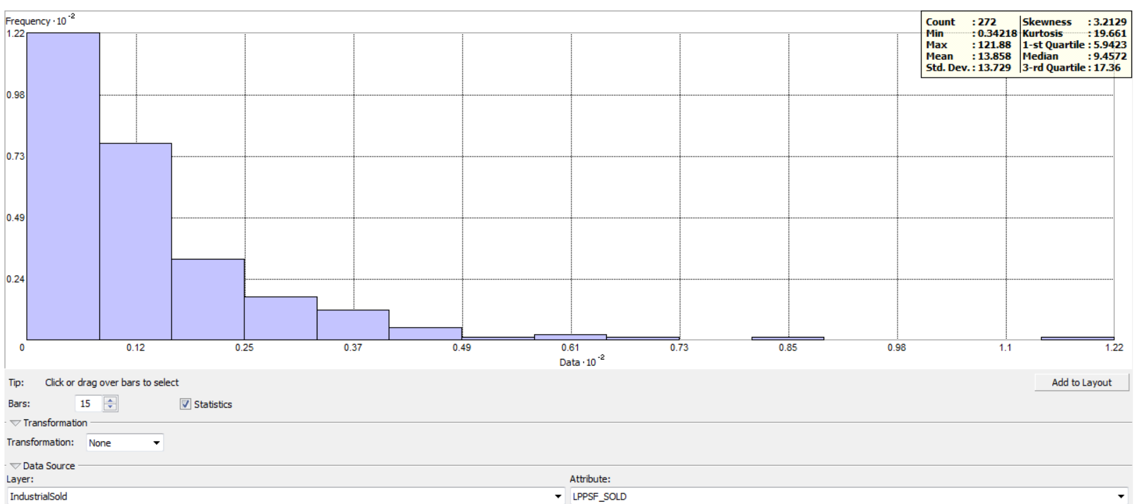

Figure B-1: Industrial Real Estate Histogram – No Transformation

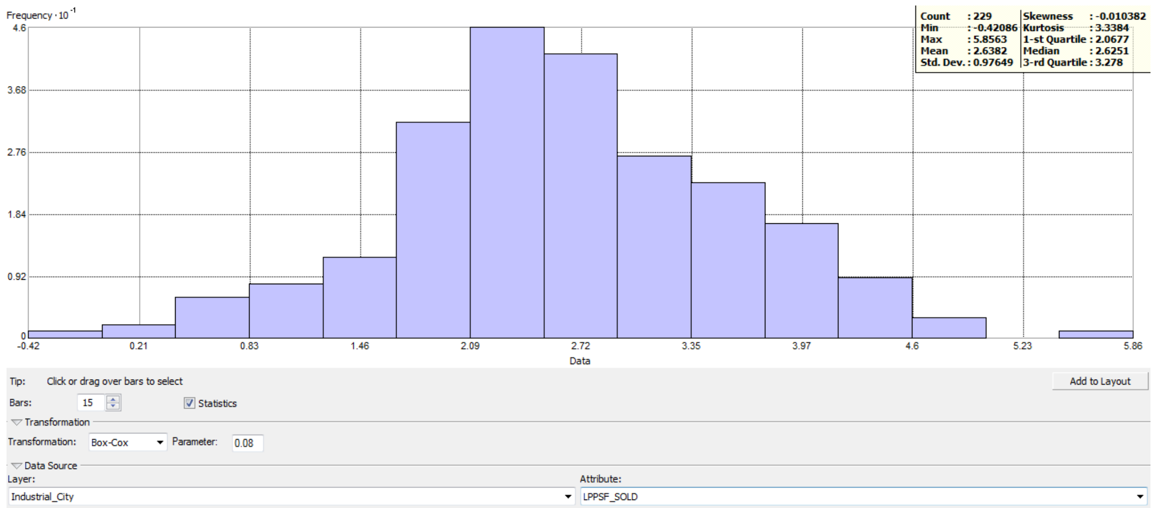

Figure B-2: Industrial Real Estate Histogram – Box Cox Transformation

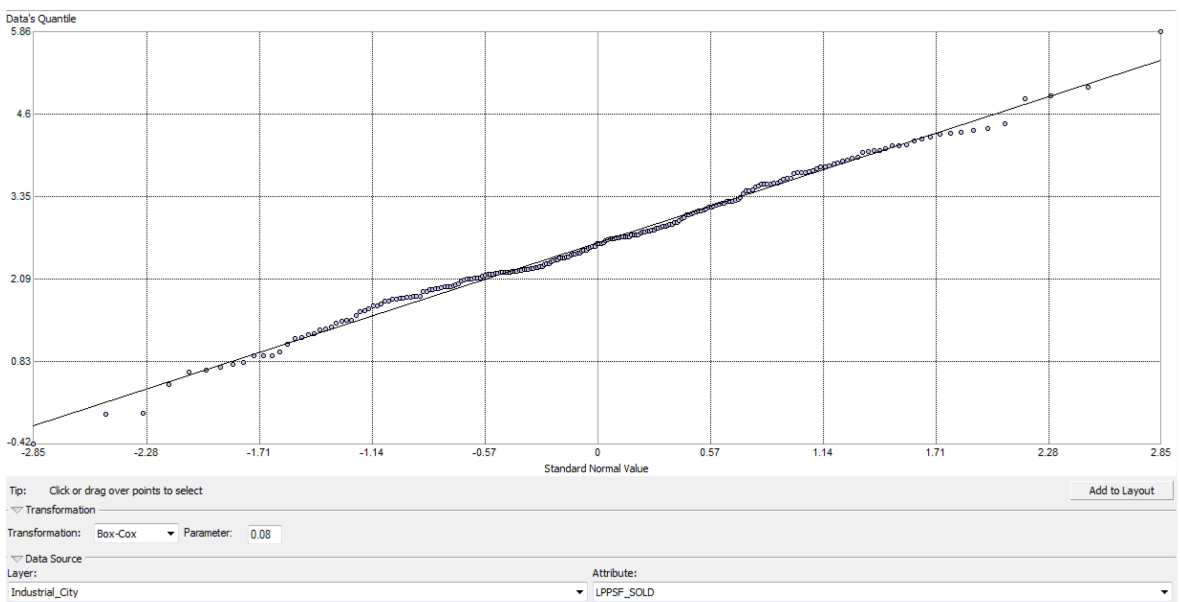

Figure B-3: Industrial Real Estate QQ Plot – No Transformation

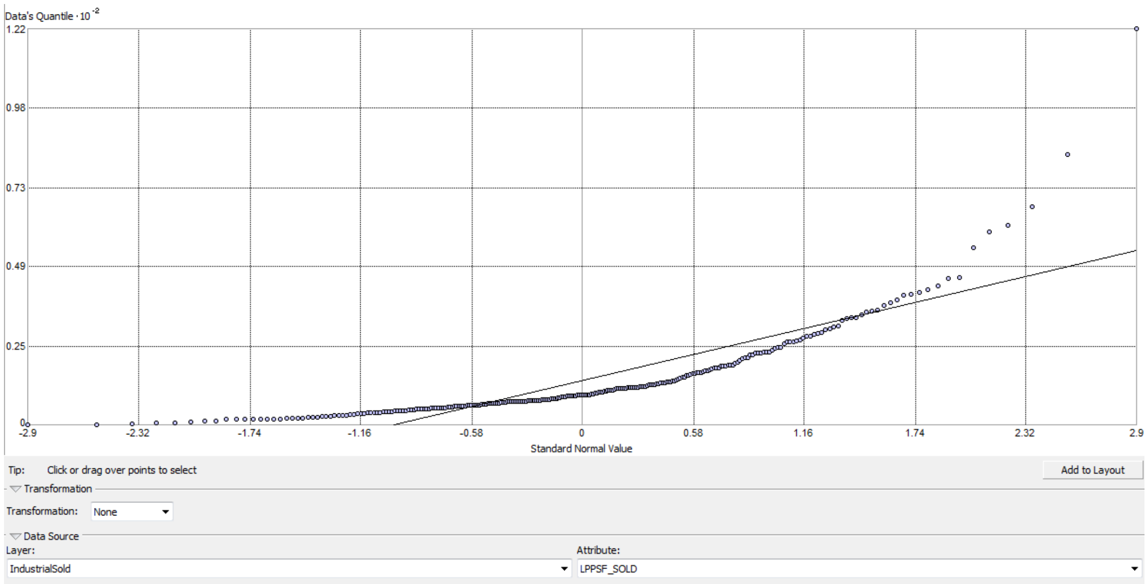

Figure B-4: Industrial Real Estate QQ Plot – Box Cox Transformation

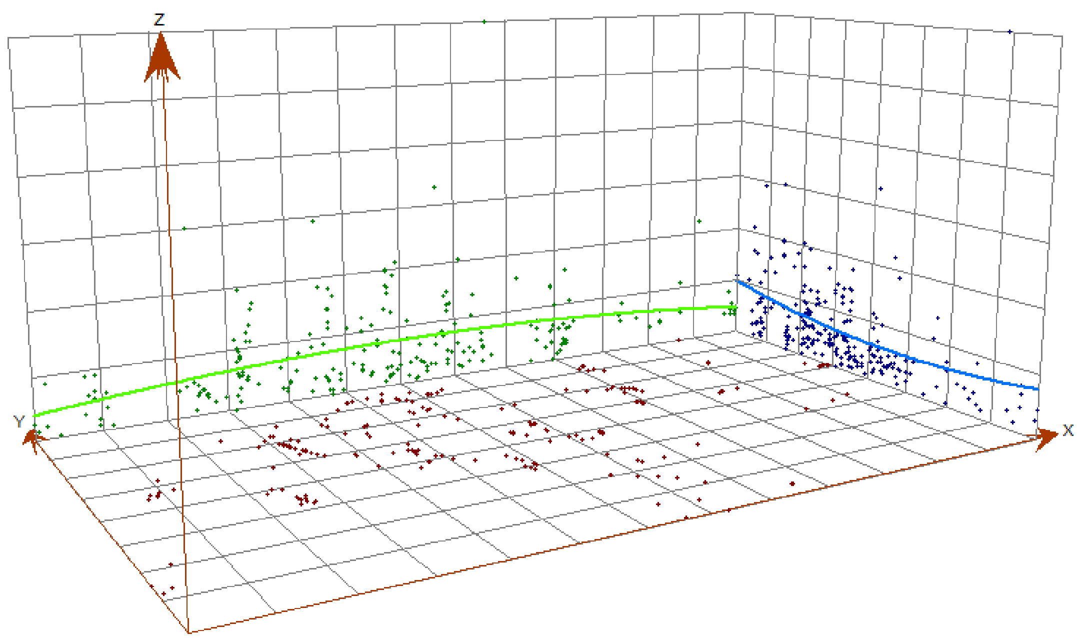

Figure B-5: City of Windsor Industrial Price Trends

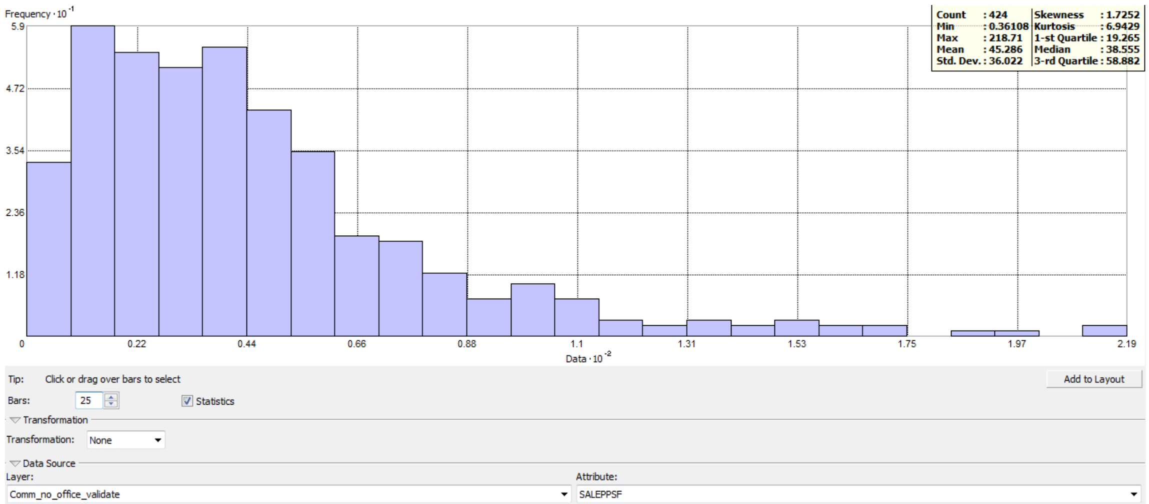

Figure B-6: Commercial Real Estate Histogram – No Transformation

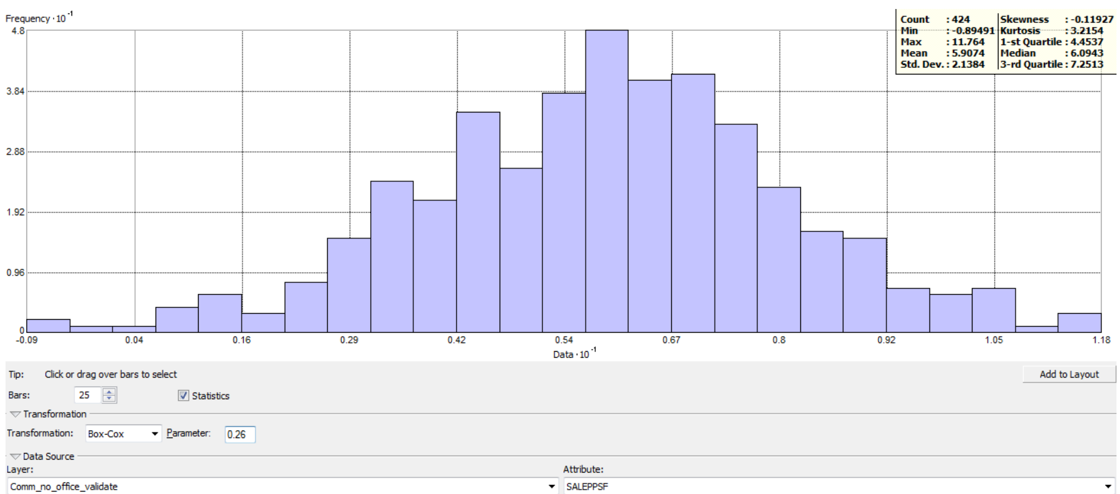

Figure B-7: Commercial Real Estate Histrogram – Box Cox Transformation

*No office prices were included to remove potential bias from vertical development

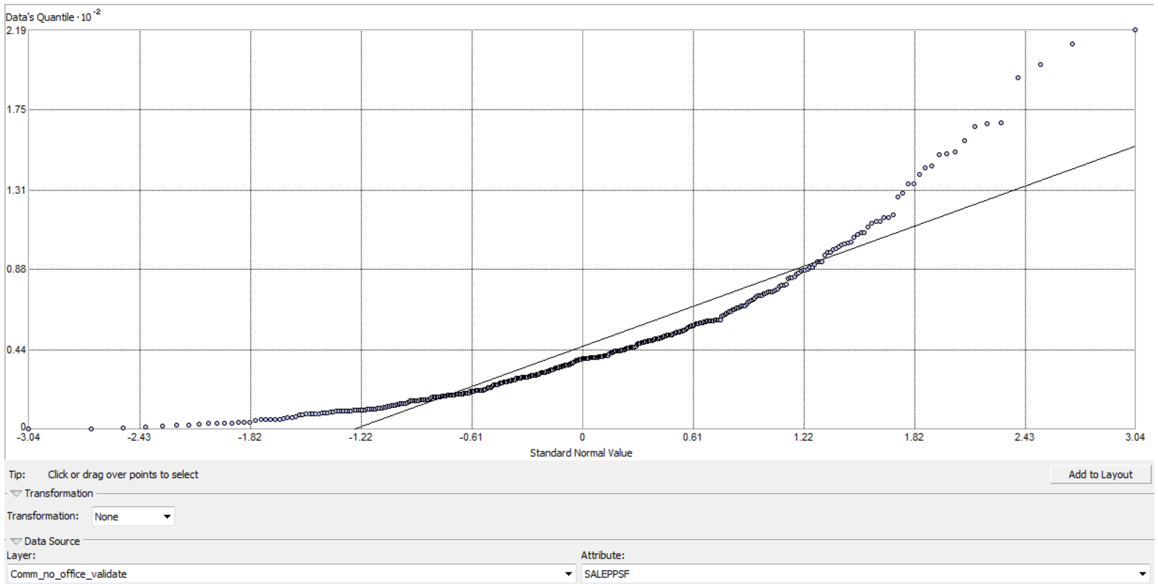

Figure B-8: Commercial Real Estate QQ Plot – No Transformation

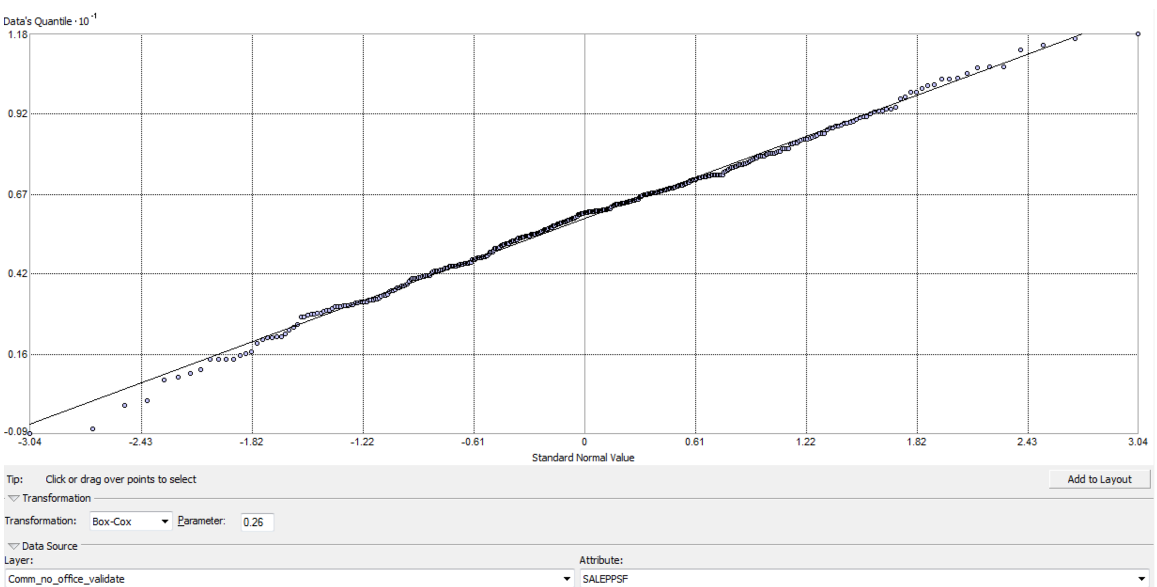

Figure B-9: Commercial Real Estate QQ Plot – Box Cox Transformation

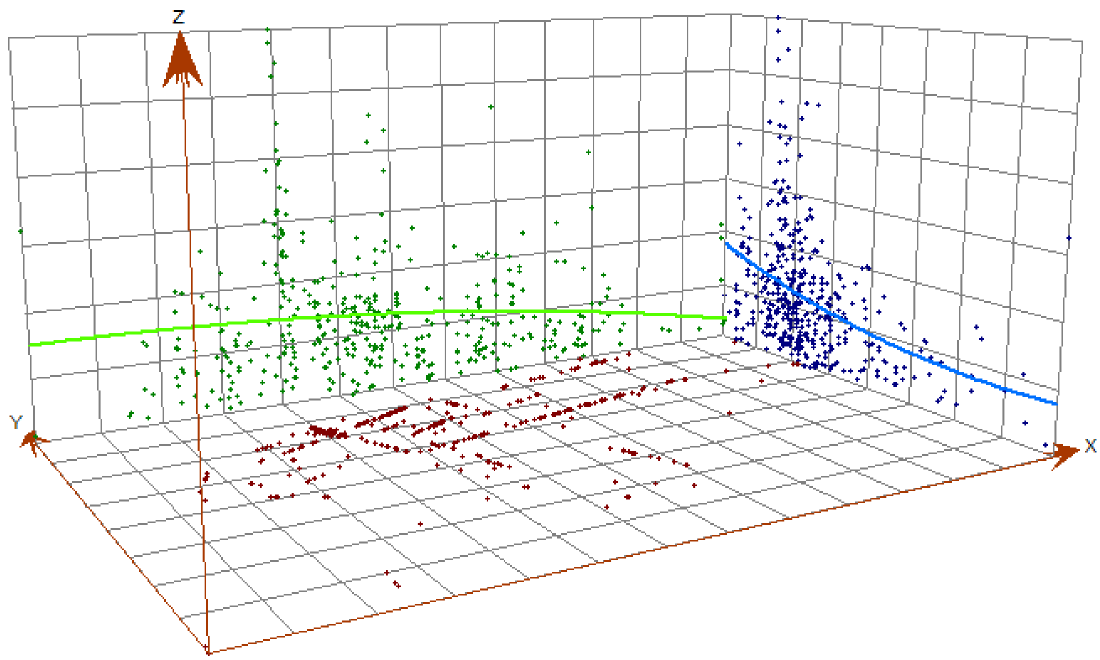

Figure B-10: City of Windsor Commercial Price Trends

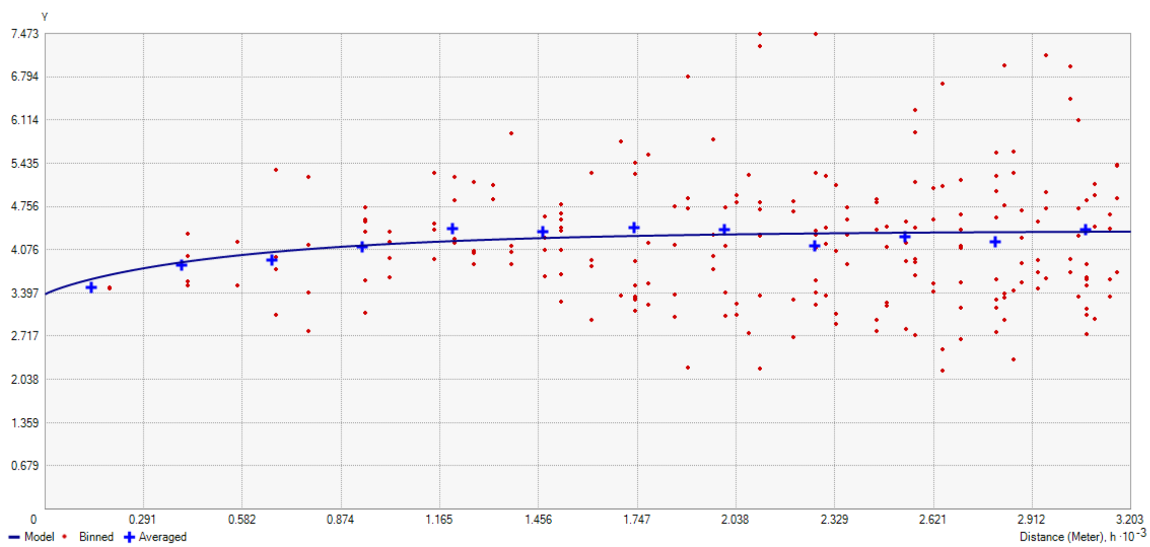

Figure B-11: Semivariogram – Commercial Properties

Kriging Surface Attributes

Industrial
Records
229
-Transformation
BoxCox
Parameter
0.1
-Trend removal
Local Polynomial Interpolation
Power
2
Output type
Prediction
Exploratory trend surface analysis
0
-Searching neighborhood
Standard
Type
Standard
Neighbors to include
10
Include at least
2
Sector type
Four and 45 degree
Angle
40
Major semiaxis
613
Minor semiaxis
920
-Variogram
Semivariogram
Number of lags
12
Lag size
76.7
Nugget
0.626
Measurement error %
100
-Model type
Stable
Parameter
2
Range
613
Anisotropy
Yes
Minor range
920
Direction
40
Partial sill
0.736

Commercial
Records
540
-Transformation
None
-Trend removal
None

-Searching neighborhood

Standard
Type
Standard
Neighbors to include
25
Include at least
2
Sector type
Full
Angle
85
Major semiaxis
442
Minor semiaxis
215
-Variogram
Semivariogram
Number of lags
12
Lag size
55.3
Nugget
761
Measurement error %
100
-Model type
Stable
Parameter
0.448
Range
442
Anisotropy
Yes
Minor range
165
Direction
85
Partial sill
851

Restaurant
Records
67
-Transformation
BoxCox
Parameter
0.32
-Trend removal
Local Polynomial Interpolation
Power
1
Output type
Prediction
Exploratory trend surface analysis
0
-Searching neighborhood
Standard
Type
Standard
Neighbors to include
10
Include at least
3
Sector type
Full
Angle
70
Major semiaxis
6080
Minor semiaxis
9130
-Variogram
Semivariogram
Number of lags
12
Lag size
760
Nugget
6.94
Measurement error %
100
-Model type
Stable
Parameter
0.2
Range
6080
Anisotropy
Yes
Minor range
9130
Direction
70
Partial sill
0

Office	**Retail**	**Commercial (no office)**
Records	Records	Records
96	135	424
-Transformation	**-Transformation**	**-Transformation**
BoxCox	BoxCox	BoxCox
Parameter	Parameter	Parameter
0.5	0.23	0.26
-Trend removal	**-Trend removal**	**-Trend removal**
Local Polynomial Interpolation	Local Polynomial Interpolation	Local Polynomial Interpolation
Power	Power	Power
2	1	1
Output type	Output type	Output type
Prediction	Prediction	Prediction
Exploratory trend surface analysis	Exploratory trend surface analysis	Exploratory trend surface analysis
0	0	0
-Searching neighborhood	**-Searching neighborhood**	**-Searching neighborhood**
Standard	Standard	Standard
Type	Type	Type
Standard	Standard	Standard
Neighbors to include	Neighbors to include	Neighbors to include
6	15	15
Include at least	Include at least	Include at least
2	2	2
Sector type	Sector type	Sector type
Full	Full	Full
Angle	Angle	Angle
0	0	0
Major semiaxis	Major semiaxis	Major semiaxis
1760	992	1210
Minor semiaxis	Minor semiaxis	Minor semiaxis
1760	992	1210
-Variogram	**-Variogram**	**-Variogram**
Semivariogram	Semivariogram	Semivariogram
Number of lags	Number of lags	Number of lags
12	12	12
Lag size	Lag size	Lag size
147	124	190
Nugget	Nugget	Nugget
28.1	1.88	3.51
-Model type	**-Model type**	**-Model type**
Stable	Stable	Stable
Parameter	Parameter	Parameter
0.2	2	2
Range	Range	Range
1760	992	1210
Anisotropy	Anisotropy	Anisotropy
No	No	No
Partial sill	Partial sill	Partial sill
0	0.883	0.73

Appendix C: Regression Results Testing Single Parameters

***Parameters are x 10^{-5}; **Parameters are x 10^{-3}; ***Parameters are x 10^{-2}**

WTM = Work Travel Mode by number of travelers
Prop = Proportion of total

Table C-1: Individual Regressors for Commercial Properties

Commercial Model	Dependent Variable: P			
	B	t-Stat	p	R^2
Spatial variables:				
Lot Size*	-17.14	-5.47	0.00	0.039
CBD Time	-0.34	-3.94	0.00	0.021
Ln(CBD Time)	-7.21	-7.46	0.00	0.071
CBD_{200m}	36.64	9.56	0.00	0.112
$Rail_{400m}$	-14.42	-4.92	0.00	0.032
Potential Accessibility*	6.96	2.66	0.01	0.010
$ECRow_{1000m}$	-12.02	-1.76	0.08	0.004
$Transit_{200m}$	9.37	3.48	0.00	0.016
$Coast_{400m}$	17.61	4.59	0.00	0.028
$Coast_{600m}$	16.92	5.75	0.00	0.043
Urban Area	6.25	2.13	0.03	0.006
$Ramp_{1000m}$	-11.74	-1.90	0.06	0.005
Census tract demographic variables:				
Median Income*	-91.56	-4.45	0.00	0.032
Average Income*	-59.90	-3.28	0.00	0.018
Males Aged 20-39**	17.95	2.65	0.01	0.012
Total Occupied Dwellings**	6.25	2.71	0.01	0.012
1996-2006 Est. Dwellings**	-16.83	-3.04	0.00	0.015
Labour - Manufacturing **	-21.95	-3.73	0.00	0.023
Labour - Retail**	-35.74	-2.40	0.02	0.010
WTM Public Transit**	92.74	3.73	0.00	0.023
WTM Public Transit Prop.	58.12	2.08	0.04	0.007
WTM Walking Prop.	170.32	9.07	0.00	0.122
WTM No Car Prop.	88.12	7.21	0.00	0.081

Table C-2: Significant Individual Regressors for Retail Properties

Retail Model	Dependent Variable: Ln(P)			
	B	**t-Stat**	**p**	**R^2**
Spatial variables:				
Lot Size*	-1.21	-6.41	0.00	0.173
CBD Time	-0.01	-2.48	0.01	0.030
Ln(CBD Time)	-0.18	-3.98	0.00	0.075
$Rail_{400m}$	-0.34	-2.45	0.02	0.030
CBD_{400m}	0.72	3.70	0.00	0.065
CBD_{200m}	0.79	3.88	0.00	0.071
CBD_{800m}	0.44	2.51	0.01	0.031
$Coast_{400m}$	0.51	2.90	0.00	0.041
$Coast_{200m}$	0.58	1.96	0.05	0.019
Median Income*	-2.35	-2.29	0.02	0.034
Census tract demographic variables:				
Total Movers Prop.	1.28	1.93	0.06	0.024
WTM Walking **	1.41	2.19	0.03	0.031
WTM Walking Prop.	2.97	3.21	0.00	0.064
WTM Public Transit Prop.	3.78	2.49	0.01	0.039
WTM No Car Prop.	1.89	3.26	0.00	0.066

Table C-3: Significant Individual Regressors for Plaza Properties

Plaza Model	Dependent Variable: Ln(P)			
	B	**t-Stat**	**p**	**R^2**
Spatial variables:				
CBD Time	-0.03	-3.71	0.00	0.330
Ln(CBD Time)	-0.58	-4.55	0.00	0.425
Leamington	-0.85	-2.17	0.04	0.144
Potential Accesibility*	1.10	4.45	0.00	0.414
$Transit_{200m}$	0.89	3.62	0.00	0.319
Urban Area	1.10	3.74	0.00	0.333
Heritage $Density_{1000m}$**	0.01	1.68	0.10	0.091
Census tract demographic variables:				
Age20-39**	-16.25	-1.87	0.07	0.127
Female 20-39**	-1.14	-1.47	0.15	0.083
Labour - Manufacturing**	-1.20	-2.07	0.05	0.151
WTM Walking Prop.	5.97	1.81	0.08	0.120
WTM Transit Prop.	7.73	2.69	0.01	0.232
WTM No Car Prop.	3.59	2.39	0.02	0.193

Table C-4: Significant Individual Regressors for Office Properties

Office Model	Dependent Variable: Ln(P)			
	B	**t-Stat**	**p**	R^2
Spatial variables:				
Lot Size*	-1.63	-2.11	0.04	0.035
CBD Time***	0.70	0.98	0.33	0.008
Ln(CBD Time)	0.09	1.21	0.23	0.012
CBD_{200m}	0.39	1.25	0.21	0.012
CBD_{400m}	-0.09	-0.31	0.76	0.001
$Coast_{400m}$	-1.36	-5.26	0.00	0.182
$Rail_{400m}$	-0.72	-2.96	0.00	0.066
Last Year	0.04	2.37	0.02	0.043

Table C-5: Significant Individual Regressors for Restaurant Properties

Restaurant Model	Dependent Variable: Ln(P)			
	B	**t-Stat**	**p**	R^2
Spatial variables:				
Lot Size*	-0.57	-5.28	0.00	0.235
CBD Time	-0.02	-2.96	0.00	0.088
Ln(CBD Time)	-0.28	-2.86	0.01	0.083
Potential Accessibility*	0.65	2.75	0.01	0.077
$Rail_{200m}$	-0.50	-1.48	0.14	0.023
Urban Area	0.73	2.94	0.00	0.087
$Transit_{200m}$	0.61	2.62	0.01	0.070

Table C-6: Significant Individual Regressors for Industrial Properties

Industrial Model	Dependent Variable: Ln(P) B	t-Stat	p	R^2
Spatial variables:				
Lot Size*	-0.35	-8.84	0.00	0.224
Ln(CBD Time)	-0.25	-4.03	0.00	0.057
Leamington	-0.23	-0.84	0.40	0.003
Sandwich	-0.26	-1.34	0.18	0.007
$Rail_{200m}$	0.31	2.78	0.01	0.028
Potential Accessibility*	0.68	4.85	0.00	0.080
$Auto_{600m}$	0.75	2.37	0.02	0.020
$Ramp_{400m}$	0.20	0.86	0.39	0.003
Urban Area	0.53	4.26	0.00	0.063
$Coast_{600m}$	0.32	1.66	0.10	0.010
$Tran_{200m}$	0.18	1.71	0.09	0.011
Census tract demographic variables:				
2006 Population*	-6.24	-2.89	0.00	0.033
Median Income*	-1.06	-1.94	0.05	0.015
Age20-39**	-0.18	-2.39	0.02	0.023
Age 20-39 Prop.	1.91	1.50	0.14	0.009
Labour - Manufacturing **	-0.35	-2.45	0.02	0.024
Labour - Manufacturing Prop	-0.69	-0.49	0.62	0.001
Labour - Retail**	-0.90	-2.61	0.01	0.028
Total Movers (Past 5 Years)**	-0.11	-2.15	0.03	0.019
WTM Public Transit Prop.	1.29	2.16	0.03	0.019

Table C-7: Significant Individual Regressors for Vacant Properties

Vacant Model	Dependent Variable: Ln(P) B	t-Stat	p	R^2
Lot Size*	-0.25	-6.22	0.00	0.388
CBD_{4000m}	1.22	4.42	0.00	0.243
Ln(CBD Time)	-0.41	-3.84	0.00	0.195
$Transit_{200m}$	1.13	5.09	0.00	0.298
Sandwich	-0.70	-1.32	0.19	0.028
$Ramp_{1500m}$	-0.62	-2.05	0.04	0.065
$Coast_{1000m}$	0.50	1.67	0.10	0.044
$ECRow_{600m}$	-0.68	-1.84	0.07	0.053
Potential Accessibility*	0.90	3.35	0.00	0.155
Heritage $Density_{1000m}$**	9.84	2.21	0.03	0.074
Urban Area	0.89	3.28	0.00	0.150

Appendix D: Price Regression Validation Charts

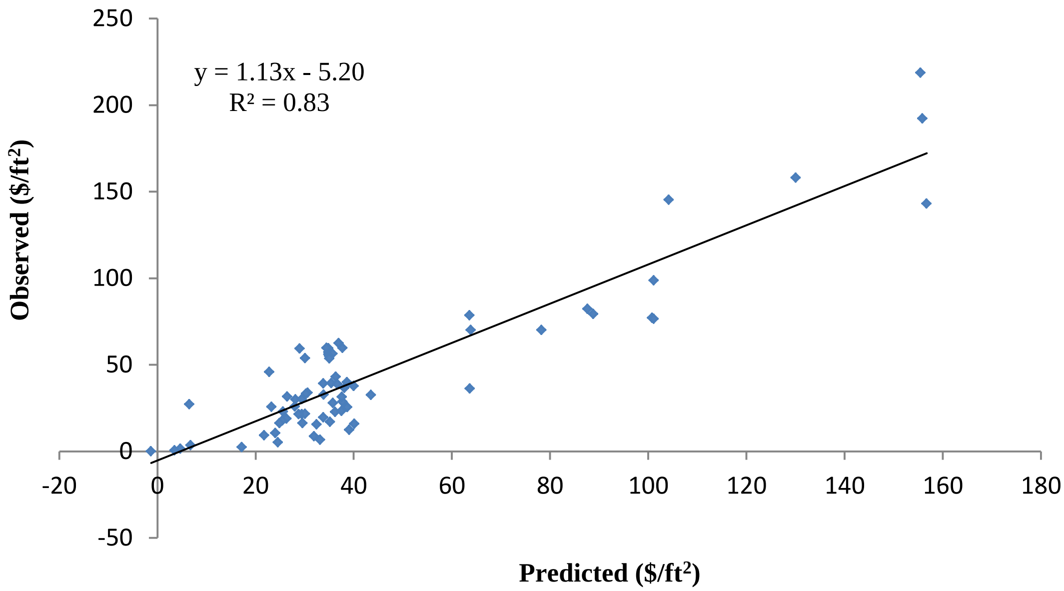

Figure D-1: Commercial Price Model Validation

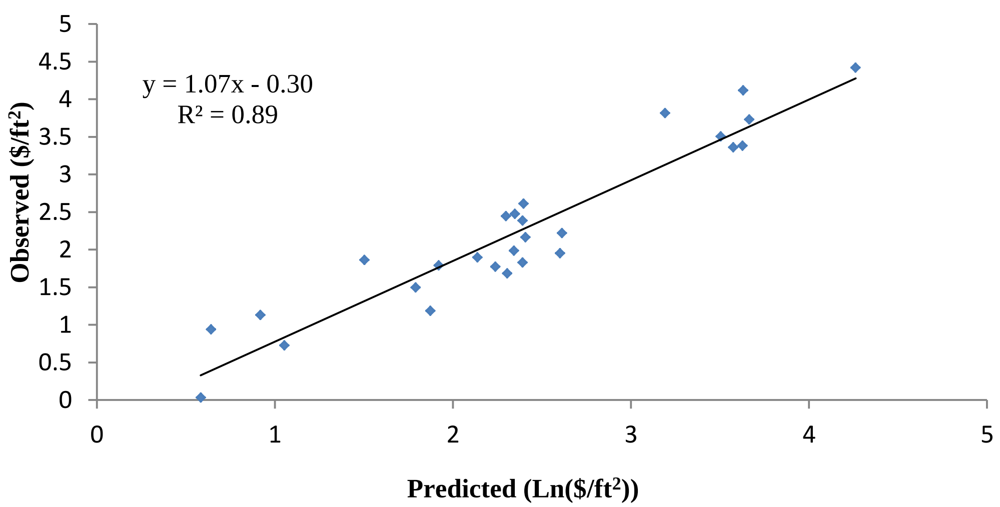

Figure D-2: Industrial Price Model Validation

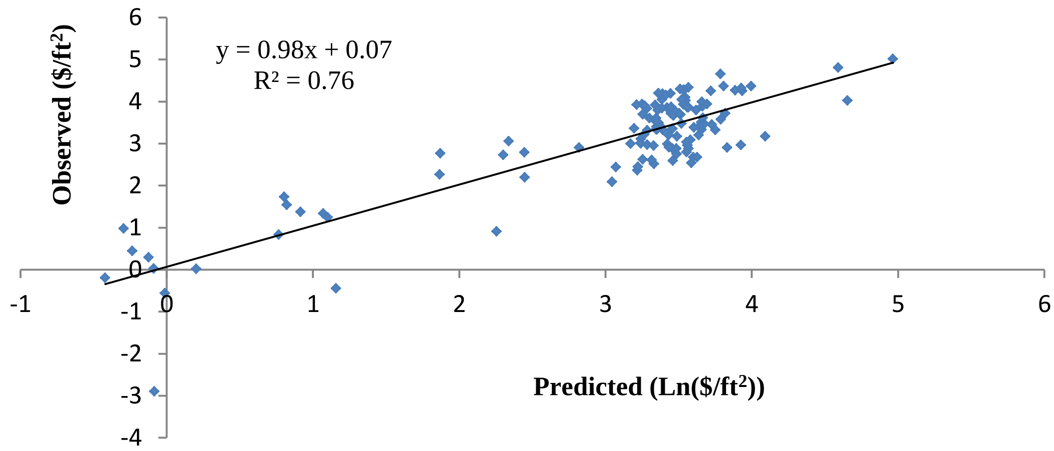

Figure D-3: Office Price Model Validation

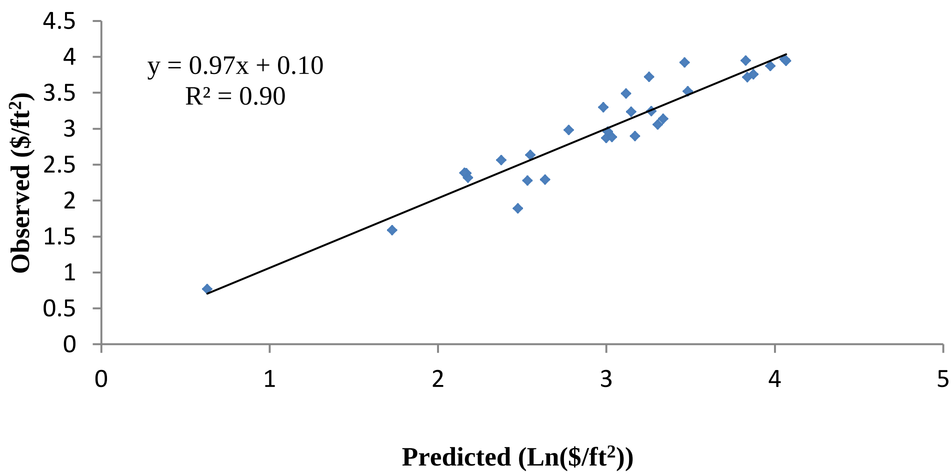

Figure D-4: Plaza Price Model Validation

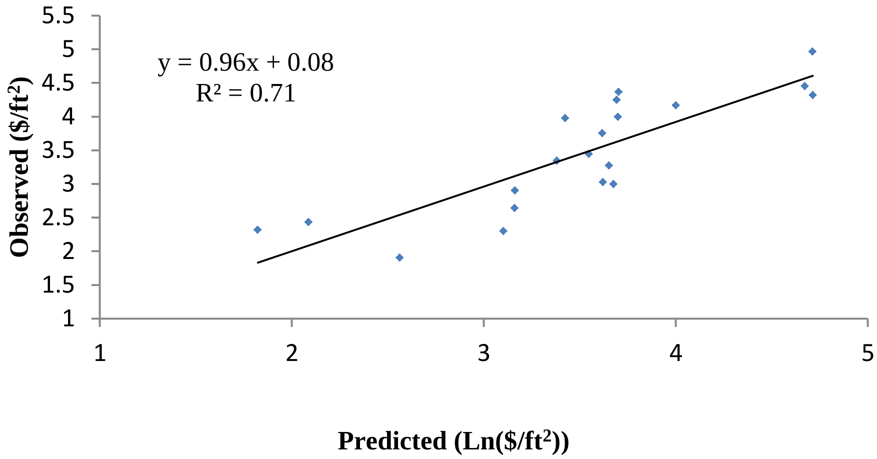

Figure D-5: Retail Price Model Validation

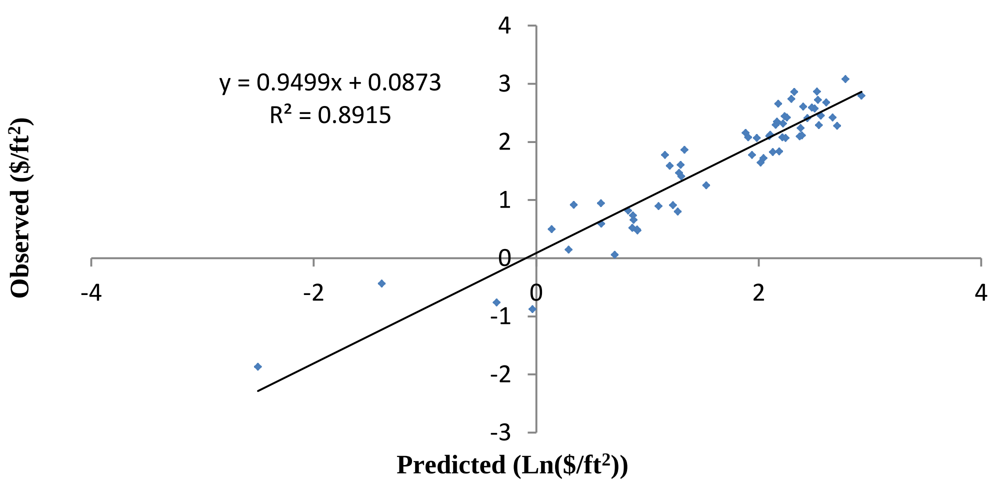

Figure D-6: Vacant Price Model Validation

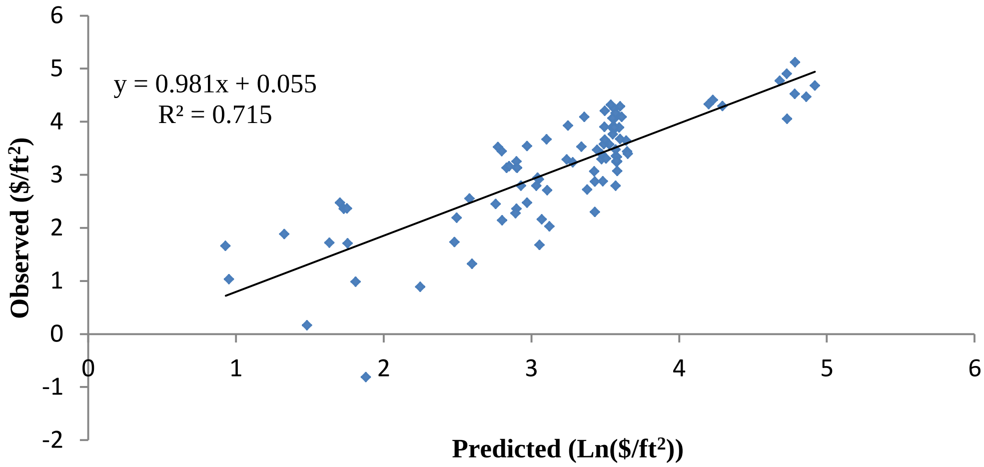

Figure D-7: Restaurant Price Model Validation

Appendix E: Logit Model Specification Codes

Commercial (Multinomial) Code

```
NLOGIT; lhs = Choice
        ; Choices = OF, R, F, OT
        ; Model:

U(OF)=  OF_Tran2H     *     Tran2h
        + OF_ATOP     *     ATOP
        + OF_rhodes   *     CL_rhode
        + MEDIAN      *     MEDIAN /
U(R) =  R_CBDTIME     *     CBDTIME
        + R_COM2HS    *     COM_2HS
        + MEDIAN      *     MEDIAN /
U(F) =   F_Tunnel     *     CL_Tunne
        + F_TEC_RD2H  *     TEC_RD2H  /
U(OT)=  OT_PPSF_AS    *     PPSF_AS $
```

Commercial (Spatial) Code

```
NLOGIT; lhs = Choice
        ; Choices = OF, R, F, OT
        ; Model:

U(OF)   = OF_Tran2H   *     Tran2h
        + OF_ATOP     *     ATOP
        + OF_rhodes   *     CL_rhode
        + MEDIAN      *     MEDIAN
        + Rho         *     RHO5/
U(R)    =  R_CBDTIME  *     CBDTIME
        + R_COM2HS    *     COM_2HS
        + MEDIAN      *     MEDIAN
        + Rho         *     RHO5/
U(F)    =  F_Tunnel   *     CL_Tunne
        + F_TEC_RD2H  *     TEC_RD2H
        + Rho         *     RHO5 /
U(OT)   =  OT_PPSF_AS *     PPSF_AS $
```

Commercial (Nested) Code

```
NLOGIT; lhs= Choice
; Choices = OF, R, F, OT
        ; Tree =  Food(F), Ret_Off(R,OF), Other(OT)
        ; ivset: (Food,Other)=[1.00]
        ; start=logit
        ; Model:

U(OF)=  OF_Tran2H     *     Tran2h
        + OF_ATOP     *     ATOP
        + OF_rhodes   *     CL_rhode
        + MEDIAN      *     MEDIAN /
U(N) =  R_CBDTIME     *     CBDTIME
        + R_COM2HS    *     COM_2HS
        + MEDIAN      *     MEDIAN /
U(F) =  Tunnel        *     CL_Tunne
        + F_TEC_RD2H  *     TEC_RD2H  /
U(OT)=  OT_PPSF_AS    *     PPSF_AS $
```

Commercial (Nested Spatial) Code

```
NLOGIT; lhs = Choice
        ; Choices = OF, R, F, OT
        ; tree = OffRet(OF,R), Food(F), Other(OT)
        ; ivset: (Food,Other)= [1.00]
        ; Model:

U(OF)   = OF_Tran2H     *       Tran2h
        + OF_ATOP       *       ATOP
        + OF_rhodes     *       CL_rhode
        + MEDIAN        *       MEDIAN
        + Rho           *       RHO5/
U(R)    = R_CBDTIME     *       CBDTIME
        + R_COM2HS      *       COM_2HS
        + MEDIAN        *       MEDIAN
        + Rho           *       RHO5/
U(F)    = F_Tunnel      *       CL_Tunne
        + F_TEC_RD2H    *       TEC_RD2H
        + Rho           *       RHO5 /
U(OT)   = OT_PPSF_AS    *       PPSF_AS $
```

Industrial (Multinomial) Code

```
NLOGIT; lhs= Choice
        ; Choices = W, F, O
        ; Model:

U(W)    =  W_tran2H     *       Tran2H
        + W_ECR4H       *       ECR4H /
U(F)    =  F_PPSF_IND   *       PPSF_IND
        + F_Rail6H      *       Rail6H
        + F_INDPR4H     *       IND_PR4H
        + F_NewCon      *       OC_9600 /
U(O)    =  O_PPSF       *       PPSF_IND $
```

Industrial (Nested) Code

```
NLOGIT; lhs = Choice
        ; Choices = W, F, O
        ; Tree = WF(W,F), Other (O)
        ; IVSET: (Other)=[1.00]
        ; Model:

U(W)    =  W_tran2H     *       Tran2H
        + W_ECR4H       *        ECR4H /
U(F)    =  F_PPSF_IND   *       PPSF_IND
        + F_Rail6H      *       Rail6H
        + F_INDPR4H     *       IND_PR4H
        + F_NewCon      *       OC_9600   /
U(O)    = O_PPSF        *       PPSF_IND   $
```

Industrial (Spatial) Code

```
NLOGIT; lhs= Choice
            ; Choices = W, F, O
            ; Model:

U(W)        = W_tran2H      * Tran2H
            + W_ECR4H       * ECR4H /
U(F)        = F_PPSF_IND    * PPSF_IND
            + F_Rail6H      * rail6H
            + F_INDPR4H     *IND_PR4H
            + F_NewCon      * OC_9600
            + Rho           * Rho2_7M  /
U(O)        = O_PPSF        * PPSF_IND $
```

Industrial (Nested Spatial) Code

```
NLOGIT; lhs = Choice
            ; Choices = W, F, O
            ; Tree = WF(W,F), Other (O)
            ; IVSET: (Other)=[1.00]
            ; Model:

U(W)        = W_tran2H      * Tran2H
            + W_ECR4H       * ECR4H /
U(F)        = F_PPSF_IND    * PPSF_IND
            + F_Rail6H      * rail6H
            + F_INDPR4H     * IND_PR4H
            + F_NewCon      * OC_9600
            + Rho           * Rho2_7M /
U(O)        = O_PPSF        * PPSF_IND $
```

Appendix F: Spatial Visualization of Parameters

Figure F-1: Automotive Plants

Figure F-2: Coastline

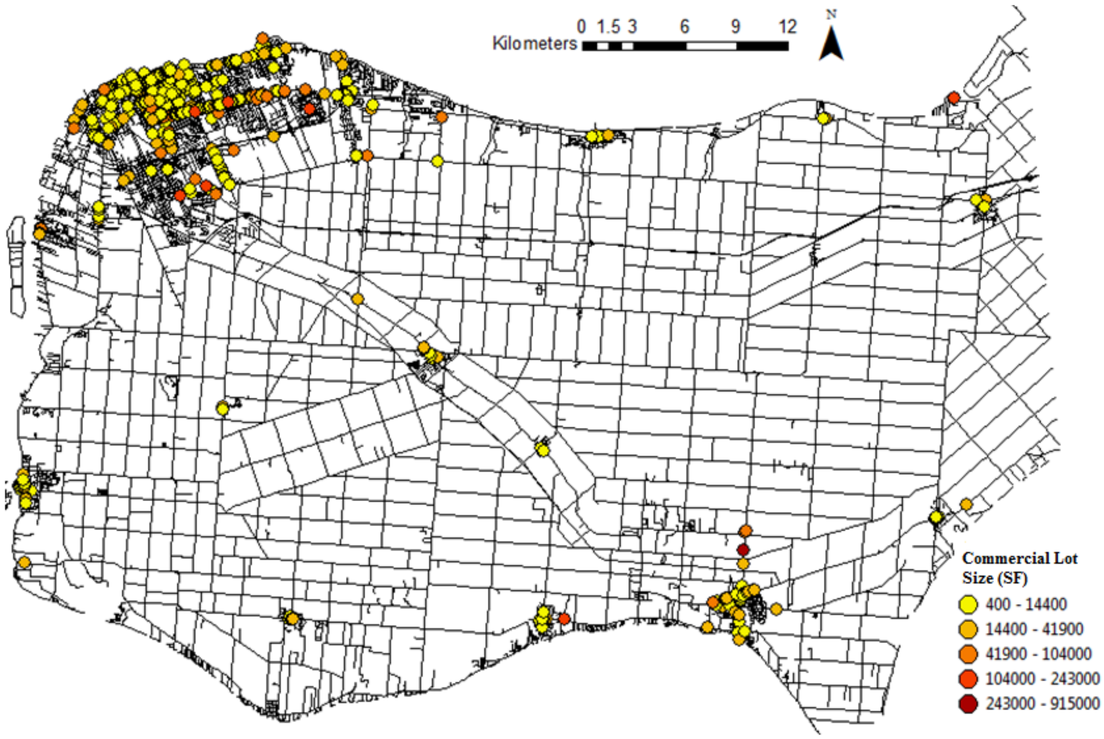

Figure F-3: Lot Size

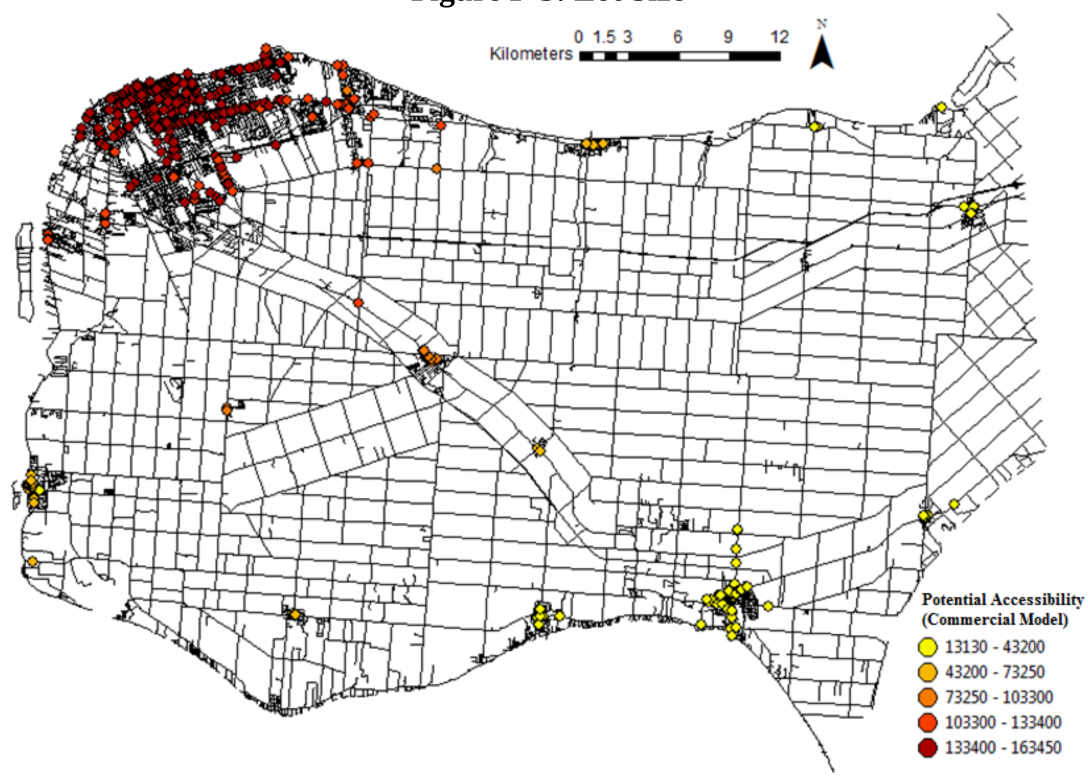

Figure F-4: Potential Accessibility

Figure F-5: Rail Corridors

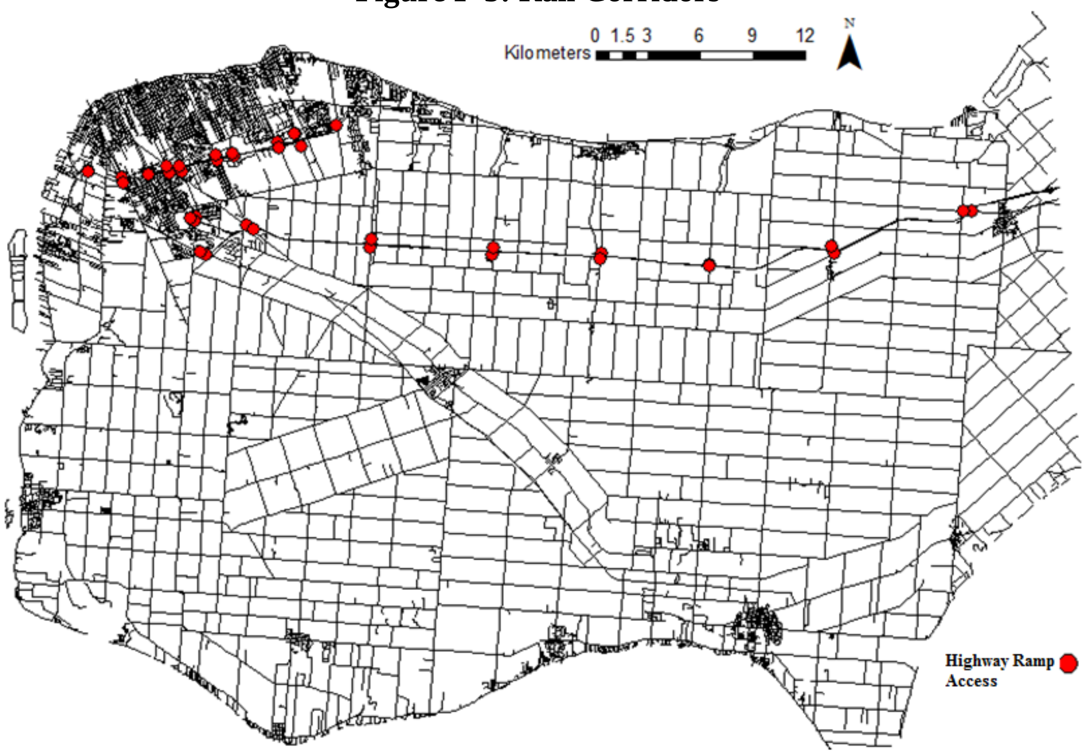

Figure F-6: Highway Ramp Access

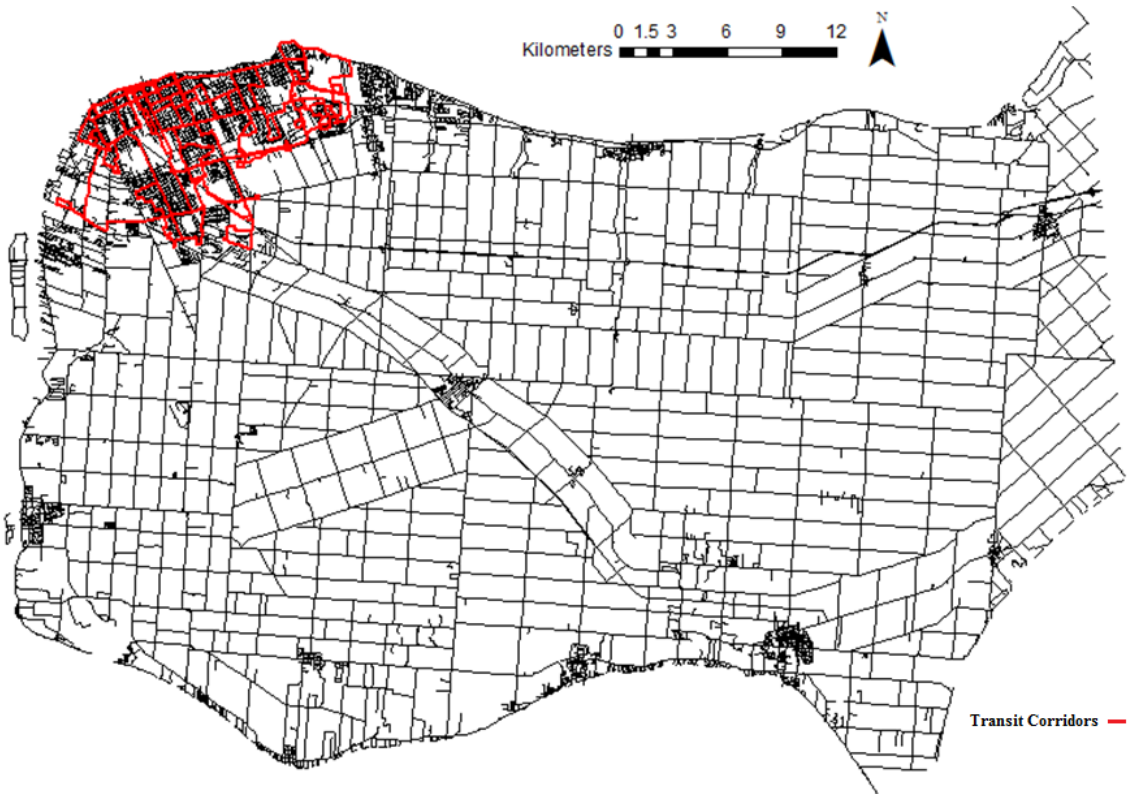

Figure F-7: Transit Corridors

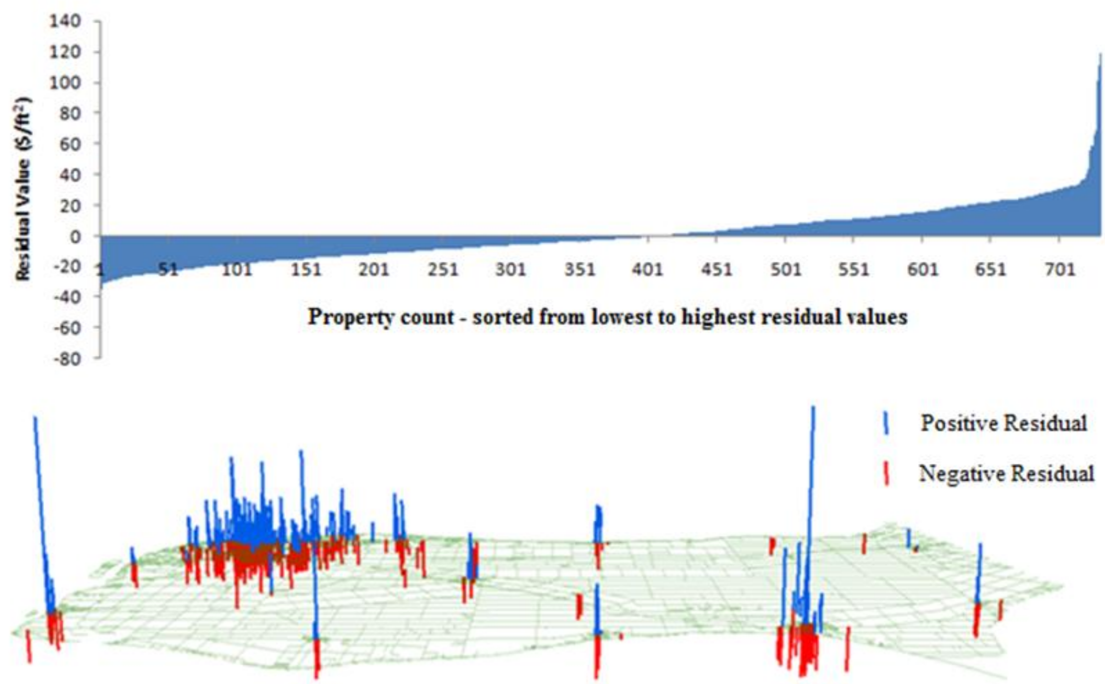

Figure F-8: Residual Values – Commercial

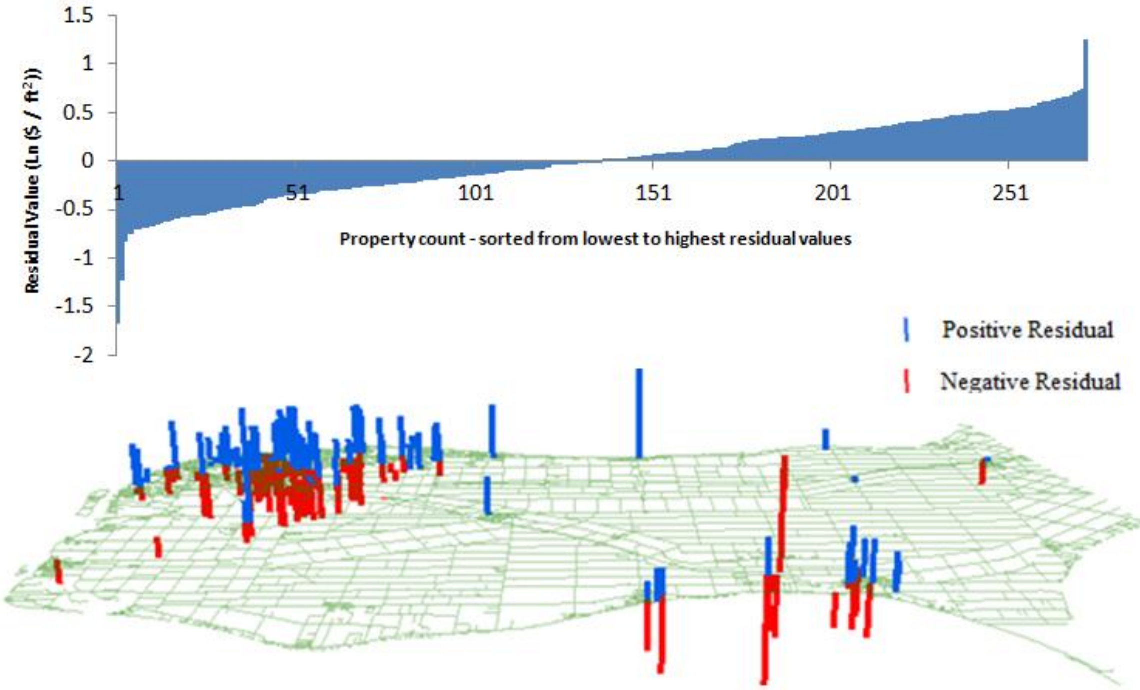

Figure F-9: Residual Values - Industrial

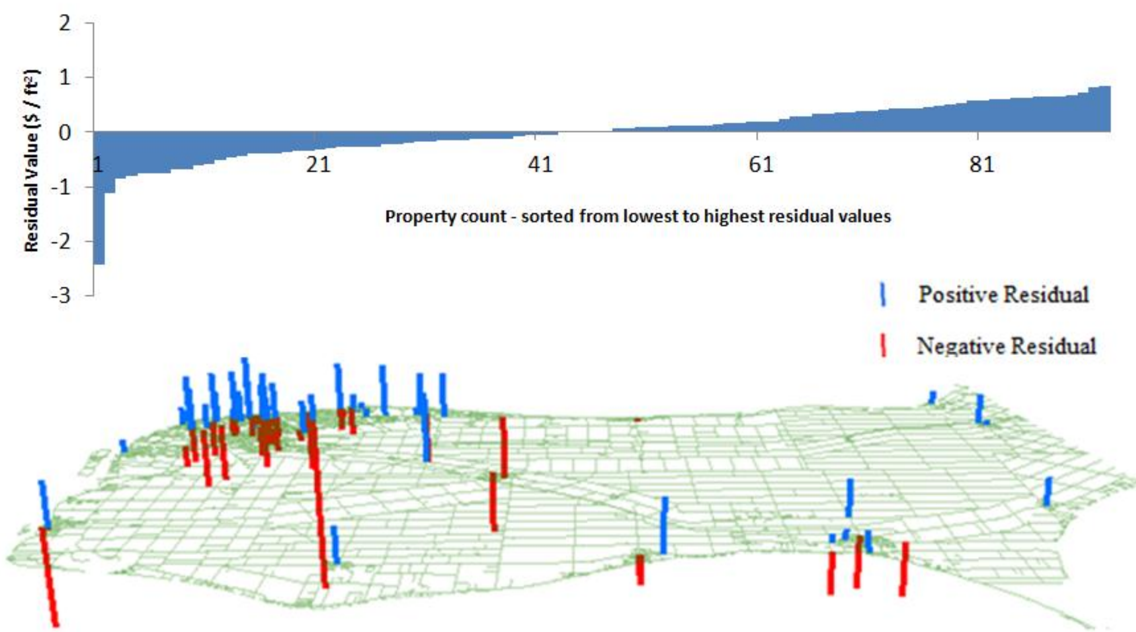

Figure F-10: Residual Values - Restaurant

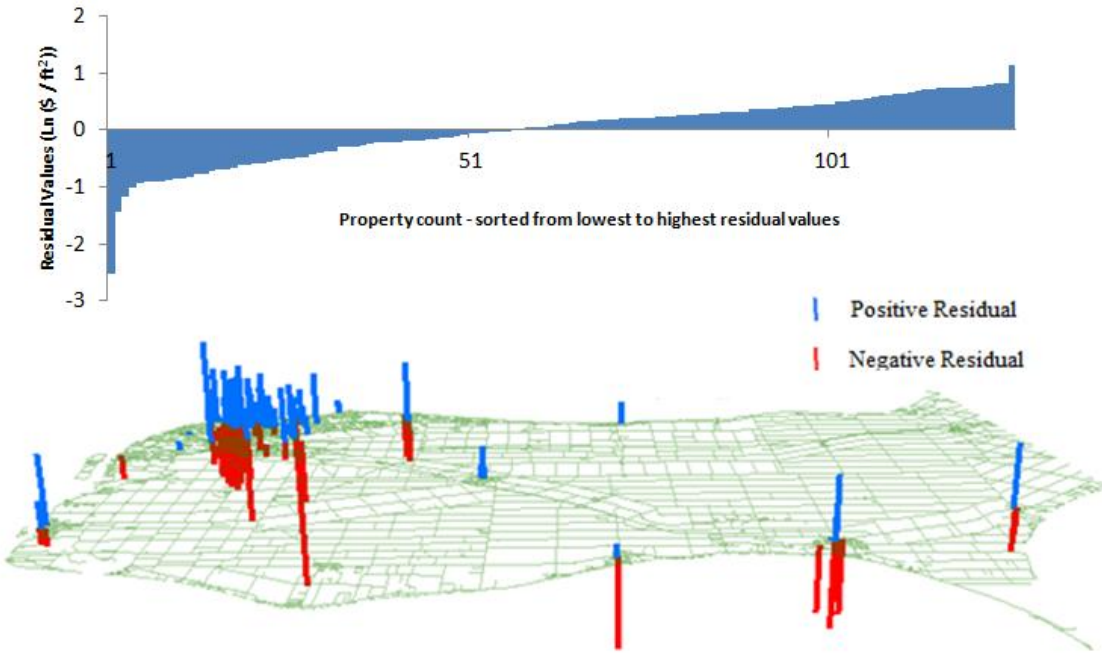

Figure F-11: Residual Values - Office

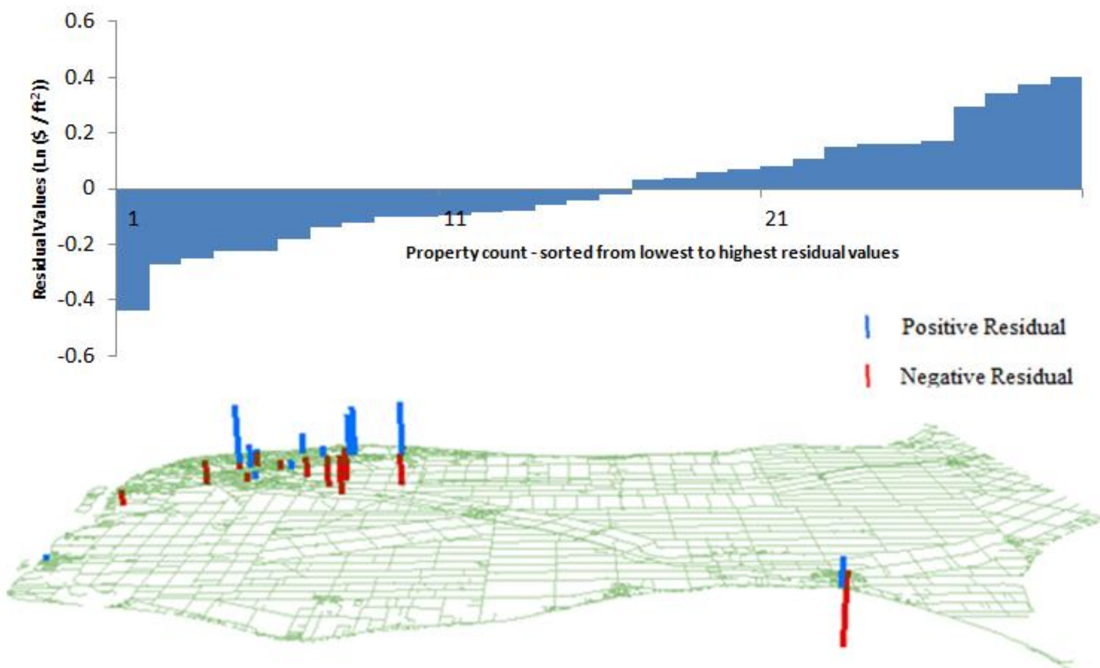

Figure F-12: Residual Values - Plaza

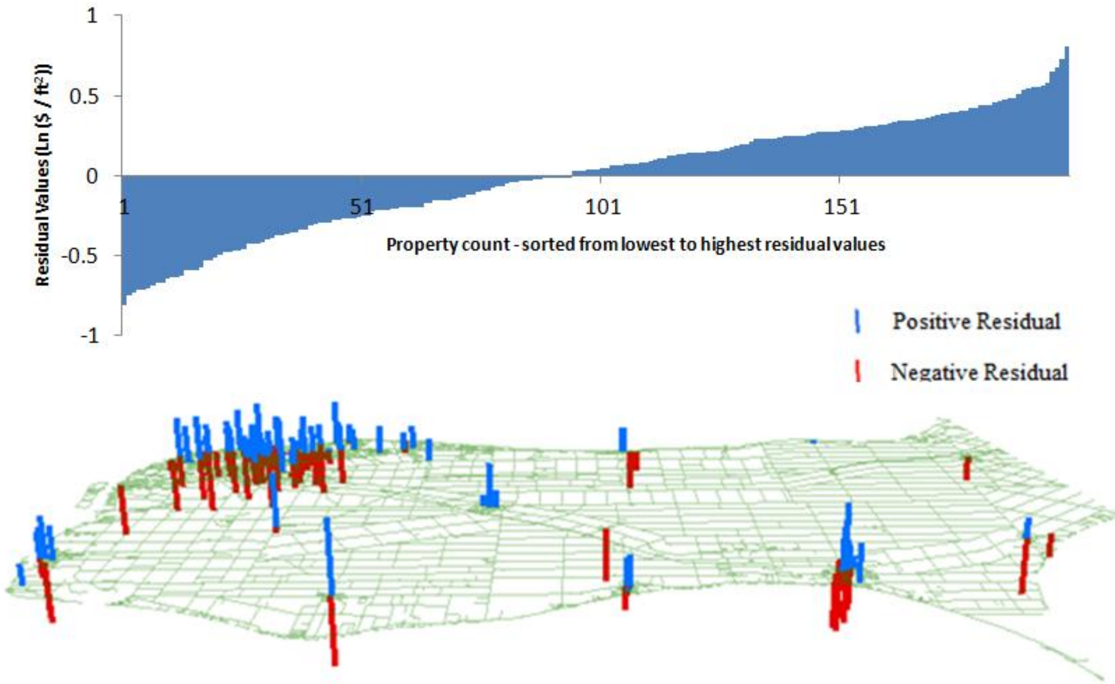

Figure F-13: Residual Values - Retail

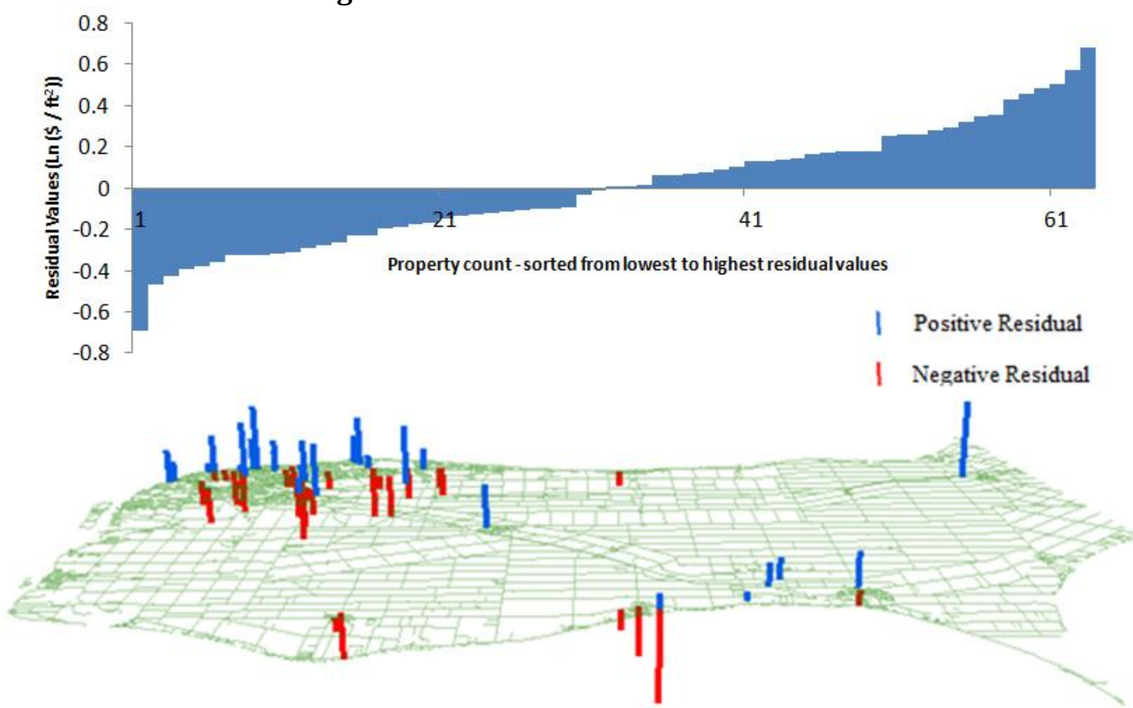

Figure F-14: Residual Values – Vacant

VITA AUCTORIS

Name: Kevin Gingerich

Place of Birth: Kitchener, Ontario

Year of Birth: 1987

Education: University of Windsor, Windsor, Ontario
2005-2009 B.A.Sc

University of Windsor, Windsor, Ontario
2011-2013 M.A.Sc

Presentations: 59th Annual NARSC Conference, Ottawa, Ontario
November 10, 2012

TRB 92nd Annual Meeting, Washington D.C.
January 15, 2013