DEMOGRAPHIC CHANGE AND SUSTAINABLE COMMUNITIES:
THE ROLE OF LOCAL FACTORS IN EXPLAINING POPULATION CHANGE

A Thesis

Submitted to the College of Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Agricultural Economics

University of Saskatchewan

By

Mark Andrew Ferguson

Permission to Use

Abstract

Ferguson, Mark A. M.Sc. University of Saskatchewan, Saskatoon, July 2005.
<u>Demographic Change and Sustainable Communities: The Role of Local Factors in Explaining Population Change.</u>

Supervisor: Dr. M.D. Partridge

Population retention and growth is a concern for cities, towns, and rural municipalities across Canada, and population change is one of the best available indicators of economic prosperity and community success. As such, it is important to understand the factors driving the location decisions of Canadians, and to use this information to help communities develop strategies to ensure their longevity and to comprehend the various features influencing future prosperity. The results of this study clearly show that local community characteristics do indeed influence local population growth. Important factors include economic indicators, the presence of different types of amenities, and the proximity of the community to urban areas.

Previous research has been completed on the topic of community population change and amenities in other countries, but Canada has not been examined until now. This study utilizes census data at the municipality level to examine these issues. The analysis consists of an econometric model with population change as the dependent variable, and a number of local factors as the explanatory variables. In general, the results of this study complied with theoretical predictions. Communities with favourable amenities and economic factors were found to have higher population growth. Also, different age groups were found to value different bundles of amenities and economic opportunities.

Amenities were found to be important factors affecting population growth. Communities with higher average housing prices and lower average incomes had higher population growth. Although this is seemingly a contradictory result, it implies that amenities have been capitalized into incomes and housing prices over time and thus more than income appears to be determining the pattern of housing values across Canada; an outcome predicted by the theoretical framework of the study. Medical amenities were found to be more important for older segments of the population, though all ages valued being near large acute care hospitals. Communities with high rates of violent crime tended to have lower population growth rates. Natural amenities such as mountains and pleasant weather, and the presence of water did not consistently result in higher community population growth.

Economic factors such as industry diversification, high local employment rates, and growing employment prospects were very important in influencing population growth, especially among younger segments of the population. However, economic and financial opportunities do not appear to affect migration decisions of the elderly, who are influenced more by medical amenities. Except for youth, *local* employment opportunities were not as important as having opportunities in *surrounding* communities. The presence of agriculture and resource extraction sectors tended to result in lower population growth. Finally, proximity to larger urban centres and population size appeared to be beneficial for communities.

Overall, the results of this study provide insight for community leaders, policy makers, and others interested in the dynamics of community population change, and will help governments efficiently allocate resources to communities and form strategies to deal with declining community populations.

Acknowledgements

I would like to thank my supervisor, Mark Partridge for his excellent ideas, motivation, and guidance throughout the development and writing of this thesis. I would also like to extend thanks to M. Rose Olfert and Ray Bollman who served as members of my graduate committee for all of their help and useful comments. I would also like to thank Larry St. Louis for his helpful comments in his role as external advisor. The combined efforts of this group of people have helped me produce a project I am very proud of.

I would also like to acknowledge the hard work that Mike St. Louis, the manager of C-RERL put into adapting and developing variables through GIS techniques, and to thank Ray Bollman and Statistics Canada for providing such a rich census database for use in this thesis. Without these important partnerships, it would not have been possible to complete this project.

Finally, I want to thank my parents Martin and Pat Ferguson for their encouragement of my academic pursuits, and the support they have given me throughout my education. I also want to thank my brother Shon for the advice and guidance he has generously provided over my university career, and Jill for her support and motivation.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.0 Introduction

Like most industrialized nations, Canada is comprised of a wide variety of different sizes of communities ranging from sprawling metropolitan areas, to smaller cities, towns, and rural communities. The economic importance of urban centres has been gradually increasing over time through the process of urbanization. Over the past century, Canadian agricultural, forestry, mining, and fishing productivity has risen to such a level that very few citizens produce their own food, and an increasing proportion of the population is choosing to live in and around cities and specialize in occupations unrelated to food (and lumber, metal and fish) production. Urban centres have become the preferred locations for much of commerce, and industry.

In general, Canadians have experienced increases in their quality of life, and business has flourished as evidenced by persistent increases in the real per capita GDP over the past century. In the midst of this success, many rural communities across Canada are in a struggle to survive. In many cases, rural areas are less fortunate than their urban counterparts in terms of their geographic location, economic growth, and level of infrastructure and public services. Although it is true that urban sprawl and congestion can inhibit productivity within cities, in many cases urban areas do have a large advantage over rural areas in terms of their business climate, availability of labour, and the amenities they provide to local citizens.

Though not to the degree of their urban neighbours, rural Canada as a whole is gaining in population. Yet, a large proportion of individual rural communities face problems due to population loss through out-migration.[1] This creates economic challenges for the communities as it becomes more difficult to sustain existing businesses and to attract new businesses and immigrants. There is a perception that many rural communities face problems related to poverty, aging populations, and retaining and attracting young people to ensure their long run vitality. Urban centres will likely continue to become more prominent in terms of their economic importance, and the percentage of Canadian residents that live in urban areas is also likely to continue to grow. However, this need not preclude rural areas thriving and growing in their own right.

One of the main indicators of economic success in regions and communities is population change. It is not a coincidence that one of the largest challenges facing rural communities is retaining their population. Evidence shows that many rural communities are successful in terms of retaining their population, while others are failing miserably, particularly in the Great Plains region of Canada. Indeed, rural and urban communities across Canada compete in attracting residents and businesses. For rural regions to thrive, they must be successful in this competition, and design strategies to level the playing field between themselves and other regions in Canada, particularly urban centres.

The key to discovering why some rural communities have shrinking populations and economies, while other communities are thriving is to determine what specifically influences people to choose one locale over another. Indeed, there are many factors which affect the decisions people make to either remain in a community, or depart. Every city, town, and rural area in Canada is unique in terms of economic conditions and

[1] Detailed definitions of "Rural will be provided and referenced in Chapter 2/

quality of life characteristics. Financial factors (i.e. wages, job availability, taxes, cost of living, etc…) are certainly very important features in determining the attractiveness of communities as places to reside.

However, researchers have shown that personal income is not the only objective influencing location decisions. In fact, people are concerned with improving their personal utility, including their quality of life. Normally, personal satisfaction includes both financial well being as well as non-pecuniary benefits from quality of life considerations. Thus, economic indicators alone do not tell the whole story, as many factors potentially affect the quality of life afforded to residents by their communities.

This perspective is supported by Partridge and Rickman (2003a), who found that income measures provided an incomplete picture of economic development in the U.S. Their findings indicate that many regions with high per capita incomes have experienced steady out-migration. Thus, "amenities", "social capital", and other non-pecuniary factors likely play a significant role in the location choices of residents. Amenities are qualities that communities possess that contribute to physical or material comfort, and subsequently increase the utility of residents. Social capital is a sense of community, goodwill, and belonging that can serve to strengthen and improve communities. Social capital has been described as the "glue that holds society together."[2]

The focus of this study will be on examining the extent to which amenity attributes and social capital act as determinants of population growth. The presence of hospitals and doctors, recreational opportunities, favourable geography, weather and the absence of crime are just a few of the amenities that may influence migration. The

[2] Detailed definitions of 'amenities' and 'social capital' will be presented and referenced in chapter 2.

importance of proximity to other types of civic infrastructure such as schools, universities, and highways can also be assessed using this type of approach.

Another important factor is that various population segments may desire different community amenities. For example, a recent survey found that young adults not only leave rural areas to pursue jobs, but also to access more entertainment, recreational, and educational opportunities. The survey also found that youth identified safety, family, and the rural lifestyle as reasons they would want to remain in rural areas (R.A. Malatest & Associates Ltd., 2002). Possible reasons for middle-aged and elderly residents to reside in rural areas are not nearly as well documented. It may be that their reasons are similar to those of young people, with availability of health care being more important and educational opportunities less important. Elderly residents may be less concerned about wage rates, and job availability.

It is vital that the role of amenities in influencing population migration be understood, especially for any amenities under government control. One example of this would be the role that health facilities and physicians play in demographic change in rural communities. A common concern voiced by rural stakeholders is that access to a nearby (usually small) local hospital is essential for the viability of rural communities. It is not known to what extent this argument has merit, because a large number of small local hospitals may actually provide an inferior level of healthcare compared with a smaller number of larger facilities located in centralized locations, as found by Liu et al. (2001).

1.1 Background

This section provides an overview of rural communities in Canada. First, a brief history of rural regions is presented to explain why Canada has a large number of small communities scattered across the country, and the challenges these communities have faced. Recent demographic statistics are then presented to describe more recent population trends across Canada.

1.1.1 A History of Rural Canada

The land now known as Canada was far from being a vast empty land when European explorers discovered North America. The land was inhabited by native people who lived in nomadic tribal groupings. Hunting and gathering was the way of life for most of Canada's First Nations. This way of life was destined to change with the arrival of the white man.

French and English settlers began to take possession of the northern parts of North America beginning around 1600. Up until the 1800's, most newcomers settled in the areas around the St. Lawrence lowlands, the Great Lakes regions, and in the Atlantic Provinces. The impetus for much of this original exploration and settlement was to support the fur trade. Due to the relatively small amount of agriculturally productive land in Eastern and Atlantic Canada, it did not take long for this land to become fully occupied by new immigrants. In the early years of settlement, Upper and Lower Canada were known for arable land, and a high capacity to sustain agricultural activity (Burnet and Palmer, 1988).

During the 1870's, the Dominion of Canada set its focus on western rural settlement. At the time, eastern agricultural land was becoming fully occupied and politicians saw the vast expanse of the prairies as an ideal location for expanding the country's agricultural production and population. Also, settlement was believed to be the best way of protecting the sovereignty of Canadian territory. The government secured the western land-base by purchasing Rupert's land from the Hudson's bay company in 1870. The Federal government then established the Dominion Lands Act, which allowed each settler to claim 160 acres of land upon which they were obliged to erect a homestead. The establishment of the North West Mounted Police in 1873, and the establishment of the Canadian Pacific transcontinental railroad in the 1880's helped to attract existing Canadian residents to the west, along with British, European, and American immigrants (Creighton, 1970).

Depots known as 'grain elevators' were constructed at regular intervals along the railway lines across the prairies to collect the primary output produced by settlers: grain, and cereal crops in particular. Small towns, which served as trade centres where early farmers and ranchers could purchase and sell goods and services were founded at these grain delivery points. Across the prairies, these communities flourished with an economy mainly supported by the grain trade (Stabler and Olfert, 1996). According to Creighton (1970), another reason for construction of the transcontinental railway and settlement of the west was to create an east-west economy. Western Canada was to provide raw materials and food to Eastern Canada, which was quickly becoming the industrial centre of the new nation. Tariffs on manufactured goods such as farm machinery were established to ensure that Western Canada would serve as a market for manufactured goods produced in the east.

Although this network of hundreds of communities across the prairies served both residents and farmers well for decades, many of them were destined to fail through no fault of their own. According to Stabler and Olfert (2002), technological advances in agricultural production, and advances in transportation, communications, and distribution infrastructure across Canada contributed to the end of prosperity in many small communities.

Technological change in the agricultural sector contributed to decreasing populations in rural areas. Mechanization of agriculture meant that fewer farmers (and farm families) were needed to produce the same output. In the absence of new economic activity to absorb the surplus labour released from agriculture, rural populations declined and thus the market area served by rural communities fell below threshold levels.

The other factor that led to declining communities in rural areas was that technological change made it more attractive for people to patronize cities. Transportation infrastructure was improved through the creation of new inter-city highways in the 1950s. This innovation gave farmers and rural residents the ability to travel to larger centres to shop. In doing so, many resident bypassed their local communities and intermediate communities in order to access a greater array of businesses in larger trade centres. As a result of this change in shopping patterns, most of the new commercial developments occurred in larger centres, while existing businesses in smaller communities closed down.

The Canadian and Saskatchewan population trends shown in figures 1.1 and 1.2 certainly support the idea that urban centres are becoming more prominent across the country, yet it is not totally clear that rural areas are struggling when examining the Canadian trends.[3] In Saskatchewan, rural areas are clearly losing their populations, while rural population continues to grow across Canada, despite the challenges rural areas face. In Saskatchewan, rural areas made up the majority of the population until 1971, while in

[3] Figures 1.1 and 1.2 utilize a different definition of 'rural' than the remainder of the study. These charts define rural as any centre under 1000 people, while the definition of rural utilized for remainder of the study is the Rural and Small Town (RST) definition described later.

Canada as a whole, rural areas made up a majority of the population only until around 1931.

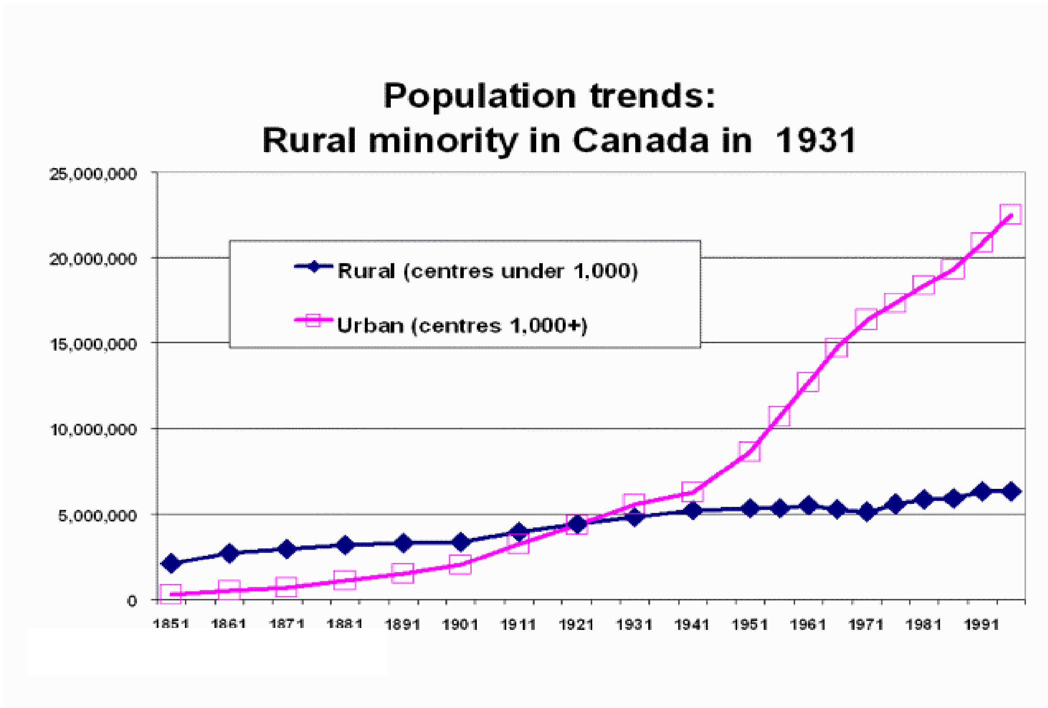

Figure 1.1: Canada: Rural and Urban Population, 1851-1996
Source: Statistics Canada Census of Population, 1851-1996, Courtesy of Ray Bollman

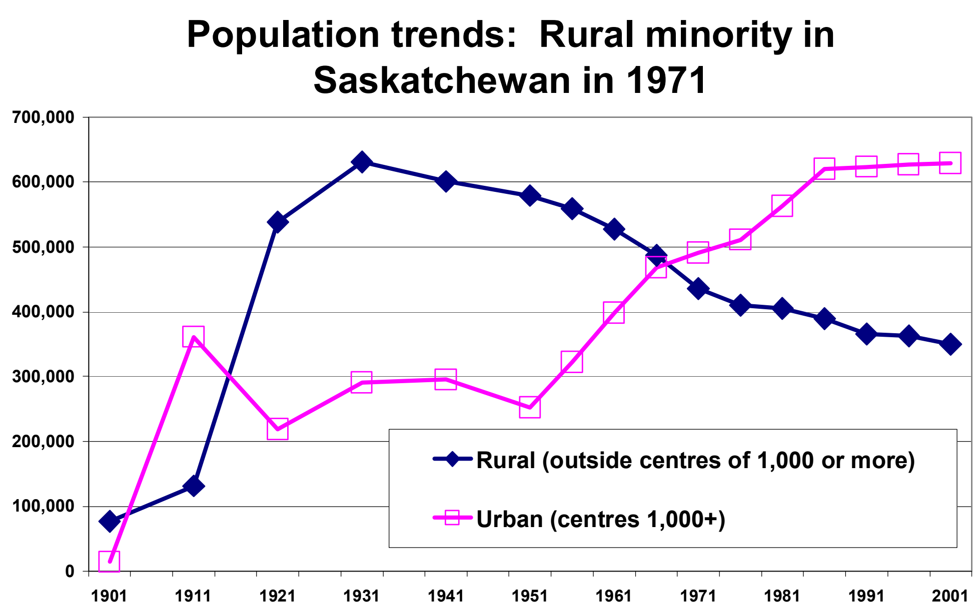

Figure 1.2: Saskatchewan: Rural and Urban Population: 1901-2001
Source: Statistics Canada Census of Population, 1901-2001, Courtesy of Ray Bollman

Given the trends presented above, the issue of maintaining the economic well-being and population levels of rural areas is of keen interest to the province of Saskatchewan. The Action Committee on the Rural Economy (ACRE) was established following a promise by the provincial government in the December 1999 throne speech to establish a committee to examine rural problems and propose solutions. ACRE was composed of a group of respected community leaders, their mandate being to develop recommendations for the government to strengthen and diversify rural communities in the province. The government has also recently (Spring, 2005) separated the Rural Issues office from the Department of Agriculture to create a new Department of Rural Development at the request of ACRE.

It is unclear what (if any) impact their recommendations have had thus far, although it should be noted that their recommendations were presented fairly recently (2002), and a number of their recommendations were never adopted. The future will show whether this type of broad government policy implementation has any hope of stemming the flow of rural residents to urban areas within the province, or even more troubling, leave the province altogether.

1.1.2 Rural Trends and Statistics

Demographic and regional trends of the past two decades provide a good picture of what is happening in rural areas in terms of population growth, and makeup. The objective of this section is to describe recent population trends in rural and small town Canada, and compare them to Canada as a whole.

In general, rural population has been growing at the national level. Between 1981 and 2001, both rural areas and urban areas in Canada have experienced population increases. Since 1981, rural population has grown by 418,255 people or 7.7 percent.[4] As a whole, Canada's population has grown by nearly six million people, or 24.6 percent since 1981. Clearly, most of the population growth in Canada can be attributed to urban areas, yet rural areas as a whole are experiencing some increases.

However, it is interesting to note that in Canada 546 of 1895 (28 percent) of rural communities lost population between 1991 and 2001. Five hundred forty six communities across Canada are losing population, but this number vanishes when you just look at the total growth in rural populations. Many of these struggling communities are located on the prairies where the problem is much more apparent. For example, Saskatchewan's rural population grew by about 5,000 people (1 percent) between 1991 and 2001, yet 179 of 274 (65 percent) rural communities lost population. Again, the average population increase would seem to indicate that rural communities in aggregate are not losing population; yet with 65 percent of communities losing population, the majority of communities are experiencing population losses.

Out-migration of young adults, and aging populations are several adverse trends often mentioned in the media when describing rural demographics. Population pyramids are useful tools to examine whether these arguments have any merit. Figure 1.3 shows a population pyramid for rural communities and Canada as a whole for the year 1991, while figure 1.4 shows the same pyramid for the year 2001. The pyramid shows the proportion of the population (in percent) that each age group occupies.

[4] Throughout this thesis, any un-referenced statistics were calculated by the author using data from the Statistics Canada census of population, and the C-RERL research lab at the University of Saskatchewan.

Regarding youth out-migration, the data does not provide clear evidence to support or reject this argument at a national level. For example, figure 1.3 shows that the percentage of young people age 20-29 in 1991 living in rural areas was smaller than the percentage of young people living in Canada as a whole. Assuming that differences in fertility between rural and urban areas were not responsible for this discrepancy, it is likely that out-migration could account for this difference.

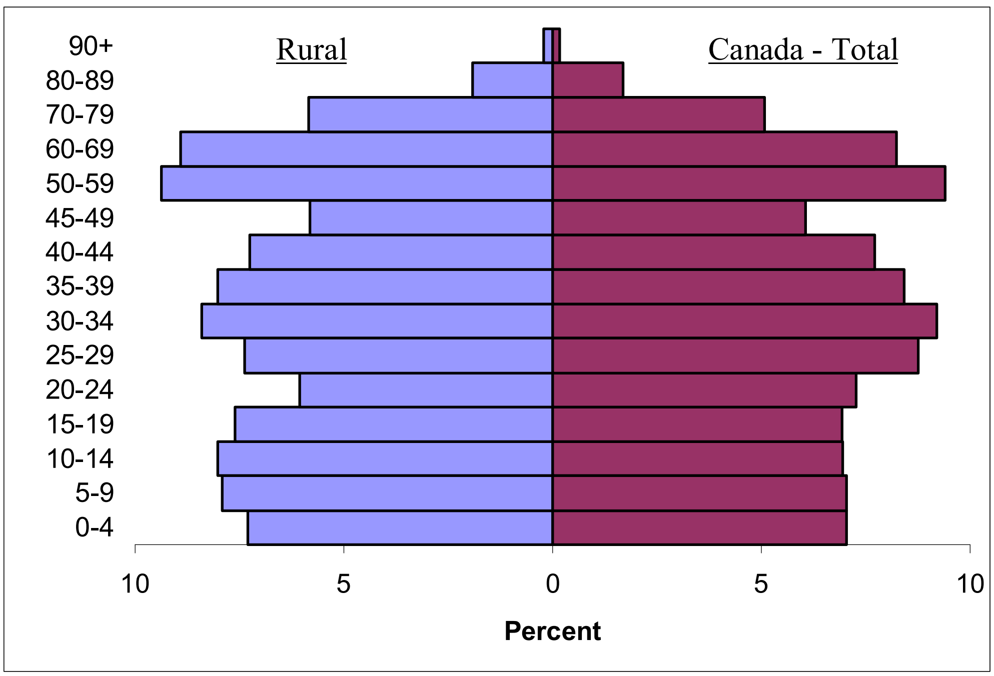

Figure 1.3: Population Pyramid Including Rural and Urban Areas in Canada for the Year 1991
Source: Statistics Canada Census of Population

However, if you examine the 1991 pyramid and compare it to the 2001 pyramid, it is not clear that youth out-migration is occurred between these two periods. In fact, between 1991 and 2001, it appears that young people remained in rural areas, and the proportion of young people is increasing. What is very noticeable is that in 2001, the age group 30-39 comprises one of the largest cohorts in rural areas, whereas in 1991, the age group 20-29 was one of the smallest. It is possible that many young adults that had perhaps not lived there ten years earlier migrated into the rural. Whether this represents young adults returning, or new migrants cannot be determined from this data.

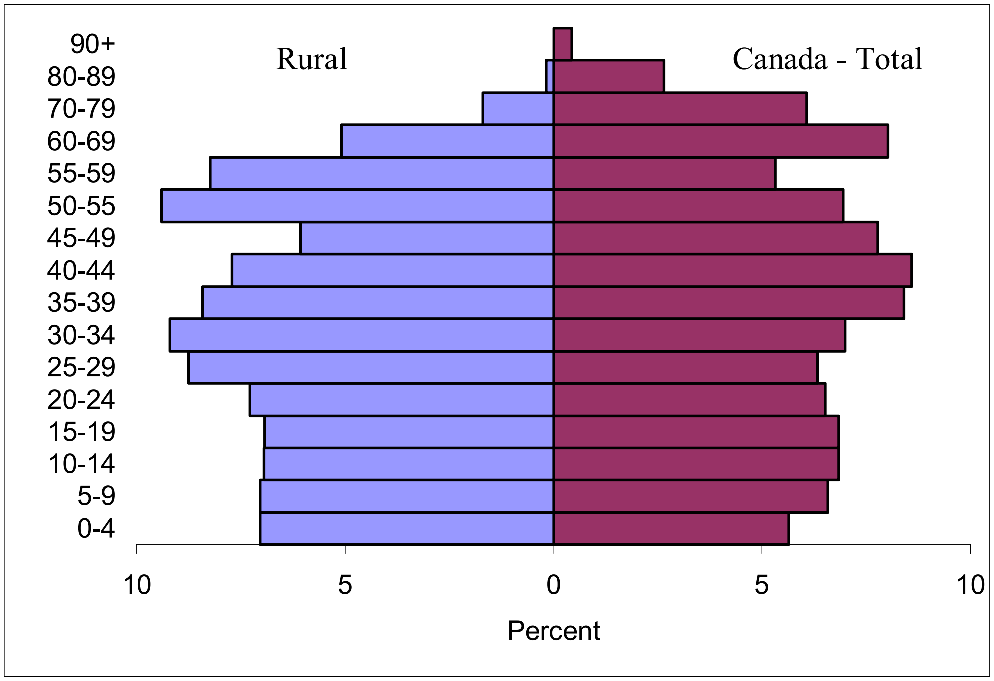

Figure 1.4: Population Pyramid Including Rural and Urban Areas in Canada for the Year 2001
Source: Statistics Canada Census of Population

One stereotype that can be clearly addressed in these diagrams is that of rural areas have aging populations. In general, the older population cohorts in rural areas are not significantly different from the Canadian totals. In fact, Canada as a whole has a larger proportion of elderly individuals than rural areas, indicating that older population cohorts are migrating into urban areas.

The statistics presented here represent national averages, but they provide an overall picture of the situation facing rural communities in the 21st century. In general, rural populations are increasing, but there are many communities waging an uphill battle against population loss, especially on the prairies. Young people may or may not be abandoning rural areas, but it appears that young people have been returning, as they get older (30-39). Finally, it does not appear to be the case in Canada that elderly people are becoming more concentrated in rural areas. The opposite appears to be the case, as elderly people appear to be moving to urban areas.

1.2 Need for the Study

There are three main reasons why this thesis is important. First, the role of amenities in influencing internal Canadian migration has never been explored. Indeed, most of the literature produced to describe the role of amenities and social capital in migration and economic development has used U.S. data, and it does not necessarily follow that these results apply to Canada. For example, Canadian migration and rural economic development may follow fundamentally different patterns than in the U.S. due to differences in industry mix, different types of social programs, and differences in the relative sizes of population centres. This research will be of particular importance to all levels of government, health care administrators, and community leaders in understanding the effects of location specific amenities and social capital on migration patterns and developing strategies to influence future population growth. It is important for all stakeholders to consider all of the forces that are driving this demographic shift.

The second reason this study is necessary is because all previous studies looking into the effects of amenities on the economic well being of specific regions have used a fairly limited set of amenities. In most cases, weather data have been the only variables used to represent amenities. In other cases, sets of amenities have been vaguely grouped into broad categories, which do a poor job of indicating the importance of different types of amenities. This thesis uses more unique measurements of amenities and other factors than have ever been utilized in a localized study of Canadian population movement. As such, this study will add to a growing body of literature on amenities, and provide economists with new variable ideas, and test the effectiveness of amenity measures constructed through GIS methods.

The third reason this thesis is important is because it will not only look at total migration; migration among different population cohorts will be examined. Other than retiree migration, past studies have not really examined these trends, and have mostly focused on total population change. In this study, population will be separated into distinct age groups such as youth, young adults, adults nearing retirement, and the elderly. Each of these groups will be examined to see how different local factors influence their migration decisions. These results will help policy makers understand how different local factors influence different age groups.

1.3 Purpose and Objectives

This thesis will provide an in-depth examination of the causes of population migration and demographic change in Canada. Using a variety of explanatory variables, various econometric models will be estimated to determine the exact causes of people moving to or departing rural areas. The primary objective of this thesis is to examine the extent to which location choices of rural residents are influenced not only by financial and economic factors, but also location-specific amenities and social capital within the community.

1.4 Hypothesis

The hypothesis of this study is that in general, the presence of amenities and social capital have significant and positive effects on population migration within rural communities[5]. The associated findings will aid policy makers in designing the proper mix in enhancing quality of life versus enhancing economic opportunities in sustaining rural Canada.

Several of the specific hypotheses to be tested are as follows: 1) lack of recreational amenities and educational opportunities are important factors in the loss of young adults from rural communities, 2) the presence of hospitals and physicians influences population migration, especially among older population segments, and 3) data generated through GIS databases and techniques can produce statistically significant variables that comply with economic theory for econometric studies.

1.5 Organization of the Thesis

This thesis is composed of six chapters. A review of relevant literature, including a summary of past theoretical and empirical models is included in Chapter 2. Chapter 3 provides the theoretical framework that is used to analyze the problem, and puts the methodology of this thesis into the context of the applicable economic theory. The methodology of the thesis is described in Chapter 4, including data sources and descriptions, and the empirical model. The results of the regression analysis are presented in Chapter 5, with conclusions following in Chapter 6.

[5] The null hypothesis is that amenities and social capital do not influence community population change.

Chapter 2: Literature Review

2.0 Introduction

The objective of this chapter is to outline the concepts utilized, and studies completed in the past that are relevant to the topic of this thesis. There are several areas of literature that provide the theory and practical empirical experience necessary to understand what factors affect migration, why amenities and social capital are indeed important factors in influencing migration, and how one should proceed with a similar study. In the first section, a broad variety of different concepts and definitions are examined to give readers the necessary analytical framework with which to look at the issue of rural population change.

Past Canadian migration studies will then be reviewed, revealing a substantial gap in the literature. It will be shown that there has been a lack of migration studies using sub-provincial level data, and the influence of amenities and social capital have never been examined using Canadian data. Next, utility theory, which forms the theoretical basis of this thesis will be examined. The evolution of theoretical models used in past migration studies will then be outlined, followed by an outline of the relevant empirical studies that have been completed to date.

2.1 Basic Concepts

Throughout this thesis, various terms will be presented which may seem familiar to readers, yet their definition is actually more complex. This section addresses the concepts “community” and “rural” and although most readers will recognize these terms, this section will show that defining and applying these concepts to empirical research terms is not simple. The importance of migration to communities, and factors influencing migration will then be discussed to give readers a framework to begin thinking about how to analyze migration patterns within Canada, and the different stimuli that affect these patterns. Within this discussion, the terms "amenities" and "social capital" will arise as important factors influencing migration. These concepts form an important part of this thesis, and are subsequently defined.

2.1.1 Defining "Communities"

A community is defined as "a body of people organized into a political, municipal, or social unity" (Oxford University Press, 2004). In order to study trends affecting communities delineated in geographic space, one must set geographical boundaries on what is believed to constitute a community. In examining Canadian population trends, researchers are limited by the fact that access to individual-level census data, and data for small geographic areas is restricted because legislation guarantees protection of the privacy of Canadians when they complete the census questionnaire. Individual-level data may be preferable for researchers in examining rural trends, yet researchers external to Statistics Canada are limited to an extent in the level of data that can be retrieved from the census database.

Data consisting of groupings of individuals identified by geographic location was chosen as the unit of observation for this study. Census data can be retrieved for a number of different territorial units ranging from enumeration areas to Census divisions. Table 2.1 describes a selection of the different levels of geography that are potentially available for research use. Since data for this thesis is based on 1996 geography, the 1996 definitions are provided here.

Table 2.1: Census Geography

Name	Description
Enumeration Area (EA)	Geographic Areas canvassed by one census enumerator ranging from a minimum of 125 dwellings in rural areas to a maximum of 650 dwellings in urban areas. There were 49,361 enumeration areas in the 1996 census.
Census Subdivision (CSD)	An area that is a municipality (such as a rural municipality, town, or city), or equivalent to a municipality for statistical reporting purposes (such as a reserve or unorganized territory). This is a level of geography in between the EA and CCS. There were 5984 CSDs in the 1996 census.
Consolidated Census Subdivision (CCS)	A grouping of adjacent census subdivisions. Generally the smaller, more urban census subdivisions (towns, villages, etc) are combined with the surrounding, larger, more rural census subdivision, in order to create a geographic level between the census subdivision and the census division. There were 2607 CCSs in the 1996 census
Census Division (CD)	Group of neighbouring municipalities, joined together for the purpose of disseminating statistical data. In Nova Scotia, New Brunswick and Ontario, they represent counties and they represent MRCs (Municipalité Régionale de Comté) in Quebec. There were 288 CDs in the 1996 census. This level of geography is between the CCS and provincial level.

Source: Adapted from definitions in the 1996 Census Dictionary- Final Edition (Statistics Canada,1999)

EAs most likely represent neighbourhoods, while CSDs and CCSs could probably be described as communities, and CDs as regions. Data suppression is a major problem when accessing census data and in general, any geographic region with less than 250 people will have some data suppressed to protect the privacy of individuals living within the region. EAs are likely too small a geographic region to be considered communities. Although CSDs could be appropriate in terms of examining communities, many rural observations would be suppressed if this level of geography were selected. Also, CSDs may not be appropriate in some situations because individuals living in rural municipalities surrounding towns may associate themselves with the community of the nearest town or village, though they may technically be in separate municipalities. CCSs combine several adjacent CSDs into larger geographic areas, which combine CSDs with smaller populations. Since CDs are too large to be considered as communities, CCSs become the obvious geographic choice to represent "communities." In this thesis, CCSs are considered to be communities.

2.1.2 Defining "Rural"

Since this study is primarily concerned with rural population migration, it is appropriate to define exactly what is meant by "rural" in the context of this thesis. The concept of "rural" will continue to be referred to within this study, so it is important to understand exactly what is being implied by this term. Du Plessis et al. (2002) have put considerable effort into developing the definition of "rural." They contend that the method researchers use to delineate rural areas will have a definite impact on research findings.

Like du Plessis et al., this thesis refers to rural as a geographical concept with identifiable boundaries. The challenge lies in setting these boundaries. For example, most people would agree that large-scale farming areas should be considered rural, but what about hamlets and villages near these farming areas? Should towns and very small cities be considered rural or urban as long as they are not near major urban centres? What is obvious is that there is no firm definition of what constitutes a rural area, and the definition of "rural" is subject to interpretation. Certainly one way of distinguishing rural areas from urban is through distance (i.e. the distance from markets and amenities), and density (a lack of agglomeration economies). Thus, farming is not the best gauge of a community being rural. Many non-farming communities can be isolated and considered to be rural, just as farming communities located adjacent to large cities may not actually be rural. In their study, du Plessis et al. provide six possible definitions of "rural." These definitions are identified below in table 2.2.

Table 2.2: Definitions of "Rural"

Term	Population Size	Description
Census "Rural Areas"	Population living outside places of 1,000 people or more OR population living outside places with densities of 400 or less people per square kilometre.	Sparsely populated lands lying outside urban areas.
Rural and Small Town (RST)	All areas with an urban core population of less than 10,000 as long as they are outside the commuting zones of areas with an urban core population of more than 10,000 people	Population living outside the commuting zones of larger urban centres. Specifically outside of Census Metropolitan Areas (CMAs) and Census Agglomerations (CAs)
Metropolitan Area and Census Agglomeration Influenced Zones (MIZ)	N/A	A refinement of the RST measure of rurality with four subgroups based upon the percent of the workforce in the CSD that commutes to any large urban centre (10,000+).
OECD Rural Community	<u>Rural Communities:</u> Population has a density of less than 150 people per square kilometre.	This measure is at the CCS level as any CCS with a population density of less than 150 people per square kilometre.
Non-Metropolitan Regions (modified Beale Codes)	Rural non-metropolitan regions are identified as CD with no urban settlements of 2,500 or more. A non-metropolitan region is a CD with less than 50,000 people living in settlements of 2,500 or more.	This measure essentially provides a scale which identifies the size of the area (divided into metropolitan, small city, small town, rural, and Northern hinterland) and its adjacency to a places with 250,000+ people.
Rural Postal Codes	N/A	Places with a zero in the second position of their postal code. Not generally applied to census geography.

Source: Adapted from du Plessis et al. (2002)

For the purposes of this thesis, the rural and small town (RST) definition is adopted, which accounts for both the size of the community (i.e. the density criteria of being rural), and commuting patterns (i.e. the distance part of being rural). Rural communities are considered to be any CCS without a component CSD that is part of a

Census Metropolitan Area (CMA) or a Census Agglomeration (CA)[6]. When "rural" is referred to in the context of this study, we are actually referring to any communities that do not have an urban core population of 10,000 or more people, and are not within the commuting zone of an adjacent community with more than 10,000 people.

2.1.3 Migration – Why is it important?

This study adapts the definition of migration proposed by Goetz (1999). Migration will be considered to be movement of people across CCS boundaries for the purpose of establishing a new place of residence.

Goetz (1999) provides numerous reasons as to why migration, and the ability to retain and attract population are important drivers of rural economic development. Migration is the result of the level of economic activity within communities. Loss of population is the most recognizable indicator that a community is in socio economic decline, with movers leaving behind infrastructure that is expensive for remaining residents to maintain, which exacerbates the problem. Also, many communities do not realize any return on investment from their education dollars when high school graduates leave the area. This trend is commonly known as "brain drain", and education investment is known as "human capital investment."

As residents (usually the more educated and employable ones) vacate rural areas, the ability of local governments to provide police, fire, health services, and education will decline. Often, provision of these types of services is comprised of both fixed and

[6] Census metropolitan areas (CMAs) and census agglomerations (CAs) are formed by one or more adjacent CSDs centred on a large urban area (known as the urban core). The census population count of the urban core must be at least 10,000 to form a census agglomeration and at least 100,000 to form a census metropolitan area. To be included in the CMA or CA, other adjacent municipalities must have a high degree of integration with the central urban area as measured by commuting flows derived from census place of work data. To be specific, neighbouring CCSs where 50 percent or more of the population commutes into the urban core are considered to be part of the CA/CMA (Statistics Canada, 2004).

variable costs. The problem facing communities losing population is that variable costs can be reduced, but fixed costs cannot be reduced. As the fixed costs are spread over fewer and fewer residents, this increases the financial burden on those left behind.

Another argument for stabilizing rural populations relates to the brain drain issue, and poverty. While policy makers may argue that it is desirable for people to move away from rural areas, they need to be mindful of who is left behind, and who will actually move away from rural areas to seek employment elsewhere. Usually, it will be the more skilled educated residents that will move first. Those leaving an impoverished area have more marketable job skills than those left behind, and in general, steady out-migration tends to lead to a greater concentration of poverty (at least in the USA, perhaps not to the same extent in Canada) that may be especially difficult to mitigate.

2.1.4 Drivers of Migration

Brewin (2004) provides a summary of the drivers of population migration, and contends there are many factors besides income that affect the movement of labour. The characteristics of the home area, the destination area, and the household itself are all important factors in influencing migration. Goetz (1999) notes that the decision process of potential migrants is likely similar to a cost-benefit analysis, where potential migrants weigh the costs of moving between two places with the benefits of moving. The costs and benefits of moving would include both pecuniary and non-pecuniary factors.

Pecuniary factors include wages, the probability of finding employment, the cost of living, and the cost of searching for new jobs and homes. Non-pecuniary factors could include proximity to relatives, or living in a pleasant climate, or near enjoyable entertainment options. Individuals are willing to forego income and incur lower incomes, or a higher cost of living for the benefit of living near amenities. Regional differences in the availability of amenities are key factors in the development and growth of regions.

Goetz (1999) established that the education of residents affects their propensity to migrate. Thus, “human capital”, or “human capacity” has an important influence on migration. Human capital also represents the ability of individuals to contribute to their community and improve their own well-being. The average level of formal education is usually used as a proxy for the level of human capital in a community (Bollman, 1999).

2.1.5 The Role of Amenities

The term “amenity” commonly refers to any feature that increases the attractiveness of a location or piece of real estate. Alternatively, amenities can be described as the pleasurable aspects of a specific locale (Oxford University Press, 2004). Economists have increasingly recognized the role of amenities in regional economic development since the 1970s, and this fact will be demonstrated throughout this chapter.

It is important to note that amenities can be categorized according to their age, and whether or not they are man-made. Bruekner, Thisse, and Zenou (1999) dissected the term “amenity”, and have classified the expression into three categories. They define *natural amenities* as those “generated by an area’s topographical features, including rivers, hills, coastline, etc.” (Bruekner, Thisse, and Zenour, 1999, p. 94). Natural amenities would also include climactic conditions. Another category these authors pioneered was that of *historical amenities*. Examples of historical amenities include monuments, buildings, and any other type of infrastructure from a past era. Both natural amenities and historical amenities can be considered exogenous to the current economic conditions of the local community such as income and employment levels. In other words, in any empirical model, historical and natural amenities are predetermined; their levels are determined outside the model. These types of amenities can be used as a causal factor in determining population location patterns without any endogeneity concerns.

The third and final category of amenities identified by Bruekner, Thisse, and Zenou are *modern amenities.* These amenities include any modern buildings, or public institutions, and could include restaurants, movie theatres, swimming pools, or even schools. Modern amenities are endogenous, meaning that their levels depend upon the current economic conditions in each community, and are perhaps even dependent upon the levels of natural and historical amenities present in the community. For example, a community with more movie theatres may attract more population, but at the same time, a community with more population may simply be attracting more movie theatres. In other words, modern amenities may be both a consequence, and a cause of the location choices of residents, or the income of residents. This endogeneity problem can complicate empirical models, making it necessary to employ some type of simultaneous equations model. Alternatively, the most common approach utilized by economists has been to simply ignore modern amenities in their empirical studies, invoking excluded variable bias.

2.1.6 The Role of Social Capital

The term “social capital” can be described as the institutions, relationships, attitudes, values, and beliefs that facilitate interaction among people and contribute to economic and social development. Social capital facilitates a sense of community, goodwill, loyalty, and a belonging. From an economic standpoint, social capital reduces transaction costs within the community by increasing trust between people in the community, and reduces the need for law enforcement and security. It is suggested that the level of social capital present in communities can be gauged through various indicators including participation in voluntary associations (churches, clubs), and voter turnout in elections (Rupasingha, Goetz, and Freshwater, 2000).

2.2 Relevant Canadian Studies

Most of the literature produced to describe the role of amenities and social capital in migration and economic development has used U.S. data, and it does not necessarily follow that these results will apply to Canada. For example, Canadian migration and rural economic development may follow fundamentally different patterns than the U.S. due to differences in industry mix, the different types of social programs, and differences in the relative sizes of population centres. Prior Canadian studies on migration have not had an explicit focus on amenities and social capital as causes of migration. What follows is a review of recent Canadian migration studies.

Day (1992) looked at average temperatures, and the level of public spending in the health, education, and social services sectors as determinants of inter-provincial migration. It could be argued that these variables represent amenities, though Day refers to them as "local public goods." She found temperatures to be highly significant, while all types of government expenditures were also found to be significant. In spite of highly significant results, this study is limited for rural-urban analysis because the analysis was done at the provincial level, and thus the results are less useful in examining sub-provincial problems such as rural development or community level problems.

Day's paper appears to be the only Canadian study that explicitly looks at the role of amenities in migration. Other papers such as those by Dickie and Gerking (1998) and Day and Winer (2001) acknowledge the theoretical importance of amenities, yet do not incorporate amenities and social capital into their empirical models. Still others such as Newbold (1996), Courchene (1970), Rosenbluth (1996), and Islam (1990) neglect to mention the theoretical importance of amenities, and do not account for amenities in their empirical migration models.

Even papers produced very recently do not utilize local amenities in their analysis. These omissions are exacerbated by the fact that the body of literature on amenities and social capital has been rapidly growing since the 1970's. For example, in a recent paper, Shearmur and Polèse (2004) attempted to explain why employment and population growth-rates differed across regions in Canada between 1971 and 2001. Their study included both rural and urban areas, and was completed at the census division level. Interestingly, the model they produced did not include one single measure of amenities, and also lacked many of the economic indicator variables one would usually expect to find in a paper on economic development.

In summary, the literature explaining Canadian regional migration has lacked a number of important explanatory variables, and many of the studies are somewhat dated. This chapter will clearly demonstrate that local amenities must be included as an integral part of any study examining regional economic development. Day's 1992 study does attempt to include local amenities but uses geographic regions that are too large to be useful for community level analysis. There clearly exists a major gap in Canadian regional migration literature.

2.3 Utility Theory

The objective of this study is to examine what influences people to choose one community over another as their place of residence. To begin this process, it is necessary to have an understanding of the concept of utility. Utility maximization is a fundamental concept used in the theoretical model of this thesis. In general, utility theory is a descriptive model of human behaviour; in other words, a method proposed to predict how people behave. Utility is defined as the level of satisfaction individuals obtain from consuming a good or service. People gain utility by consuming things that provide pleasure, and avoiding things that produce unhappiness (Pindyck and Rubinfeld, 1998).

Utility was originally viewed as a *measurable* level of satisfaction dating back to the late 18th century when Jeremy Bentham argued that society should attempt to achieve the "greatest good for the greatest number" of people by maximizing an aggregate social utility function. Bentham believed it was possible to measure, or assign numerical value to utility. The problem with the *Cardinal Utility* approach is that utility is not observable, and it is impossible to measure how much enjoyment different individuals get from different items, and compare these measurements (Binger and Hoffman, 1998.)

In modern economic analysis, an *Ordinal Utility* approach is most often utilized. This approach assumes that the concept of utility is only valuable as a representation of a consumer's preference over consumption bundles. Modern utility theory is most often used to summarize preference rankings of different market baskets. Individuals yield information on their preferred basket through their choices, either in experiments or in reality. If we have access to information on the choices individuals make when choosing among various market baskets, we can rank different market baskets based upon individual preferences. This concept is known as "revealed preference" (Pindyck and Rubinfeld, 1998).

2.4 Theoretical Significance of Amenities

In the realm of economic literature, the study of amenities is relatively new, spanning only about 30 years. Knapp and Graves (1989) argue that a common omission in past migration models was that they did not account for location-specific amenities. They contend that amenities are a critical factor in determining mobility behaviour, and must be used in combination with economic indicators to form reliable models.

It is interesting to examine the evolution of the theoretical foundation of amenity-based papers. There have been several different avenues for modeling the theory behind the role of amenities in migration. One approach has been to use a cost-benefit framework to explain migration, while another has been to use a gravity model. The dominant approach in the literature has been to utilize a framework based upon utility maximization.

In perhaps the earliest study examining migration and quality of life, Cebula and Vedder (1973) used a cost benefit framework to validate their empirical model. Benefits to migration include higher wages, and more employment opportunities. Costs would include the monetary cost of relocation. They consider amenities to be "psychic benefits" that can be expressed in pecuniary terms. Residents will migrate if the benefits of migration between two areas exceed the cost of migration. Although no mention of utility is made, this model basically amounts to a utility model where utility measurements are assumed quantified into dollar values.

Porell (1982) utilized a generalized systemic gravity model of migration to relate his empirical model to economic theory. A system of equations is used to predict migration between different places. Amenities are incorporated into the model through a scoring system that incorporates quality of life and economic factors. This "score" could be interpreted as a cardinal utility measurement. This type of gravity model seems to be unnecessarily complicated, and not needed to explain migration when a simple utility maximization framework would suffice.

The majority of amenity studies have used some type of utility maximization framework to relate migration and amenities to economic theory. Most of the papers described in the next section use a simple objective function based upon utility maximization. Other authors have included more complex theoretical models that relate amenities to wages, housing costs, and amenities. Other models include both firms and individuals. The most influential theoretical model in the area of amenities and migration was offered by Roback (1982). Roback assumed that utility maximization drives the location choices of workers, whereas profit maximization determines firm location. In her model, levels of wages, land prices, and amenities are determined simultaneously in the labour and land markets. Many authors have followed Roback's framework, and her model is examined in detail and adapted in the theoretical framework of this thesis.

2.5 Empirical Amenity Studies

Most empirical research that has examined the value of amenities has utilized an econometric model with some indicator of economic prosperity (population growth, employment growth, income growth, etc…) as the dependent variable. An array of variables detailing the economic conditions and amenity levels within each region are invariably used as the explanatory variables. Most authors have chosen to utilize data for specific geographic regions, while other authors were fortunate to discover household-level data that suited their research needs. Most of the early papers that examined amenities took a very limited view of what constituted an amenity, utilizing a limited number of natural and modern amenities

Weather variables, crime data, and some environmental factors were the main amenities used in these early models. Also, many early papers used fairly limited data, which only included a small number of metropolitan areas. Around the year 2000, authors began creating more sophisticated models with more complete specifications. We will first examine these early models, and then move onto more modern studies.

2.5.1 Early – Urban amenity Studies

During the early 1970s, economists became more interested in the economic impact of quality of life (QOL) factors on urban areas. In one of the earliest studies on this topic, Cebula and Vedder (1973) note that all previous studies in the relatively new field of QOL were primarily concerned with the effects of pollution and environmental factors on QOL. Among others, Getz and Huang (1978) attempted to estimate hedonic prices for urban amenities. These early environmental papers, and papers attempting to measure wage and rent differentials between cities with different QOL factors appear to be the point of departure for the study of the economic impact of amenities.

In their unique study, Cebula and Vedder (1973) attempted to discover whether metropolitan migration within the U.S. could be explained by a variety of economic, social, and environmental variables.[7] The authors' model ultimately found all variables except pollution to be significant, with some of the amenity variables significant at the two percent level. Cebula and Vedder concluded that their model supported the notion that migrants are responsive to QOL factors.

Getz and Huang (1978) used a rich set of amenities (for the period) to estimate the impact of amenities on the wages of white males in 39 metropolitan areas. Using individual-level data, and a list of variables that included cost of living, migration rates, commuting time, violent crime, air pollution, weather, employment, health facilities, and pollution, the authors estimate a fairly detailed econometric model, and obtain good results. In general, the model explains a high proportion of the variation in wage rates between individuals. The most important determinants of the wage rate were the cost of living, healthcare, education, and crime.

Graves and Linneman (1979), and in a similar paper, Graves (1980) examined the role of amenities in influencing migration. In these papers, the authors examined how differences in income, unemployment, race, and climate in urban centres in the U.S. affected migration between urban areas. Graves showed that climate amenity variables tended to be significant.[8] Furthermore, when the climate variables were removed from the model, the income variable tended to be insignificant with an unexpected sign. These results suggest that location-fixed amenities (primarily climate) are likely more important than income or unemployment as determinants of migration.

---

[7] The explanatory variables utilized by Cebula and Vedder (1973) included the economic variables income, the unemployment rate, and changes in income. Quality of life variables were physicians per 100,000, population, crime rate, percent of non-white population, and days below freezing. The sole environmental variable utilized was the amount of pollution in the air.

[8] Graves (1980) specifically utilized heating degree days, cooling degree days, annual temperature variance, relative humidity, and wind speed.

Graves contends that his results show income and unemployment rates represent the compensation required for residents to be indifferent between different locations with varied levels of amenities. It is evident that Graves supports the notion of a long-run equilibrium theory of migration. He suggests that there is a long-run relationship between amenities and economic opportunity variables, and this is the reason why traditional migration studies (that do not utilize amenity variables) failed to find significance on economic variables. Graves also argued that there may be no *correct a priori* sign for economic variables because they are simply adjusting for different levels of amenities.

Graves readily acknowledges that his model is likely under-specified, and suggests that the existence of other natural features such as mountains, or recreational water could be considered as amenities. Furthermore, man-made location-specific goods such as sporting events or symphonies could arguably be included in future models. This seemingly infinite number of pool of potential amenity variables, and possible correlation problems with these variables was the topic of another paper by Graves (1983). He hypothesised that rents (housing costs) could be used to proxy for all interrelated amenity variables. His model featured migration for various age cohorts as the dependent variable with three independent variables: unemployment, income, and rental rates. Overall, he found positive and significant coefficients on the rent variables and interpreted this result to lend support to his hypothesis.

Graves (1980) was followed by Porell (1982), who developed a similar yet more sophisticated model to assess the relative importance of economic and QOL factors as determinants of migration.[9] Porell's model was very detailed and well specified even by modern standards, including a large variety of amenities. He utilized principal-component analysis in the formation of many of his independent variables. The analysis was completed using data from a sample of U.S. urban areas. From his results, Porell concluded that both economic and QOL factors were important causes of migration. However, he found that migration was more responsive to changes in economic factors than changes in the QOL factors. Also, Porell was not supportive of the long-run equilibrium hypothesis of migration, and believed that migration occurs in disequilibrium, with both favourable job incentives and amenity bundles attracting residents to areas.

At this point, Roback (1982) published a paper that is now widely known for its theoretical contribution. The empirical portion of this paper computes implicit prices for amenities in metropolitan areas in the U.S., with the aim of calculating the values of specific amenities. The value of an amenity is represented by the wages an individual is willing to forego plus the additional housing price individuals are willing to pay.

Roback estimates two regressions, with wages being the dependent variable in her first model, and land prices the dependent variable in the second model. She utilized "amenity" variables such as population density, heating degree days, snowfall, crime rates, cloudy days, and pollution rates. The unemployment rate was used as an economic indicator. Many of the amenity variables in both regressions had the correct *a priori* signs, but finding significance on these variables was an issue. Roback ultimately

---

[9] Porell (1982) considered economic variables to be the total employment, percent of employment in agriculture, wage rate, unemployment rate, layoff rate, and a family allowance. QOL variables were formed through principal component analysis, and included indices for temperature, rainfall, parks and water, major sports events, crime, pollution, presence of physicians, and government expenditures on welfare payments.

concludes that her model demonstrates that the value of amenities is reflected in both wages and land prices (or rental rates). The results of this paper challenge the hypothesis of Graves (1983) that land rental rates alone can be used as a proxy for an array of amenity measures.

Greenwood and Hunt's (1989) paper is a response to Graves' earlier work. As with Porell, they attempt to assess the relative importance of economic opportunity variables versus amenity variables. The dependent variable in the study was net migration, while the independent variables included economic variables and the same set of climate variables utilized by Graves. Greenwood and Hunt utilize a different set of economic variables in their study, abandoning aggregate unemployment rates and income, while using employment growth, lagged average earnings, and lagged migration. Ultimately, the authors found that the presence of jobs and wages was significantly more important in affecting migration than the climate variables. They also agree with Porell (1982) that a disequilibrium model more appropriately describes the forces that determine metropolitan migration.

Herzog and Schlottmann (1993) examined the migration of white working males aged 19-55 in U.S. metropolitan areas to identify the economic and amenity factors which contributed to their decision to move between metropolitan areas. To be specific, the authors attempt to calculate a willingness-to-pay value for different amenity characteristics. The analysis was completed using a rich data set with measurements at the individual level, and personal characteristics data were matched up with metropolitan area characteristics to complete the data set.[10] Climate was not significant in their

[10] In Herzog and Schlottmann (1993), personal characteristics included education, work experience, whether the individual was married or had children, whether the individual had a disability, and their

regression, while population levels and population density were somewhat significant. The authors conclude that population levels can be considered a net amenity in cities with populations of less than 4.6 million people – a finding that certainly challenges the notion of agglomeration effects at the largest urban scale. Overall, this study included a fairly limited set of amenities, which may limit its usefulness.

Clark and Knapp (1996) attempted to examine elderly interstate migration using a variety of economic and amenity variables. The authors examined two different age cohorts: 55-64, and 75+. Their paper provides a major contribution to the literature because Clark and Knapp demonstrated through their regression analysis that the explanatory power of factors such as amenities or fiscal measures tends to decline with age. Overall, the authors found that local amenities actually had a relatively minor impact on migration when compared with fiscal factors – especially the tax burden.

In a study prepared in the United Kingdom, Wall (2000) used a panel data set with cross-migration rates as the dependent variable to estimate the standard of living in 10 different geographic regions. Wall suggests that the fixed effect estimates from his panel regression are estimates of the relative standard of living in different U.K. geographic regions, and by extension, these fixed effects essentially measure the net level of amenities present in the region. Wall controlled for local economic conditions such as wages (income per capita), job availability (job vacancy rate and unemployment), housing costs (mortgage down-payments), and the cost of moving (distance in the out-migration model).

---

occupation. The characteristics of the metropolitan area economic conditions included the unemployment rate, and the population. The amenities included were a climate index and population density. The authors also consider total population to be an amenity/disamenity.

Using the fixed effects coefficients, Wall ranked the different regions based upon the level of amenities in each region. Overall, the author did find significance on many of his variables, and a fairly high R-squared, leading him to believe his model did a fair job explaining the overall QOL in each region. He suggests that the logical course for further study would be to calculate the value of specific amenities rather than the value of the entire set of amenities to the population as was done in this paper. Wall's suggestion came almost 20 years too late, as Graves, Porell, Getz and Huang, Roback, and others pioneered measuring specific amenities in the early 1980's!

Knapp, White, and Clark (2001) examined U.S. intra-metropolitan and inter-metropolitan migration patterns at the household level. In this study, the authors found that amenities such as police spending (a proxy for community safety) and the amount of sunshine in the area had a large effect on influencing the location choices of households, while other amenities such as low crime rates, and temperature variability had a much smaller effect.

2.5.2 Modern – Regional Amenity Studies

There have been a number of recent empirical studies completed to examine the role of amenities in regional economic development. Dellar et al. (2001) constructed a structural model of regional economic growth, taking into account the effects of amenities and the willingness of residents to relocate to experience "location-specific" amenities to improve their quality of life. Using population growth, employment growth, and income growth as dependent variables, the authors utilize two-stage least squares procedure where the endogenous dependent variables enter as explanatory variables for the other dependent variables. Utilizing an econometric model with data from 2243 nonmetro United States counties, the authors employed a number of economic and amenity type variables in their analysis.[11]

The authors cite numerous sources suggesting that quality of life (via amenities) plays an increasingly important role in community economic growth. They find a significant and positive relationship between growth and many of the amenity variables, suggesting that amenity characteristics of communities can be effectively used in empirical models.

In a study looking at rural population growth, Huang, Orazem, and Wohlgemuth (2002) attempt to establish which factors caused rural communities to grow or decline between 1950 and 1990. The dependent variable in the study was the population growth rate, while independent variables included measures of income, human capital, local amenities, cost of living, government taxes and expenditures, commuting costs, and job search costs. From an amenities standpoint, the authors included local government expenditures on items such as highways, education, and public welfare. However, none of these variables were significant, but the authors were not concerned since their hypothesis was more concerned with the impacts of education and job market attributes on growth than with amenities.

Brewin (2004) examined in-migration and out-migration from rural counties in the U.S. The econometric model utilized in this study is very similar to that of Deller et al. (2001), where a two-stage least squares procedure utilized. Endogenous dependent variables enter as explanatory variables for other dependent variables, and principal-

[11] Deller et al (2001) utilizes indices developed through principal-component analysis where many specific variables are captured in one variable. Principal component analysis was used to form indices for climate, recreational infrastructure, land, water, and winter. Other amenity variables included the crime rate, number of physicians, property taxes, and government expenditures in the region.

component analysis was used in the formation of many of the independent variables.[12] An interesting departure in Brewin's analysis was that models for in-migration and out-migration were estimated separately instead of using one model with net migration as the dependent variable. This approach was taken because he argues that theory predicts that a net-migration model can cause important explanatory forces to cancel out.

Ultimately, the results of the empirical model lent support to the notion that in-migration and out-migration should be estimated separately rather than using a net migration variable. This conclusion was supported because the out-migration model was not significantly impacted by eight variables that significantly affected the in-migration model, suggesting that different factors drive in-migration and out-migration. The empirical model did lend some support to the notion that amenities affect migration. Crime and climate were both significant (at the 10 percent level) in the out-migration model, while climate, crime, urban adjacency, and mild winters were significant in influencing in-migration.

In a recent paper considering U.S. rural areas, Wojan and McGranahan (2004) test the hypothesis that regional economic development depends largely on the local concentration of workers specializing in fields related to creativity, knowledge, and ideas. To further their hypothesis, the authors contend that this "creative class" of workers tends to locate itself in metropolitan areas with high levels of local amenities. The authors therefore test two separate ideas: First, that natural amenities affect the net migration of

[12] Brewin (2004) utilized principal component indices for climate, crime, land, degree of rurality, water, and winter.

the creative class, and secondly that the presence of the creative class has an effect on rural economic development in the form of greater employment.[13]

Using a three stage least squares model to address issues of simultaneity, the authors model suggests that the creative class is drawn to high natural-amenity areas, and areas with larger population densities – but not areas adjacent to metropolitan areas. Also, the authors found that the presence of college graduates and an initially high population of creative class workers were highly significant in attracting new creative class workers. The authors' model also found that the initial levels and growth of the creative class were highly significant in increasing employment (their measurement of rural economic development). Overall, this paper provides an interesting hypothesis that is ultimately confirmed through an econometric model in perhaps the best analysis relating natural amenities to migration and economic development to date.

In another recent study examining migration and quality of life, Rappaport (2004a) examines the effects of weather, coastal proximity, and topography on the location choice of U.S. Residents. Rappaport's empirical analysis is at the U.S. County level, and he uses two dependent variables: the level of population density, and the growth rate of population density over time. He examines the total population, and elderly migration. The author ultimately concludes that local population growth in the U.S. is highly correlated with warmer winter weather, and cooler, less humid summer weather.

The development of air conditioning was not found to be the primary driver of migration towards nice weather, nor could all of the migration be attributed to the elderly.

[13] Wojan and McGranahan (2004) made the decision to only utilize natural amenities (population density, climate, landscape) and exclude other local amenities because they assume that rural areas are naturally devoid of man-made amenities.

Rappaport concludes that increases in per-capita income are most likely driving this trend. One limitation of Rappaport's model is that it does not include any measurement of the economic conditions within each county, or measurements of modern amenities. Lack of these variables is likely causing omitted variable bias in his models.

2.6 Explaining Future Growth – Initial Conditions

In examining persistent population flows in U.S. counties, Rappaport (2004b) found that areas experiencing changes in local characteristics such as productivity and quality of life experienced population flows that are proportional to such changes. Furthermore, changes in local characteristics will influence population flows over several decades. The implication is that regressions of local population growth on local characteristics can help identify the determinants of migration. Rappaport's paper supports the methodology of this thesis, which assumes that initial local conditions impact migration in the subsequent decade.

2.7 Social Capital in Economic Development

Glaeser, Laibson, and Sacerdote (2002) provide an excellent summary of the growing body of research into social capital, and how the literature has developed since its empirical inception in 1993. Early empirical research into social capital actually had little to do with economic development, as authors examined how social capital affected efficiency of governments and judicial systems. This paper also explains that a commonly accepted theoretical framework to explain social capital does not exist.

Goetz (1999) argues that social capital is a barrier to migration. He contends that people "invest" in social capital, and in fact will lose this investment if they decide to move. This is known as cumulative inertia in the migration literature. Durlauf (2002) provides a compelling argument that social capital should be utilized in empirical studies with extreme caution. His point is emphasized by the fact that the definition of social capital varies across studies, and includes a number of different ideas. Durlauf concludes that the concept itself is vague, and is not up to the high standards expected in the field of economics. Also, he argues that the observational measures currently used in econometric literature to identify forms of social capital are not properly measuring social capital.

There have been several empirical studies completed to determine the economic impact of social capital. Rupasingha, Goetz, and Freshwater (2000, 2002) developed an empirical model to assess differences in social capital on income growth at the U.S. county level. Ultimately, the authors found their aggregate measure of social capital was a meaningful factor in explaining differences in income growth. Higher levels of social capital had a positive effect on income growth. The theoretical framework presented in this paper suggests that increases in amenities and social capital should actually decrease the wage rate but in terms of increasing incomes, one possible explanation is that social capital helps increase firm productivity by reducing transaction costs, which could actually increase the wage rate.

Glaeser, Laibson, and Sacerdote (2002) have produced a simple economic model of social capital investment. Their study utilizes membership in different types of social organizations (churches, veterans groups, fraternities, farm organizations, unions, etc...) as their measurement of the stock of social capital. Data used in this analysis was at the individual level, and came from the U.S. General Social Survey in the U.S., which is a longitudinal annual cross-section of 1,200 to 2,500 respondents. The authors use econometric models to test the effects that age, mobility, investment, home ownership, and several other variables have on social capital investment.

Overall, the authors found that membership in social organizations appears to peak between people aged 35-55. Their model predicted a negative relationship between social capital and expected mobility. However, their measurement of mobility was not very good, as they used marital status and age to proxy for expected mobility. The authors also found that social capital was correlated strongly with home ownership. Overall, the contribution of this study is that it demonstrates different types of variables that can potentially proxy for social capital in other empirical models. In particular, home ownership and age were highly correlated with participation in volunteer organizations.

Other than these two variables, the results of the empirical model were not overly conclusive. A possible implication of this relationship between mobility and social capital investment is that the level of social capital may be high in declining areas because newcomers (migrants) typically would invest in lower levels of social capital that existing residents. Declining areas would have few new migrants, and many existing residents, so they could be expected to have higher levels of social capital.

The most recent paper examining social capital was written by Guiso, Sapienza, and Zingales (2004). The objective of their paper was to investigate the link between levels of social capital and financial development and economic prosperity. The authors develop their theoretical model by assuming that there is a probability that brokers/fund managers will flee with investor's money, and this probability is impacted by the extent of social networks in the area along with other factors like the quality of law enforcement and other individual factors.

For their empirical model, they utilized data from Italy as their sample, and utilize several financial indicators as the dependent variable.[14] They measured social capital using voter turnout numbers and participation in blood donation and included these as explanatory variables along with other household-level indicators. Overall, the authors found that the two measures of social capital did play an important role in financial development in different parts of Italy. Social capital seemed to matter most where education levels were low and law enforcement was weaker, which is the situation in many developing countries.

2.8 Chapter Summary

This chapter has examined several areas of economic literature relevant to this thesis. We first learned that Canadian migration literature is deficient in terms of utilizing amenities, and many of the studies are older. Utility theory, the theoretical basis for migration, and the role of amenities in migration were examined. The chapter continued by scrutinizing the growing body of empirical literature examining amenities, and social capital. Finally, a brief overview of spatial econometrics was provided. The next chapter provides the theoretical model of this thesis.

[14] In this study, the authors utilized the use of cheques, portfolio allocation, availability of loans, and reliance on friends and family for loans as their dependent variables.

Chapter 3: Theoretical Framework

3.0 Introduction

The purpose of this chapter is to examine relevant economic theory relating to migration, and adapt existing theoretical models to build the theoretical framework of this thesis. This framework is used in the remaining chapters to develop the methodology, and analyze the results of this thesis. The theoretical model presented here is particularly useful because it allows an accurate prediction of the expected relationship between migration levels and the explanatory variables in the empirical model. It provides particular insight into the interaction between migration, wages, and housing costs, and provides useful results regarding these variables.

A simple model of migration is initially provided, which supplies the basic framework to explain migration choices made by utility maximizing individuals. Next, a general equilibrium model of migration is presented, which adds more detail to the simple model, and explains the endogenous relationship between migration, amenities, wages, and housing cost. This expanded model is then separated to show the same results in disequilibrium. The final section of this chapter addresses the controversial issue of modeling the effects of social capital on migration.

3.1 Simple Model of Migration

Utility maximization is the central behavioural criterion of this model. Individuals are hypothesized to weigh both the pecuniary and non-pecuniary benefits of moving against the pecuniary and non-pecuniary costs of moving before making a migration decision. As suggested by Goetz (1999), the following assumptions are made:

1. Individuals maximize utility (U_i). Utility values are not assigned to individuals; they are defined for every location to which individuals can possibly migrate, such that i=1,2,3,…n (there are n potential locations).
2. Individuals are able to rank any two locations using the location's utility value.
3. Preferences over locations are transitive. In other words, if place P_1 is preferred over place P_2, and P_2 is preferred over P_3, then P_1 is also preferred over P_3.
4. Individuals derive utility from three things – wages (w), lower housing costs (r), and amenities (a). These factors are all location specific.

Potential migrants compare the expected utility of residing in different communities (U_1,U_2…U_n) with the utility at their current location (U_0) accounting for the cost of migrating m_i. The cost of migrating could include the transactions cost of purchasing a new home, and finding a new job, costs related to transporting belongings, and the social costs of moving and developing new relationships. This migration decision is shown in equation 1 as:

$$D = U_i(w_i, r_i, a_i, m_i) - U_1(w_1, r_1, a_1) \quad (1)$$

If D>0, utility maximizing individuals will migrate. If D<0, individuals will stay where they are. In making a migration decision and choosing their preferred location, individuals are essentially voting with their feet, and revealing their preference rankings for different locations. If an individual chooses location m over location 0, we can conclude that $U_m > U_{0,1,2\ldots i}$. Even though we cannot actually assign a numerical value to these utilities, revealed preference allows us to rank the different communities based upon their utility levels.

3.2 General Equilibrium Model of Migration

The simple model outlined above is insufficient in explaining one important detail. If all individuals maximize utility, and utility is based upon the characteristics of places (because utility is assumed to be homogeneous among all individuals living in a place), why don't all individuals end up moving to the same location – the location with the highest overall utility? The answer is that the model presented above does not account for adjustments in economic conditions and utility derived from places as people migrate over time. As people migrate into an area (or leave an area) they impact the labour and land markets, and subsequently affect the overall utility residents and other potential migrants can get by living in the community. This interaction between the influencers of migration (i.e. amenities, wages, and the cost of living) in the long run is what makes the following model desirable.

This framework is based on a general equilibrium model developed by Roback (1982). The point of departure for her model was to assume that utility maximization drives the location choices of workers, whereas profit maximization determines firm location. This model has been adapted and used in several subsequent papers. Bloomquist, Berger, and Hoehn (1988) used a similar model for a framework used in creating a quality of life index. Beeson and Eberts (1989) constructed a comparable model to examine the relative importance of amenity and productivity differences in explaining wage differentials across metropolitan areas.

In adapting this classic model, the assumptions outlined in section 3.1 still hold, and it will be assumed there are *n* communities across Canada, each endowed with a fixed set of location-specific amenities (denoted by the vector a_n). The residents of each community work in firms that produce a numeraire consumption good, *X*. It is assumed that residents can move freely between communities at no cost. Because our focus is on migration, I will assume, for simplicity, that workers cannot commute between communities for work or to consume the amenities of another community.

3.2.1 Residents

The residents of each community are assumed to have identical tastes and skills. Each resident supplies a single unit of labour, which is not affected by the wage rate. Residents attempt to maximize their well-being through their choice of location. They gain utility through consumption of good *X*, the purchase and use of local residential land, and the consumption of local amenities. The preferences of each individual are assumed to be homogeneous, meaning that every individual prefers the same amenities, and their preferences do not vary over their life cycle, or depend up their family situation.[15]

The only way for a resident to access local amenities is to rent land (or housing) L_n at the rental rate r_n. Residents in each community sell their labour to local firms at the wage rate w_n and use these wages to purchase residential land and the consumption good *X*. The supply of land in each community is assumed to be finite in the short-run, but can adjust in the long-run. Every individual living in a community attains an identical level

[15] Roback (1988) removes the assumption of homogeneous individual preferences by introducing household heterogeneity.

of utility due to the homogeneity assumption.[16] The level of Utility each individual attains through living in a community *n* is defined by the indirect utility function $V_n()$.:

$$V_n = V_n(w_n, r_n; a_n) \tag{2}$$

which is increasing in w_n and a_n, and decreasing in r_n. The objective function for residents is to maximize equation 2.

In the short run, individuals maximize equation 2 by relocating to locations that provide the highest possible level of utility. In the long run, there is an equilibrium among places, (which would ensure that no resident has an incentive to relocate to any other community). That is, there must be a level of utility (for example, V_0) that all residents have the ability to achieve regardless of their location in the long run. Wages and the land rental rate must adjust to ensure that the indirect utility function equals V_0 in every community based upon each location's site-specific attributes. This long-run relationship is shown in equation 3, and when this equality holds, it ensures that current residents do not have an incentive to leave the community, nor do outsiders have an incentive to relocate to the community.

$$U_n(w_n, r_n, a_n) = U_0 \tag{3}$$

3.2.2 Businesses

The second part of this model focuses on the location decisions of businesses that employ residents and produce the consumption good *X*. The model assumes that the cost of producing *X* is equal across all locations in the long run, and equal to the price of purchasing *X* (in other words, the production of *X* is perfectly competitive in all firms across the country). Since *X* is a numeraire good, its price is equal to 1, and in the long

[16] Implies that $\partial V_n/\partial w_n>0$, $\partial V_n/\partial a_n>0$, $\partial V_n/\partial r_n<0$. V is derived by maximizing utility subject to a budget constraint.

run, its cost will be equal to 1. In the short-run, costs are allowed to vary across locations. These assumptions yield the following unit production function and cost function C(). in equations 4 and 5:

$$X = f(L_n, P) \tag{4}$$

$$C(r_n, w_n) = 1 \tag{5}$$

where P is the population size of the entire community, and hence represents the work force of the community, while L_n represents available land in the community. In the short run, firms seek to minimize the left hand side of equation 5 by relocating to communities with the lowest costs. In the long run, wages and the land rental price adjust to ensure that the equality depicted by equation 5 holds. This ensures that the production costs of businesses regardless of their locations are identical, and no firm has an incentive to relocate in the ling run. C(.) is increasing in w_n and r_n, which are assumed to be the sole costs of production for the firms. For simplicity, it is assumed that firms sell the numeraire good in a frictionless environment and do not incur marketing, transportation, or transaction costs.

Roback allowed amenities to affect the costs of firms C(.), but unlike Roback, I assume that the level of amenities does not play a part in the production function or cost functions of firms. In this model, it is assumed that any amenities present in communities do not have any influence on production. While this is done for simplicity here, in practice, the empirical measures being considered as amenities and social capital in this study should have very little impact on firm productivity in general.[17]

[17] Although it is assumed here that amenities have no impact on production, Wojan and McGranahan (2004) found that high-amenity areas attract a "creative class" of workers, and an influx of this class of highly educated and creative labour could in fact potentially increase firm productivity.

3.2.3 Equilibrium

As residents and businesses relocate throughout the country to different communities, wages and rental prices constantly adjust. The final wage rate and land rental rate in each community is determined through the interaction of the supply of labour (residents) and demand for labour (businesses) in the labour market, and the demand for land. Equilibrium is attained when all firms in the country have equal production costs, and all households have an equal level of utility.

Figure 3.1 can be used to help understand the effects of different quantities of amenities on wages and rents within the community in wage-rent space. The equilibrium values are depicted by the intersection of the upward sloping isoutility curves with the downward sloping isocost curve. At every point on the isoutility curve, individuals obtain the same level of utility. The curve is upward sloping because if land prices increase, wages must also increase in order for individuals to continue to receive the same level of utility. At every point on the isocost curve, businesses incur identical costs. This curve is downward sloping because if the price of land increases, wages must decrease in order for businesses to have the same level of costs.

Figure 3.1 shows the final equilibrium when there is a community with a high level of amenities (a_2) versus a community with a low level of amenities (a_1). It is assumed that an increased level of amenities has a positive effect on the utility of residents and the diagram captures the previously specified behavioural assumptions.

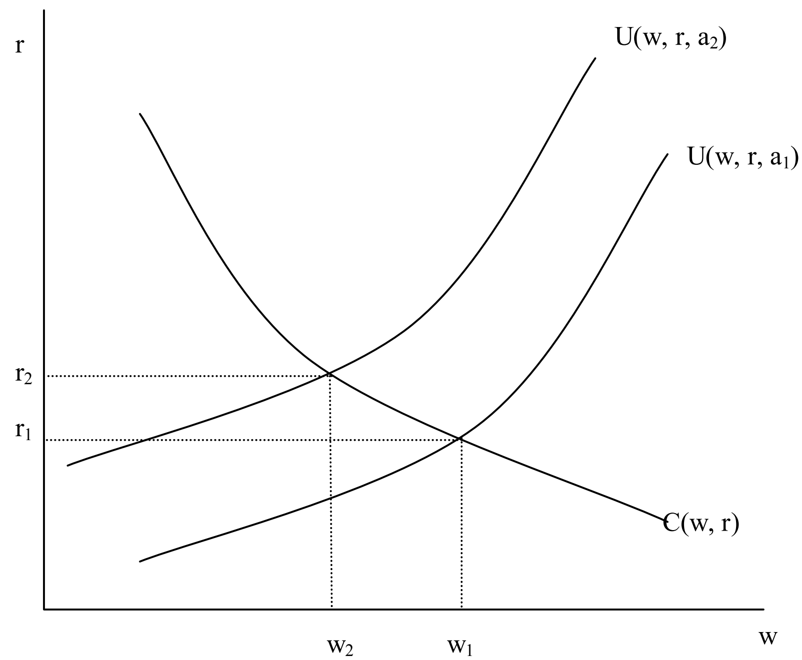

Figure 3.1: General Equilibrium for Wages and Rental Prices When the Level of Amenities is varied

In equilibrium, the wages and rents in the community with a higher level of amenities (a_2) are w_2 and r_2. The wages and rents in the community with a lower level of amenities (a_1) are w_1 and r_1. If more amenities are offered within a community, economic theory predicts that lower wage rates and a higher land rental rates within that community will typically exist.

3.3 Disequilibrium Model of Migration

The model of migration discussed in section 3.2 assumed that individuals react to initial disequilibria in factors such as wages, housing prices and amenities. In the process of migrating, according to this model, they restore an equilibrium across all communities as the levels of wages, and housing costs increase. In the long run, no individual has any incentive to relocate. Another way to demonstrate the same principals is in a disequilibrium setting using several different markets. What follows is a demonstration of how an increase in amenities affects labour markets and housing markets.

The purpose of this particular model is to adapt Roback's theoretical model, and show the results in wage-employment space. Following Roback, this model shows how an increase in amenities within a community will affect wages and cause in-migration for the community. The relationship between amenities, wages, and the population of communities can be outlined using a regional labour market, with the price of labour (w) on the vertical axis and the quantity of labour (P) on the horizontal axis in figures 6 and 7. The downward sloping long-run labour demand curve is shown by D_0^{LR}, while S_0^{LR} and S_0^{SR} show the long-run and short-run labour supply curves respectively.[18]

Both the long-run and short-run labour supply curves are upward sloping. The long-run supply curve is upward sloping to reflect congestion and other disamenities associated with urban scale. In other words, congestion, crime, and other inconveniences associated with living in a densely populated area compel residents to seek higher wages as the region becomes more crowded. The long-run labour supply curve is more elastic than the short-run curve to reflect the delayed response of potential migrants to changes in wage rates in the community. Figure 3.2 depicts the labour market in a long-run equilibrium before amenity levels are increased. The labour market is in a long-run equilibrium at the intersection of S_0^{LR}, S_0^{SR}, and D_0^{LR} (Point A), with the wage rate at w_0 and the quantity of labour at P_0

[18] Assume the short-run and long-run demand curves are equivalent for simplicity.

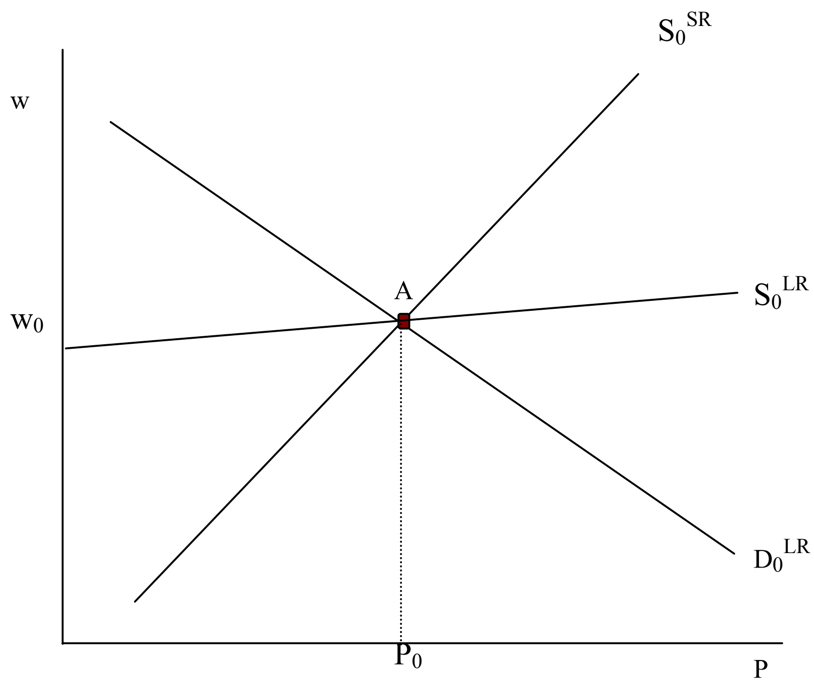

Figure 3.2: Labour Demand and Supply Framework

Now, suppose the community depicted in figure 3.2 initiates some type of strategy to provide more amenities to local residents. We know that residents gravitate towards high-utility communities, and the level of amenities within the community has a positive relationship with regional utility levels. Economic theory indicates that an increase in population size raises labour supply, and results in a rightward shift in the labour supply curve.

An increase in amenities would bring about a population influx for the community and shift the long-run supply curve out to a new level (S_1^{LR}) in figure 3.3. The lower wage w_1 reflects the willingness of the population to trade lower wages for higher amenities. In the long-run, the equilibrium shifts to point B, so an adjustment process must ensue where the short-run labour supply curve shifts out to bring the labour market back to a long-run equilibrium at point B. Note that wage rate decreases from w_0 to w_1, and the quantity of labour increases from P_0 to P_1

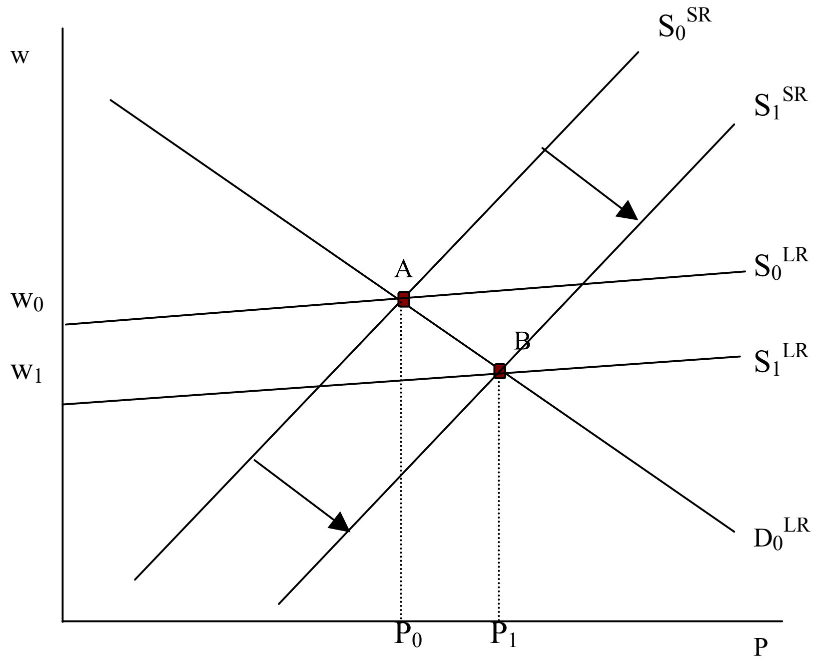

Figure 3.3: Labour Demand and Supply Framework –Labour Supply Adjustment Resulting From Increased Amenities

An increase in demand for the community's products or an increase in the productivity of the region's labour force will similarly shift out D_0^{LR} (not shown). This would instead yield increases in long-run wages and employment through in-migration. This model indicates that an increase in community amenities will unambiguously decrease wage levels, and increase the population size of the given community through in-migration. These results are consistent with the results of Roback's amenity model, but present the results in wage-employment space.

The general equilibrium model presented in section 3.2 also addresses land prices, or the land rental rate. Figure 3.4 shows these results in terms of the local market for land in a community that experiences an increase in amenities that causes population to grow. The price of land (r) is shown on the vertical axis, while the supply of land (L) is shown on the horizontal axis. The short run supply of land S_0^{SR} is perfectly inelastic to reflect delays in zoning, and construction time, while the long run supply of land S_0^{LR} is upward sloping because the quantity of land is assumed to adjust in the long run in response to demand.

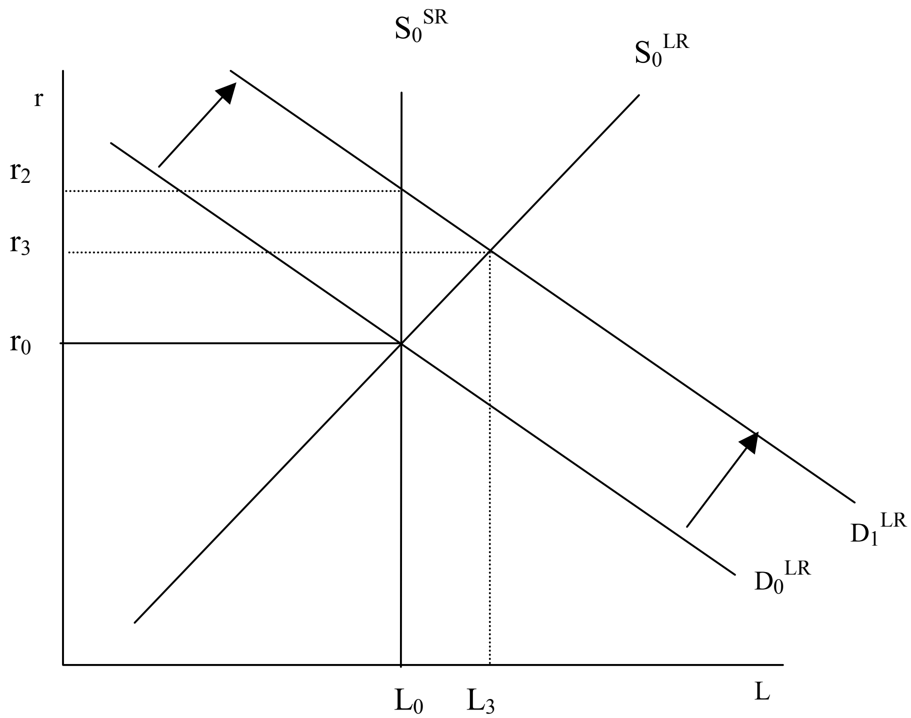

Figure 3.4: Land Demand and Supply Framework – Demand Shift Resulting From an Increase in Amenities

An increase in amenities causes an outward shift in the demand for housing from D_0^{LR} to D_1^{LR} as residents migrate into the area and require shelter. In the short run, the quantity of land does not increase, but the price increases to r_2. In the long run, the quantity of land can adjust, and it increases to L_3, while the price of land decreases from its short run level from r_2 to r_3. This model indicates that an increase in amenities in a community will increase both the price of land and the quantity of land in the community in the long run.

3.4 Social Capital

As explained in chapter 2 of this thesis, there is no commonly accepted theoretical model to apply to concept of social capital. This thesis assumes that accumulated levels of social capital in communities act as non-rival amenities. In other words, social capital is a pleasurable thing that everyone can enjoy without cost, and is specific to individual communities. Existing residents are attracted to this “social glue” because of the time they have invested in developing social capital, and because they simply enjoy it. Potential migrants are attracted by the opportunity of being able to reap the benefits of residing in an area rich with social capital.

3.5 Chapter Summary

This chapter provided a theoretical framework in which to analyze the effects of amenities on migration. Overall, we can form several judgements regarding the effects of amenities on migration from this framework. Individuals maximize utility, and the level of amenities in communities is one component of utility, along with wages, and the cost of living (i.e. housing/land costs). Economic theory predicts that high levels of amenities within communities will cause in-migration of residents. The theory presented in this chapter also indicated that the level of wages and the price of housing in communities compensate for different levels of amenities in different communities. Thus, we can expect that increases in population will be positively correlated amenities, but the relationship among wages and housing costs, and migration is not clear, as wages and land prices represent monetary differentials that compensate for regional differences in amenities to a certain extent.

Information contained in this chapter is used in forming expectations for the relationship between different variables in the empirical modeling process, and identifying potential issues of endogeneity. The next chapter describes the methodology of the empirical analysis.

Chapter 4: Methodology

4.0 Introduction

This chapter outlines the details of the econometric model that forms the basis for testing the hypotheses of this study. First, a brief outline of the model is presented to give readers a general overview of the econometric model, and the general types of variables that will be employed. Next, the various data sources for the dependent and independent variables are outlined, along with several key issues relevant to the model. A discussion of the specific variables used to test the hypotheses follows, which details exactly how the variables are derived, and what each variable is attempting to measure. The foundation for using these particular variables was presented in chapters 2 and 3. Chapter 4 concludes with a section outlining the different equations that were estimated, the rationale behind these estimations, and the expected results.

4.1 The Basic Model

The empirical portion of this study will provide an in-depth examination of the influencers of population migration and demographic change throughout Canada. Using a variety of financial, economic, and amenity indicators, the model will be used to model population change in urban and rural Canada. The model will examine the period between 1991 and 2001 using community-level data. The unit of observation chosen for this thesis is the census consolidated subdivision (CCS). From a theoretical perspective, household-level data would provide the best unit for measuring migration; however, this data is not readily available. Instead, it is assumed that the average CCS variable levels constitute a representative household.

The dependent variable in this model is the net population change for each CCS over the 10-year period. Population change for several different population cohorts, as well as the total population change will be utilized in the model. Equations will be re-estimated for the various dependent variables to examine the effects of the explanatory variables on different segments of the population. These dependent variable choices will be regressed on initial levels of the explanatory variables to the extent that data are available for the initial period levels. Otherwise, current levels of the explanatory variables will be utilized. The basic model will be set up similar to equation 6:

$$\text{MIGRATION} = \alpha + \beta_1 (\text{ECON}) + \beta_2 (\text{AMENITIES}) + \beta_3 (\text{SC}) + e \qquad (6)$$

where MIGRATION represents the net population change for each CCS, while ECON, AMENITIES, SC, and *e* represent vectors of economic indicators, amenities, social capital variables, and a stochastic error term respectively.

The purpose of this model is to examine the effects of local community characteristics on the location choices of residents. If easy access to location specific amenities, jobs, and wages is important to residents, then areas with poor access to these attributes should have larger population losses over time as residents relocate to communities that provide them with higher utility bundles. Conversely, communities with access to a superior utility bundle will experience stronger population growth, ceteris paribus. Of course, different population segments may respond favourably to different bundles of community attributes. Since it takes time for individuals to adjust to changing local factors, I follow Rappaport (2004b) in examining the “change” in population over time rather than simply population “levels.”

In creating a model to explain migration, it is necessary to control for all factors affecting migration including economic conditions, differences in population, and differences in amenities and social capital between CCSs. By separately considering different age cohorts, the model should explain most factors underlying the out-migration of young people from communities, as well as answer important policy questions such as whether increased distance to hospitals causes seniors to abandon their respective communities, and what factors attract middle-aged individuals to communities. Population retention and growth is an important factor in future community success., particularly for rural communities. Once population falls below threshold levels necessary to provide enough clients for any given enterprise, these businesses may fail (Stabler and Olfert, 2002). The loss of services provided by the businesses contributes to further out-migration and could render the community unsustainable, leading to even more future business closures and population declines.

4.2 Avoiding Endogeneity – Initial Conditions

As noted by Rappaport (2004a), any attempt to compute population change in specific regions in response to local stimuli requires a relatively long time-span. This study measures population change over a ten-year period (1991-2001). These specific years are utilized because they are census-years, and because it is felt that ten years is a long enough period to allow for residents to respond to local factors.

The decision to analyze migration over a ten-year period creates a definite issue that must be addressed through the modeling process. As described by Partridge and Rickman (2003a), there exists a long-standing question in regional economics as to whether people move to access new employment opportunities, or whether jobs actually move to where established workforces already exist. By extension, this "chicken-egg" conundrum may in fact exist in the relationship between modern amenities and people. Do people follow modern amenities into communities, or do modern amenities soon follow newly established immigrants?

To deal with possible endogeneity, the model will utilize the 1991 level of explanatory variables to predict post-1991 migration. In other words, the level of all explanatory variables will be predetermined in the model. It is necessary to utilize the bundle of local factors each community possessed in the past (i.e. 1991) because it is hypothesized that it is the characteristics of each community in the past that drives subsequent population change into the future – not the current levels (or not "expected" current levels). The implicit assumption being that individuals examined the 1991 level of local factors for their community and all other Canadian communities, and based their post-1991 location on the initial 1991 levels. If individuals decided to move away from their original CCS, this is attributed to them being able to attain a higher level of utility through their bundle of amenities and economic factors in their new CCS.

As shown by Rappaport (2004b), in empirical analysis, it is appropriate to utilize initial conditions to predict subsequent migration, because initial population distribution is far from its long-run equilibrium. Rappaport demonstrated the dramatic effects that the development of air conditioning had on U.S. population change over time, and effectively demonstrated that technological developments, as the prices of new technology falls over time, can affect population change for decades following their implementation. From a Saskatchewan perspective, the farm economy is continuing to react to on-going price trends due to on-going technical advances since the time of settlement. The predetermined nature of the specification will mitigate or eliminate any endogeneity. Initial conditions should not be endogenous with future growth, which eliminates the need for simultaneous equation models.[19]

By definition, endogeneity exists in models where the stochastic disturbance term of the equation is correlated with one or more independent variables. If an independent variable is correlated with the error term, it is said to be endogenous. If it is not correlated with the error term, it is said to be exogenous. Because predetermined explanatory variables are utilized, there is little chance that the explanatory variables are correlated with the error term.

There are two caveats to the suggestion that predetermined variables cannot possibly be correlated with the error term. First, endogeneity can arise if the disturbance term consists of omitted variables, and these omitted variables are correlated with one or more of the independent variables. For example, if an observable determinant of population change was omitted from the model, and the omitted variable is correlated with one or more explanatory variables, these explanatory variables could be endogenous. Basically, if the shocks generated by the omitted variable affected the levels of both the dependent and one or more independent variables, the independent variables may be endogenous. However, the empirical model of this thesis contains many different control variables, and it is not very likely that any such variable has been omitted from the model.

Another way endogeneity could be present in the model is if individuals form long-term expectations regarding expected future levels of local factors. For example, if

[19] One potential problem is that economic agents have been forward looking in the past such that "initial" conditions have been affected (i.e. they may be discounted values of the "expected" long-run equilibrium values). That is one reason why a long 10-year window has been selected to examine this issue.

they expect local services to be discontinued or jobs to be terminated over the 10-year period, and form their location decisions based upon these “long-term expectations” rather than the “actual” levels of local factors in the initial period, endogeneity bias could exist if this somehow enters the error term. As an example of something that is not necessarily endogenous, suppose common (but inestimable) knowledge exists (and existed prior to the initial period) that the agricultural sector will decline in the future, and it causes people to exit the community. This immeasurable future would be correlated with the prominence of the agricultural sector in each community. However, the agriculture share regression coefficient would capture this effect, and it is not clear it would be reflected in the residual. Long-term expectations are really an extension of the omitted variable problem discussed above, and the econometric model includes a large number of control variables to eliminate this problem, and the author does not consider endogeneity to be a significant problem due to the steps taken to mitigate its influence.

Obtaining 1991 data for many variables was fairly simple, as the census of population provides socio-economic data for each CCS. Natural amenities can be measured at any point in time, since they are comprised of time-invariant natural geographic and environmental features, so these amenities did not pose a problem. It ultimately proved to be impossible to locate 1991-level data for several modern amenity variables, as a database comprised of the geographic position of interesting institutions and businesses like hospitals, schools, retirement homes, universities, and golf courses was largely absent in 1991. In one case, a government agency was reluctant to disclose the past location of hospitals and long-term care facilities for privacy reasons, even though they possess a database containing the information.

In order to address this dilemma, there were two options. First, the modern amenity variables could be removed from all estimations. However, this action would invoke excluded variables bias, which is obviously an undesirable consequence. The second possible solution was to include all of the modern amenity measures (mostly for 2001), ignoring any possible endogeneity. The adopted solution was ultimately a compromise; the 2001 levels of selected variables believed not to be significantly different from their 1991 levels were utilized.

For example, the location of highways between 1991 and 2001 would not have changed, nor would the location of universities across Canada, so these variables are used. The presence of schools, churches, police stations, and hospitals may have changed to some extent between 1991 and 2001, but across Canada the difference is likely very small. For these variables, it is assumed that the 2001 level is equivalent to the 1991 level, and thus the variable is assumed to remain predetermined. Many variables were excluded from the model because it would have been unreasonable to assume that their 1991 levels equalled their 2001 levels. For example, all measurements of specific businesses (restaurants, stores, gas stations, bars, gyms, professional sport franchises, etc…) were excluded in the econometric model for this reason. It is felt that this compromise will not be detrimental to the methodology of the study, and provides enough benefits to make the risk worthwhile.

4.3 How to Proxy for Access? - Several Approaches

One of the main issues that had to be tackled in developing a framework to assess the determinants of migration was selecting a metric with which to measure the relationship between each community, and the various amenities the community had access to in 1991. Conducting surveys to gauge access to amenities was neither feasible nor possible given the national scope of the thesis. Even at the provincial level, it would have proven impractical to assemble a dataset through surveys. Therefore, secondary or existing data sources at the community level proved to be the only viable data sources for this thesis.

Following an exhaustive search of data sources, it was determined that the only way to determine where amenities were located in 1991 was to find data revealing their location coordinates and utilize this data in conjunction with GIS software to link location-specific amenities with different CCSs. Consider the example of an acute-care hospital. The only data available denoting the presence of hospitals is the address of the hospital, which can ultimately be converted to point coordinate locations (latitude/longitude). But how does one relate this series of point locations to each of the CCSs being considered in the dataset? Basically, there are four ways to relate point data to the CCS to measure access: 1) measure the distance from the centroid of the CCS to the nearest point of interest, 2) count the number of points of interest lying within the boundaries of the CCS, 3) compute dummy variables that indicate whether or not the CCS possesses a particular amenity, or 4) the per-capita number of points of interest in the CCS.[20] The advantages and disadvantages of each of these measures will now be considered.

[20] Alternatively, one could calculate the number of amenities within the boundaries of the CCS plus the number of amenities within adjacent CCSs. Another alternative is to calculate the number of amenities within a radius (i.e. 100 or 200 kilometres) of the CCS centroid. A centroid is an area's geographic centre defined as the halfway point on its east-west and north-south boundaries.

4.3.1 Distance to the Nearest Amenity

The first way to measure access to different amenities is to calculate the distance between the CCS centroid and the closest point of interest. This "as the crow flies" distance is intended to provide a measure of travel time to the nearest amenity. In theory, the closer the CCS is to the nearest amenity, the less time it will take for residents in the CCS to access the amenity. In urban centres, the distance to amenities will be close to zero, while in remote northern areas, the distance could be hundreds of kilometres. Variables utilizing a distance calculation will be prefaced by **DIST_**. Generally, it is felt that a distance measurement provides the best measure of access, though the measure does have several shortcomings.

The first deficiency of the distance measurement is that the CCS geographic centroid may not accurately represent the location where the majority of the population of the CCS resides. For example, a town located within a rural CCS may be positioned along the border of the CCS (not at the geographic centroid). The majority of the amenities in the CCS will be located at the coordinates of the actual town site. This problem leads to distance calculations that may not represent the distance the majority of the residents in the CCS have to travel to access the amenities. Such measurement error will likely bias the regressions coefficients to zero.

The second issue is that an "as the crow flies" distance measure may not accurately gauge travel time to amenities. For example, travelling ten kilometres in a rural area will take significantly less time than traveling the same ten kilometres in downtown Vancouver or Toronto due to traffic congestion. Also, the calculated linear distances do not account for geographic barriers to travel and the presence of roads. For example, a CCS may be physically close to an amenity "as the crow flies," yet a river may impede direct travel to the amenity in question, and residents must navigate around this obstacle, increasing the actual travel time substantially. Furthermore, roads may not be present that take a direct route towards the points of interest, so in most cases the calculated linear distances provide a downward biased measure of the actual distance. These issues can be accounted for in the modelling process as measurement error in the independent variable, which produces a downward bias in the regression slope coefficients. This type of measurement error can be mitigated through using a weighted least squares estimation procedure if it can be reasonably assumed that the degree of error is related to population. This possibility will be explored later in the chapter.

4.3.2 The Number of Amenities

Using the resources of the C-RERL research lab, it was possible to calculate the number of different individual amenities located within geographic areas related to each specific CCS across Canada. The three geographic areas for which numbers of specific amenities were calculated were: 1) within the boundaries of the specific CCS, 2) within the boundaries of the CCS plus within the boundaries of all adjacent CCSs, and 3) within a specific radius (100 kilometres and 200 kilometres) of the CCS centroid. Counts were generated for each of these geographic areas for every CCS across Canada. As with the distance measurement, these tabulations also have shortcomings that will now be addressed.

Tallies of the number of amenities within the specific CCS and within surrounding CCSs have one main limitation: CCSs across Canada are comprised of various sizes. Generally, CCS sizes are much larger in Alberta and northern regions, while Southern Ontario and Quebec in particular have tiny CCSs. These differences mean that obviously there is more likelihood of a larger CCS containing specific amenities than a smaller CCS. For example, an urban suburb that exists as a stand-alone CCS has a lower chance of being associated with a hospital than if the same urban suburb was combined with the downtown core of its neighbouring city. For this reason, counts utilizing the CCS alone may not provide an accurate picture of access. Utilizing a count of the hospitals located within the CCS plus surrounding CCSs, or a count of the hospitals within 100 kilometres of the CCS centroid would likely provide a more realistic picture of the amenities available to that CCS.

The final issue with utilizing counts in any form is that they are highly correlated with the population level of the given CCS. For example, an urban centre may have hundreds of doctors when compared with a neighbouring rural area, but the total number of doctors is not indicative of the access individuals have to doctors. The more people living in a specific area, the higher the sheer number of amenities will be in that area, and the total number of specific amenities does not really provide a good measurement of an individual's level of access to amenities. From an econometric standpoint, the total number of specific amenities is highly correlated with the total population of each CCS. For these reasons it is not appropriate to utilize total counts in the empirical model. Rather, this thesis proposes using dummy variables and the per-capita number of amenities to measure access, as opposed to total counts.

4.3.3 Dummy Variables as Amenity Indicators

Under this method, CCSs (or related geographic regions described earlier) containing at least one specific amenity (i.e. a hospital) would be assigned a 1, while

CCSs that do not have a specific amenity would be assigned a 0. In essence, this method indicates whether a CCS has access to at least one point of interest, and does not attempt to measure the "level" of access. In many ways, this measurement makes sense because individuals probably do not consider having access to 5 different hospitals or 900 doctors to be an asset. Having access to "at least one hospital" or "at least one doctor" may be a more realistic measure of access from a "rural" concept of access. When dummy variables are used for amenity measures in this thesis, the variable name is prefaced by **D_**.

4.3.4 Per-Capita Counts as Amenity Indicators

The final method of measuring access is to divide the number of location specific amenities in a CCS (or associated geographic region) by the population of the CCS. This per capita measurement specifies the number of amenities per person in the CCS. However, does a higher per capita measurement actually mean that an individual has a greater level of access? For example, if there is one hospital per 100,000 people in one CCS compared with 2 hospitals per 100,000 in another CCS, it doesn't necessarily mean that access to hospitals in the second CCS is twice as good. The first CCS may have twice as many doctors and beds, and better equipment in its hospital than the first CCS, and provide an identical level of care and access as the second facility. For this reason, per-capita counts on buildings may not provide an accurate measure of access. When per-capita counts are utilized for variables in this study, the variable name is prefaced by **PERCAP_**.

4.4 Data Sources

A number of different data sources were utilized in the formation of the variables used in this thesis. The purpose of this section is to detail these sources.

4.4.1 Census of Population (CoP)

A number of variables were drawn from the Census of Population (CoP) dataset, which was acquired directly from Statistics Canada. The census is a national survey that currently covers nearly 12 million households, and is completed every five years. This thesis draws upon data from the 1991 and 2001 "major censuses." The purpose of the census is to gather information on demographic, social, and economic conditions across Canada, and an attempt is made to obtain data for every household, and by extension every person in Canada. A 20 percent sample of households are selected to complete a detailed questionnaire, while 80 percent of households are sent a shorter version of the questionnaire (Statistics Canada, 2005).

Census data is not released at the household level, and is therefore tabulated at various geographic and socio-economic levels. The particular data used in this thesis is tabulated at the CCS level, using 1996 CCS boundaries. Due to the 20 percent sample of households on the long census questionnaire, CCSs with a population of less than 250 people are suppressed by Statistics Canada to address data accuracy issues and/or privacy concerns. Since important variables were not available for these communities, the decision was made to delete CCSs with missing data from the sample. There was also a problem with the 2001 tabulation because several CCSs were missing, and were deleted from the dataset. Finally, the six CCSs located in the Northwest Territories, Yukon, and Nunavut were deleted due to data accuracy issues.

In all, 205 CCSs were removed from the dataset (which ultimately meant that the data contained 205 fewer units of observation). Although this is an unfortunate outcome, these 205 observations only accounted for 119,786 people in 1991, and even fewer in 2001. In all, the deleted observations account for less than one half of a percent of Canada's entire population. Even after the removal of these CCSs, the dataset had 2402 observations. Due to the small population and large geographic area of the deleted the CCSs, they would likely have contributed a degree of measurement error to the model, which is an undesirable consequence.

CoP population counts are utilized in the formation of the dependent variable of this thesis. Measurement error in the dependent variable is an area of concern in this model, though it is less serious than measurement error in the independent variables. Statistics Canada randomly rounds off population counts in each CCS to five persons, and this rounding will ultimately impact the dependent variable of this thesis. Random measurement error in the dependent variable does not bias the slope coefficient, but it does lead to larger standard errors, and by extension, lower t-statistics.

4.4.2 Canada Rural Economy Research Lab (C-RERL)

One of the main issues with compiling data for a study of this magnitude is that it is necessary to bring together data from various diverse sources, and make all of this data compatible by modifying it in many cases. In this thesis, an early decision was made to adopt the CCS as the unit of observation. CCSs are geographic areas with boundaries determined by Statistics Canada. However, many of the variables ultimately employed in this study were not tabulated by Statistics Canada, and therefore did not assimilate easily with the CCS geography. The Canada Rural Economy Research Lab (C-RERL) and its geographic information system (GIS) facilities and expertise made it possible to generate variables from previously unusable data sources.

Through C-RERL, approximately half of the variables utilized in this study were formed. Basically, any data obtained that corresponded with a particular geographic location could be modified at C-RERL, and converted into CCS units. For example, climate data in Canada comes from Environment Canada, and is available for each of the weather stations in Canada. Obviously, weather station data is not helpful for this study because the unit of observation of this thesis is the CCS, not weather stations. At C-RERL, GIS software was used to calculate which particular weather station was closest to each CCS, and assign the nearest weather station's data to the CCS. There are numerous examples of C-RERL expertise and software being used to modify diverse data into usable units provided in section 4.5.

4.4.3 DMTI Spatial Inc.

GIS methods and data are becoming more accepted and utilized in the construction of variables for economic and regional development studies. In the search to identify location-specific amenities for use in this thesis, a GIS company called DMTI Spatial Inc. was discovered. DMTI is a software and data company specializing in address and geospatial related applications and data. DMTI has created a national Enhanced Points of Interest (EPOI) database that contains over one million Canadian businesses, recreational, and other points of interest identifiable by coordinate location, SIC code, name, address, and other interesting attributes. The EPOI database is compiled through formal agreements with business partners such as telecommunications companies that send DMTI verified records from their client lists or directory listings across Canada.

DMTI utilizes other sources to verify the addresses and locations of points of interest. The EPOI database was fortunately available for academic use through the Data Liberation Initiative (DLI). DMTI's EPOI database provided the foundation for most of the location-specific modern amenity variables utilized in this study.

4.4.4 Other Miscellaneous Sources

In addition to Census and DMTI data, this thesis has also drawn upon a number of other sources that were ultimately adapted through C-RERL and incorporated into the dataset. Voter Turnout data at the federal electoral district level was obtained through Elections Canada. Digital elevation data containing point elevations, and Land cover data was acquired through Natural Resource Canada. Physician location data was obtained through the Canadian Institute for Health Information (CIHI). Weather station data was acquired through Environment Canada, while crime data was obtained from the Canadian Centre for Justice Statistics and Statistics Canada. This short list provides a general concept of the considerable number of sources from which the dataset used in the empirical model was assembled.

4.5 Data – Specific Variables

This section details the specific variables utilized in the econometric model. Specifically, the source of each variable, what each variable is hypothesized to measure, and any modifications carried out by the author or C-RERL to form the variables are detailed here. The discussion starts by defining the dependent variables, and moves on to the independent or explanatory variables. The dependent variable utilized in the study is net population change between 1991 and 2001 for several age cohorts, and net population change for all age groups.

The explanatory variables utilized in this study will be comprised of economic indicators, variables indicating the presence of human capital, demographic indicators, regional dummy variables, social capital indicators, variables that measure proximity to urban centres, natural amenity variables, and modern amenity variables each CCS. A summary of all of the variables and their sources is provided in appendix A, while descriptive statistics for this data are presented in Appendix B.

4.5.1 Population Change

The dependent variable used in the econometric model is net "population change." An actual migration variable would have been preferable, but net migration cannot be calculated because data for out-migration does not exist. In-migration can be determined, but it is impossible to track the individuals departing each CCS. Therefore, population change (POPCHG) is used out of necessity, and is calculated according to the basic formulas outlined in equations 7 and 8

$$POPCHG = POP_{2001} - POP_{1991} \tag{7}$$

$$POPCHG = POP_{2001} - POP_{1991} + D \tag{8}$$

where POP_{2001} is the population of the CCS in 2001, POP_{1991} is the total population of the CCS in 1991, and D is the estimated number of deaths that occurred in the CCS between 1991 and 2001. The 1991 and 2001 population numbers originated from the CoP data, while the expected number of deaths originated from two different sources. An average 1994-1996 measurement of age standardized mortality rates (ASMRs) for each census division were taken from a Statistics Canada report entitled “Vital Statistics Compendium, 1996” (Duchesne et al., 1996).[21]

ASMRs were used to calculate D to adjust the total population for mortality, but different numbers had to be used to adjust specific cohorts. National age-specific mortality rates (AMRs) from 1994-1995were used to calculate D for the cohort analysis, and these numbers were taken from “Births and Deaths” (Statistics Canada, 1997). These mortality rate measures used to estimate deaths in each CCS are not the most ideal measurements, but it was not possible to obtain better mortality rates, or the actual number of deaths occurring in each CCS between 1991 and 2001.

Note that two equations are given to calculate population change; equation 6 accounts for deaths, while equation 7 does not. Equation 6 implies that all individuals that died between 1991 and 2001 are migrants, while equation 7 implies that no persons who died between 1991 and 2001 are migrants. Thus, equation 6 provides an upward biased estimate of POPCHG (because deaths are considered to be migrants), and equation 7 provides a downward-biased estimate of POPCHG (because people may have migrated, and then died in another CCS). These estimates likely bound what happened in reality. The difference between these two variables is negligible for younger population cohorts, but significant for older cohorts (age 50+). This thesis takes the approach of accounting for mortality.

In calculating total population change for CCSs, natural increases in population are not accounted for in the analysis. For any CCS, a proportion population change will

---

[21] An age- standardized mortality rate (ASMR) is a weighted average of age-specific mortality rates calculated for a particular time frame and geographic location. ASMRs weight mortality rates by a standardized population. ASMRs enable comparisons to be made between time periods and/ or across regions.

be due to births occurring in the CCS. In other words, births occurring in CCSs between 1991 and 2001 are actually accounted for in the study as in-migration. The assumption is made that every CCS has an identical birth rate. As with deaths, a scalar could be utilized to subtract a national birth rate from the population growth rate in each CCS, but it was not done in this case. This deficiency does not affect the cohort analysis portion of this study.

The composition of the total population change variable **(POPCHG_TOT)** will now be addressed. This variable was calculated according to equation 9:

$$POPCHG = \log(POP_{2001} + D) - \log(POP_{1991}) \quad (9)$$

where $ASMR*POP_{1991}*10$ equals D. The mortality rate is multiplied by the 1991 population of the CCS to calculate the expected number of annual deaths. This number is subsequently multiplied by 10 because the time period being examined is a ten-year period. Equation 9 uses a logarithmic difference to calculate the percentage change in population between the two time periods using the geometric mean as the mid-point. Percentage change in population is used rather than total population change to mitigate differences in the scale of CCSs.

In addition to total population change, this study utilizes a selection of other variables to assess how the explanatory variables affect different population cohorts. This thesis uses a method known as “artificial cohort analysis” to track cohorts over the ten year period of this study. For example, individuals born between 1976 and 1986 will be aged 5-20 years in 1991, but by 2001, these exact same individuals will be 15-30 years old. One can reasonably assume that the vast majority of the individuals born between 1976 and 1986 are accounted for in 1991 as 5-20 year old individuals, and then again in

2001 as 15-30 year old individuals. Since the objective of this thesis is to account for the behavioural decision of different cohorts, it makes sense to track the cohort by identifying them at different life-stages in 1991 and 2001.

In all, this thesis examines five different population cohorts, each at different phases of their lives. The first segment examined is individuals born between 1971 and 1986. These individuals were aged 5-20 in 1991, and are intended to represent young people. The second cohort examined is that of individuals born between 1956 and 1971 who were 20-35 years old in 1991. These individuals represent young adults, who are presumably entering a stage of their life where they will be starting to have families. The third cohort examined is that of middle aged adults who were born between 1941 and 1956. These individuals were aged 35-50 in 1991.

Early retirement aged individuals make up the fourth demographic group. This cohort was born between 1931 and 1941, and was aged 50-60 in 1991. The final demographic group examined is that of older retired individuals. These people were born before 1931, and were aged 60+ in 1991. A summary of the different cohorts is provided in table 4.1.

Table 4.1: Summary of Cohorts to be Analyzed

Cohort Name	**Years Born**	**Age - 1991**	**Age - 2001**
POPCHG_YOUTH	1971-1986	5-20	15-30
POPCHG_YOUNG_ADULT	1956-1971	20-35	30-45
POPCHG_ADULT	1941-1956	35-50	45-60
POPCHG_EARLY_RETIRE	1931-1941	50-60	60-70
POPCHG_ELDERLY	Before 1931	60+	70+

The method for calculating population change for each of these individual cohorts is now examined. The equation used to calculate population change is identical to equation 10, except that expected deaths are calculated differently. In this case, expected deaths (D) for each cohort are calculated in equation 9 as:

$$D_i = POP_{i,1991}*AMR_i + EST\_POP_{i,1992}*AMR_i + EST\_POP_{i,1993}*AMR_i + \ldots + EST\_POP_{i,2001}*AMR_i \qquad (10)$$

where *i* is the specific cohort being examined, D_i is the total number of expected deaths over the 10-year period, AMR_i is the national age specific mortality rate of the cohort being examined, and $POP_{i,1991}$ is the initial 1991 population of the cohort. $EST\_POP_{i,}$ is the estimated population of the CCS in each year when the estimated deaths that occurred in all previous years are subtracted from the initial 1991 population. This is done so that people that are already alleged to have died are not multiplied by the mortality rate a second time.

4.5.2 Economic Indicators

Economic indicators form the basic explanatory variables of the empirical model. These variables are essential in testing the hypotheses of this study. The main variables included are employment rates, an income indicator, industry composition indicators, and housing prices. Most of these variables are taken directly from the 1991 CoP, though several required minor modification.

Job availability in each CCS is measured through employment rates. The 1991 employment rate of individuals aged 15+ **(EMPLOYMENT_RATE)** is the first variable utilized. Also, two industry mix employment growth variables were created. Industry mix employment growth **(INDMIX_EMPGROW)** is calculated by multiplying "national" employment growth rates for each industry by the CCS employment shares. INDMIX_EMPGROW is a good exogenous measure of shifts in labour demand in specific locations over time (Partridge and Rickman, 1999).[22] It is exogenous because although it measures employment growth over time (clearly not an initial condition), it only measures the national industry growth rate, and applies the national rate to specific regions. It is does not measure what actually happened in terms of employment growth in each CCS; it measures the national growth of the bundle of industries held by the CCS. In this variable, every industry grows at the same rate, regardless of community, and it is the "mix" of industries present in the communities that essentially determine growth.

Also used in this study is weighted surrounding CCS industry mix employment growth variable **(INDMIX_EMPGROW_SURR)** intended to capture shifts in labour demand in surrounding CCSs, and give a higher weight to closer CCSs. SP_INDMIXGROWTH is calculated by multiplying the CCS industry mix employment growth rate by a spatial weight matrix. Essentially, this variable equals the distance-weighted average of neighbouring CCS industry mix employment growth.[23]

Several other measurements of the business climate of each CCS are included in the empirical model. The percentage of the workforce that is employed in agriculture in 1991 (PER_EMPLOY_AGRIC) is included as a variable to measure the influence of the agricultural sector concentration on population change, given the fact that the agriculture sector is declining in importance in terms of employment. PER_EMPLOY _PRIMARY measures the percentage of the population that is employed in a "primary industry other than agriculture." This includes people employed in mining, forestry, and petroleum

---

[22] Partridge and Rickman (1999) calculate industry mix employment growth for U.S. states as: INDMIX_EMPGROW = $\sum(g_{USit}E_{sit-1})/E_{st-1}$, with g_{USit}, $E_{sit-1,}$ E_{st-1} defined as the national growth rate in industry i, the employment of state *s* in industry i in year t-1, and the total employment of state *s* in year *t*-1 respectively.

[23] A spatial weight matrix is used to calculate the spatial lag of explanatory variables. The spatial weight matrix W is defined as $w_{ij}=1/d_{ij}^2$ where d_{ij}^2 is the squared distance in kilometres between the centroids of CCS i and j.

extraction sectors. Also included in the model is a variable intended to account for entrepreneurial spirit in each CCS: The percentage of the workforce that is non-farm self-employed in 1991 **(PER_SELFEMPLOY)**.[24]

The Herfindahl Concentration Index is also included to measure the degree of industry concentration in each CCS. The Herfindahl index is calculated by summing the squares of the initial-year industry employment shares. A CCS with a high number on the Herfindahl index has a high percentage of its citizens employed in a small number of industries.

The theoretical model of this paper hypothesized that housing costs and wages both impact migration. The 1991 average market value of dwellings in the CCS **(AVG_VALUE_HOME)** is utilized as this measure. The 1991 average per-capita income **(INCOME)** was included to gauge the average wages in each CCS, and indicate which communities have higher wages, and which communities have lower wages on average. The percentage of the community living in households with income below the median national household income **(PER_BEL_MEDIAN)** is also included to provide a variable that indicates the proportion of resident in each community that have low incomes. Due caution is utilized when employing these variables in the modeling process due to potential high correlation between incomes, housing prices, and the various amenities included in the model as suggested in chapter 3.

The final economic indicator included in the model is the distance of each CCS to the nearest "national highway" **(DIST_NATLHWY).** Transport Canada has identified all highways that connect major population/commercial centres, provide major routes

[24] While it is the author's belief that the majority of self-employed individuals are entrepreneurial, some may simply be self-employed because there are no other opportunities available, or their skill-set is not desired by employers.

between provinces, or connect cities to a major port of entry into the U.S. (Transport Canada, 2004). This "National Highway System" was used to create a map using the resources of C-RERL, and then the distance from each CCS centroid to the nearest point on the National Highway system was calculated to form this variable. It is hypothesized that distance to a major trade route is inversely related to community success.

4.5.3 Human Capital

Human capital is usually measured by the education level of individuals within communities. It is generally accepted that more educated individuals are more mobile, so it is essential that the initial education levels of communities be included in the model. This study includes three measurements of human capital: the 1991 percentage of individuals aged 15+ that did not complete high school **(PER_NO_HSGRAD)**, the 1991 percentage of individuals aged 15+ that completed a university degree **(PER_UNIVERSITY)**, and the 1991 percentage of individuals 15+ that completed a post-secondary diploma or certificate program as their highest level of educational attainment **(PER_CERTIFICATE)**. All of these variables were taken from the CoP database.

4.5.4 Demographic Indicators

Demographic indicators are included in the model to measure the degree of ethnicity and age patterns in the CCS. All of these variables were taken directly from the CoP database. The percentage of the population that self-identified as aboriginal on the 1991 census **(PER_ABORIGINAL)** is included as a demographic indicator, as well as the percentage of people that immigrated to the CCS from another country in the last 10 years **(PER_IMMIG_10)**. Both of these variables are utilized to measure the degree of ethnicity in each CCS. Two variables measuring population ages are included in the model as demographic indicators. **PER_OLD** measures the percentage of the population that was aged 60+ in 1991, while **PER_YOUNG** measures the percentage of the population that was aged 5-20 in 1991.

4.5.4 Regional Dummy Variables

Several regional dummy variables are included in the empirical model to distinguish between geographic regions that may be fundamentally different from other regions, and therefore have different migration tendencies that cannot be explained by other explanatory variables. A good example of one region being intrinsically different than all others is the case of the province of Quebec. Since a large proportion of that province speaks only French, Quebec residents have often been found to be less likely to migrate than residents of other provinces (Dickie and Gerking, 1998). This factor would be left unexplained if regional dummy variables were not employed. Often, studies utilize dummy variables at the provincial level to distinguish fundamentally different regions.

This study does not utilize provincial dummies; rather, a regional approach is employed. Dummy variables were created that included all Atlantic Provinces **(D_ATLANTIC),** the province of Quebec **(D_QUEBEC),** and the province of Ontario **(D_ONTARIO).** A western dummy is excluded. Provincial dummies were calculated from the CCS ID number, which designates the province of origin for each CCS. **D_NORTHERN**, which indicates whether the CCS is located in a northern census division of a province. Regional dummy variables are utilized primarily as control variables to account for differences between provinces that are not accounted for by other variables in the model. For example, Quebec is different because its residents primarily speak French, and it is necessary to account for this in the model in case French-speaking Canadians have different behaviour than the rest of Canada.

4.5.5 Social Capital

Five variables attempting to measure the level of social capital, or the level of social cohesion in the community are included in the model. Specifically, I include the presence of religious institutions, the percentage of the households which own their own dwellings, the percentage of the households that have had the same address for five years, the per capita number of volunteer organizations, and the percentage of the electorate that turned out to vote. Several other possible measures social capital were considered such as the presence of community halls, curling rinks, amateur sport, and per-capita charitable donations, but data for these features could not be found.

Several of the variables measuring social capital were available from the CoP, including the percentage of individuals living in an owned home **(PER_OWN_HOME)**, and the percentage of individuals living at the same address as they had lived 5 years ago **(PER_SAMEADDRESS).** The number of volunteer organizations in each census division in 1998 was obtained from "Canadian Business Patterns" available via Statistics Canada. A per-capita number of volunteer organizations (number per 100,000 residents) in each census division was calculated, and applied to every CCS lying within each census division. The variable is defined as **(PERCAP_VOL).**

A variable indicating the presence of churches near each CCS was derived from the DMTI data. **DIST_RELIG** measures the distance from the CCS centroid to the nearest religious institution (churches, temples, synagogues, etc...). A measure of federal election participation at the electoral district level was also included as a measure of social capital. Voter-turnout data was available for each electoral district in Canada, and this data was applied to each of the CCSs falling within a particular electoral district by C-RERL. The **PER_VOTE** variable represents the percentage of eligible voters that actually voted in each CCS in the 2000 general election.

4.5.6 Urban Proximity

The presence of a large population base in a nearby community, or the businesses and the services that only a metropolitan area can provide may be important factors in influencing migration. Also, whether or not the CCS itself is an urban area may influence migration patterns. A variety of variables were formed to account for the proximity of the CCS to an urban area. The first variable utilized in the econometric model is a dummy variable indicating whether the CCS itself is located within a CA or CMA (**D_CCSINCMA).** This variable was formed using the resources of the C-RERL lab, and any CCS located totally or partially within the boundaries of a CA or CMA was considered to be part of a CA or CMA. It should be noted that this variable forms the basis for splitting the dataset into two parts in the empirical model – rural (and small town) and urban.

POP_91 is defined as the population of the CCS, and is used to detect whether the scale of the community has any effect on migration. The data used to form this variable is the CoP database. **POP91_100k** is the second measure of population, and includes the population of the CCS plus the population of all CCSs with centroids within a radius of 100k of the CCS. The distance (in kilometres) to the nearest CMA/CA is also utilized in the study as **CMA_CA_DIST**. The latter two variables were formed using data from the CoP, and the resources of C-RERL, and are used to indicate adjacency to large centres.

4.5.7 Natural Amenities

The empirical model utilized in this thesis includes measurements of the various natural amenities present in each CCS. Natural amenities are the topographical and climatic conditions of the CCS that have existed for thousands of years, and are almost certainly unaffected by human presence. The natural amenities utilized in this study include seven different measures of climactic conditions and indicators of the presence of forest, water, and interesting terrain.

COVER_FOREST is a variable that indicates the presence of forested area in the CCS. The dataset used to create this variable was the1996 AVHRR Land Cover Data produced by Natural Resources Canada's Canada Centre for Remote Sensing. The source data itself consists of national digital vector data describing the topographical makeup of every square kilometre across Canada.[25] The AVHRR land cover data was used to calculate the percentage of each land-cover type comprising each CCS using the resources of C-RERL. In order to form COVER_FOREST, the percentages of each of the individual forest types were summed to obtain a "total forest" percentage for each CCS. Therefore, COVER_FOREST represents the percentage of the total geographic area in each CCS that is forested.

Mountains and hills are hypothesized to be desirable topographical features that enhance recreation and provide pleasurable scenery for residents. Unfortunately, it is

[25] The AVHRR land cover data identifies the following land cover types: mixed forest, deciduous forest, water, transitional forest, coniferous forest, tundra, barren land, permanent ice or snow, agriculture – cropland, agriculture – rangeland, built-up area.

very difficult to systematically decide where mountains and hills exist, and where they do not exist. Indeed, no set index of rugged or scenic terrain in each CCS exists, so such a measurement had to be created for the purposes of this thesis. A digital elevation model was obtained Natural Resource Canada and the resources of C-RERL were used in the creation of this variable. The digital elevation model utilized consisted of an elevation point for every square kilometre in Canada. In order to measure variation in the landscape, or variation in elevation, the standard deviation of the elevation data points lying within each CCS is used to reveal the degree of rugged terrain. This variable is defined as **ELEV_STD_DEV**.

Two variables were constructed to indicate the presence of different types of water in the CCS. First, a dummy variable indicating whether the CCS lies adjacent to the coastline of an ocean or a great lake was generated by C-RERL **(D_ADJ_COAST).** Secondly, another dummy variable was formed to indicate the presence of any body of water adjacent to or within the CCSs boundaries **(D_ANYWATER).** This variable was formed by combining D_ADJ_COAST with the AVHRR land cover data. If the AVHRR land cover data indicated that water comprised greater than zero percent of the CCS, or the CCS was adjacent to a coastline, the CCS was assigned a value of 1.

Climatic conditions comprise the final set of natural amenities utilized in the study. Weather data for every weather station across Canada was obtained from Environment Canada. The problem that had to be overcome with this data is that weather data had to be assigned to specific CCSs because the unit of observation of this thesis is the CCS, not weather stations. C-RERL GIS software was used to calculate which particular weather station was closest to the centroid of each CCS, and the nearest weather station's data (with at least 20 years of observations) was assigned to the CCS.

The weather variables included in the study are: average annual precipitation in millimetres **(WEATH_AVE_PRECIP)**, the average annual snowfall in the CCS in millimetres **(WEATH_AVE_SNOW)**, the average January sunshine measured in hours **(WEATH_JAN_SUNSHINE)**, the average January temperature in degrees Celsius (**WEATH_JAN_TEMP)**, and the average July humidity (**WEATH_JULY_HUMID).** January measurements are used in several cases because summer sunshine and temperatures are very homogeneous across Canada, and it is the winter weather that varies across the country. Humidity is a factor that is differs across Canada in the summertime.

4.5.8 Modern Amenities

The modern amenities utilized in this study include a variety of different man-made institutions that are hypothesized to influence individual's location choices. Violent crime rates and property crime rates **(CRIME_VIOLENT** and **CRIME_PROPERTY)** were included as modern amenities. Crime is an undesirable amenity (disamenity) that people try to avoid. Crime rates were obtained the Canadian Centre for Justice Statistics at Statistics Canada.

Municipal and CMA data was assigned to all appropriate CCSs using GIS software. Data from 26 CMAs was assigned to 260 CCSs, while crime data from approximately 500 municipal police forces was assigned to 310 CCSs. This left approximately 1800 predominately rural CCSs with no crime observations. In order to correct this shortcoming, provincial crime rates were modified for use in these vacant observations. CCSs lacking crime observations were assigned modified provincial level crime rates. The provincial rates were modified by subtracting crimes already accounted for in each province (in the municipal and CMA data) from the total number of crimes committed at the provincial level, and the rates were re-calculated. As long as rural CCSs have relatively uniform crime rates, this is likely a reasonable approach, though there is a possibility of attenuation bias resulting from measurement error from the relatively homogeneous sample. This may bias the estimates towards zero. The crime data consists of 580 unique observations assigned to 2402 observations. The 580 unique observations account for 85 percent of the total Canadian population.

Records of the number of physicians practicing in each CCS were obtained from the Canadian Institute for Health Information (CIHI) Southam Medical Database, which tracks the address of every physician in Canada. CIHI was provided with a Postal Code Conversion File (PCCF) that their technicians used to link the physician address database, and assign the total number of physicians to each individual CCS. The data accounts for 96 percent of the physicians practising in Canada in 1991.

Three variables were calculated from this physician data: a variable measuring the distance from the centroid of every CCS to the nearest CCS centroid having at least one physician **(DIST_PHYS)**, a variable measuring the number of physicians per 100,000 residents in each CCS **(PERCAP_PHYS)**, and a dummy variable indicating the presence of at least one physician in each CCS **(D_PHYS).**

A number of other interesting modern amenity variables were generated from a dataset produced by DMTI spatial Inc. Location data for hospitals, long-term care facilities, bowling alleys, movie theatres, colleges, universities, schools, police stations, and ski hills were available from DMTI. The DMTI Enhanced Points of Interest (EPOI) database uses data from 2003, and consists of the longitude and latitude for over one million points of interest across Canada, along with Standard Industry Classification (SIC) codes and various other pieces of information for these points of interest. The amenities mentioned above were extracted from the database via SIC codes.

C-RERL calculated the distance between each CCS centroid and the nearest individual amenities, as well as counts of the amenities within each CCS and within surrounding CCSs. These total counts were used to compute per-capita counts of the number of amenities in each CCS plus surrounding CCS, and dummy variables indicating whether or not each CCS plus surrounding CCSs possess or do not possess specific amenities. Table 4.2 provides a description of these amenity variables. Three versions of each of these variables were created – A distance variable, a per-capita variable (number per 100,000 citizens), and a dummy variable indicating presence. As such, each of these variables will be prefaced by DIST_, PERCAP_, and D_ respectively.

Table 4.2: Modern amenity variables generated via DMTI data

Variable Name	Description
ACUTE	Acute Care Hospitals
BOWL	Bowling Alleys
CIN	Movie Theatres
COLLEGE	Colleges
GOLF	Golf Courses
LACUTE	Large Acute Care Hospitals >200 beds
LTERM	Long-term Care Facilities
POLICE	Police Stations
SCHOOL	Schools
SKI	Ski Resorts
UNIV	Universities

4.6 Econometric Estimation

An econometric model is the best tool available to researchers to separate and determine the influence that multiple explanatory variables have on a single dependent variable. In this study, the objective is to separate the effects that individual economic factors and amenities have on population change.

The problem with the econometric model utilized in this study is that there are a massive number of variables to contend with, and an almost infinite number of possible specifications, and it is difficult to decide exactly what equations to estimate. Therefore, for organizational purposes, the methodology and results of the econometric model are presented in distinct sections that examine specific issues in the modelling process, and the specific hypotheses outlined in chapter 1.

First, potential problems regarding the linear regression model are outlined. A reasonable expectation regarding whether these problems actually exist are formed, and the steps taken to mitigate these issues are assessed. Common econometric problems that can cause a violation of the fundamental assumptions of regression modelling for cross sectional data include multicollinearity, heteroskedasticity, and measurement error. In the results chapter, exploratory regressions considering a number of different specifications will be estimated to determine the most robust model, and to mitigate econometric and theoretical issues.

The exploratory regressions are utilized to identify the most robust set of explanatory variables. This set of variables is utilized in the "benchmark model", which form the basis of the results. Next, the cohort-specific models are estimated, and the factors affecting these different cohorts are examined through estimation of a number of regressions to compare population change among various population cohorts. Finally, the results as they relate to the hypotheses of the thesis will be outlined.

4.6.1 Potential Econometric Problems

A number of critical assumptions necessary to conduct multiple regression analyzes must be adhered to in order to obtain good unbiased estimators. Here, the most common assumptions, and steps taken to adhere to these assumptions are outlined. First, it is assumed that the relationship between the independent and dependent variables is linear. In practice, it is difficult to confirm this assumption. One can examine previous research to determine how other researchers have specified their models, or examine bivariate scatterplots of the variables. Most past migration research has used linear models to estimate the parameters, so this thesis simply follows past research. It is also assumed that the regression variables disturbances are normally distributed. This assumption was tested through examining data plots, and appears to be true for most variables.

It is assumed that the variables are homoscedastic; in other words, the variance of the error terms is evenly distributed. In the presence of heteroskedasticity, OLS provides consistent parameter estimates, but the usual OLS standard errors will be incorrect and should not be used for inference. To mitigate any unknown heteroskedasticity in the model, White's heteroskedasticity consistent covariance matrix estimator is utilized, which provides correct estimates of the coefficient covariances in the presence of heteroskedasticity of unknown form.

One assumption that must be made is that the variables are measured without error. Earlier, it was argued that CCSs with small populations and larger land areas will have more measurement error in the independent variables than CCSs with large populations and small geographic areas. Therefore, to mitigate any measurement error in the independent variables, a weighted least squares (WLS) estimation procedure is utilized, weighting the observations with respect to the initial population of the CCS. This option assigns a lower weight in the model to observations that are likely to have a high degree of measurement error.

Finally, it is assumed that no exact linear relationship between the independent variables exists. In practise, this assumption is achieved through examining pair-wise correlations for all of the variables. In general, the only variables that are highly correlated in this thesis are the "distance-based" amenity measures, and different specifications of the same variables. Care was taken not to include any two independent variables in the model with a pair-wise correlation greater than 0.8.

4.6.2 Determining the Most Robust and Unbiased Model

The first step in addressing the hypotheses is to establish the set of variables that provide the most robust specification, while minimizing potential theoretical and econometric concerns. Robustness has a variety of definitions, but for the purposes of this thesis, the factors used to determine robustness are: 1) the "fit" of the overall model as represented by the F-Statistic and R-squared values, 2) the level of significance of the individual explanatory variables as revealed by the coefficient t-statistics, 3) whether or not groupings of similar variables are jointly significant, and 4) whether or not the individual variables exhibit the direction of influence on the dependent variable predicted by the literature and the theoretical model (as shown in table 4.3).

The dependent variable in these initial models will be the total population change for the entire national dataset comprised of both rural and urban observations. Initially, the model will only include economic indicators, but regional dummy variables, urban adjacency variables, human and social capital, natural amenity variables, and modern amenity variables will be incrementally added to the estimation to determine whether such variables improve robustness, and how adding different variables affects previously added variables. Furthermore, a number of other concepts will be tested including:

- Estimating WLS regressions versus OLS regressions to mitigate measurement error.[26]
- The effects of including income and housing prices in the model, because the theoretical model predicts these variables are endogenous with amenity measures.
- Assessing the three different possible amenity access measures (Distance, presence, and per-capita number) to determine whether they all provide the same results, and which measures correspond with theoretical expectations.
- Interacting water and temperature variables to examine whether it is a combination of water and nice weather that attract residents.
- The effectiveness of separating the sample into rural and urban observations, and how this affects the results.

Ultimately, the results of these experiments will lead to the estimation of the benchmark model, which is useful in analyzing how the explanatory variables impact total population change. Minor variations to this benchmark may be utilized in the cohort analysis.

26 Two WLS specifications will be generated; one weighting observations by the log of the initial population, and one weighting by initial population.

4.6.3 Cohort Analysis

The next step in the modelling process is to assess differences between the five different age cohorts. To accomplish this, five dependent variables representing population change for different age-cohorts will be substituted into the model. In general, the estimated equation will include the benchmark model explanatory variables, though several key variables will vary for the different population cohorts. For example a number of education and healthcare variables are highly correlated, so more specific education variables will be included for younger cohorts, while more specific health care variables will be included for older cohorts.

4.6.4 Expected Direction of Influence

Table 4.3 illustrates the expected direction of influence for the different variables utilized in this study. The expected signs reflect theoretical expectations, and the results of past empirical studies. Most of the expected signs are self explanatory, though several will require additional rationalization.

Expectations regarding AVG_VALUE_HOME and INCOME cannot be formed since the levels of these variables are highly dependent upon the level of amenities in the community. Housing prices and income may be fundamentally related with levels of community amenities, along with many other factors. The theoretical model predicts that housing prices are positively related with the level of amenities, while incomes should be negatively related. Since amenities, housing costs, and incomes are all inter-related, it is difficult to predict the effect that one of these variables will have on population change.

Proximity to urban centres is assumed to be an amenity, which is why the direction of influence is expected to be positive. Amenities are generally expected to cause more people to be living in a particular area, and given our assumption that population migration is slow and long process, we expect population growth to be associated with the level of amenities, which is why expect signs for these variables to be positive. The exception is when amenity access is measured through distance. For distance measurements, the expected sign is negative because the greater the distance, the poorer the level of access.

Table 4.3: Expected Direction of Influence

Variable Name	Expected Direction of Influence
Economic Indicator Variables	
AVG_VALUE_HOME	?
DIST_NATLHWY	-
EMPLOYMENT_RATE	+
HERF_INDEX	-
INCOME	?
INDMIX_EMPGROW	+
INDMIX_EMPGROW_SURR	+
PER_BEL_MEDIAN	-
PER_EMPLOY_AGRIC	-
PER_EMPLOY_PRIMARY	-
PER_SELFEMPLOY	+
Human Capital	
PER_UNIVERSITY	+
PER_CERTIFICATE	+
PER_NO_HSGRAD	-
Demographic Indicators	
PER_ABORIGINAL	+
PER_IMMIG_10	?
PER_OLD	-
PER_YOUNG	+
Regional Dummy Variables	
D_ATLANTIC	?
D_QUEBEC	+
D_ONTARIO	?
D_NORTHERN	-
Social Capital	
DIST_RELIG	-
PER_OWN_HOME	+
PER_SAMEADDRESS	+
PERCAP_VOL	+
PER_VOTE	+
Urban Proximity	
D_CCSINCMA	+
POP_91	+
POP91_100k	+
CMA_CA_DIST	-
Natural Amenities	
COVER_FOREST	?
D_ANYWATER	+
ELEV_STD_DEV	+
WEATH_AVE_PRECIP	-
WEATH_AVE_SNOW	-
WEATH_JAN_SUNSHINE	+
WEATH_JAN_TEMP	+
WEATH_JULY_HUMID	-

Table 4.3 continued on next page

Table 4.3: Expected Direction of Influence Continued

Variable Name	Expected Direction of Influence
Modern Amenities	
CRIME_VIOLENT	-
CRIME_PROPERTY	-
DIST_PHYS, DIST_ACUTE, DIST_BOWL, DIST_CIN, DIST_COLLEGE, DIST_GOLF, DIST_LACUTE, DIST_LTERM, DIST_POLICE, DIST_SCHOOL, DIST_SKI, DIST_UNIV	-
D_PHYS, D_ACUTE, D_BOWL, D_CIN, D_COLLEGE, D_GOLF, D_LACUTE, D_LTERM, D_POLICE, D_SCHOOL, D_SKI, D_UNIV	+
PERCAP_PHYS, PERCAP_ACUTE, PERCAP_BOWL, PERCAP_CIN, PERCAP_COLLEGE, PERCAP_GOLF, PERCAP_LACUTE, PERCAP_LTERM, PERCAP_POLICE, PERCAP_SCHOOL, PERCAP_SKI, PERCAP_UNIV	+

4.7 Chapter Summary

The procedure for testing the hypotheses of the thesis was described in this chapter. The general model was initially described, followed by a detailed description of data sources, and a discussion of several unique modeling issues. The data sources and description of the specific variables utilized were then detailed. Finally, a description of the procedure methods used to estimate the regressions was outlined. The next chapter presents the results of these regressions.

Chapter 5: Results

5.0 Introduction

This chapter presents the results of the estimated econometric models. The chapter is organized as follows: first, the hypotheses tests utilized to assess the statistical significance of the estimates are reviewed. Next, the results of a number of exploratory regressions are outlined to explain the different model specifications that were tested, why different variations of the model were rejected, and why the benchmark model was ultimately adopted.

The chapter follows by presenting the results of the benchmark model and the cohort-specific models. The results are presented with the primary objective of addressing the hypotheses of this thesis, though other interesting results will be outlined. Due to the massive amount of results generated during the course of this study, only the most relevant and interesting results are presented and analyzed.

5.1 Hypothesis Tests: Are the Coefficients Significant?

Two different statistical tests are utilized to establish the validity of the estimated regression coefficients: the standard t-test, and a restricted least squares F-test. Basically, the t-test is used to test whether the individual coefficients are significant, while the F-test is used to determine whether variable groupings are jointly significant.

5.1.1 Standard t-Test

The t-test is one of the most commonly used methods of determining whether specific coefficient estimates are statistically significant. For the purposes of this study, the null hypothesis H_0 is that the estimate equals zero, while the alternate hypothesis H_1 is that the absolute value of the coefficient is not equal to zero.[27] However, if a coefficient is not significant at (say) the10 percent level, it does not necessarily mean that the relationship does not exist; it simply means that the probability of that estimate reflecting reality is not high enough to have complete confidence in it.

5.1.2 The F-Test Approach: Restricted Least Squares

An alternative but complementary approach to the t-test method of testing significance is to use a restricted least squares, or grouped approach, and test whether a subset of variables in the equation jointly have coefficients equal to zero and might thus be deleted from the equation. In my case, examples of variable groupings include economic indicators, natural amenities, and modern amenities. The null hypothesis H_0 is that the particular subset of coefficients all equal zero, and are thus not jointly significant, while the alternative hypothesis is that the coefficients are not all equal to zero.

5.2 Exploratory Regressions – Identifying the Best Model

The dataset utilized in this thesis is comprised of a large number of explanatory variables. As such, there are many different possible ways to specify this model. It is important to test the effects of inserting different variables into the model to determine which variables provide the optimal results, and to understand how the model responds as small changes are made. This section examines several of the different models utilized to find an optimal set of explanatory variables to use in the analysis. Many more series of

[27] Though one could argue that a one-tailed t-test may be appropriate for variables where a strong *a priori* expectation of the direction of influence exists, the two-tailed t-test is used for all coefficients as a cautious effort to avoid type I error. For a detailed description on the calculation and interpretation of t-statistics and F-statistics, refer to Gujarati (2003).

models were estimated, but most will not be shown here. This optimal set of variables is referred to as the "benchmark model."

5.2.1 Adding Groups of Variables in Stages

In order to assess the effects that different groups of variables have on the overall model and to separate the effects that different groups of variables have on population change, several models were estimated where groups of variables were incrementally added to the model. In general, it was found that when one individual group of variables was utilized as the only regressors in the model, the t-statistics were larger for the majority of the variables in the grouping. In other words, when utilized alone, most variable groups had more significant coefficients than when utilized in conjunction with additional variable groups.

As additional variable groups were added to the model, coefficient signs did not generally change, though the t-statistics of many of the variables did decline. Overall, this behaviour could be indicative of some collinearity between the different variable groups, or simply that other variables were superior in explaining the variation in the dependent variable, thus rendering previous variables less significant. Appendix C provides evidence of these phenomena.

5.2.2 Weighted vs. Unweighted Regressions

The justification for utilizing regressions weighted by population was discussed in chapter 4. The basic reasoning for weighting is to mitigate measurement error resulting from smaller sample sizes in CCSs with less population. However, the objective of this study is to determine why different communities have experienced different rates of population change, and not just to explain population change in the places where the most people live. Essentially, the objective in choosing amongst weighted models was to find a model that does mitigate measurement error, yet provides roughly the same results as an unweighted model because what happens in smaller rural communities is essentially what this study is really trying to expose. Ultimately, the decision was made to weight all of the regressions presented in this thesis by the log of the initial population due to the results of exploratory regressions (presented in appendix D), and the fact that rural observations are given some influence in the model.

5.2.3 Assessing the Various Modern Amenity Measures

There are three different ways to measure the level of access a community has to modern amenities: the distance to the nearest amenity, a dummy variable indicating whether or not the amenity exists in the CCS, and a per-capita number of amenities that exist in each CCS. In order to test these three different types of variables, several models were estimated in which the different amenity measures were substituted in for one another. In weighing which group of amenity measures to use, the following criteria were considered: 1) whether the group of variables were jointly significant, 2) which group provided the highest R^2 value, and 3) which group corresponded most closely with the theoretical sign predictions in table 4.3.

Modifying the regression results by inserting different groups of modern amenity measures did not appear to have a huge impact on the other variable groups. In fact, deleting the amenity measures altogether did not have a significant impact on the other results.[28] It was also found that the coefficients for the three different amenity measures were very inconsistent when compared with one another, and the one thing that is abundantly clear is that the groups are certainly not interchangeable. The grouped F-test results suggest that all three groups provide coefficients that are jointly significant at the 1 percent level. Of the three groups, the distance variables returned the highest R^2 value, but the difference was not very big. The first two criteria do not clearly favour one amenity grouping over another.

The third criteria of deciding which group of modern amenities corresponded most closely with theoretical predictions proved to be the deciding factor for which set of variables to utilize for the remainder of the study. It was found that the distance regressors corresponded most closely with a priori expectations, and had more significant t-statistics than either the dummy variable or per-capita coefficients. Four distance coefficients were significant at the 10 percent level compared with zero for the dummy variable set, and three for the per-capita group. With the distance variable grouping, six variables corresponded with a priori predictions of the sign, while only four variables complied in the dummy variable grouping, and zero with the per-capita grouping.

For these reasons, the decision was made to proceed with using the distance variables for the remainder of the study, as they provide the best compromise with respect to significant results and the connection with theory. Overall, the results suggest that dummy variables can function well as estimates of access to amenities, but per-capita counts of amenities within a particular geographic region do not provide robust results. Further work comparing different amenity variable specifications is warranted, as the

[28] The results when the Modern amenity variables are dropped from the benchmark regression are shown in appendix C, table C-2.

dummy variable specification did show some promise.

5.2.4 Adjacency to Water

Initial estimations of the model indicated that the adjacency to water variable (D_ANYWATER) was typically insignificant, and always returned the opposite sign that theory and previous research predicted. Initial results showed that areas located adjacent to water had lower population growth, which implies that water is a disamenity. As such, some work was undertaken to use different variations and interactions with this variable to determine whether the chosen variable was the cause of this strange result, or whether the presence of water is consistently a disamenity and/or statistically insignificant as different water measurements are utilized. The initial variable, D_ANYWATER is a dummy variable indicating the presence of any water within the CCS and/or adjacency to the ocean or large lakes.

Dummy variables separating D_ANYWATER into its components (adjacency to ocean, adjacency to large Canadian lakes, and presence of lakes within the CCS) were utilized within the regression, but none provided significant results that complied with theory. Some work was completed attempting to interact the water variables with temperature as well, but none of these specifications provided improved results.[29] Because no superior specification was found, D_ANYWATER was used in the benchmark model.

[29] The reasoning behind interacting January temperature and water presence is because it was hypothesized that it is a combination of warm weather and water that is considered an amenity. For example, being near the coast of Labrador or Hudson's Bay may not be considered a 'pleasurable' amenity.

5.2.5 *Capitalization of Amenities Into Housing Costs and Income?*

At numerous stages of this thesis, the assertion has been made that wages, housing costs, and amenities are determined simultaneously. The theoretical model of this thesis presented in chapter 3 demonstrates that the levels of income and housing prices in each CCS are dependent upon the level of amenities present in the CCS and each other. If this is indeed true, it was hypothesised that utilizing these two variables in conjunction with each other, and with other amenity variables in the study could produce adverse effects, as this is the textbook definition of endogeneity. Although all of these variables are predetermined in the model, it is possible that including all of these variables at the same time in the model could produce inconsistent estimates

For these reasons, it is important to test the relationship between housing costs, income, and all other amenity variables. The effects of removing the INCOME and AVG_VALUE_HOME variables from the model, and the effects on these two variables when all of the other amenity measures are removed are shown in Appendix E, table E-1. If endogeneity is a severe problem, removing one of these variables from the model may drastically affect the remaining regression coefficients.

Model 1 consists of the benchmark model, while AVG_VALUE_HOME has been deleted from model 2. Overall, this action does not have a large effect on the other variables. Income surprisingly loses its significance (but the coefficient is largely unchanged), and the human capital grouping becomes less significant. POP_91_100K also becomes significant, while JAN_SUNSHINE loses its significance. Overall, the effects of these actions are relatively minor, and perhaps most surprisingly, the deletion of AVG_VALUE_HOME from the model does not have a large effect on the amenity variables present in the model.

In model 3, INCOME was deleted from the model. As with the housing cost variable, this action did not appear to have a large effect on the remaining estimated coefficients. About the only major consequence of deleting this variable is that the human capital grouping of variables becomes more significant. In model 4, all of the amenity variables were removed from the benchmark model to assess the impact on AVG_VALUE_HOME, and INCOME. This action leaves these two variables virtually unchanged.

Overall these tests appear to indicate that inserting both AVG_VALUE_HOME and INCOME into the model does not severely impact the results, especially amongst the amenity variables. Perhaps the particular amenities that are capitalized into these two variables are simply not present in this particular model. If these two variables were highly endogenous with the amenity measures, one would expect the deletion of AVG_VALUE_HOME and INCOME to drastically impact the amenity estimates, which clearly doesn't happen here. Also, one would expect deletion of the amenity variables to impact these two variables, and that is not the case. Overall, these results indicate that including AVG_VALUE_HOME and INCOME in the model does not produce and adverse effects, and it actually increases the explanatory power of the model. As such, both of these controversial variables will be included in the benchmark model.

5.2.6 Urban Differences

In estimating the results of this thesis, it was noted that the results differed when the dataset was split into rural and urban components. This was originally done because it was recognized that urban and rural communities are vastly different entities, and that perhaps if the analysis was completed separately on these two groups, it would provide better results. In appendix E, table E-2 shows the results of the cohort analysis when only the urban CCSs were utilized in the model, and table E-3 shows the results when the rural CCSs were utilized in the model.

Basically, the urban models did not work overly well, with a high number of insignificant variable groupings, and a small number of individually significant variables. The results of the rural models were very similar to the results when the total sample was utilized. One of the main reasons for the insignificant results of the urban sample was that when the rural observations are removed from the dataset, a large number of the variables ended up being highly correlated, particularly any variables utilizing a distance measurement.

The fact is that this dataset was compiled with the intent of examining rural areas and differences between rural and urban areas. As such, the geography utilized was not fine enough to capture differences between urban areas. For example, most CA/CMA have nearby hospitals, so the distance the nearest hospital should be close to zero for every urban CCS. In order to facilitate comparisons between urban areas, a finer level of geography would be needed. Although the results of the rural/urban analysis will not be discussed in detail here, readers are encouraged to study these results and compare them with the results when the total dataset is utilized.

5.3 The Benchmark Model

Table 5.2 (presented at the end of this section) outlines the results when the total population change in each CCS was regressed on the explanatory variables. The explanatory variables utilized in this model were selected through a rigorous process described in detail in section 5.3. The model explains 78.4 percent of the variation in CCS population change between 1991 and 2001. The benchmark model is presented to give readers an overall representation of the results, and to get readers familiarized with them, as more complicated tables are presented later in the chapter.

The high R^2 depicted in this model is complemented by the fact that all of the variable groupings with the exception of the human capital variables were found to be jointly significant in the model at the 1 percent level. The human capital variables were significant at the 10 percent level. A large number of variables were individually significant in this model, and most exhibit the direction of influence that was predicted in table 4.3.

Particularly influential in this model are the economic indicator variables. Interestingly, the sign for AVG_VALUE_HOME is positive, meaning people are gravitating towards areas with a high cost of living, and the sign for income is negative, indicating that population growth rates are higher where average incomes are lower. Although one would expect Canadians to prefer areas with a lower cost of living and higher incomes, the evidence presented here clearly refutes what would normally be considered a logical conjecture of human behaviour.

In fact, this result is not surprising because it is consistent with the theoretical model of this thesis. It is likely that amenities have been capitalized into both of these regressors over time, as was predicted in the theoretical framework. That is, amenities caused past population increases that raised housing costs and lowered average incomes, cetaris paribus. Past housing values may only partially reflect the level of amenities, and subsequent migration flows reflect a disequilibrium adjustment, as people flock to communities where amenities are under-valued in terms of housing prices. In other words, local factors (including all those measured in this model and other factors not measured here) influence population movement, leading to adjustments in housing values over time (due to demand). As long as housing values and income adjustments lag behind the perceived community amenity levels, *subsequent* population growth will ensue.[30]

One final explanation is that universal income increases and changing tastes in Canada continue to favour amenities, and people are moving to areas that historically have had a favourable bundle of amenities and economic opportunities. At any rate, these results can be interpreted as strong evidence that Canadians consider factors other than earning power and cost of living in choosing their preferred location.

CCSs with a high industry concentration tend to exhibit a much lower population growth rate, as shown by the negative sign on the HERF_INDEX variable. This provides evidence that communities with diversified business climates have a decided advantage in attracting population compared with non-diversified communities. People may prefer diversified economies due to a high number of alternative job opportunities in the event that one industry weakens. There is evidence that the presence of highly concentrated primary industry sectors (PER_EMPLOY_PRIMARY) and agriculture (PER_EMPLOY_AGRIC) sectors negatively impact population growth. It appears that

[30] The hypothesis that amenities and other local factors influence dwelling prices over time was tested by regressing the log of the 2001 average housing price on the selection of explanatory variables (including the log of the 1991 average housing price). The results clearly show that future housing values are heavily influenced by the initial housing values. However, initial levels of amenities and economic factors also play a very strong role in influencing future housing values. This confirms that average dwelling values fluctuate over time in response to favourable local conditions, and areas with historically high housing prices can have their values reduced if they do not maintain favourable local factors and amenities.

there is something unique about the presence of these two industries that cannot be explained simply by initial-employment, employment growth, or industry concentration measures.

The presence of these sectors has had a negative impact on population growth between 1991 and 2001. For example, a CCS where 100 percent of the workforce was employed in primary industry would have 18 percent lower population growth than an identical community with zero employment in the resource sector, *cetaris paribus*. One possible explanation for these results is that technological change in the agriculture and resource sector that transpired decades ago continues to affect these sectors, and influence the communities where agriculture and resource extraction are/were prominent.[31] High employment shares in these sectors could be an indicator of an expection of fewer jobs in the future, as residents may be anticipating that machines will continue to substitute for labour in these sectors.

Also, initial employment rates and increases in employment demand (measured by INDMIX_EMPGROW and INDMIX EMPGROW_SURR) were significant factors in influencing population growth; though the industry mix employment growth rate was only significant when the surrounding CCSs were included in the measure.[32] These results are indicative of the importance of jobs, especially in surrounding communities in influencing population growth. Places with a higher initial demand for labour, or a

[31] Although the percentage employment in agriculture is not statistically significant in this model, it was highly significant in other runs, particularly when the urban observations are excluded from the model, and for younger population segments.

[32] INDMIX_EMPGROW was negative and insignificant, while INDMIX_EMPGROW_SURR was positive and significant, indicating that job growth in surrounding regions is a more important factor than local growth in influencing population growth rates.

growing need for labour will invariably attract more residents than communities where a larger proportion of the community is looking for work.

Demographic factors also appear to be important influencers of migration. For example, areas with a higher initial percentage of aboriginal people tended to experience high population growth rates. This could be due to high birth rates among the aboriginal population. This statistically significant trend could be viewed as an area for future economic growth in many prairie and remote regions. Places with a high concentration of youth in 1991 exhibited higher population growth. Though not statistically significant in this model, areas with a high concentration of senior citizens tended to have lower population growth rates.

The urban indicators utilized in this model were not individually statistically significant, yet they do provide some interesting results worthy of mention. The implication of non-significant t-statistics could definitely mean these coefficients equal zero and are thus irrelevant, yet the high F-statistic indicates that these variables are important as a group. One possible explanation is that there may be some collinearity among these indicators and other variables in the model. The coefficients themselves indicate that overall, population growth tended to increase when CCSs were closer to CMAs, yet being a part of a CMA or CA tended to result in lower population growth. CCSs with higher initial populations tended to have lower population growth, while CCSs with high populations in neighbouring CCSs tended to have higher population growth. These results appear to suggest that in general, it is more advantageous for CCSs to be near a large urban centre than to actually be an urban centre, though as previously mentioned, these results are not individually significant.

In terms of modern amenities, the results suggest that close access to common acute care hospitals may actually function as a disamenity. Places that were close to smaller hospitals tended to have a lower population growth rate than places that were further away, when all other factors are held constant. At the same time, proximity to large hospitals (with more than 200 beds) appears to function as an amenity, as CCSs closer to large hospitals tended to have higher population growth rates. What makes these results even more intriguing is that the distance to the nearest CMA/CA, and population sizes have been accounted for in the model, so this result is not simply picking up the fact that people like to locate near towns and cities. Size and capabilities of hospitals may play an important role in the economic development of communities, not just the fact that a facility exists.

The results indicate that CCSs with high rates of violent crime experienced lower growth, and areas that were closer to police stations had higher growth rates. It seems as though community security is a feature valued by residents. College distance was significant but had the opposite sign as predicted by theory. Most of the other modern amenities, while insignificant, did exhibit the predicted sign. The only significant variables in the natural amenity grouping were the July humidity and January sunshine indicators.

While only a small number of amenity and social capital variables were individually significant, it is important to note that groupings of these variables were jointly significant. The presence of multicollinearity amongst these variables is one likely explanation for this discrepancy. However, the fact that modern amenity, natural amenity, and social capital groupings were jointly significant at a 1 percent level combined with the results of the direction of influence of the INCOME and AVG_VALUE_HOME variables provides strong evidence that amenities play at least some role in influencing population growth rates.

Table 5.2: Benchmark Model - Dependent Variable: Percentage Change in Total Population

Variable Name	Coefficient	Grouping	F-Statistic
C	5.036 (0.634)		
AVG_VALUE_HOME	0.025*** (3.03)	Economic	49.2 (p=.000)
DIST_NATLHWY	0.005* (1.77)		
EMPLOYMENT_RATE	0.124*** (3.06)		
HERF_INDEX	-18.134*** (-5.31)		
INCOME	-0.00021* (-1.74)		
INDMIX_EMPGROW	-0.025 (-1.03)		
INDMIX_EMPGROW_SURR	0.382*** (5.03)		
PER_BEL_MEDIAN	-0.06*** (-2.9)		
PER_EMPLOY_AGRIC	-0.05 (-1.32)		
PER_EMPLOY_PRIMARY	-0.18*** (-5.16)		
PER_SELFEMPLOY	0.252*** (5.24)		
PER_CERTIFICATE	-0.07 (-1.61)	Human Capital	2.3 (p=.075)
PER_NO_HSGRAD	-0.043 (-1.14)		
PER_UNIVERSITY	-0.089 (-1.41)		
PER_ABORIGINAL	0.086*** (2.78)	Demographic	30.7 (p=.000)
PER_IMMIG_10	-0.023 (-0.293)		
POP_PER_OLD	-0.062 (-1.1)		
POP_PER_YOUNG	0.139** (2.11)		
D_ATLANTIC	0.134 (0.16)	Regional	6.5 (p=.000)
D_NORTHERN	-1.018 (-1.1)		
D_ONTARIO	-1.558** (-2.32)		
D_QUEBEC	-0.854 (-1.06)		
DIST_RELIG	0.013 (0.834)	Social Capital	34.3 (p=.000)
PER_OWN_HOME	0.028* (1.8)		
PER_SAMEADDRESS	-0.086 (-1.44)		
PERCAP_VOL	-0.899 (-1.61)		
PER_VOTE	0.007 (0.284)		
D_CCSINCMA	-0.505 (-1.45)	Urban Scale	5.2 (p=.000)
POP91_100K	0.00028 (1.58)		
POP_91	-0.001 (-1.19)		
CMA_CA_DIST	-0.007 (-1.53)		

Table 5.2 Continued on next page

Table 5.2: Benchmark Model - Dependent Variable: Percentage Change in Total Population Continued

Variable Name	Coefficient	Grouping	F-Statistic
CRIME_PROPERTY_RATE	0.00015 (0.973)	Modern Amenities	5.9 (p=.000)
CRIME_VIOLENT_RATE	-0.001*** (-2.73)		
DIST_ACUTE	0.019** (2.32)		
DIST_LACUTE	-0.011*** (-3.2)		
DIST_COLLEGE	0.009** (2.51)		
DIST_GLF	-0.004 (-0.534)		
DIST_PHYS	-0.003 (-0.349)		
DIST_POLICE	-0.015* (-1.72)		
DIST_SCHOOL	-0.015 (-0.955)		
DIST_SKI	0.005 (1.05)		
DIST_UNIV	-0.001 (-0.47)		
COVER_FOREST	-0.005 (-1.2)	Natural Amenities	10.0 (p=.000)
D_ANYWATER	-0.316 (-0.925)		
ELEV_STD_DEV	0.003 (1.18)		
WEATH_AVE_PRECIP	0.00025 (0.444)		
WEATH_AVE_SNOW	-0.003 (-1.39)		
WEATH_JAN_SUNSHINE	0.017* (1.69)		
WEATH_JAN_TEMP	-0.079 (-1.45)		
WEATH_JULYHUMID	-0.049*** (-3.67)		
R^2	0.784		
N	2402		

This regression is weighted by the log of the initial-year CCS population. The territories and CCSs with a population of less than 250 are excluded from this regression. T-Statistics and f-statistic p-values are reported in parenthesis. Model uses White's Heteroskedasticity-Consistent Standard Errors and Covariance. Regression is weighted by log(POP_91). See Appendix A for variable definitions, and table 4.3 for theoretical expectations of directional influence. *, **, and *** denote significance at 10%, 5%, and 1% levels respectively.

5.4 Cohort Estimations

Table 5.3 (presented at the end of this section) contains the results of five different models, each utilizing a different dependent variable representing a unique segment of the population. These models form the core of the results, and are used to examine how local factors affect different segments of the population. The benchmark model presented above shows us what the population as a whole appears to value, but is somewhat limited in that it is not helpful in showing the community attributes that young people tend to value, or the factors that make a community attractive to retired people, which may completely offset one another or be masked if one simply examines total population change.

Model 1 utilizes the percentage change in population for the *youth* cohort (aged 5-15 in 1991) while models two thru five examine progressively older age cohorts. The cohort-specific dependent variable definitions are outlined in Appendix A. The majority of the explanatory variables used in the models are identical to the benchmark model variable selection, except that the older-cohort models do not include educational variables, but do include a variable indicating access to long-term care facilities. Educational institution (distance) variables were dropped to mitigate the multicollinearity that exists between schools and long-term care facilities.

Although it is difficult to articulate what amounts to a massive quantity of information, results are discussed in a logical manner by examining each population cohort separately, and noting interesting differences between theses results, and the results from alternate population segments. Following the cohort-specific analysis, the results are summarized to denote different trends that appear to exist between the different population segments. The results specifically as they pertain to the hypotheses of this thesis will be examined at the conclusion of this section. For interested readers, cohort estimations for both "urban" and "rural" subsets of the data are presented in Appendix E.

5.4.1 Factors Influencing Youth

Model 1 utilizes population change for what this thesis classifies as *youth* as the dependent variable. In 1991, these individuals would have been aged 5-20, and thus most were still in school, presumably still living with their parents during the initial period. By 2001, most of these individuals will have graduated from high school, and reached the age of majority and the ability to make their own location decisions. This segment of the population is perhaps the most interesting, as many jurisdictions may feel the key to future success starts with retaining and attracting young people.

The model explains 73.8 percent of the variation in the population growth rate of the *youth* cohort in different CCSs across Canada, which puts this model among most robust of the cohort-specific models in terms of explanatory power. Results of the grouped F-test indicate that all of the groups utilized in this model are jointly significant at a 1 percent level with the exception of the urban indicator group, which is significant at the 10 percent level.

The results of this regression appear to indicate that amenities play an integral role in determining population growth for the *youth* cohort. As with the benchmark model presented earlier, the results suggest that young people are moving to places with higher housing costs. One rational explanation for this phenomenon is that amenities are being capitalized into housing costs. The results show that CCSs with $10,000 higher average housing costs have .39 percent higher population growth rates for young people.

To further this point, a number of the modern and natural amenity variables utilized in the model are individually significant. One interesting result is that young people appear to be gravitating into communities with high property crime rates, but away from those with high violent crime rates. One possible explanation for this phenomenon is that high rates of property crime only exist in high-amenity areas. Young people were also not attracted to areas adjacent to bodies of water, but did prefer communities with lower precipitation.

Young people appear to be drawn to areas that are closer to large hospitals, but further away from smaller hospitals. Proximity to schools, and universities were not significant in predicting the location choices of youth. This is unexpected, considering that many members of the *youth* cohort surely require education, and must live where these schools are located to obtain education. Even more surprising is the result that communities closer to colleges did not experience higher *youth* population growth. There are several explanations for these phenomena. Students are typically quite nomadic, residing in different places at different times of the year, and certificates and degrees only take two to four years to obtain. As a result, many students may have completed their education and subsequently moved away over the ten-year period. Others students (like the author of this thesis) may keep their parent's address as their permanent address and thus may never be on record as ever having moved to the city in which they are being educated.

Another interesting result is that young people do not simply appear to be gravitating towards other large groups of young people. A common notion is that young people locate in places where they can congregate for recreational purposes and also meet more members of the opposite sex. The results of this study are inconsistent with this theory, as a higher percentage of "young" people in 1991, appeared to result in a reduced growth rate for the *youth* cohort. This could be because young people are at more of a risk of moving. Young people also tended to avoid areas that had a high percentage of older people in 1991, which is not surprising.

Overall, a large number of economic variables were significant in explaining why some communities had high youth population growth rates, while others fared worse. Most obvious is the Herfindahl index measuring industry concentration in each CCS. Young people were also attracted to areas that had high employment rates in the initial period, and CCSs with a high industry mix employment growth rate. Communities that had a 10 percent higher employment rates than neighbouring communities could be expected to have a youth population growth rate 3.5 percent higher than neighbouring communities, *cetaris paribus*. Interestingly, the distance-weighted industry mix growth rate for surrounding CCSs was not significant in this particular model.

Other key results are that the level of income was not significant in explaining youth population growth rates, though it is negative (as was the case in the benchmark model). The *youth* cohort also tended to avoid areas where a high percentage of employment opportunities were in the agriculture and other primary industry sectors. Also, according to these results, young people have tended to gravitate towards areas where a higher percentage of residents have lower incomes, and where fewer people are non-farm self-employed (this may indicate areas with less entrepreneurship and more large companies).

5.4.2 Factors Influencing Young Adults

Model 2 utilizes the percentage change in the *young adult* demographic between 1991 and 2001 as the dependent variable. This group of people would have been aged 20-35 in 1991. It is important to examine this cohort in order to assess what factors are responsible for bringing *young adults* back to rural communities. In this model, all variable groups were jointly significant at a 1 percent level with the exception of the human capital variables, which were jointly significant at a 10 percent level. Many of the results of this model are similar to the benchmark and *youth* models presented previously with some intriguing differences.

The first major difference between the behaviour of *young adults* versus *youth* is that the urban indicator grouping is highly significant. The results suggest that CCSs that were part of CMAs or CAs actually experienced one percent lower population growth than rural CCSs. Other non-significant results for the urban proximity grouping indicate that young adults preferred to live within communities with lower populations, but where the surrounding communities had higher populations. This is consistent with "return" migration to rural areas.

Like *youth*, *young adults* are attracted to areas with higher housing costs, which again may indicate that amenities have been capitalized into housing costs over time. Young adults generally appear to respond to similar types of modern amenities as youth, the exception being that young adults do not appear to value proximity to police stations. Young adults exhibit the same unexpected statistically significant pattern of migrating towards large acute care hospitals, but away from smaller facilities.

This age grouping appears to be particularly receptive to natural amenities, as they preferred communities with lower winter temperatures, more sunshine, more precipitation, and less humidity. They also did not appear to value forest, as they tended to locate in areas with a lower percentage of forested area. From the results of all of the cohort groupings, it is definitely unclear as to whether forest (and by extension, wildlife and nature) represent an amenity for any segment of the population.

Social capital was again highly significant as a group, as indicated by the F-statistic located at the bottom of table 5.3. However, only two of these variables were individually significant. *Young adults* tend to locate in communities with a higher percentage of home ownership, and a lower percentage of volunteer organizations per-capita. Proximity to religious institutions was also not a significant factor in explaining the location decision of this cohort.

Young adults also appear to be heavily influenced by economic factors, though there are a few subtle differences from the *youth* cohort. One interesting difference is that for the young cohort, local community employment and local industry mix employment growth were significant in increasing population growth. For the *young adult* cohort, local industry mix employment growth was not significant, and local industry mix employment growth actually served to decrease population growth! For *young adults*, employment growth in the surrounding CCSs (INDMIX_EMPGROW_SURR) was significant in influencing population change.

These results appear to suggest that this cohort prefers to live in areas with lower local growth, but close to adjacent communities with high growth. This could be indicative of a preference for bedroom communities or suburban areas outlying high-growth urban areas as opposed to living right where the growth is occurring.

Other notable differences are that highly concentrated agricultural sectors do not appear to influence migration. *Young adults* do tend to migrate towards areas with fewer impoverished residents (indicated by the negative sign on PER_BEL_MEDIAN), and also tend to have higher population growth in areas with a high percentage of non-farm self-employed individuals.

5.4.3 Factors Influencing Adults

The behaviour of the *adult* cohort (individuals aged 35-50 in 1991) appears to be the most difficult to explain using the dataset compiled for this thesis. Variables measuring the distance to universities and colleges were removed from this model. The DIST_SCHOOL variable was left in because this cohort may have children, and the needs of their children may affect their location decision. Model 3 only explains 47.2 percent of the variation in population change among CCSs. This particular cohort likely has a number of different obligations, including family and jobs that impact their decisions to a higher degree than younger cohorts. A larger number of commitments are one explanation for the decrease in explanatory power.

Overall, the factors influencing the behaviour of the *adult* cohort appear to be very similar to that of the young adult cohort. The INCOME variable is negative and significant, which separates it from the younger cohorts where income was negative but insignificant. This result implies that the *adult* cohort migrated into areas with lower average incomes; one would expect people to move to communities with higher average incomes, so one explanation for this result is that amenities are being capitalized negatively into wages, as predicted in the theoretical model.

Another interesting result shown in model 3 is that the *adult* cohort had a lower population growth if the community was located within an urban area (demonstrated by the negative sign on D_CCSINCMA), but at the same time, population growth was higher for communities that were closer to urban areas (demonstrated by the negative sign on (CMA_CA_DIST). The *adult* cohort also appeared to exhibit a preference for employment growth in surrounding areas as opposed to local areas. An explanation for both of these results could be that the adult cohort, like the young *adult* cohort, prefers to live near CCSs with economic growth but not within them, perhaps preferring bedroom communities or suburbs as opposed to commercial or industrial hubs of growth. Commuting appears more important as people age, and develop lasting relationships. Perhaps one explanation is that among spouses, one spouse works nearby, while the other spouse commutes.

The *adult* cohort had similar results for the amenities variables as the younger cohorts. Though significant as a group, most of these variables were not individually significant. This could be due to multicollinearity between the amenity measures. As with the younger cohorts, adults preferred to be close to large acute hospitals, but further away from smaller facilities. Other significant results were that adults appeared to consider crime and humidity as disamenities.

5.4.4 Factors Influencing Early Retirees

Model 4 contains the results when the dependent variable is the percentage change in population for the *early retiree* cohort. These individuals were aged 50-60 in 1991, and many of the individuals in this cohort may be entering a stage of their life where their children are moving away, they are retiring from their careers, and thus becoming more mobile. They are also an age group with a lot of disposable income, so many communities may be interested in attracting these individuals. Other than the human capital grouping, all of the variable groupings were statistically significant. This cohort appears to closely follow the behaviour of the adult cohort with several exceptions.

Economic factors appear to play less of a role in the location decisions of early retirees. The Herfindahl index, which was very important in the explaining the population change for the younger cohorts, is no longer significant for this cohort. Local economic growth/local employment rates are also not a factor in explaining population growth for the *early retiree* cohort. Job growth in surrounding areas was significant, again indicating that this cohort prefers to live near but not within CCSs that are experiencing a high degree of economic success. This cohort also preferred to live in places with smaller populations, as evidenced by the negative sign on the POP_91 variable.

This cohort tended to migrate towards areas with a high percentage of both young and old people. These results combined with the results from the *young adult* and *adult* cohorts indicate that having an initial population with a high percentage of elderly people does not necessarily result in a stagnant population in the future. For aging communities, this result could be looked upon as a positive.

For the first time in these regressions, the DIST_RELIG variable, which measures the distance to the nearest religious institution, was negative and significant, indicating that the *early retiree* cohort preferred communities that were closer to churches and other religious organizations. All younger cohorts appeared to prefer living further away from churches, though the coefficients were not highly significant.

This difference may be indicative of a divide between preferences of younger and older cohorts. Prior to the examining the results, the presence of churches was considered to be the most important measurement of social capital in this study. Churches are significant for the fact that they facilitate community interaction and volunteer activity, and many are associated with community halls. It is very interesting that older cohorts tend to value being close to these institutions, while younger cohorts do not appear to consider it an important community attribute.

The *early retiree* cohort also appeared to value communities with a high percentage of home-ownership, which is another important measure of social capital. Taken together, these results seem to indicate that social capital plays a more important role in the location choices for older segments of the population than for the younger cohorts.

In terms of modern amenities, the *early retiree* cohort, like all others, preferred to be close to large acute care hospitals, but for the first time in these regressions, the DIST_ACUTE variable had a negative sign, though it was not significant. This is a shift in preferences from the younger cohorts, which actually appeared to view close proximity to smaller hospitals as a disamenity. Also notable is that the distance to a long-term care facility did not appear to impact the location choices of this cohort.

Early retirees were the only cohort to place a positive and significant value on the presence of water within or adjacent to the CCS. They also viewed humidity as a disamenity. As with other cohorts, a large number of the amenities variables were not individually significant, yet as a group they were jointly significant. Combined with the theory that amenities are capitalized into housing values and incomes, there is still a strong case to suggest that amenities are an important factor in explaining population change for the *early retiree* cohort.

5.4.5 Factors Influencing the Elderly

Model 5 examines the factors affecting population change for the retired segment of the population (aged 60+ in 1991). This model explained 83.9 percent of the variation in population change for the *elderly* cohort, making it the best of the cohort models in terms of explanatory power. The main feature that separates the elderly cohort from younger cohorts is that many of the economic factors that proved to be so important in influencing population change for the younger cohorts were no longer significant, and economic factors are negatively correlated with elderly population growth in several cases.

For example, both the local and surrounding CCS industry mix employment growth coefficients had negative signs (indicating a negative relationship between employment growth and population growth), though these coefficients were not statistically significant. The Herfindahl index was also not significant, and it is interesting to note that the coefficient size of the HERF_INDEX variable declined as older cohorts were estimated, indicating that industry concentration does have a smaller effect on population change for older cohorts.

This cohort tended to locate in areas where a higher percentage of the population had lower incomes. Also, communities with a highly concentrated agricultural and primary industry sectors tended to have lower *elderly* population growth. One explanation is that older people are selling their farms and moving away from rural areas. Surprisingly, communities closer to the national highway system tended to have lower population growth than more isolated communities, as indicated by the positive sign on DIST_NATLHWY. This segment of the population tended to have higher growth rates in communities with smaller populations, and communities that were located further away from urban centres.

All of the variable groupings were jointly significant in this model, including the largely ineffective human capital grouping. Interestingly, the elderly cohort tended to have higher population growth in areas with a high percentage of university graduates, and a lower percentage of individuals that did not graduate from high school.[33] This is different from conventional human capital theory, which has established that in general, areas with lower education levels will have higher population growth because less education means less mobility.

As with the early retiree cohort, the elderly cohort also preferred communities that were closer to religious organizations. In fact, for every 10 kilometres further away a community is from a church, elderly population growth declines by 2.4 percent *cetaris paribus*. This cohort also preferred to be nearer to CCSs with more volunteer organizations. Oddly, the elderly cohort preferred communities where a lower percentage of residents owned their own home, which is odd because home ownership is supposed to be gauge higher levels of social capital.

Medical amenities appear to be of the utmost importance to older segments of the population. The elderly cohort is the only segment where increased distance to physicians had a significant and negative impact on population growth. For every 10-kilometre increase in the distance to the nearest physician, elderly population growth is expected to decrease by 1.6 percent, *cetaris paribus*.

---

[33] University graduates themselves may be considered amenities to elderly people. Medical professionals are usually university graduates. Another explanation is that university graduates may live in areas with more amenities.

Also for the first time, the distance to smaller acute care hospitals is negatively related with population growth, and statistically significant. This means that elderly citizens tend to have higher growth rates in communities closer to any hospital, regardless of its size or capabilities! Also, this cohort, like all others, preferred to be closer to larger acute care hospitals. This is a sharp contrast to the younger cohorts who preferred to be located further away from smaller acute care facilities, but closer to the larger facilities.

Table 5.3: Cohort Analysis - Dependent Variable Varies – Total Sample

Variable Name	Model 1 Dependent Var: POPCHG_ YOUTH	Model 2 Dependent Var: POPCHG_ YOUNG_ ADULT	Model 3 Dependent Var: POPCHG_ ADULT	Model 4 Dependent Var: POPCHG_ EARLY_ RETIREE	Model 5 Dependent Var: POPCHG_ ELDERLY
C	40.593***	-19.149**	5.446	-21.811**	7.716
	(2.89)	(-2.11)	(0.754)	(-2.53)	(1.5)
AVG_VALUE_HOME	0.039***	0.038***	0.024***	0.022***	0.01
	(3.13)	(3.3)	(3.23)	(2.85)	(1.2)
DIST_NATLHWY	-0.00041	0.007*	0.003	0.005	0.006**
	(-0.081)	(1.68)	(1.28)	(1.41)	(2.19)
EMPLOYMENT_RATE	0.346***	0.066	-0.002	0.005	0.038
	(5.17)	(1.37)	(-0.054)	(0.111)	(1.26)
HERF_INDEX	-22.605***	-18.26***	-15.973***	-3.565	-5.329
	(-3.36)	(-3.21)	(-3.71)	(-0.678)	(-1.61)
INCOME	-0.0003	-0.00016	-0.00032**	-0.00057***	-0.00006
	(-1.3)	(-0.713)	(-2.46)	(-4.32)	(-0.56)
INDMIX_EMPGROW	0.124**	-0.13***	-0.006	0.029	-0.039
	(2.46)	(-2.8)	(-0.171)	(0.854)	(-1.62)
INDMIX_EMPGROW_SURR	-0.088	0.797***	0.371***	0.335***	-0.019
	(-0.713)	(6.95)	(4.94)	(3.9)	(-0.326)
PER_BEL_MEDIAN	0.146***	-0.231***	-0.113***	-0.11***	0.051***
	(2.89)	(-4.17)	(-4.54)	(-5.09)	(3.18)
PER_EMPLOY_AGRIC	-0.118*	-0.013	0.011	-0.111**	-0.104***
	(-1.86)	(-0.268)	(0.288)	(-2.49)	(-3.99)
PER_EMPLOY_PRIMARY	-0.272***	-0.205***	-0.086**	-0.157***	-0.075**
	(-4.45)	(-3.97)	(-2.26)	(-3.27)	(-2.44)
PER_SELFEMPLOY	-0.155*	0.436***	0.272***	0.473***	0.075
	(-1.84)	(4.88)	(4.54)	(5.79)	(1.54)
PER_CERTIFICATE	-0.214***	-0.062	-0.055	0.152**	0.066
	(-2.99)	(-0.715)	(-1.01)	(1.99)	(1.17)
PER_NO_HSGRAD	-0.098	-0.013	-0.044	0.114**	-0.066*
	(-1.33)	(-0.181)	(-1.04)	(2.24)	(-1.81)
PER_UNIVERSITY	0.075	-0.156	-0.074	0.095	0.152**
	(0.63)	(-1.23)	(-1.2)	(1.32)	(2.54)
PER_ABORIGINAL	0.073	0.042	0.082***	0.028	-0.115***
	(1.36)	(0.779)	(2.76)	(1)	(-5.57)
PER_IMMIG_10	0.295**	-0.179	-0.125	-0.116	-0.165**
	(2.14)	(-1.49)	(-1.52)	(-1.53)	(-2.31)
PER_OLD	-0.567***	0.365***	0.388***	0.871***	0.041
	(-5.95)	(4.21)	(5.57)	(12.2)	(0.745)
PER_YOUNG	-1.142***	0.618***	0.008	0.687***	0.2***
	(-7.97)	(4.46)	(0.101)	(8.61)	(3.07)
D_ATLANTIC	-0.573	-1.365	-0.338	1.24	-1.303**
	(-0.354)	(-1.05)	(-0.462)	(1.27)	(-1.97)
D_NORTHERN	2.608	-3.706***	-1.219	-2.49***	0.222
	(1.46)	(-2.74)	(-1.41)	(-2.66)	(0.276)
D_ONTARIO	-2.92**	-3.83***	-1.328*	-0.948	-1.492***
	(-2.54)	(-3.92)	(-2.1)	(-1.28)	(-2.85)
D_QUEBEC	-3.59**	-0.204	-1.072	1.303	-2.841***
	(-2.51)	(-0.142)	(-1.47)	(1.41)	(-4.4)
DIST_RELIG	0.016	0.031	0.013	-0.036***	-0.024**
	(0.788)	(1.18)	(1.01)	(-2.77)	(-2.32)
PER_OWN_HOME	-0.102**	0.112*	0.043*	0.042***	-0.022*
	(-2.38)	(2.22)	(2.29)	(3.34)	(-1.85)
PER_SAMEADDRESS	-0.148	-0.033	-0.063	-0.096	-0.035
	(-1.54)	(-0.931)	(-1.35)	(-1.61)	(-1.32)
PERCAP_VOL	0.753	-3.2***	-0.016	-0.382	0.906**
	(0.761)	(-3.5)	(-0.032)	(-0.631)	(2.29)
PER_VOTE	0.046	-0.033	0.039	-0.005	0.006
	(1)	(-0.78)	(1.39)	(-0.168)	(0.247)
D_CCSINCMA	-0.169	-1.018*	-0.739*	-0.433	0.17
	(-0.28)	(-1.88)	(-2.15)	(-0.874)	(0.55)
POP91_100K	0.00015	0.00039	0.00004	-0.00019	-0.00001
	(0.588)	(1.61)	(0.239)	(-1.12)	(-0.104)
POP_91	0.00093	-0.002	-0.001	-0.002*	-0.002**
	(0.5)	(-1.16)	(-0.891)	(-1.95)	(-2.02)
CMA_CA_DIST	-0.012	-0.008	-0.01**	0.006	0.006
	(-1.6)	(-1.28)	(-2.54)	(0.872)	(1.6)

Table 5.3 Continued on next page

Table 5.3: Cohort Analysis - Dependent Variable Varies – Total Sample Continued					
Variable Name	Model 1	Model 2	Model 3	Model 4	Model 5
	Dependent Var: POPCHG_ YOUTH	Dependent Var: POPCHG_ YOUNG_ ADULT	Dependent Var: POPCHG_ ADULT	Dependent Var: POPCHG_ EARLY_ RETIREE	Dependent Var: POPCHG_ ELDERLY
CRIME_PROPERTY_RATE	0.0006*	0.00004	-0.00006	-0.00003	0
	(2.18)	(0.178)	(-0.46)	(-0.224)	(0.014)
CRIME_VIOLENT_RATE	-0.003***	-0.002**	-0.001**	-0.001**	-0.00066*
	(-3.57)	(-2.02)	(-2.24)	(-2.36)	(-1.8)
DIST_ACUTE	0.024*	0.04***	0.014*	-0.008	-0.018**
	(1.82)	(2.95)	(1.81)	(-0.771)	(-2.42)
DIST_LACUTE	-0.011*	-0.015***	-0.007**	-0.011***	-0.008***
	(-1.75)	(-3.29)	(-2.48)	(-3.21)	(-3.05)
DIST_COLLEGE	0.011	0.004			
	(1.63)	(0.653)			
DIST_GLF	0.014	-0.02	0.00036	0.016	0.009
	(1.06)	(-1.58)	(0.055)	(1.5)	(1.42)
DIST_PHYS	0.002	0.011	-0.002	0.009	-0.016**
	(0.16)	(0.73)	(-0.244)	(0.739)	(-2.35)
DIST_POLICE	-0.047***	-0.002	-0.00047	0.017	0.013*
	(-3.04)	(-0.122)	(-0.053)	(1.47)	(1.66)
DIST_SCHOOL	-0.008	-0.027	-0.009		
	(-0.321)	(-1.08)	(-0.702)		
DIST_SKI	0.014*	-0.00025	0.002	-0.01**	0.00087
	(1.9)	(-0.037)	(0.426)	(-2.39)	(0.271)
DIST_UNIV	-0.001	0.0006			
	(-0.254)	(0.145)			
DIST_LTERM				-0.002	-0.008
				(-0.201)	(-1.11)
COVER_FOREST	0.002	-0.012*	-0.003	-0.008	-0.007*
	(0.27)	(-1.69)	(-0.627)	(-1.55)	(-1.9)
D_ANYWATER	-1.308***	-0.537	0.038	0.749**	-0.236
	(-2.7)	(-0.999)	(0.127)	(2.03)	(-0.885)
ELEV_STD_DEV	0.001	0.004	0.00063	0.003	0.002
	(0.369)	(1.23)	(0.334)	(0.958)	(1)
WEATH_AVE_PRECIP	-0.002*	0.002**	0.00029	0.00016	-0.00059
	(-1.75)	(2.31)	(0.616)	(0.292)	(-1.48)
WEATH_AVE_SNOW	-0.002	-0.004	-0.002	0.00086	-0.00095
	(-0.588)	(-1.16)	(-0.813)	(0.321)	(-0.67)
WEATH_JAN_SUNSHINE	0.013	0.039***	0.004	0.007	-0.003
	(0.864)	(2.62)	(0.479)	(0.609)	(-0.482)
WEATH_JAN_TEMP	0.06	-0.155*	-0.064	0.019	0.03
	(0.612)	(-1.77)	(-1.27)	(0.306)	(0.648)
WEATH_JULYHUMID	-0.016	-0.1***	-0.044***	-0.029*	-0.02**
	(-0.776)	(-5.28)	(-4.19)	(-2.2)	(-2.07)
F-Econ	38.6***	35.5***	39.5***	32.2***	14.7***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-Human Capital	8.4***	2.4*	1.5	4.3***	15.3***
	(0.000)	(0.064)	(0.218)	(0.005)	(0.000)
F-Demographic	82.8***	28.1***	55***	75***	17.7***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-Regional	10.1***	15.2***	4***	5.9***	9***
	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)
F-Social Capital	59.23***	30.9***	21.5***	23.5***	11***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-Urban Scale	2.06*	4.7***	4.22.***	2.72**	3.9***
	(0.083)	(0.000)	(0.002)	(0.028)	(0.003)
F-Modern Amenities	6.7***	4.4***	3.7***	5.3***	5***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-Natural Amenities	3.9***	13.6***	5***	3***	3.4***
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)
R^2	0.738	0.611	0.472	0.614	0.839
N	2402	2402	2402	2402	2402

This regression is weighted by the log of the initial-year CCS population. The territories and CCSs with a population of less than 250 are excluded from this regression. T-Statistics and f-statistic p-values are reported in parenthesis. Model uses White's Heteroskedasticity-Consistent Standard Errors and Covariance. Regression is weighted by log(POP_91)., See Appendix A for variable definitions, and table 4.3 for theoretical expectations of directional influence. *, **, and *** denote significance at 10%, 5%, and 1% levels respectively.

5.5 How Different Factors Affect Different Age Groups

This section summarizes the results of the cohort analysis, and to addresses important trends and differences between the factors impacting different age groups.

Every cohort with the exception of the *elderly* tended to have higher population growth in communities with higher housing values. Also, the *adult* and *early retiree* groupings tended to have higher growth in areas with lower incomes. All other cohorts had the same signs on INCOME, though the results were not statistically significant. One possible explanation for this trend is the one put forth in the theoretical framework of this thesis: amenities have been capitalized into income and housing values over time. A favourable bundle of amenities and other local factors leads to subsequent population growth, which subsequently leads to further increases in housing values. Another explanation is that increasing incomes across Canada and changing tastes favour amenities, and people are moving to places that have had high levels of amenities.

In terms of the specific amenity variables included in the model, only a small proportion of these variables were significant for any given cohort, and the particular variables that were individually significant was variable between cohorts. However, at the same time, results of the grouped F-tests indicated that groupings of social capital, modern amenity, and natural amenity variables were jointly significant for every single population segment. In other words, as a group, these variables increased the explanatory power of the model. A possible explanation for this discrepancy is that there is some degree of collinearity, either amongst the amenity variables themselves, or amongst the amenity variables and other non-amenity variables. Another possible explanation is that measurement error in the amenity variables may be biasing the coefficient standard errors upward, thus making the individual coefficients appear to be insignificant when in fact they would be significant if measured correctly.

Proximity to large acute care hospitals always increased population growth. However, close access to "just any" hospital was not always beneficial for communities, as younger cohorts viewed this as a disamenity.[34] Also, the *elderly* were the only population cohort that tended to have higher population growth in communities located closer to doctor's offices. These results lend strong support to the notion that older people consider medical accessibility to be more important than younger residents. It also brings into question the effectiveness of small hospitals as engines of economic growth.

Other interesting results include the fact that all cohorts tended to exhibit lower population growth rates in regions with high violent crime rates, while high property crime rates did not seem to inhibit growth. Also, the educational amenity variables included in the study were never significant, even for the younger population cohorts. The coefficients for these variables were simply not statistically significant, implying that they may have no effect on population growth, though the sign of the DIST_SCHOOL variable always indicated that shorter travel times to schools increased growth for the *youth*, *young adult*, and *adult* cohorts.

Although the social capital variables were always significant as a group, individual coefficients did not always comply with a priori expectations. There may be some deficiencies with variable choices in this case, or it could simply be that different cohorts value different aspects of social capital more than others. For example, the youth and

[34] It should be noted that the hospital variables were calculated from 2002 data, and therefore this result is not simply an effect of residents anticipating eventual hospital closures.

elderly cohorts did not tend to have higher growth rates in areas with more home ownership, while the other cohorts did appear to value home ownership. One interesting result was that the presence of churches only appeared to be an amenity for older generations, and there is no indication that younger people prefer living near religious institutions. People aged 50+ did have higher growth rates in communities with churches.

Economic opportunity variables tended to be most important for younger people, and gradually became insignificant for older citizens. Also, there appears to be an interesting discrepancy between whether economic opportunity within the home CCSs increases population growth, or whether opportunities in surrounding CCSs are more important. The *youth* cohort definitely valued a high degree of economic opportunity within CCSs, while older cohorts tended to have higher population growth in communities with lower employment demand shifts locally, but higher demand for jobs in outlying communities. Perhaps young people tend to locate in close proximity to employment opportunities, while older segments of the population commute to adjacent centres of economic opportunity. Several other interesting trends that were noted are:

- The presence of highly concentrated primary industry and agricultural sectors tended to have a negative effect on population growth for most cohorts. There is something unique about these sectors that cannot simply be explained through industry concentration or employment growth variables.

- Young and elderly population segments tended to locate in areas where a larger percent of the population had low incomes.

- Human capital variables were largely ineffective in explaining community population growth. They often did not exhibit the correct direction of influence, and they were rarely significant.

- It is not necessarily true that communities with a high initial population of young people attracted more *youth*, or that the presence of more *elderly* people will automatically result in lower population growth for other cohorts. In some cases higher percentages of both young and elderly tended to result in population increases.

5.6 Addressing the Hypotheses of the Thesis

At the outset of this thesis, the stated objective was to address some very specific statements contained within the hypotheses. Though a wealth of valuable information has been generated and reported through this process, it is important to bring the discussion back to the motivation behind the study, and address the specific hypotheses of the study and determine the validity of these statements.

Is a lack of recreational amenities and educational opportunities important factors in the loss of young people from rural areas? The results are inconclusive regarding this specific hypothesis. Variables measuring access to educational amenities were not significant in any of the models, so no conclusionscan be drawn with any certainty regarding the effects of schools, universities and colleges on youth population change.

In general, economic factors appeared to play a much larger role in population growth among the *youth* and *young adult* cohorts than amenities. There is no strong evidence to suggest that recreational amenities are any more important for young people as opposed to older people. However, it should be noted that the model did not include a large number of recreational amenities; distance to golf courses, ski hills, and of course the presence of nature itself were the only recreational amenities included. It is possible that recreational amenities were capitalized into housing values, and it should be noted that young people did have higher population growth in communities with higher housing costs.

In general, amenities were jointly significant in explaining the population growth of young people, but without conclusive individual coefficient results, it is impossible to conclude that education and recreation are important.

Does the presence of hospitals and doctors influence population change, especially among older segments of the population? The results indicate that proximity to large acute care hospitals is an important factor in explaining population growth for Canadians of all ages. Communities located closer to large hospitals consistently had higher population growth rates.

However, the issue of the importance of smaller acute care facilities and doctor's offices is not as clear. Reduced distance to smaller hospitals definitely does not appear to be beneficial to communities. All age cohorts except possibly the *early retiree* and definitely the *elderly* cohorts had lower population growth in communities that were closer to smaller acute care facilities. This raises interesting policy questions regarding the perceived importance of smaller hospitals to community vitality. The *elderly* were the only cohort that appeared to value close access to physicians.

In conclusion, proximity to hospitals and doctors definitely appeared to be more important to older cohorts, as suggested by the hypothesis. This is not to say that heath care is not important to younger people. Younger people clearly value being close to larger health facilities. However, as the *elderly* cohort appeared to be the only one exhibit higher population growth in communities closer to smaller hospitals and physicians, it is evident that they find these services more important.

Can data generated through GIS databases and techniques produce statistically significant variables that comply with economic theory? The results ultimately have shown that many of the GIS-produced variables were significant, both individually and as part of a group. The vast majority of these variables did exhibit the direction of influence predicted by theory, so there appears to be some merit to utilizing GIS modifications to create datasets for this type of study. At the same time, researchers must be cautious to test their variables before use, as several different modifications of the GIS variables that did not utilize distance measurements were experimented with, and these variables did not function nearly as well. However, high correlations between the "distance" variables did prove to be problematic. Although many distance variables are statistically significant, there are perhaps superior specifications for GIS variables that have yet to be discovered.

Does the presence of amenities and social capital have significant and positive effects on population migration? The results of the grouped F-tests provide strong evidence that the amenity and social capital groupings of variables are significant in influencing population change. However, results for the individual coefficients provide somewhat weaker results. The results indicate that while some social capital indicators can account for increased population, social capital does not appear to play a positive effect on population change in all cases, or for all population segments.

Regarding modern and natural amenities, it is clear that as a group, they play an important role in migration. Property crime, and distance to smaller acute care hospitals, colleges, and ski facilities actually had a negative effect on population growth in some instances. The majority of the variables considered amenities tended to have a positive effect on population change, but as noted, several had a negative effect. Natural amenity variables were not typically individually significant, though many exhibited the correct sign. In general, the assertion can be made that amenities and social capital do play a significant and positive role, though the results vary depending upon the population cohort being examined.

5.7 Chapter Summary

In this chapter, the results of the econometric analysis examining the factors affecting population change in communities across Canada were presented and critically analyzed. The results were then used to address the hypothesis of the thesis in section 5.5. The results highlight that fact that amenities and social capital did play an important role in the location choices of Canadian citizens between 1991 and 2001, though they did so in combination with other socio-economic factors. It is apparent that different population segments placed more importance on different local factors. The next chapter concludes the thesis by briefly summarizing the results, and discussing the policy implications of these results.

Chapter 6: Conclusions

6.0 Introduction

The purpose of this chapter is to present conclusions based upon the results and analysis. First, a summary of the results is provided along with any relevant policy implications that can be drawn from this research. This is followed by a brief discussion of the limitations of the study. The chapter concludes by making several recommendations for further research.

6.1 Conclusions and Policy Implications

6.1.1 Amenities and Social Capital

The null hypothesis of this thesis was that in general, the presence of amenities and social capital do not affect community population change. This hypothesis was rejected in favour of the alternative hypothesis that amenities and social capital have a significant and positive impact, because as a group, natural amenities, modern amenities, and social capital variables were found to be statistically significant factors in explaining community success in terms of population change. However, some types of amenities were far more influential in explaining population change than others, and different age groups were found to favour different bundles of local amenities.

Some of the most interesting results relate to health care amenities and how different population segments value these institutions. This study found that all population segments valued close access to larger acute care hospitals, while close access to smaller acute care hospitals only increased population among older segments of the population. Being located close to doctor's offices was only important for elderly segments of the population. The results of this study provide strong evidence that smaller hospitals and doctor's offices should not be considered engines of population growth unless a community aspires to be a retirement destination. Even so, any citizens a community can attract are beneficial to the community in terms of reaching a population threshold. However, all segments of the population tended to grow in communities that were close to larger acute care hospitals, so there may be a case for the role of large hospitals in influencing economic development.

Other amenity measures were important in influencing population growth among all segments of the population. For example, communities with lower crime rates had lower population growth, while the results show that increased rates of property crime did not appear reduce population growth. Also, communities with nearby police stations generally had higher population growth, especially among younger people. Clearly, communities can realize benefits from controlling more serious crimes. Natural amenities such as mountains and pleasant weather, and water did not consistently result in higher community population growth.

Interestingly, community proximity to amenities such as golf courses and ski facilities were not typically statistically significant in terms of influencing population change. The importance of educational institutions to communities was also somewhat ambiguous. Communities that were closer to colleges actually tended to have lower population growth. The results for universities and high schools/elementary schools showed close access increased population growth, though these results were not statistically significant. Therefore, no conclusions can be drawn from these results. One important statistic is that only 30 communities across Canada were located more than 50 kilometres from the nearest school, and the average distance to the nearest school was only 8.6 kilometres. This indicates access to schools is fairly homogeneous across the country, and may explain why distance to schools isn't very important in explaining population change. Technically, virtually all communities have relatively good access to schools.

Housing values and incomes were very important in explaining population growth, but not for the reasons one might expect. The results establish that in Canada, people tended to move where it's expensive to live, and where the average citizen earns less. In other words, population growth rates are higher in communities with high housing costs and lower average incomes. On the surface, Canadians appear to be exhibiting irrational behaviour until you contemplate why this is happening. People are not moving to these types of communities because they dislike money; they are moving there because over time, *amenities* have likely been capitalized into housing prices and incomes.

For example, in remote northern regions of Canada, employers are forced to pay high wages because the communities themselves lack amenities, and are therefore not desirable places to reside. High wages are the only means businesses have to attract skilled labour. Communities with favourable amenities tend to have higher housing costs because people wish to live there. A favourable bundle of amenities and other local factors leads to subsequent population growth, which leads to increases in housing values. Another explanation is that general increases in incomes across Canada and changing tastes may also favour amenities, and as a result, people are moving to places that have historically had high levels of amenities. These results form much of the basis for the argument that amenities are just as important as financial factors when residents form their location decisions.

Community trust and cohesion (also known as “social capital”) also appears to play a key role in influencing population change among older segments for the population. Close access to churches and volunteer organizations only appeared to influence older cohorts, and it is interesting that people under 50 years of age do not appear to be attracted by these types of community amenities. Communities with higher voter turnout rates did not appear to have higher population growth among any segment of the population, which contradicts previous findings in this area. Communities with a high proportion of home-ownership (and by extension, more deep-rooted communities) tended to have higher growth among middle-aged people, but young people and the elderly had higher growth in areas where rental dwellings were more common.

6.1.2 Economic and Other Factors

Although amenities were found to be important influencers of community population change, they are certainly not the only factors. The results indicate that economic factors are very important in explaining population change in addition to demographic features and the proximity of the community to larger population centres. Important economic factors include diversified economies, and the presence and growth of jobs in the area.

Diversification of the business climate in communities appears to be particularly important, as communities where employment was highly concentrated into one or two industries had lower population growth among most population segments. Increased business diversification appeared to be most important in attracting younger people, and less important for attracting and retaining older people (age 50+). Also, communities that were highly reliant on agriculture and resource extraction for employment tended to have lower population growth rates. The undesirable effects of these two sectors are not simply an industry concentration problem, as this was accounted for in the model. One explanation is that past technological change in these sectors continues to impart economic hardship on communities heavily involved in these sectors. Clearly, communities can benefit by employing strategies to attract new types of business, and to avoid reliance on any one industry – especially agriculture and other primary industries.

Employment and an increasing demand for labour are very important in terms in attracting residents. However, the results show that a community doesn't necessarily need to possess the hot job market *themselves*. They simply need to be located *near* other communities that have a shortage of labour, or an abundance of available jobs. In fact, every population segment other than young people (aged 5-20) had lower population growth in areas with flourishing job markets.

Young people tended to locate in high growth areas, but other segments of the population preferred to be located not within, but near these communities. One explanation is that middle-aged individuals favour commuting to their jobs from bedroom communities, or suburbs. Population growth among the elderly segment of the population was not affected at all by employment demand indicators. These results indicate that cities, towns, and rural municipalities should work together to foster business growth, as the benefits derived from attracting new jobs do in fact appear to be regional, not just local.

6.1.3 A Future for Rural Communities?

Though rural areas have received a great deal of attention in terms of declining populations, the results of this study indicate that urban centres do not always have higher population growth. In fact, 75 percent of urban communities lost youth population between 1991 and 2001.[35] Attracting and retaining residents is therefore not just a problem for rural areas. The results show that older segments of the population (50+) tended to avoid highly populated communities, but in general, residents preferred to be located *near* large population centres, but not necessarily *inside* these highly populated communities. The results support the notion that middle-aged adults are returning to rural areas, perhaps after moving into cities when they were younger. Certainly, areas located near large centres appeared to have an advantage in attracting population, but cities are definitely not guaranteed to grow by nature, and must also take steps to diversify and expand employment opportunities, just as rural regions must strive for this.

One myth that can be broken by this study is that communities with a high proportion of elderly people are destined for economic failure. Middle-aged people tended to move to areas with a high proportion of elderly people, though young people avoid these areas. Also, the results support the idea that elderly people are not remaining in rural areas; many are moving into urban areas, presumably to enjoy the amenities and health benefits these communities provide.

In developing strategies to retain and attract residents, communities should consider which age groups they wish to attract, and tailor strategies to meet these goals.

---

[35] Across Canada, urban CCSs had an average youth population growth rate of –4.83 percent, while rural CCSs had an average youth population growth rate of –13.36 percent. The majority of both rural and urban CCSs are losing youth, and only a minority of CCSs actually had increasing youth population between 1991 and 2001.

Communities wishing to attract young people need a different approach than if they are trying to attract or retain retired people. Communities wishing to attract or retain young people should be less concerned about the availability of a local hospital or doctors, and these communities need to have an abundance of employment opportunities within their community. It is also advantageous for communities wishing to attract young people to have higher populations and be located closer to urban centres.

Communities wishing to attract middle-aged individuals should be located near employment opportunities, but these opportunities need not be within the community. Middle-aged people appear to prefer communities with smaller populations that are located near larger centres. Communities wishing to attract retired or retain people must have nearby hospitals and doctors, but do not need to have an abundance of jobs.

People of all ages are not necessarily abandoning “rural areas” for “urban areas.” The results of this study confirm that the problem is much more complicated than that. Cities are not guaranteed to experience growth; favourable amenities and economic conditions are important for every community, and population growth (or at least avoiding future losses) is achievable for both rural and urban communities. People are not just interested in living where it is cheap to live, and where they can earn a high wage. In fact, the results show the opposite trend. People value excellent and stable employment prospects, and the chance to live in a nice place.

More often than not, urban centres do feature a favourable mix of amenities and economic factors that people value, and that is why people move there. This study has demonstrated that simply being near other people is a relatively minor factor in influencing population change. Certainly being located near large population centres is an advantage, yet there are many other ways to influence growth.

Investment in expensive infrastructure is not necessarily the best way to induce growth in smaller rural communities. It appears that the best potential for future growth lies with communities located near urban centres, and all of the amenities, services, and jobs that the urban centres provide. For communities not lucky enough to be near urban centres, the most cost-effective solution may be to work together with surrounding communities and consolidate resources to develop "hubs" of employment and amenities that many different communities can draw upon. This way, many communities can draw upon employment and resources but are not solely responsible for funding their development and maintenance.

6.2 Limitations of the Study

There are several limitations of this study. One weakness is that spatial autocorrelation, or spatial dependence may be biasing the results. The empirical model also assumed that the explanatory variables all exhibit a linear relationship with the dependent variable (population change). This assumption was made to simplify the model, yet it is likely that a number of the variables may actually have a non-linear relationship.

Some of the variables measured by distance are highly correlated, and as such, collinearity between these variables makes it difficult to disentangle the influence of individual explanatory variables. Finally, there are undoubtedly numerous missing explanatory variables that perhaps could help explain variations in population change among communities. It was impossible to account for the impact of personal relationships, family issues, and individual situations. Unfortunately, given that individual-level data was not available, it is impossible to gather information on these types of variables. There may also be missing amenity variables, as data on past modern amenities was not readily available.

6.3 Recommendations for Further Research

There are many different opportunities for further research in the areas of community population change, amenities, and regional economic development. A study similar to this one could be recreated at a future data as more data sources and different GIS techniques become available to measure local characteristics. Utilizing population-weighted CCS centroids instead of the simple geographic centroids utilized in this study is just one example of a new GIS process that has resulted in more accurate variable measurements. As time goes by, coordinate data pinpointing the exact locations of more amenities will become more accurate and accessible, and databases of the historical locations of points of interest will begin to accumulate. Eventually, researchers will not simply be limited to cross-sectional data and large panel datasets measuring changes over time will be available. The 2006 census of population will also provide new opportunities for research.

The dataset assembled for the purposes of this study is not only valuable for examining population change; many different questions in regional economic development could potentially be examined. For example, researchers could look at the factors underlying employment growth, industry diversification, or housing values. Future studies could also utilize spatial econometrics to provide a different empirical specification and eliminate spatial autocorrelation.

Finally, the possibility of using census tract data or some finer level of geography to examine urban areas and their access to amenities and other local factors could be examined. The level of geography utilized in this thesis was not suitable to compare the advantages and amenities of different metropolitan areas.

References

Beeson, P.E., R.W. Eberts. 1989. Identifying Productivity and Amenity Effects in Interurban Wage Differentials. *The Review of Economics and Statistics*. 71(3): 443-452.

Binger, B.R., E. Hoffman. 1998. *Microeconomics With Calculus.* Second Edition. Addison-Wesley Educational Publishers Inc. Reading, Massachusetts.

Blanchard, O., L. Katz. 1992. Regional Evolutions. Brookings Papers on Economic Activity. 1:1-61.

Bloomquist, G.C., M.C. Berger, J.P Hoehn. 1988. New Estimates of Quality of Life in Urban Areas. The American Economic Review. 78(1): 89-107.

Bollman, R.D. 1999. Working Paper #39 Human Capital and Rural Development: What are the Linkages?. Agriculture Division, Statistics Canada. Catalogue no. 21-601-MIE99039.

Brewin, D.G. 2004. *Three Essays in Regional Economics That Consider the Importance of Space, Agglomeration, and Income.* Unpublished PhD Dissertation. **Pennsylvania State University**

Bruekner, J.K., J.F. Thisse, Y. Zenou. 1999. Why is Central Paris Rich and Downtown Detroit Poor? An Amenity-Based Theory. *European Economic Review.* 43: 91-107.

Bumsted, J.M. 1992. The Peoples of Canada A Post ~ Confederation History. Oxford University Press. Toronto.

Burnet, J.R., H. Palmer. 1988. “Coming Canadians” An Introduction to a History of Canada’s Peoples. McClelland and Stewart. Toronto.

Cebula, R.J., R.K. Vedder. 1973. A Note on Migration, Economic Opportunity, and the Quality of Life. *Journal of Regional Science.* 13(2): 205-211.

Clark, D.E., T.A. Knapp. 1996. Personal and Location-Specific Characteristics and Elderly Interstate Migration. Growth and Change. 27(3): 327-354.

Courchene, T.J. 1970. Interprovincial Migration and Economic Adjustment. The Canadian Journal of Economics. 3(4): 550-576.

Creighton, D. 1970. Canada's First Century: 1867-1967. Toronto: Macmillan of Canada.

DaVanzo, J. 1978. Does Unemployment Affect Migration? Evidence from Micro Data. *The Review of Economics and Statistics*. 60(4): 504-514.

Day, K.M., S.L. Winer. 2001. *Policy-induced Migration in Canada: An Empirical Study.* Applied Research Branch, Strategic Policy, Human Resources Development Canada. Paper W-01-2E. November.

Day, K.M. 1992. Interprovincial Migration and local Public Goods. The Canadian Journal of Economics. 25(1): 123-144

Deller, S.C., T. Tsung-Hsiu, D.W. Marcouiller, and D.B.K. English. 2001. The Role of Amenities and Quality of Life in Rural Economic Growth. *American Journal of Agricultural Economics*. 83(2): 352-365.

Duchesne, D., F. Nault, H. Gilmour, R. Wilkins. 1999. *Vital Statistics Compendium, 1996*. Statistics Canada. Catalogue no. 84-214-XPE. Ottawa.

Dickie, M., S. Gerking. 1998. Interregional Wage Disparities, Relocation Costs, and Labor Mobility in Canada. *Journal of Regional Science.* 38(1): 61-87.

Du Plessis, V., R. Beshiri, R.D. Bollman, H. Clemenson. 2002. *Agriculture and Rural Working Paper Series Working Paper No.61 : Definitions of "Rural".* Statistics Canada, Agriculture Division. Ottawa. December.

Durlauf, S.N. 2002. On the Empirics of Social Capital. *The Economic Journal.* 112: F459-F479.

Getz, M., Y.C. Huang. 1978. Consumer Revealed Preference for Environmental Goods. *The Review of Economics and Statistics.* 60(3) :449-458.

Glaeser, E.L., D. Laibson, B. Sacerdote. 2002. An Economic approach to social capital. *The Economic Journal*. 112: F437-F458.

Goetz, S.J. 1999. Migration and Local Labor Markets. *The Web Book of Regional Science.* Ed. Scott Loveridge. Morgantown: Regional Research Institute, West Virginia University. Accessed online at: http://www.rri.wvu.edu/WebBook/Goetz/contents.htm.

Greenwood, M.J., G.L. Hunt. 1989. Jobs Versus Amenities in the Analysis of Metropolitan Migration. *Journal of Urban Economics.* 25: 1-16.

Graves, P.E. 1980. Migration and Climate. *Journal of Regional Science.* 20(2): 227-237.

Graves, P.E., P.D. Linneman. 1979. Household Migration: Theoretical and Empirical Results. *Journal of Urban Economics*. 6(3): 383-404.

Guiso, L., P. Sapienza, L. Zingales. 2004. The Role of Social Capital in Financial Development. *American Economic Review.* 94(3):526-556.

Gujarati, D.N. 2003. *Basic Econometrics*. Fourth Edition. McGraw-Hill, Inc. New York.

Herzog, H.W., A.M. Schlottmann. 1993. Valuing Amenities and Disamenities of Urban Scale: Can Bigger be Better?. *Journal of Regional Science.* 33(2): 145-165.

Huang, T.L., P.F. Orazem, D. Wohlgemuth. 2002. Rural population growth, 1950-1990 : The Roles of Human Capital, Industry Structure, and Government Policy. *American Journal of Agricultural Economics*. 84(3): 615-627.

Industry Canada. 2005. *Strategis – Trade Data online.* Accessed online at http://strategis.ic.gc.ca/.

Islam, M.H. 1990. Self Selection and Intermunicipal migration in Canada. *Regional Science and Urban Economics*. 20:459-472.

Knapp, T.A., P.E. Graves. 1989. On the Role of Amenities in Models of Migration and Regional Development. *Journal of Regional Science*. 29(1): 71-87.

Knapp, T.A., N.E. White, D.E. Clark. 2001. A Nested Logit Approach to Household Mobility. *Journal of Regional Science*. 41(1): 1-22

Liu, L., J. Hader, B. Brossart, R. White, S. Lewis. 2001. Impact of rural hospital closures in Saskatchewan, Canada. *Social Science and Medicine*. 52: 1793-1804.

Natural Resources Canada. 2004. *Important Facts on Canada's Natural Resources (as of November, 2004.* Accessed online at: www.nrcan.gc.ca/statistics.

Newbold, K.B. Income, Self Selection, and Return and Onward Interprovincial Migration in Canada. *Environment and Planning*. 28: 1019-1034.

Ng, E., Wilkins, R. Perras, A. 1993. How Far Is It to the Nearest Hospital? Calculating Distances Using the Statistics Canada Postal Code Conversion File. *Health Reports*. Statistics Canada catalogue number 82-003. 5(2): 179-188.

Oxford University Press. 2004. *Oxford English Dictionary*. Second Edition. Oxford University Press. Online.

Partridge, M.D., D.S. Rickman. 1999. A Note on the Benefits to Current Residents of State Employment Growth: Is there an Industry Mix Effect on Migration?. *Journal of Regional Science*. 39(1): 167-181

Partridge, M.D., D.S. Rickman. 2003a. Do We Know Economic Development When We See It?. *The Review of Regional Studies*. 33(1): 17-39.

Partridge, M.D., D.S. Rickman. 2003b. The Waxing and Waning of Regional Economies: The Chicken-Egg Question of Jobs Versus People. *Journal of Urban Economics.* 53(1): 76-97.

Pindyck, R.S., D.L. Rubinfeld. 1998. *Microeconomics.* Fourth Edition. Prentice-Hall, Inc. Upper Saddle River, New Jersey.

Pissarides, C.A., I. McMaster. 1990. Regional Migration, Wages and Unemployment: Empirical Evidence and Implications for Policy. *Oxford Economic Papers.* 42(4): 812-831.

Porell, F.W. 1982. Intermetropolitan Migration and Quality of Life. *Journal of Regional Science.* 22(2) :137-158.

R.A. Malatest & Associates Ltd. 2002. Rural Youth Study, Phase II, Rural Youth Migration: Exploring the Reality Behind The Myths. Retrieved from the World Wide Web on August 16, 2004 at: http://www.rural.gc.ca/researchreports/youth2002/cover_e.phtml.

Rappaport, J. 2004a. *Moving to Nice Weather.* Working Paper. Research Division, Federal Reserve Bank of Kansas City.

Rappaport, J. 2004b. Why are Population Flows So Persistent?. *Journal of Urban Economics.* 56:554-580.

Roback, J. 1988. Wages, Rents, and Amenities: Differences among Workers and Regions. *Economic Inquiry*. 26(1): 23-41

Roback, J. 1982. Wages, Rents, and the Quality of Life. *The Journal of Political Economy*. 90(6): 1257-1278.

Rosenbluth, G. 1996. Interprovncial Migration and the Efficacy of Provincial Job Creation Policies. *Canadian Business Economics*. 4(2): 22-35.

Rupasingha, A., S.J. Goetz, D. Freshwater. 2000. Social Capital and Economic Growth: A County-Level Analysis. *Journal of Agricultural and Applied Economics*. 32(3): 565-572.

Rupasingha, A., S.J. Goetz, D. Freshwater. 2002. Social and Institutional Factors as Determinants of Economic Growth: Evidence from the United States Counties. *Papers in Regional Science*. 81: 139-155.

Shearmur, R., M. Polèse. 2004. *Do Local Factors Explain Local Employment Growth?: Evidence from Canada, 1971-2001.* INRS-Urbanisation Culture et Societe

Stabler, J.C., M.R. Olfert. 1996. *The Changing Role of Rural Communities in an Urbanizing World* - S*askatchewan – An Update to 1995.* Canadian Plains Research Center. Regina.

Stabler, J.C., M.R. Olfert. 2002. *Saskatchewan's Communities in the 21st Century – From Places to Regions.* Canadian Plains Research Center. Regina.

Statistics Canada. 2005. *History of the Census of Canada.* Accessed online at: http://www12.statcan.ca/english/census01/Info/history.cfm

Statistics Canada. 1999. *1996 Census Dictionary – Final Edition*. Catalogue No. 92-351. Accessed online at: www.statcan.ca/english/IPS/Data/92-351-UPE.htm.

Statistics Canada. 2004a. CANSIM II database: Canadian socio-economic information management system. Table 510005. *Estimates of population, Canada, Provinces and Territories*. Ottawa.

Statistics Canada. 2004b. Census of population data set. Ottawa.

Statistics Canada. 1997. *Births and Deaths 1995.* Catalogue No. 84-210-XMB. Ottawa.

Transport Canada. 2004. *National Highway System.* Accessed online at: http://www.tc.gc.ca/mediaroom/backgrounders/b04-R003e.htm.

Wall, H.J. 2001. Voting With Your Feet in the United Kingdom: Using Cross-Migration Rates to Estimate Relative Living Standards. *Papers in Regional Science.* 80: 1-23

Wilkins, R. 1993. Use of Postal Codes and Addresses in the Analysis of Health Data. *Health Reports*. Statistics Canada catalogue number 82-003. 5(2): 157-177.

Wojan, T.R., D.A. McGranahan. 2004. *Recasting the Creative Class to Identify Rural Potential.* Presented at the 51st Annual Meetings of the North American Regional Science Council, Seattle, WA, November 11-14.

Appendix A: Variable Definition and Data Sources

Table A-1: Variable Definition and Data Sources

Variable Name	Description	Source[1]
Dependent Variables		
POPCHG_TOT	Percentage change in the total population accounting for mortality between 1991 and 2001	CoP, Auth
POPCHG_YOUTH	Percentage change in young people aged 5-15 in 1991 (born 1976-1986) accounting for mortality between 1991 and 2001	CoP, Auth
POPCHG_YOUNG_ADULT	Percentage change in young adults aged 20-35 in 1991 (born 1956-1971) accounting for mortality between 1991 and 2001	CoP, Auth
POPCHG_ADULT	Percentage change in adults aged 35-50 in 1991 (born 1941-1956) accounting for mortality between 1991 and 2001	CoP, Auth
POPCHG_EARLY_RETIREE	Percentage change in early retirees aged 50-60 in 1991 (born 1931-1941) accounting for mortality between 1991 and 2001	CoP, Auth
POPCHG_ELDERLY	Percentage change in late retirees aged 60+ in 1991 (born before 1931) accounting for mortality between 1991 and 2001	CoP, Auth
Economic Indicator Variables		
AVG_VALUE_HOME	The 1991 average market value of dwellings in the CCS	CoP
DIST_NATLHWY	The Distance (km) between the CCS centroid and the nearest 'national highway' as defined by Transport Canada	C-RERL, Auth
EMPLOYMENT_RATE	1991 employment rate for individuals age 15+	CoP
HERF INDEX	1991 Herfindahl Industry Concentration Index at the CCS level. Calculated as: $\sum_{i=1}^{n} s_i^2$ where s is the share of employment in industry i	CoP
INCOME	The 1991 average per-capita income of individuals 15+ in the CCS	CoP
INDMIX_EMPGROW	Industry mix employment growth, calculated by multiplying each industry's national employment growth (between 1991 and 2001) by the initial period (1991) industry employment shares in each CCS	CoP, Auth
INDMIX_EMPGROW_SURR	A distance-weighted measure of employment growth between 1991 and 2001 in surrounding CCSs. Calculated by multiplying INDMIX_EMPGROW by a spatial weight matrix (W)	CoP, C-RERL, Partridge
PER_BEL_MEDIAN	The 1991 percentage of the households in each CCS that have an income below the national median level	CoP
PER_EMPLOY_AGRIC	The percentage of the workforce in each CCS that is employed in the agricultural sector	CoP
PER_EMPLOY_PRIMARY	The percentage of the workforce in each CCS that is employed in primary industry other than agriculture (natural resource extraction)	CoP
PER_SELFEMPLOY	The percentage of the workforce in each CCS this is self-employed	CoP
Human Capital		
PER_CERTIFICATE	Percentage of individuals 25-54 that have attained a post-secondary certificate or diploma	CoP
PER_NO_HSGRAD	Percentage of individuals 25-54 that did not attain a high school diploma	CoP
PER_UNIVERSITY	Percentage of individuals 25-54 that have attained a university degree	CoP
Demographic		
PER_ABORIGINAL	The 1991 percentage of the population in the CCS that is aboriginal	CoP
PER_IMMIG_10	Percentage of the population that has immigrated from outside of Canada in the last 10 years	CoP
PER_OLD	Percentage of the population aged 60+ in 1991	CoP
PER_YOUNG	Percentage of the population aged 5-20 in 1991	CoP
Regional Dummy Variables		
D_ATLANTIC	Dummy variable; 1if the CCS is in either Newfoundland, P.E.I., Nova Scotia, or New Brunswick	StatsCan, Auth
D_QUEBEC	Dummy variable; 1 if the CCS is located in Quebec	StatsCan, Auth
D_ONTARIO	Dummy variable; 1 if the CCS is located in Ontario	StatsCan, Auth
D_NORTHERN	Dummy variable; 1 if the CCS is located in a remote northern region	StatsCan, Auth
Social Capital		
DIST_RELIG	Distance to the nearest religious institution (km)	DMTI, C-RERL
PER_OWN_HOME	Percentage of Individuals living in an owned home	CoP
PER_SAMEADDRESS	Percentage of Individuals living at the same address as 5 years ago	CoP

Table A-1 continued on next page

Table A-1: Variable Definition and Data Sources Continued

Variable Name	Description	Source[1]
PERCAP_VOL	Number of volunteer organizations per 100,000 people, measured at the CD level	CBP, C-RERL
PER_VOTE	Percentage of eligible voters that voted in the 2000 federal election., measured at the Electoral District level	Ecan, C-RERL
Urban Proximity		
D_CCSINCMA	Dummy variable; 1 if the CCS is located either partially or entirely within the boundaries of a CMA or CA	StatsCan, Auth
POP_91	The total population of the CCS in 1991	CoP
POP91_100k	The total population of the CCS in within 100k of the CCS centroid.	CoP, C-RERL
CMA_CA_DIST	Distance to the centre of the nearest CMA or CA	C-RERL
Natural Amenities		
COVER_FOREST	% of the total geographic area of the CCS covered by forest	NRCan, C_RERL, Auth
D_ANYWATER	Dummy variable; 1 if the CCS is located adjacent to the coastline of an ocean or one of the great lakes OR if water area comprises >0 of the CCSs land area	C-RERL, NRCan, Auth
ELEV_STD_DEV	Standard deviation of the elevation points located within the CCS - suggests variation in terrain - mountains, hills	NRCan, C-RERL
WEATH_AVE_PRECIP	Average annual precipitation (mm) (20+ year average)	EnvCan,C-RERL
WEATH_AVE_SNOW	Average annual snowfall (mm) (20+ year average)	EnvCan,C-RERL
WEATH_JAN_SUNSHINE	Average January sunshine (hours) (20+ year average)	EnvCan,C-RERL
WEATH_JAN_TEMP	Average January temperature (degrees Celsius) (20+ year average)	EnvCan,C-RERL
WEATH_JULY_HUMID	Average July Humidity (20+ year average)	EnvCan,C-RERL
Modern Amenities		
CRIME_VIOLENT	Violent crime rate (number of violent crimes per 100,000 people)	CCJS, StatsCan, Auth
CRIME_PROPERTY	Property crime rate (number of property crimes per 100,000 people)	CCJS, StatsCan,
ACUTE[2]	Acute Care Hospitals	DMTI, C-RERL
LACUTE[2]	Large Acute Care Hospitals >200 beds	DMTI, C-RERL
COLLEGE[2]	Colleges	DMTI, C-RERL
DIST_PHYS	Distance between the CCS centroid, and the nearest CCS centroid with at least one physician (km)	CIHI, C-RERL
PERCAP_PHYS	Number of physicians per 100,000 residents	CIHI, C-RERL
D_PHYS	Dummy Variable; 1 if the CCS possesses at least one physician	CIHI, C-RER
BOWL[2]	Bowling Alleys	DMTI, C-RERL
CIN[2]	Movie Theatres	DMTI, C-RERL
GOLF[2]	Golf Courses	DMTI, C-RERL
LTERM[2]	Long-term Care Facilities	DMTI, C-RERL
POLICE[2]	Police Stations	DMTI, C-RERL
SCHOOL[2]	Schools	DMTI, C-RERL
SKI[2]	Ski Resorts	DMTI, C-RERL
UNIV[2]	Universities	DMTI, C-RERL

1. CoP: Census of population 1991, Auth: the author of this thesis, C-RERL: the Canada Rural Economy Research Lab, StatsCan: Statistics Canada, NRCan: Natural Resources Canada, EnvCan: Environment Canada, CCJS: Canadian Centre for Justice Statistics, CIHI: Canadian Institute for Health Information, DMTI: DMTI Spatial Inc.'s EPOI database, CBP: Canadian Business Patterns (Statistics Canada product), Ecan: Elections Canada.

2. Three versions of this variable were created – A distance variable, a per-capita variable (number per 100,000 citizens), and a dummy variable indicating presence. As such, the variables are prefaced by DIST_, PERCAP_, and D_ respectively when used in the study.

Appendix B: Descriptive Statistics

Table B-1: Variable Descriptive Statistics (Means) – Unweighted

Variable Name	Total Sample	Rural Sample	Urban Sample
POPCHG_TOT	3.37	2.44	6.84
POPCHG_YOUTH	-11.55	-13.36	-4.83
POPCHG_YOUNG_ADULT	3.12	2.35	6.01
POPOPCHG_ADULT	2.28	2.42	1.73
POPCHG_EARLY_RETIRE	6.17	6.50	4.94
POPCHG_ELDERLY	6.08	5.63	7.75
AVG_VALUE_HOME	79.57	69.04	118.73
DIST_NATLHWY	38.93	43.83	20.70
EMPLOYMENT_RATE	56.52	55.10	61.81
HERF_INDEX	0.19	0.19	0.17
INCOME	17515.67	16503.36	21280.53
INDMIX_EMPGROW	10.47	10.11	11.82
INDMIX_EMPGROW_SURR	10.42	10.23	11.13
PER_BEL_MEDIAN	42.33	43.81	36.85
PER_EMPLOY_AGRIC	14.06	16.75	4.06
PER_PRIMARY	2.87	3.14	1.87
PER_SELFEMPLOY	8.31	8.28	8.40
PER_CERTIFICATE	14.46	14.03	16.07
PER_NO_HSGRAD	27.90	28.67	25.02
PER_UNIVERSITY	5.35	4.48	8.58
PER_ABORIGINAL	4.09	4.24	3.50
PER_IMMIG_10	0.82	0.59	1.64
PER_OLD	16.87	17.82	13.32
PER_YOUNG	30.59	30.65	30.39
D_ATLANTIC	0.14	0.15	0.12
D_NORTHERN	0.02	0.02	0.02
D_ONTARIO	0.21	0.17	0.33
D_QUEBEC	0.42	0.43	0.37
SC_DIST_RELIG	10.02	10.79	7.13
SC_PER_OWN_HOME	67.44	65.25	75.58
SC_PER_SAMEADDRESS	63.28	66.05	53.01
SC_PERCAP_VOL	0.79	0.83	0.61
SC_ PER_VOTE	62.63	62.77	62.11
U_D_CCSINCMA	0.21	0.00	1.00
U_POP_91	11.17	2.91	41.88
U_POP91_100K	752.49	600.19	1318.89
CMA_CA_DIST	52.70	60.92	22.15
Table B-1 continued on next page			

Table B-1: Variable Descriptive Statistics (Means) – Unweighted Continued

Variable Name	Total Sample	Rural Sample	Urban Sample
A_COVER_FOREST	54.50	56.87	45.66
A_D_ANYWATER	0.70	0.67	0.83
A_ELEV_STD_DEV	50.07	49.22	53.25
A_WEATH_AVE_PRECIP	944.40	931.42	992.71
A_WEATH_AVE_SNOW	217.70	221.52	203.50
A_WEATH_JAN_SUNSHINE	90.96	91.10	90.43
A_WEATH_JAN_TEMP	-11.63	-12.09	-9.93
A_WEATH_JULYHUMID	58.47	58.23	59.40
A_CRIME_PROPERTY_RATE	3641.35	3329.99	4799.31
A_CRIME_VIOLENT_RATE	869.27	859.54	905.49
A_DIST_ACUTE	22.17	23.99	15.43
A_DIST_BOWL	26.21	29.02	15.78
A_DIST_CIN	45.14	51.76	20.51
A_DIST_COLLEGE	46.37	52.85	22.26
A_DIST_GLF	20.76	23.24	11.54
A_DIST_LACUTE	54.39	62.06	25.88
A_DIST_LTERM	19.89	21.47	14.01
A_DIST_PHYS	10.63	12.49	3.72
A_DIST_POLICE	19.80	21.70	12.72
A_DIST_SCHOOL	8.63	9.27	6.23
A_DIST_SKI	53.62	57.76	38.24
A_DIST_UNIV	86.15	95.56	51.16
PERCAP_ACUTE	0.68	0.76	0.39
PERCAP_BOWL	0.70	0.76	0.45
PERCAP_CIN	0.07	0.05	0.13
PERCAP_COLLEGE	0.16	0.15	0.21
PERCAP_GLF	1.12	1.20	0.80
PERCAP_LTERM	1.54	1.73	0.87
PERCAP_PHYS	10.37	8.78	16.29
PERCAP_POLICE	0.87	0.91	0.73
PERCAP_SCHOOL	7.54	7.98	5.91
PERCAP_SKI	0.21	0.24	0.12
PERCAP_UNIV	0.02	0.01	0.05
D_ACUTE	0.74	0.69	0.93
D_BOWL	0.79	0.75	0.94
D_CIN	0.28	0.17	0.69
D_COLLEGE	0.42	0.32	0.80
D_GLF	0.84	0.80	0.97
D_LTERM	0.89	0.86	0.97
D_PHYS	0.95	0.93	0.99
D_POLICE	0.58	0.51	0.83
D_SCHOOL	1.00	1.00	1.00
D_SKI	0.87	0.86	0.92
D_UNIV	0.13	0.05	0.43

Appendix C: Incremental Variable Addition

Table C-1 provides just one example of how incrementally adding variables to the model to assess significance can be useful. In model 1, natural amenities were the sole regressors, while modern amenities and social capital variables were added in models 2 and 3. Note that the t-statistics decrease for the natural amenities group as additional variables are added. This behaviour is not isolated, and exists for most of the variable groupings. The exception to this trend is the economic variable grouping, where most t-statistics remained relatively consistent as additional variables were added.

In the models presented below, the dependant variable is POPCHG_TOT. In model 1, six variables were significant, but if one refers to the benchmark model presented in table 5.2, only two of these natural amenity variables remain significant – WEATH_JULY_HUMIDITY and WEATH_JAN_SUNSHINE. It is interesting to note that although these regressions are missing many key variables, their explanatory power is relatively large as indicated by a high R^2 values, and that the direction of influence for most of these variables is as predicted in table 4.3.

Table C-1: Incremental addition of Variables, Dependent Variable: Percentage Change in Total Population

Variable Name	Model 1 Nat. Amen.	Model 2 Add Mod. Amen.	Model 3 Add SC
C	13.659***	12.366***	14.455***
	(7.31)	(5.75)	(3.08)
COVER_FOREST	-0.028***	-0.022***	-0.013***
	(-6.06)	(-4.73)	(-3.25)
D_ANYWATER	-0.498	-0.464	-0.428
	(-1.06)	(-0.962)	(-0.992)
ELEV_STD_DEV	0.006***	0.005*	0.003
	(2.64)	(1.75)	(1.38)
WEATH_AVE_PRECIP	-0.0001	-0.00087	-0.00037
	(-0.133)	(-1.23)	(-0.612)
WEATH_AVE_SNOW	-0.007***	-0.006**	-0.006***
	(-2.74)	(-2.34)	(-2.62)
WEATH_JAN_SUNSHINE	0.013	0.019	0.017
	(1.03)	(1.6)	(1.49)
WEATH_JAN_TEMP	0.208***	0.215***	0.135**
	(3.63)	(3.46)	(2.44)
WEATH_JULYHUMID	-0.07***	-0.058***	-0.058***
	(-3.92)	(-3.21)	(-3.79)
CRIME_PROPERTY_RATE		0.00061***	0.00023
		(4.09)	(0.885)
CRIME_VIOLENT_RATE		-0.002***	-0.002***
		(-4.13)	(-4.17)
DIST_ACUTE		0.048*	0.051***
		(2.22)	(4.44)
DIST_LACUTE		-0.004	-0.003
		(-1.08)	(-0.727)
DIST_GLF		-0.003	-0.013
		(-0.182)	(-1.29)
DIST_PHYS		-0.019*	0.014
		(-1.65)	(0.908)
DIST_POLICE		-0.013	-0.025*
		(-0.721)	(-1.86)
DIST_SCHOOL		0.018	-0.029
		(1.03)	(-1.28)
DIST_SKI		-0.014***	-0.015***
		(-2.7)	(-3.05)
DIST_UNIV		-0.00079	-0.00081
		(-0.229)	(-0.237)
DIST_RELIG			0.059***
			(2.93)
PER_OWN_HOME			0.047***
			(4.86)
PER_SAMEADDRESS			-0.126*
			(-1.69)
PERCAP_VOL			-2.16***
			(-2.71)
PER_VOTE			0.05**
			(2.09)
R^2	0.566	0.602	0.678
N	2402	2402	2402

This regression is weighted by the log of the initial-year CCS population. The territories and CCSs with a population of less than 250 are excluded from this regression. T-Statistics are reported in parenthesis. Model uses White's Heteroskedasticity-Consistent Standard Errors and Covariance. Regression is weighted by log(POP_91). See Appendix A for variable definitions, and table 4.3 for theoretical expectations of directional influence. *, **, and *** denote significance at 10%, 5%, and 1% levels respectively.

Appendix D: Weighted Vs. Unweighted Regressions

Three different possibilities were examined: utilizing no weight, utilizing the initial population as a weight, and weighting the regression by the log of the initial population. These results are shown in table D-1. In general, weighting by the log of the initial population provided approximately the same results as the unweighted model in terms of the size of the t-statistics and the coefficient signs. It is easy to see why, because when weighting by the log of population (model 2), the city of Calgary and Saskatoon receive only 1.6 and 1.4 times respectively the amount of weight in the model as the Rural Municipality of North Qu'Appelle, a small rural CCS located in Saskatchewan.

When total population is used as the weight (model 3), it drastically impacts the regression results. In this model, the cities of Calgary and Saskatoon receive 179 and 46 times the weight as North Qu'Appelle in the model. Accordingly, this heavily modifies the results, as only larger urban centres are given a great deal of weight in the regression. Very few of the coefficients that were significant in the unweighted model are significant when weighting by total population, though the direction of influence appears to remain unchanged.

Table D-1: Examining Weighted Specifications, Dependent Variable = POPCHG_TOT, Total Sample

Variable Name	Model 1 No Weight	Model 2 Weight: Log(POP)	Model 3 Weight: POP
C	-4.374 (-1.14)	5.036 (0.634)	43.44*** (3.71)
AVG_VALUE_HOME	0.029*** (4.7)	0.025*** (3.03)	0.034*** (4.00)
DIST_NATLHWY	0.00006 (0.021)	0.005* (1.77)	0.006 (1.08)
EMPLOYMENT_RATE	0.084*** (5.06)	0.124*** (3.06)	-0.024 (-0.234)
HERF_INDEX	-14.053*** (-6.37)	-18.134*** (-5.31)	-38.922*** (-3.31)
INCOME	0.00001 (0.129)	-0.00021* (-1.74)	-0.00049*** (-2.58)
INDMIX_EMPGROW	0.001 (0.077)	-0.025 (-1.03)	-0.228*** (-3.44)
INDMIX_EMPGROW_SURR	0.386*** (8.14)	0.382*** (5.03)	0.437*** (4.32)
PER_BEL_MEDIAN	-0.01 (-1.18)	-0.06*** (-2.9)	-0.122* (-1.87)
PER_EMPLOY_AGRIC	-0.045*** (-2.85)	-0.05 (-1.32)	0.131* (1.74)
PER_EMPLOY_PRIMARY	-0.139*** (-6.57)	-0.18*** (-5.16)	-0.139* (-1.88)
PER_SELFEMPLOY	0.073*** (3.26)	0.252*** (5.24)	-0.032 (-0.203)
PER_CERTIFICATE	-0.017 (-0.684)	-0.07 (-1.61)	0.025 (0.181)
PER_NO_HSGRAD	-0.012 (-0.627)	-0.043 (-1.14)	-0.055 (-0.571)
PER_UNIVERSITY	-0.012 (-0.3)	-0.089 (-1.41)	0.05 (0.514)
PER_ABORIGINAL	0.141*** (8.03)	0.086*** (2.78)	0.077 (1.08)
PER_IMMIG_10	0.098 (1.26)	-0.023 (-0.293)	-0.081 (-1.23)
PER_OLD	-0.037 (-1.23)	-0.062 (-1.1)	-0.012 (-0.134)
POP_PER_YOUNG	-0.027 (-0.725)	0.139** (2.11)	-0.076 (-0.705)
D_ATLANTIC	-0.44 (-0.595)	0.134 (0.16)	0.384 (0.252)
D_NORTHERN	-1.018 (-1.41)	-1.018 (-1.1)	-3.155* (-1.9)
D_ONTARIO	-1.242** (-2.26)	-1.558** (-2.32)	-0.453 (-0.528)
D_QUEBEC	0.039 (0.052)	-0.854 (-1.06)	-1.122 (-0.741)
DIST_RELIG	0.026** (2.44)	0.013 (0.834)	-0.007 (-0.325)
PER_OWN_HOME	0.008 (1.41)	0.028* (1.8)	0.132*** (3.08)
PER_SAMEADDRESS	-0.032 (-1.23)	-0.086 (-1.44)	-0.47*** (-9.76)
PERCAP_VOL	-0.172 (-0.543)	-0.899 (-1.61)	0.314 (0.22)
PER_VOTE	0.057*** (2.93)	0.007 (0.284)	-0.007 (-0.21)

Table D-1 continued on next page

Table D-1: Examining Weighted Specifications, Dependent Variable = POPCHG_TOT, Total Sample Continued

Variable Name	Model 1 No Weight	Model 2 Weight: Log(POP)	Model 3 Weight: POP
D_CCSINCMA	-0.049	-0.505	-1.457**
	(-0.171)	(-1.45)	(-2.39)
POP91_100K	0.00048***	0.00028	0.00000
	(3.72)	(1.58)	(-0.02)
POP_91	-0.006**	-0.001	0.00001
	(-2.15)	(-1.19)	(0.003)
CMA_CA_DIST	-0.007**	-0.007	-0.026***
	(-2.00)	(-1.53)	(-3.78)
CRIME_PROPERTY_RATE	0.00015	0.00015	-0.00014
	(1.24)	(0.973)	(-0.801)
CRIME_VIOLENT_RATE	-0.00019	-0.001***	-0.002*
	(-0.459)	(-2.73)	(-2.05)
DIST_ACUTE	0.004	0.019*	0.052***
	(0.593)	(2.32)	(2.91)
DIST_LACUTE	-0.008***	-0.011***	-0.002
	(-3.26)	(-3.2)	(-0.436)
DIST_COLLEGE	-0.001	0.009**	0.012
	(-0.366)	(2.51)	(1.47)
DIST_GLF	-0.004	-0.004	-0.021
	(-0.724)	(-0.534)	(-1.12)
DIST_PHYS	-0.004	-0.003	-0.00029
	(-0.415)	(-0.349)	(-0.017)
DIST_POLICE	-0.00096	-0.015*	-0.026
	(-0.124)	(-1.72)	(-1.32)
DIST_SCHOOL	-0.013	-0.015	0.018
	(-1.13)	(-0.955)	(0.65)
DIST_SKI	0.004	0.005	0.019*
	(1.25)	(1.05)	(2.19)
DIST_UNIV	-0.001	-0.001	0.004
	(-0.59)	(-0.47)	(0.918)
COVER_FOREST	0.004	-0.005	-0.01
	(1.44)	(-1.2)	(-1.19)
D_ANYWATER	0.136	-0.316	-1.052**
	(0.649)	(-0.925)	(-1.99)
ELEV_STD_DEV	-0.002	0.003	0.00083
	(-1.31)	(1.18)	(0.262)
WEATH_AVE_PRECIP	-0.00049	0.00025	0.002**
	(-0.685)	(0.444)	(2.18)
WEATH_AVE_SNOW	-0.00069	-0.003	0.002
	(-0.454)	(-1.39)	(0.337)
WEATH_JAN_SUNSHINE	0.008	0.017*	0.038***
	(1.24)	(1.69)	(2.99)
WEATH_JAN_TEMP	-0.112***	-0.079	-0.107
	(-2.68)	(-1.45)	(-1.42)
WEATH_JULYHUMID	-0.029***	-0.049***	-0.075***
	(-2.82)	(-3.67)	(-4.67)
R^2	0.511	0.784	0.982
N	2402	2402	2402

The territories and CCSs with a population of less than 250 are excluded from this regression. T-Statistics and f-statistic p-values are reported in parenthesis. Model uses White's Heteroskedasticity-Consistent Standard Errors and Covariance. See Appendix A for variable definitions, and table 4.3 for theoretical expectations of directional influence. *, **, and *** denote significance at 10%, 5%, and 1% levels respectively.

Appendix E: Additional Regression Results

Table E-1: Examining the Effects of Removing AVG_VALUE_HOME and INCOME From the Model, Dependent Variable = POPCHG_TOT, Total Sample

Variable Name	Model 1 Benchmark Model	Model 2 Remove AVG_ VALUE_HOME	Model 3 Remove INCOME	Model 4 Remove amenities
C	5.036 (0.634)	7.725 (0.963)	2.85 (0.381)	0.202 (0.026)
AVG_VALUE_HOME	0.025*** (3.03)		0.021** (2.44)	0.022*** (2.91)
DIST_NATLHWY	0.005* (1.77)	0.007** (2.21)	0.005* (1.83)	0.006** (2.06)
EMPLOYMENT_RATE	0.124*** (3.06)	0.104** (2.55)	0.097** (2.3)	0.141*** (3.53)
HERF_INDEX	-18.134*** (-5.31)	-18.773*** (-5.4)	-18.104*** (-5.34)	-16.4*** (-4.77)
INCOME	-0.00021* (-1.74)	-0.00003 (-0.254)		-0.00022* (-1.78)
INDMIX_EMPGROW	-0.025 (-1.03)	-0.014 (-0.529)	-0.016 (-0.676)	-0.022 (-0.812)
INDMIX_EMPGROW_SURR	0.382*** (5.03)	0.404*** (5.23)	0.383*** (5.1)	0.288*** (4.5)
PER_BEL_MEDIAN	-0.06*** (-2.9)	-0.062*** (-2.98)	-0.048*** (-2.83)	-0.062*** (-3.13)
PER_EMPLOY_AGRIC	-0.05 (-1.32)	-0.032 (-0.857)	-0.034 (-0.887)	-0.051 (-1.26)
PER_EMPLOY_PRIMARY	-0.18*** (-5.16)	-0.215*** (-6.06)	-0.203*** (-6.13)	-0.182*** (-4.79)
PER_SELFEMPLOY	0.252*** (5.24)	0.331*** (6.09)	0.258*** (5.33)	0.271*** (5.31)
PER_CERTIFICATE	-0.07 (-1.61)	-0.095** (-2.04)	-0.08* (-1.82)	-0.078* (-1.7)
PER_NO_HSGRAD	-0.043 (-1.14)	-0.068* (-1.73)	-0.042 (-1.12)	-0.068* (-1.78)
PER_UNIVERSITY	-0.089 (-1.41)	-0.102 (-1.64)	-0.147*** (-2.69)	-0.066 (-1.06)
PER_ABORIGINAL	0.086*** (2.78)	0.094*** (2.87)	0.089*** (2.92)	0.077** (2.53)
PER_IMMIG_10	-0.023 (-0.293)	0.074 (1.04)	0.008 (0.098)	-0.021 (-0.274)
PER_OLD	-0.062 (-1.1)	-0.078 (-1.38)	-0.067 (-1.19)	-0.042 (-0.749)
PER_YOUNG	0.139** (2.11)	0.118* (1.79)	0.144** (2.22)	0.156** (2.4)
D_ATLANTIC	0.134 (0.16)	0.003 (0.003)	0.374 (0.437)	-1.075 (-1.44)
D_NORTHERN	-1.018 (-1.1)	-1.085 (-1.15)	-1.109 (-1.21)	-2.496** (-2.48)
D_ONTARIO	-1.558** (-2.32)	-1.077* (-1.65)	-1.611** (-2.45)	-1.675** (-2.55)
D_QUEBEC	-0.854 (-1.06)	-1.573* (-1.95)	-0.828 (-1.03)	-0.571 (-0.873)
DIST_RELIG	0.013 (0.834)	0.011 (0.623)	0.009 (0.565)	0.0007 (0.077)
PER_OWN_HOME	0.028* (1.8)	0.027* (1.8)	0.021* (1.69)	0.028** (1.97)
PER_SAMEADDRESS	-0.086 (-1.44)	-0.086 (-1.42)	-0.085 (-1.45)	-0.086 (-1.44)
PERCAP_VOL	-0.899 (-1.61)	-0.93 (-1.64)	-0.852 (-1.52)	-0.86 (-1.44)
PER_VOTE	0.007 (0.284)	-0.01 (-0.419)	0.007 (0.271)	0.047* (1.77)
D_CCSINCMA	-0.505 (-1.45)	-0.473 (-1.36)	-0.565 (-1.64)	-0.267 (-0.764)
POP91_100K	0.00028 (1.58)	0.00054*** (3.35)	0.00027 (1.56)	0.00033** (2.22)
POP_91	-0.001 (-1.19)	-0.001 (-1.18)	-0.001 (-1.2)	-0.00001 (-0.01)
CMA_CA_DIST	-0.007 (-1.53)	-0.006 (-1.5)	-0.006 (-1.52)	-0.006 (-1.47)

Table E-1 Continued on Next Page

Table E-1: Examining the Effects of Removing AVG_VALUE_HOME and INCOME From the Model, Dependent Variable = POPCHG_TOT, Total Sample Continued

Variable Name	Model 1	Model 2	Model 3	Model 4
	Benchmark Model	Remove AVG_ VALUE_HOME	Remove INCOME	Remove amenities
	0.00015	0.00018	0.00016	
CRIME_PROPERTY_RATE	(0.973)	(1.09)	(1.05)	
	-0.001***	-0.002***	-0.001***	
CRIME_VIOLENT_RATE	(-2.73)	(-2.87)	(-2.89)	
	0.019**	0.022**	0.018**	
DIST_ACUTE	(2.32)	(2.53)	(2.17)	
	-0.011***	-0.012***	-0.012***	
DIST_LACUTE	(-3.2)	(-3.37)	(-3.33)	
	0.009**	0.011***	0.009***	
DIST_COLLEGE	(2.51)	(2.94)	(2.71)	
	-0.004	-0.003	-0.002	
DIST_GLF	(-0.534)	(-0.427)	(-0.296)	
	-0.003	-0.00064	-0.00058	
DIST_PHYS	(-0.349)	(-0.067)	(-0.062)	
	-0.015*	-0.017*	-0.015*	
DIST_POLICE	(-1.72)	(-1.94)	(-1.69)	
	-0.015	-0.014	-0.014	
DIST_SCHOOL	(-0.955)	(-0.87)	(-0.859)	
	0.005	0.003	0.004	
DIST_SKI	(1.05)	(0.632)	(1)	
	-0.001	-0.002	-0.002	
DIST_UNIV	(-0.47)	(-0.784)	(-0.667)	
	-0.005	-0.004	-0.005	
COVER_FOREST	(-1.2)	(-0.794)	(-1.12)	
	-0.316	-0.292	-0.326	
D_ANYWATER	(-0.925)	(-0.843)	(-0.952)	
	0.003	0.001	0.002	
ELEV_STD_DEV	(1.18)	(0.587)	(1.07)	
	0.00025	0.00045	0.00027	
WEATH_AVE_PRECIP	(0.444)	(0.766)	(0.475)	
	-0.003	-0.003	-0.003	
WEATH_AVE_SNOW	(-1.39)	(-1.22)	(-1.22)	
	0.017*	0.011	0.017*	
WEATH_JAN_SUNSHINE	(1.69)	(1.13)	(1.66)	
	-0.079	-0.057	-0.08	
WEATH_JAN_TEMP	(-1.45)	(-1.06)	(-1.47)	
	-0.049***	-0.048***	-0.047***	
WEATH_JULYHUMID	(-3.67)	(-3.48)	(-3.5)	
F-Econ	43.9***	48.1***	53.0***	53.9***
	(0.000)	(0.000)	(0.000)	(0.000)
F-Human Capital	0.85	3.8***	7.4***	2.6**
	(0.466)	(0.010)	(0.000)	(0.050)
F-Demographic	23.3***	31.2***	34.1***	29.0***
	(0.000)	(0.000)	(0.000)	(0.000)
F-Regional	16.9***	5.2***	8.0***	13.9***
	(0.000)	(0.000)	(0.000)	(0.000)
F-Social Capital	9.5***	33.4***	32.4***	34.5***
	(0.000)	(0.000)	(0.000)	(0.000)
F-Urban Scale	14.7***	14.8***	5.3***	6.0***
	(0.000)	(0.000)	(0.000)	(0.000)
F-Modern Amenities	3.03***	7.6***	6.5***	
	(0.000)	(0.000)	(0.000)	
F-Natural Amenities	8.69***	7.2***	9.4***	
	(0.000)	(0.000)	(0.000)	
R^2	0.784	0.779	0.783	0.768
N	2402	2402	2402	2402

This regression is weighted by the log of the initial-year CCS population. The territories and CCSs with a population of less than 250 are excluded from this regression. T-Statistics and f-statistic p-values are reported in parenthesis. Model uses White's Heteroskedasticity-Consistent Standard Errors and Covariance. Regressions are weighted by log(POP_91). See Appendix A for variable definitions, and table 4.3 for theoretical expectations of directional influence. *, **, and *** denote significance at 10%, 5%, and 1% levels respectively.

Table E-2: Cohort Analysis - Dependent Variable Varies – Urban Sample

Variable Name	Model 1 Dependent Var: POPCHG_ TOT	Model 2 Dependent Var: POPCHG_ YOUTH	Model 3 Dependent Var: POPCHG_ YOUNG_ ADULT	Model 4 Dependent Var: POPCHG_ ADULT	Model 5 Dependent Var: POPCHG_ EARLY_ RETIREE	Model 6 Dependent Var: POPCHG_ ELDERLY
C	38.208*** (3.85)	76.696*** (4.65)	6.524 (0.415)	40.762*** (4.33)	18.104* (2.08)	24.805*** (3.09)
AVG_VALUE_HOME	0.017* (1.83)	0.032** (2.06)	0.02 (1.36)	0.019** (2.39)	0.012 (1.61)	0.01 (1.09)
DIST_NATLHWY	0.014** (2.45)	0.015 (1.54)	0.024*** (2.67)	0.012** (2.54)	-0.003 (-0.341)	-0.001 (-0.287)
EMPLOYMENT_RATE	0.047 (0.668)	0.292** (2.51)	-0.027 (-0.261)	-0.087 (-1.29)	0.027 (0.404)	0.017 (0.303)
HERF_INDEX	-24.359*** (-2.82)	-6.947 (-0.463)	-24.474* (-1.89)	-26.399*** (-3.02)	-15.679 (-1.52)	-14.794** (-2.08)
INCOME	-0.0006*** (-3.17)	-0.00066* (-1.92)	-0.00033 (-1.04)	-0.00083*** (-4.51)	-0.001*** (-6.28)	-0.0005*** (-2.68)
INDMIX_EMPGROW	-0.055 (-0.948)	0.101 (1.07)	-0.07 (-0.738)	-0.061 (-0.956)	-0.008 (-0.135)	-0.057 (-1.12)
INDMIX_EMPGROW_SURR	0.294*** (2.64)	-0.081 (-0.475)	0.673*** (3.94)	0.238** (2.28)	0.238** (2.4)	0.036 (0.474)
PER_BEL_MEDIAN	-0.152*** (-3.6)	0.152** (2.31)	-0.276*** (-3.68)	-0.251*** (-5.28)	-0.193*** (-4.58)	-0.097*** (-2.62)
PER_EMPLOY_AGRIC	0.185** (2.39)	0.154 (1.24)	0.263** (2.22)	0.228*** (3.28)	0.066 (1.04)	-0.129** (-2.11)
PER_EMPLOY_PRIMARY	-0.166*** (-2.9)	-0.293*** (-2.9)	-0.265*** (-2.93)	-0.036 (-0.631)	-0.04 (-0.441)	-0.107* (-1.79)
PER_SELFEMPLOY	0.237* (1.76)	-0.199 (-1.04)	0.67*** (3.11)	0.351*** (2.9)	0.523*** (4)	0.08 (0.631)
PER_CERTIFICATE	-0.117 (-1.21)	-0.299* (-1.93)	-0.142 (-0.898)	0.003 (0.035)	-0.022 (-0.212)	-0.002 (-0.024)
PER_NO_HSGRAD	-0.046 (-0.694)	-0.243** (-2.14)	-0.019 (-0.153)	0.011 (0.164)	0.028 (0.38)	-0.247*** (-3.57)
PER_UNIVERSITY	-0.002 (-0.024)	0.11 (0.683)	-0.211 (-1.1)	0.066 (0.831)	0.146 (1.55)	0.174** (2.02)
PER_ABORIGINAL	0.012 (0.206)	-0.138 (-1.29)	-0.033 (-0.382)	0.087* (1.72)	-0.014 (-0.186)	-0.189*** (-3.14)
PER_IMMIG_10	0.049 (0.578)	0.172 (1.21)	0.054 (0.388)	-0.047 (-0.61)	-0.007 (-0.105)	-0.204*** (-2.7)
PER_OLD	-0.171* (-1.69)	-0.653*** (-3.83)	-0.026 (-0.166)	0.263** (1.99)	0.717*** (7.48)	0.375*** (4.43)
PER_YOUNG	-0.116 (-0.91)	-1.283*** (-6.38)	0.23 (1.17)	-0.217* (-1.67)	0.373*** (3.38)	0.446*** (4.14)
D_ATLANTIC	1.219 (1.16)	-0.629 (-0.297)	0.524 (0.28)	1.118 (1.25)	2.511** (1.97)	-0.158 (-0.173)
D_NORTHERN	-0.701 (-0.362)	1.42 (0.461)	-1.789 (-0.744)	-0.61 (-0.293)	-2.422* (-1.66)	-1.234 (-0.836)
D_ONTARIO	0.527 (0.649)	-1.842 (-1.36)	-0.134 (-0.107)	0.406 (0.544)	1.736** (2.14)	-0.025 (-0.037)
D_QUEBEC	0.698 (0.624)	-4.003** (-2.22)	1.876 (1.01)	0.382 (0.398)	2.435** (2.12)	-2.927*** (-3.4)
DIST_RELIG	-0.028* (-1.89)	-0.00046 (-0.017)	-0.035 (-1.58)	-0.007 (-0.517)	-0.058*** (-3.16)	-0.024 (-1.42)
PER_OWN_HOME	0.141*** (4.92)	-0.064 (-1.47)	0.289*** (6.02)	0.103*** (4.29)	0.122*** (4.98)	-0.012 (-0.582)
PER_SAMEADDRESS	-0.396*** (-10.2)	-0.514*** (-9.25)	-0.275*** (-4.55)	-0.291*** (-8.48)	-0.379*** (-11)	-0.237*** (-8.34)
PERCAP_VOL	0.557 (0.532)	-0.417 (-0.25)	-0.974 (-0.604)	1.299 (1.59)	2.041* (1.93)	3.434*** (4.45)
PER_VOTE	0.005 (0.138)	0.05 (0.851)	-0.051 (-0.693)	0.024 (0.67)	-0.048 (-1.62)	-0.002 (-0.058)
POP91_100K	0.00008 (0.391)	0.00036 (1.15)	-0.00007 (-0.224)	-0.00026 (-1.47)	-0.00013 (-0.743)	0.00013 (0.926)
POP_91	0.00015 (0.107)	0.001 (0.623)	0.00008 (0.026)	-0.00074 (-0.558)	-0.002** (-2.38)	-0.00066 (-0.618)

Table E-2 continued on next page

Table E-2: Cohort Analysis - Dependent Variable Varies – Urban Sample Continued

Variable Name	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Dependent Var: POPCHG_ TOT	Dependent Var: POPCHG_ YOUTH	Dependent Var: POPCHG_ YOUNG_ ADULT	Dependent Var: POPCHG_ ADULT	Dependent Var: POPCHG_ EARLY_ RETIREE	Dependent Var: POPCHG_ ELDERLY
CRIME_PROPERTY_RATE	-0.00006	-0.00012	0.00006	-0.00013	-0.00005	-0.00009
	(-0.435)	(-0.464)	(0.265)	(-0.984)	(-0.394)	(-0.77)
CRIME_VIOLENT_RATE	-0.00073	-0.00065	-0.002	-0.00094	-0.001**	0.00011
	(-1.16)	(-0.58)	(-1.5)	(-1.6)	(-2.38)	(0.212)
DIST_COLLEGE	0.013*	-0.01	0.023**	-0.002	0.031***	0.008
	(1.86)	(-0.779)	(2.08)	(-0.373)	(2.93)	(1.24)
DIST_LACUTE	-0.011**	-0.005	-0.024***	-0.008	-0.008*	-0.005
	(-2.05)	(-0.578)	(-3.3)	(-1.56)	(-2.03)	(-1.24)
DIST_SKI	0.005	0.018	0.002	0.005	0.001	-0.00087
	(0.744)	(1.44)	(0.24)	(1.04)	(0.253)	(-0.21)
A_DIST_UNIV	0.002	0.005	0.005	-0.00046	-0.006*	0.006*
	(0.431)	(0.674)	(0.983)	(-0.133)	(-1.75)	(1.85)
COVER_FOREST	-0.007	0.003	-0.017*	-0.007	-0.014**	-0.008*
	(-1.18)	(0.321)	(-1.91)	(-1.45)	(-2.49)	(-1.88)
D_ANYWATER	-1.01**	-1.403**	-1.372*	-0.638	0.146	-0.32
	(-2.03)	(-2.18)	(-1.86)	(-1.57)	(0.334)	(-0.959)
ELEV_STD_DEV	0.002	-0.006	0.006	0.003	0.011***	0.003
	(0.838)	(-1.06)	(1.41)	(1.07)	(3.94)	(1.21)
WEATH_AVE_PRECIP	0.00051	-0.00093	0.003***	0.00008	0.0007	0.00019
	(0.89)	(-0.848)	(2.64)	(0.138)	(1.08)	(0.331)
WEATH_AVE_SNOW	0.00037	-0.003	0.00066	0.003	0.004	-0.003
	(0.095)	(-0.543)	(0.11)	(0.961)	(1.01)	(-0.955)
WEATH_JAN_SUNSHINE	0.024	0.014	0.053**	0.014	0.007	-0.01
	(1.48)	(0.584)	(2.28)	(0.946)	(0.504)	(-0.898)
WEATH_JAN_TEMP	-0.012	0.084	-0.033	0.023	-0.021	-0.074
	(-0.174)	(0.695)	(-0.27)	(0.361)	(-0.274)	(-1.24)
WEATH_JULYHUMID	-0.048***	-0.026	-0.096***	-0.051***	-0.024*	-0.01
	(-3.39)	(-1.11)	(-4.36)	(-4.15)	(-1.71)	(-0.865)
F-Econ	11***	8.4***	14.91***	17.7***	11.6***	3.58***
	0.000	0.000	0.000	0.000	0.000	0.000
F-Human Capital	0.635	4.53***	1	0.324	1.4	13.3***
	0.593	0.004	0.392	0.808	0.246	0.000
F-Demographic	0.715	16.82***	0.792	11.2***	15.5***	11.8***
	0.582	0.000	0.531	0.000	0.000	0.000
F-Regional	0.375	1.96*	0.628	0.425	2.7**	4.37***
	0.826	0.099	0.642	0.791	0.030	0.001
F-Social Capital	39.6***	42.5***	16.5***	27.6***	38.6***	29.1***
	0.000	0.000	0.000	0.000	0.000	0.000
F-Urban Scale	0.15	1.27	0.047	2	2.03	0.924
	0.860	0.281	0.954	0.139	0.132	0.398
F-Modern Amenities	1.9*	1.17	2.57**	2.15**	5.49***	1.166
	0.078	0.322	0.018	0.046	0.000	0.323
F-Natural Amenities	2.77***	1.124	5.1***	3.19***	3.57***	0.844
	0.005	0.346	0.000	0.002	0.000	0.564
R^2	0.695	0.748	0.469	0.504	0.605	0.748
N	507	507	507	507	507	507

This regression is weighted by the log of the initial-year CCS population. The territories and CCSs with a population of less than 250 are excluded from this regression. T-Statistics and f-statistic p-values are reported in parenthesis. Model uses White's Heteroskedasticity-Consistent Standard Errors and Covariance. Regressions are weighted by log(POP_91). See Appendix A for variable definitions, and table 4.3 for theoretical expectations of directional influence. *, **, and *** denote significance at 10%, 5%, and 1% levels respectively.

Table E-3: Cohort Analysis - Dependent Variable Varies – Rural Sample

Variable Name	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Dependent Var: POPCHG_ TOT	Dependent Var: POPCHG_ YOUNG	Dependent Var: POPCHG_ YOUNG_ ADULT	Dependent Var: POPCHG_ ADULT	Dependent Var: POPCHG_ EARLY_ RETIREE	Dependent Var: POPCHG_ ELDERLY
C	1.552 (0.274)	13.94 (1.16)	-14.569 (-1.6)	3.678 (0.576)	-32.313*** (-3.78)	15.688*** (2.97)
AVG_VALUE_HOME	0.045*** (5.06)	0.064*** (4.02)	0.043*** (3.06)	0.044*** (4.57)	0.065*** (5.4)	0.009 (1.11)
DIST_NATLHWY	0.002 (0.593)	-0.004 (-0.652)	0.003 (0.551)	-0.00046 (-0.166)	0.005 (1.56)	0.004 (1.26)
EMPLOYMENT_RATE	0.045 (1.64)	0.249*** (4.92)	0.005 (0.106)	-0.062* (-1.7)	-0.11** (-2.38)	-0.004 (-0.133)
HERF_INDEX	-18.026*** (-5.52)	-27.127*** (-3.58)	-19.022*** (-3.34)	-13.414*** (-2.79)	-0.865 (-0.145)	-9.404** (-2.44)
INCOME	-0.00003 (-0.264)	-0.00013 (-0.638)	0.00013 (0.673)	-0.00012 (-0.89)	-0.00019 (-1.19)	-0.00003 (-0.202)
INDMIX_EMPGROW	0.014 (0.64)	0.13*** (2.72)	-0.094** (-2.17)	0.04 (1.08)	0.065 (1.6)	-0.018 (-0.701)
INDMIX_EMPGROW_SURR	0.419*** (5.17)	-0.002 (-0.011)	0.734*** (5.5)	0.443*** (4.7)	0.48*** (4.22)	-0.047 (-0.601)
PER_BEL_MEDIAN	-0.007 (-0.604)	0.055 (1.5)	-0.075*** (-3.13)	-0.044*** (-2.98)	-0.092*** (-3.51)	0.072*** (3.66)
PER_EMPLOY_AGRIC	-0.009 (-0.353)	-0.053 (-1.12)	0.001 (0.034)	0.047 (1.38)	-0.035 (-0.778)	-0.053** (-2.17)
PER_EMPLOY_PRIMARY	-0.148*** (-4.87)	-0.26*** (-4.64)	-0.148*** (-3.03)	-0.075* (-1.71)	-0.178*** (-3.01)	-0.045 (-1.29)
PER_SELFEMPLOY	0.154*** (3.72)	-0.053 (-0.599)	0.093 (1.05)	0.123* (1.92)	0.395*** (3.91)	0.095** (2)
PER_CERTIFICATE	-0.026 (-0.636)	-0.041 (-0.526)	-0.087 (-0.958)	-0.093 (-1.47)	0.199* (1.89)	0.118* (1.68)
PER_NO_HSGRAD	-0.012 (-0.366)	-0.045 (-0.672)	0.044 (0.703)	-0.037 (-0.789)	0.151** (2.44)	0.019 (0.501)
PER_UNIVERSITY	-0.068 (-0.983)	-0.227* (-1.68)	0.082 (0.65)	-0.008 (-0.089)	-0.053 (-0.42)	0.252*** (3.1)
PER_ABORIGINAL	0.114*** (4.28)	0.191*** (4.02)	0.031 (0.783)	0.078*** (2.86)	0.071** (2.44)	-0.06*** (-3.08)
PER_IMMIG_10	0.28* (1.78)	0.887*** (2.72)	-0.482* (-1.92)	0.245 (1.55)	-0.447 (-1.58)	-0.163 (-1.04)
PER_OLD	-0.068 (-1.27)	-0.513*** (-4.38)	0.413*** (4)	0.418*** (6.14)	0.946*** (10.3)	-0.164*** (-2.73)
PER_YOUNG	-0.075 (-1.13)	-1.092*** (-7.06)	0.221* (1.67)	-0.162* (-1.91)	0.613*** (5.38)	-0.095 (-1.4)
D_ATLANTIC	-1.58 (-1.64)	2.941 (1.62)	-7.203*** (-3.99)	-3.031*** (-2.68)	0.819 (0.604)	-1.868* (-1.93)
D_NORTHERN	-1.991** (-2.17)	-1.019 (-0.737)	-3.408** (-2.47)	-1.316 (-1.34)	-1.811 (-1.39)	0.562 (0.652)
D_ONTARIO	-3.28*** (-3.54)	-3.111** (-2.07)	-6.726*** (-4.68)	-2.848*** (-3.16)	-3.852*** (-3.25)	-2.699*** (-3.5)
D_QUEBEC	-0.902 (-0.805)	1.18 (0.596)	-3.568* (-1.95)	-1.927 (-1.54)	1.77 (1.16)	-2.253** (-2.27)
DIST_RELIG	0.021* (1.87)	0.018 (0.963)	0.041** (2.34)	0.014 (1.14)	-0.014 (-0.845)	-0.007 (-0.702)
PER_OWN_HOME	-0.004 (-0.512)	-0.027 (-0.96)	-0.0003 (-0.02)	0.004 (0.631)	0.033** (1.96)	-0.021 (-1.4)
PER_SAMEADDRESS	-0.026 (-1.07)	-0.051 (-1.15)	-0.004 (-0.437)	-0.02 (-0.966)	-0.037 (-1.54)	0.004 (0.815)
PERCAP_VOL	-0.735 (-1.41)	0.606 (0.633)	-2.742*** (-3.26)	0.093 (0.167)	-0.009 (-0.013)	-0.198 (-0.449)
PER_VOTE	0.048 (1.61)	0.123** (2.33)	0.055 (1.05)	0.068** (2.07)	0.031 (0.751)	-0.001 (-0.054)
POP91_100K	0.00029 (1.4)	-0.00073** (-2.26)	0.001*** (3.96)	0.00044* (1.92)	-0.00057** (-2.04)	-0.00095*** (-5.16)
POP_91	0.086** (2.53)	0.112** (2.15)	0.051 (1.07)	0.034 (1.09)	0.091*** (2.81)	0.103*** (4.15)
CMA_CA_DIST	-0.014*** (-2.92)	-0.008 (-0.976)	-0.019*** (-2.59)	-0.014*** (-2.75)	-0.01 (-1.29)	-0.003 (-0.841)

Table E-3 continued on next page

Table E-3: Cohort Analysis - Dependent Variable Varies – Rural Sample Continued						
Variable Name	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Dependent Var: POPCHG_ TOT	Dependent Var: POPCHG_ YOUNG	Dependent Var: POPCHG_ YOUNG_ ADULT	Dependent Var: POPCHG_ ADULT	Dependent Var: POPCHG_ EARLY_ RETIREE	Dependent Var: POPCHG_ ELDERLY
CRIME_PROPERTY_RATE	0.00002	0.002***	-0.00046	-0.00067**	-0.00014	0.00016
	(0.084)	(2.8)	(-0.966)	(-2.43)	(-0.524)	(0.761)
CRIME_VIOLENT_RATE	-0.00052	-0.003***	-0.0001	-0.00001	-0.00058	-0.002***
	(-0.804)	(-2.79)	(-0.106)	(-0.011)	(-0.692)	(-3.02)
DIST_ACUTE	0.015	0.021	0.029*	0.002	-0.019*	-0.022***
	(1.53)	(1.19)	(1.88)	(0.157)	(-1.7)	(-3.02)
DIST_LACUTE	-0.003	-0.00014	-0.003	-0.002	-0.007	-0.01***
	(-0.916)	(-0.026)	(-0.581)	(-0.618)	(-1.29)	(-2.93)
DIST_COLLEGE	0.007	0.024***	-0.007			
	(1.6)	(2.87)	(-1.15)			
DIST_GLF	-0.003	0.002	-0.006	0.001	0.018*	0.009
	(-0.455)	(0.178)	(-0.601)	(0.16)	(1.72)	(1.48)
DIST_PHYS	-0.005	-0.004	0.01	0.00095	0.009	-0.008
	(-0.558)	(-0.271)	(0.621)	(0.102)	(0.735)	(-0.981)
DIST_POLICE	-0.005	-0.033*	0.006	0.007	0.018	0.00024
	(-0.516)	(-1.78)	(0.465)	(0.648)	(1.35)	(0.029)
DIST_SCHOOL	-0.021	-0.022	-0.032	0.006		
	(-1.17)	(-0.677)	(-1.15)	(0.305)		
DIST_SKI	0.009**	0.01	0.008	0.005	-0.006	0.006
	(2.16)	(1.35)	(1.07)	(1.08)	(-1.03)	(1.46)
DIST_UNIV	-0.004	-0.008	-0.004			
	(-1.26)	(-1.49)	(-0.785)			
DIST_LTERM					-0.007	-0.005
					(-0.808)	(-0.816)
COVER_FOREST	-0.003	-0.001	0.006	0.008	-0.006	-0.012**
	(-0.433)	(-0.114)	(0.6)	(1.28)	(-0.8)	(-2.23)
D_ANYWATER	-0.119	-1.592***	-0.01	0.301	1.152**	-0.634**
	(-0.366)	(-2.84)	(-0.018)	(0.768)	(2.21)	(-2.06)
ELEV_STD_DEV	-0.00074	-0.00037	0.00046	-0.003*	-0.004*	-0.00091
	(-0.443)	(-0.117)	(0.165)	(-1.95)	(-1.71)	(-0.552)
WEATH_AVE_PRECIP	-0.001	-0.003**	-0.00071	-0.00065	-0.00071	-0.00097
	(-1.45)	(-2.02)	(-0.65)	(-1.03)	(-1.01)	(-1.44)
WEATH_AVE_SNOW	-0.00051	0.002	0.00072	-0.00092	0.001	0.00095
	(-0.27)	(0.612)	(0.217)	(-0.529)	(0.496)	(0.63)
WEATH_JAN_SUNSHINE	0.006	0.008	0.022	-0.005	-0.001	0.009
	(0.519)	(0.428)	(1.15)	(-0.439)	(-0.066)	(1.13)
WEATH_JAN_TEMP	-0.13*	-0.145	-0.084	-0.056	0.05	0.032
	(-1.95)	(-1.24)	(-0.778)	(-0.873)	(0.58)	(0.553)
WEATH_JULYHUMID	-0.024	-0.005	-0.029	-0.022	-0.03	-0.005
	(-1.61)	(-0.188)	(-1.21)	(-1.47)	(-1.39)	(-0.296)
WEATH_JULYHUMID	-0.049***	-0.016	-0.1***	-0.044***	-0.029**	-0.02**
	(-3.67)	(-0.776)	(-5.28)	(-4.19)	(-2.2)	(-2.07)
F-Econ	43.9***	25.6***	19.1***	20.1***	26***	15.4***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-Human Capital	0.85	2.4*	1.9	1.9	7.51***	10.9***
	(0.466)	(0.067)	(0.110)	(0.130)	(0.000)	(0.000)
F-Demographic	23.3***	66.7***	16.3***	85.4***	67.9***	9.3***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-Regional	16.9***	8.8***	23.1***	7.4***	10.3***	5.7***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-Social Capital	9.5***	8.7***	6.8***	3.5***	4.8***	2.3**
	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(0.039)
F-Urban Scale	14.7***	7.2***	16***	6.5***	5.2***	25.4***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
F-Modern Amenities	3.03***	6.9***	2.06**	2.6***	1.45	5.6***
	(0.000)	(0.000)	(0.020)	(0.005)	(0.160)	(0.000)
F-Natural Amenities	8.69***	6.5***	2.5***	2.9***	2.3**	3.52***
	(0.000)	(0.000)	(0.010)	(0.004)	(0.019)	(0.000)
R^2	0.695	0.748	0.469	0.504	0.605	0.748
N	1893	1893	1893	1893	1893	1893

This regression is weighted by the log of the initial-year CCS population. The territories and CCSs with a population of less than 250 are excluded from this regression. T-Statistics and f-statistic p-values are reported in parenthesis. Model uses White's Heteroskedasticity-Consistent Standard Errors and Covariance. See Appendix A for variable definitions, and table 4.3 for theoretical expectations of directional influence. *, **, and *** denote significance at 10%, 5%, and 1% levels respectively.