Understanding the Commute Pattern in the Greater Golden Horseshoe: Where are the long commutes and what is causing them?

by

Coco (Yi Ning) Wang

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of B.A. Honours in Urban Studies

Department of Geography
McGill University
Montréal (Québec) Canada

April 2021

ACKNOWLEDGEMENT

I would first like to thank my supervisor, Kevin Manaugh for his valuable support for the past year. I would not have been able to complete the project without his dedication and patience. I am incredibly honoured to have had the opportunity to work with and learn from him.

I'd also like to thank Professor Ahmed El-Geneidy for being my reader and the time he has dedicated towards reviewing this paper.

Lastly, I'd like to thank Professor Sarah Turner and my honours thesis cohort for their support and encouragement throughout this process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF MAPS

LISTS OF TABLES

ABSTRACT

Canadian commutes are getting longer and is especially prevalent in large cities such as Toronto. Long commutes have serious impacts socially, environmentally, and financially. I focused my research on the Greater Golden Horseshoe (GGH), the fastest growing region in Canada, centering Toronto. I used the 2016 Census to conduct analysis on the commute pattern in the GGH with a focus on long commutes. My results identified that this region is greatly interconnected as around 20% of workers from other CMAs commute to Toronto, and most long commutes over 70 kilometers end in Toronto. Moreover, Census Tracts inhabited by households making median wages have the longest commute. It is certain that there is a degree of spatial mismatch for many workers living in the region, causing them hours of commute every day. Thus, it is crucial for the region to evaluate the diversity of housing and job to curtail extreme commutes.

CHAPTER 1: INTRODUCTION

Canadian cities in the 20th century witnessed rapid suburbanization following the popularization of personal vehicles and preferences for more spacious single-unit houses, as people were no longer constrained by mobility to live near their workplace that largely remained in urban regions (Axisa et al., 2012). Therefore, the daily commute between home and work in Canadian metropolitan cities continuously lengthened both in terms of time and distance. The 2016 Canadian Census revealed that there has been a 5% increase in the number of commuters that travel over 60 minutes by car since the 2011 Census (Government of Canada, 2019). Recent planning discourse and transportation literature, which are currently centered on sustainable growth and urban revitalization principles, has thus put more emphasis on understanding the daily commute as well as how to decrease it.

Literature has recognized the detrimental impacts of long commutes. Long car commutes causes many health issues such as obesity and cardiovascular illness, while congestion and the unpredictability of commuting contributes to stress (Lyons & Chatterjee, 2008). Economically, the assumption that time has monetary value determines that commuting is wasteful, and further, those with long commutes are found to be less productive (van Ommeren & Gutiérrez-i-Puigarnau, 2011). Moreover, long commutes put pressure on existing transportation infrastructure, increasing infrastructure investment and traffic congestion. Long commutes by car are also much worse than commuting by active modes of transportation – walking, biking, and public transit – in terms of sustainability. They create pollution, contribute greatly to oil dependency and climate change, and overall, threatens the sustainable growth of urban development.

Geographers have been using data to analyze commute patterns in various cities around the world. As each has its unique infrastructure, growth pattern, planning policy, and other factors that shape commute patterns, it is difficult to extract a common model that could sufficiently be applied to every metropolitan region. As such, this research focuses on the case of long commute pattern in the context of the Greater Golden Horseshoe (GGH) in Southern Ontario, Canada.

1.1 Geographic Context

The GGH is the largest and fastest-growing metropolitan region in Canada as it is home to a quarter of the total Canadian population and over 4.5 million jobs (Government of Ontario, 2020). The region also has some of the longest commutes in Canada (Yaropud, 2019). The largest city in Canada, Toronto, is located at the center of the GGH. It has strong economic ties with its neighbouring Census Metropolitan Areas (CMAs) in the GGH, including Barrie, Brantford, Guelph, Hamilton, Kitchener-Cambridge-Waterloo, Oshawa, Peterborough, and St. Catharines-Niagara. Many of these neighbouring CMAs developed to be predominantly residential communities that offer affordable and comfortable housing but lack sufficient and diverse employment opportunities or urban benefits for their residents (Government of Canada, 2019). Therefore, the workers who are employed in the central city but reside in the periphery cities are most likely to be the ones with the longest commute in Canada.

The development of the GGH is guided by three levels of plans and policies: provincial, municipal, and project based. At the provincial level, the growth of GGH is guided by *The Places to Grow Act* of 2005, the *Provincial Policy Statement* of 2020*, A Place to Grow Plan* of 2020, and other provincial plans such as the Greenbelt plan. *A Place to Growth* is the specific growth plan for the GGH, establishing a long-term planning framework to regulate the region's development to combat various urban issues, such as the region's urban sprawl and demographic shift (Government of Ontario, 2020). The Plan guides government investment and land-use planning policies to assist decision-makers to understand how to implement policy, such that it establishes densification and intensification targets.

There are also provincial level transportation plans that guide the region's transportation development, including Metrolinx's *Regional Transportation Plan* and the GGH Transportation Plan, *The Big Move.* On a municipal level, each municipal has their own official plan, transportation master plan, active transport plan, and zoning by-law. Municipalities in the GGH includes upper-municipality (such as the Region of York) and lower-municipality (such as City of Richmond Hill). There are also single-tiered municipalities, including City of Toronto.

The current public transport infrastructure in the GGH include the Toronto subway, which is one of the only three rapid transit systems in Canada, and the GO Transit, the regional transportation system. Ontario transportation framework favours policy and plans that decrease car dependency in order to meet the climate target. Thus, the GGH also has projects to construct a new Light Rail Transit (LRT) and adding additional stops to current transit lines.

My research investigates the commute pattern of GGH with a focus on long commutes. Moreover, my research emphasizes the spatial distribution of employment opportunities and housing options as the main driver of the commute pattern. I hope that the results of this research can contribute to understanding commutes in a Canadian context and identify key problems to be addressed by future urban transportation planning initiatives.

1.2 Purpose: Aim and Research Questions

This thesis aims to conduct a spatial analysis of the commute flow in the Greater Golden Horseshoe with a focus on patterns of long commutes. To realize this research aim, I explore three research questions:

1. What is the commute flow and where are the long commutes?
    a. What is the average commute distance per Census Tract (CT) and CMA?
    b. Which CMA has the highest rate of long commutes?
2. To what extent does the GGH have a balanced jobs-housing distribution?
    a. What is the relationship between work opportunities and housing option in each CMA?
3. How do socio-demographic factors influence commute distance in the GGH?

The first research question serves to construct and visualize the commute flow in the GGH by understanding its geographical patterns. The second research question builds on the first research question and explores the jobs-housing distribution in the GGH. It analyzes the spatial distribution of employment opportunities versus housing options across the area to determine if there is a balance between the two. The last research question explores the other socio-demographic factors that may have an impact on commute distance, including income, dwelling type, and mode.

My thesis begins with a review of existing literature, focusing on location choice theory and determinants and impacts of long commutes (Chapter 2). The methodology chapter describes the study area and unit of analysis, research data, the variables, and methods of analysis (Chapter 3). The results of the analysis are presented in Chapter 4. Finally, this thesis ends in Chapter 5 with the discussion of research results, limitations, and further research possibilities.

CHAPTER 2: LITERATURE REVIEW

This chapter outlines the grounding conceptual framework and summarizes the existing literature concerning long commutes and their relationship with the spatial distribution of employment and residential locations (Figure 2.1). This thesis is rooted in two strands of literature: location choice theory and long commute literature. I will begin by investigating the key factors that shape urban geography through location choice literature, focusing on business location choice and residential location choice. Then I will examine determinants and impacts of long commutes, including spatial distribution of job-house locations, spatial mismatch theory and modal choice, and the modern perception of commute. The combination of these literature will help me navigate to find the answers to my research questions.

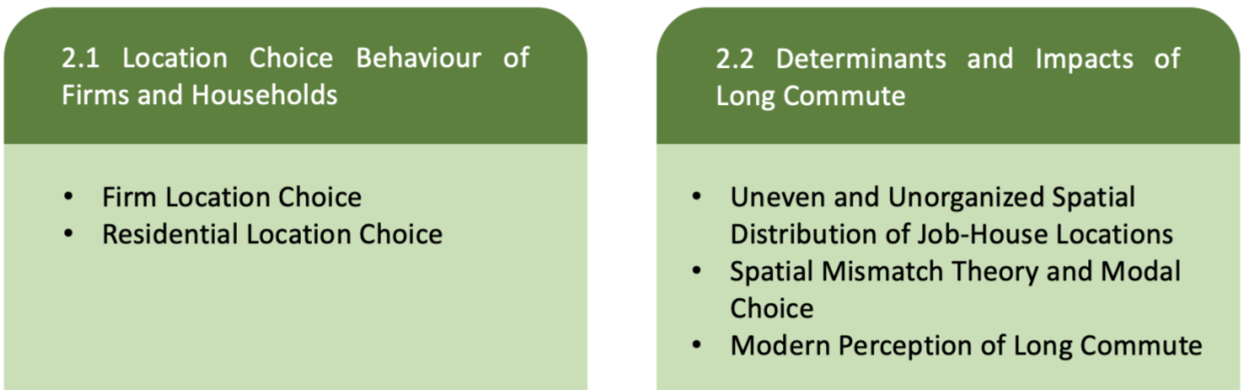


Figure 2. 1 Research Conceptual Framework

2.1 Location Choice Behaviour of Firms and Households

Location choice theory is rooted in economic geography as scholars attempt to identify a generalized model that can summarize the structures of cities. Early examples that incorporate transportation as part of the model include Burgess's Concentric Ring Model (1925) which describes the city as a series of rings with each ring having a specific land-use; Hoyt's Sector Model (1939) views different land-use as sections that surround a Central Business District (CBD); Harris and Ullman's Multiple Nuclei Model (1945) understands city patterns as layers of historical development. The limitations of these models are the oversimplification of spatial patterns, and geographers have since recognized a need for more sophisticated models that use micro data and can account for contextual variations. The scope of this thesis does not go into details of the specific modelling theory and method but will outline the most important factors that influence location choice behaviours for firms and households.

2.1.1 Firm Location Choice

The essence of firm location choice in a capitalist society is profitability — a firm will choose a location that can yield the maximum economic benefits. I begin by reviewing literature that explores location choice on a national or regional level. Early literature explored firm location choice for the manufacturing sector. Carlton (1983) applied a logit model and used data on new branch plants and region-specific economic variables to find determinants of new plant location. The study revealed that external factors including wage, energy cost, agglomeration economies, and the unemployment rate were found to be significant in predicting a firm's location; internal factors such as firm size and its technology level were also important in the decision-making process (Carlton, 1983). Determinants of location choice for firms in the quaternary sector — information and communication technologies industry—are similar to the manufacturing industry. Siedschlag et al. (2013) used data of knowledge-based foreign firms that are located in the European Union to find that key factors include market size and potential, agglomeration economies, income tax rate, skilled labour pool, labour cost, and the size of the service sector in the neighbouring regions.

A common factor across all industries' location choice is agglomeration economies. Agglomeration effect is the clustering of business activities, a phenomenon that can be observed at intra-urban, regional, and international levels. Marshall (1920) was the pioneer to identify that firms clustered for reduced transport cost, namely the cost of moving goods, people, and ideas. For example, a firm that locates close to its industry cluster will have more access to skilled-workers, and knowledge spillover allows them to observe and learn from other firms' operations. Indeed, there are many examples of agglomeration in North America such as technology firms in Silicon Valley, financial firms on Wall Street, and movie production firms in Hollywood. The advantages from the clustering of firms within the same industry can be understood as *localization economies*, on the other hand, *urbanization economies* refer to the benefits derived from the clustering of all firms in the area regardless of industry (Meyer, 2020). Indeed, economic agglomeration is the key driver to urbanization, however, it can also leads to the uneven distribution of economic activities at both international and national levels (Henderson et al., 2001). Henderson et al. (2001) summarized previous studies of geography and development, stating that centers of agglomeration are unlikely to change after it has been established as it is difficult for firms to cut ties with benefits, however, new centers will develop when these agglomerations grow.

Moreover, the size of agglomeration is directly proportional to workers' wages, cost of living, and commuting time (Henderson et al., 2001).

2.1.2 Residential Location Choice

Similar to business location choice, individuals and families choose the location of their home that will maximize utility function, given their resources. Building an empirical residential location choice model is a difficult and complicated task that has been and still is being tackled by many scholars. Early examples of modern residential location choice include Alonso (2013) who applied the Bid-Rent theory to residential location choice, Lowry (1963) who adopted spatial interaction principles to his model, and Hansen (1959) who incorporated the concept of accessibility. Numerous factors have been explored and tested in research, which I broadly summarize into two categories: the first is the home' attributes, including housing type, quality, price, and neighbourhood quality and attractiveness; the second is the transport from home to points of interests, such as employment, services, and education. In short, people favour affordable housing that is located in an accessible location to their employment and urban amenities. The two categories are also interconnected since the price of housing should naturally capture the accessibility of the location (Pagliara et al., 2010). The following part of the literature review focuses on two aspects of residential location choice: socioeconomic status and choice behaviour, and household composition and housing suitability. The commute is also one of the key determinants of residential location choice and will be further explored as its own topic in the subsequent section of the literature review.

Firstly, one's socioeconomic status greatly affects the way they experience the city and how they choose their home location. Florida (2002) argued about the differences of the residential location decision process between populations that work in knowledge-based occupations as compared to traditional occupations. Knowledge-based employees have abundant disposable income which allows them to develop a greater attraction towards urban areas with rich cultural, educational, and leisure activities as they desire a high-quality life (Yigitcanlar et al., 2007). On the other hand, those without the means have a very different set of considerations when choosing a home location and put more emphasis on factors related to price (i.e., cost of housing, cost of transportation). Cultural preferences also stand to be influential to location choice.

Secondly, the household composition has a great effect on residential location choice behaviour. Household size limits the suitable housing stock that each household would consider, for example, a household of four would consider bigger houses over small condominiums. As such, residential location choice is largely dependent on the location of suitable housing stock. As I previously examined, employment clusters typically have a higher land value that leads to the development of high-rise, and combined with neoliberal policies that favour densification, suitable and affordable housing stock for larger households are generally more difficult to find near these agglomerations (Moos et al., 2018). Lastly, multi-worker households have a more complicated process to determine a residential location that balances between the value and needs of several people (Acheampong, 2018).

2.2 Determinants and Impacts of Long Commute

2.2.1 Uneven and Unorganized Spatial Distribution of Jobs-Housing Locations

I have briefly alluded to the existence of uneven spatial distribution of employment and housing locations in the previous sections. Firms exhibit clustering patterns for the benefits of agglomeration economies, while land-value near the workplace is connected to the distinct geographic split of different housing types. The issue of the uneven distribution or the spatial imbalance of jobs and homes has been one of the key urban challenges concerning urban planners and academia. The consequences of such an imbalance are traffic congestion, accessibility issues, long commutes and other negative by-products (Cervero, 1989; Peng, 1997). However, the uneven spatial distribution of job-house locations various in degrees of influence. As such, spatial balance is not the definitive influence to commute length in every context but should be examined along with spatial organization as well.

Long commutes can be understood as a result of entropy in urban form. Scholars have developed sets of mathematical formulas that calculate the additional commute by comparing the observed, or actual commute patterns with the theoretical commute patterns derived from the spatial organization (Buliung & Kanaroglou, 2002). In other words, the equation explores the efficiency and optimization of existing commute patterns, and this additional commute was coined as “wasteful” commute by Hamilton (1982), but more commonly referred to as “excess” commute. In this sense, a long-distance between home and work locations that indicates a sprawled urban form does not necessarily conclude in a long commute (O’Kelly & Niedzielski, 2009). However, one of the biggest limitations of these theories is their lack of analysis on racial and socioeconomic constraints as well as modal differences. Therefore, the following section of the literature review focuses on the theory of spatial mismatch and modal mismatch as a determinant of long commutes.

2.2.2 The Spatial Mismatch Theory and Modal Choice

The spatial mismatch hypothesis originated as a theory to understand the effect of racial segregation on unemployment as Kain (1968) explored the spatial distribution and connection of African American housing and employment in Chicago and Detroit. His research confirmed that housing segregation reduces the minority groups' access to employment, which shined a light on the effect of the jobs-housing spatial distribution on economic mobility (Kain, 1968). As a result, minority groups in urban metropolitans are hypothesized to have longer commutes, and commute length are used as an indicator to test the spatial mismatch hypothesis. Grengs (2010) incorporated another layer of modal differences in his research on the spatial mismatch theory and found that those without automobiles are at the most disadvantage: those who take public transportation have a harder time reaching jobs, in addition to more travel time, less convenience and flexibility.

Transport geographers have also explored this phenomenon from the opposite perspective, investigating the determinants of commute length with socioeconomic status, race, and other demographic information as factors. Axisa et al. (2012) explored the different factors that influenced commute distance in Toronto using the 2006 Census Canada Master File and a multiple linear regression model. This study found a positive relationship between commute distance and migration status, such that new migrants have a longer commute as compared to long-term residents, among other statistically less influential factors.

2.2.3 Modern Perception of Long Commute

The last section of the literature review views long commutes through a cognitive lens. Empirical evidence attests to the negative consequences of long commute. Economically, the assumption that time has monetary value determines that commuting is wasteful; healthily, long car commutes cause many health issues, many of which stems from stress (Redmond & Mokhtarian, 2001; Lyons & Chatterjee, 2008). However, there are benefits of commuting in relations to positive utility. This means that commuting itself can be viewed as an enjoyable activity and productive time as opposed to understanding commuting as a mere act to reach the destination. The positive utility of time is especially relevant as the information age transforms the commuting experience (Lyons & Chatterjee, 2008). This perspective on long commutes also connects back with the residential location choice. Those who view commute as an enjoyable or necessary activity in their daily life may be more willing to trade-off a longer commute with other considerations of their choices, which explains its moderate impact on housing location choice on Hunt (2010)'s research in

Edmonton, Canada. People are getting more accustomed to longer commutes, which is a problem of its own. Planners now have the paradox to either focus their efforts on make the commute experience better, which would make people more willing to commute longer , or reduce the length of commute all-together (Lyons & Chatterjee, 2008).

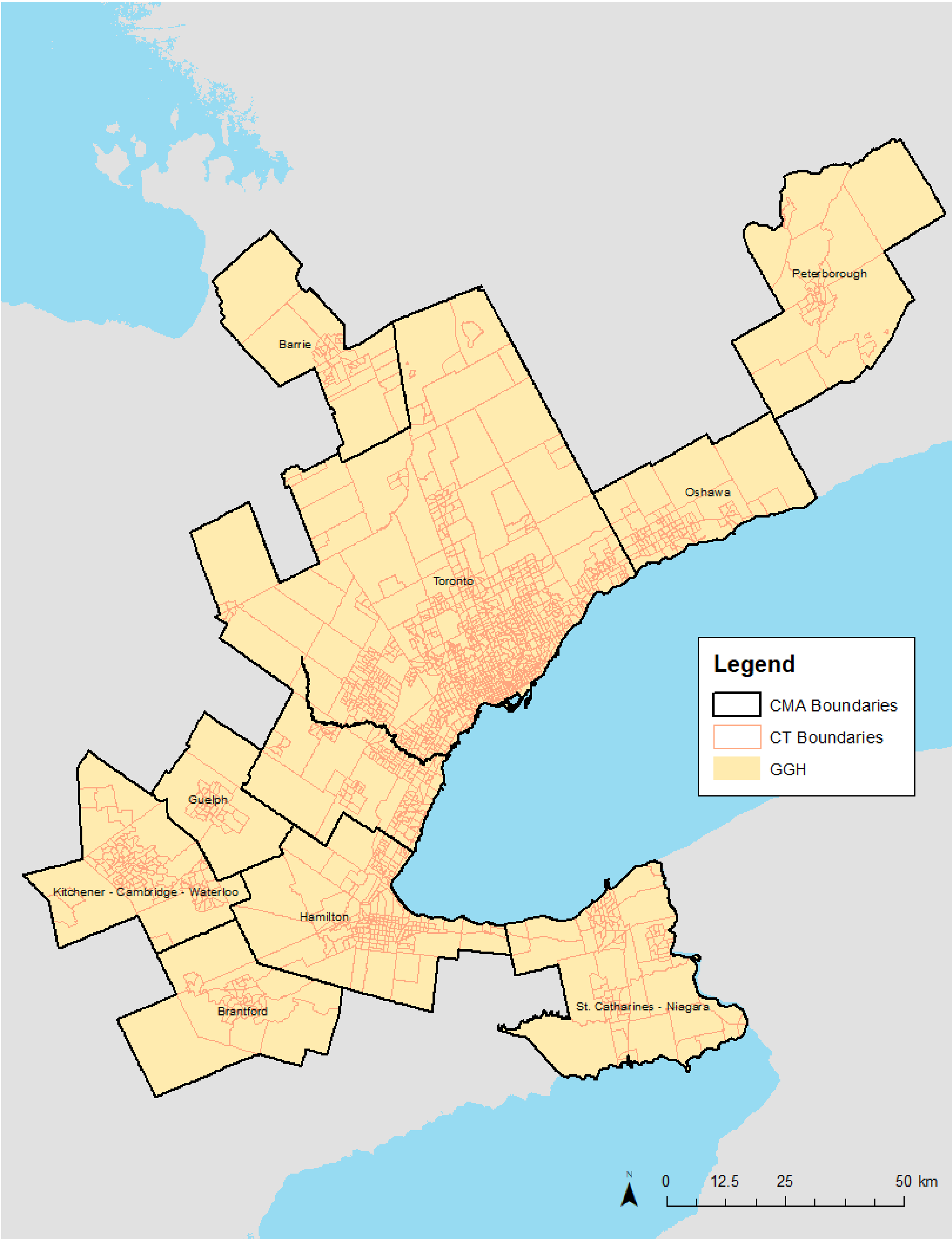
CHAPTER 3: METHODOLOGY

In this chapter, I discuss the quantitative methodology used to answer the research questions. I begin this chapter in Section 3.1 by defining the study area and unit of analysis. In Section 3.2, I describe the databases used in this thesis, which includes the main databases extracted from the 2016 Canadian Census. In Section 3.3, I describe the dependent and the independent variables, which are average commute distance, median household income, dwelling type, and mode share. I end this chapter in Section 3.4, where I introduce the various data manipulation and analysis techniques, including elementary data analysis using Excel and R, spatial analysis using ESRI ArcMap, and the statistical analysis using STATA.

3.1 Study Area and Unit of Analysis

The nine CMAs within GGH were chosen as the study area to capture most of the commutes between Toronto and its neighbouring municipalities. I conducted a quantitative cross-sectional study, meaning that the data were taken at a specific point in time, more specifically, at the time of the 2016 Canadian Census. As individual data is not publicly distributed by Statistics Canada, I chose to use Census Tract (CT) as the main geographic unit of analysis in this thesis. They are defined by Statistic Canada as "small, relatively stable geographic areas that usually have a population between 2,500 and 8000 persons", and geographically, they are "located in census metropolitan areas and in census agglomerations that had a core population of 50,000 or more in the previous census" (Statistics Canada, n.d.-a). There are a total of 1751 CTs in the GGH that are used in this research.

I chose Census Metropolitan Area (CMA) as the secondary unit of analysis in this thesis. Statistics Canada defines a CMA as "formed by one or more adjacent municipalities centered on a population center (known as the core)", and that "the CMA must have a total population of at least 100,000 of which 50,000 or more must live in the core" (Statistics Canada, n.d.-b). There are nine CMAs in the GGH, which includes Barrie, Brantford, Guelph, Hamilton, Kitchener-Waterloo-Cambridge, Oshawa, Peterborough, St. Catherine-Niagara, and Toronto. Map 3.1 exhibits the boundaries of the CMAs and CTs, while Table 3.1 presents basic characteristics about each CMA.



Map 3. 1 CMA and CT boundaries in the GGH

Table 3. 1 CMAs in the study area and their characteristics

CMA Name	2016 Population	Number of CTs
Barrie	197,059	42
Brantford	134,203	29
Guelph	151,984	30
Hamilton	747,545	189
Kitchener-Waterloo-Cambridge	523,894	107
Oshawa	379,848	84
Peterborough	121,721	30
St. Catherine-Niagara	406,074	94
Toronto	5,928,040	1146
Total	**8,590,368**	**1751**

(Source: Government of Canada, 2017)

3.2 Databases Used

This thesis relied on two databases derived from the 2016 Canadian Census, containing the characteristics of commutes and socio-demographic information, aggregated at the CT level. The 2016 Census had a 98.4% response rate as the entire Canadian population is legally required to answer the questionnaire, thus, this dataset is very rich and representative of the population (Government of Canada, 2018). The data were randomly rounded either up or down to a multiple of '5' or '10' to preserve confidentiality, which minimally affect the accuracy of the analysis in this research (Government of Canada, 2017a).

The first database extracted from the Census is a data package that includes commute flow characteristics, which was purchased from Statistics Canada in 2020 by my thesis supervisor, Kevin Manaugh. Each row in this database represents the number of people that commute from a certain CT of residence to a certain CT of work, broken down by travel mode. This database does not reflect the mixed-modal nature of many commutes because the respondents were directed to answer about the mode that they use for most of the travel distance in the case that their commutes involve more than one. Appendix A lists the survey questions used in the Census to produce this database.

The second database contains economic and typology characteristics at a CT level, including median household income and dwelling type. It was accessed through the Computing in the Humanities and Social Sciences (CHASS) Data Centre at the University of Toronto. Table 3.2 presents a summary of the database used and the specific variables.

Many shapefiles were also used in this thesis to construct the final maps shown in the following Chapter. I used the Ontario Street Network dataset to derive commute distance in ArcMap. The reasoning behind calculating this variable in this thesis rather than using the self-reported commute time collected by the Census is explained further in Section 3.3.1, where I elaborate about the dependent variable. This network dataset was extracted from the CanMap street file obtained from DMTI Spatial.

Boundary shapefiles used in this thesis include the Canadian CT boundary and Provincial Boundary in 2016, the Great Lakes Boundary, and the US State Boundary at 500k resolution. I provide the descriptions and sources of these databases and shapefiles in Table 3.3.

Table 3. 2 Description of the databases used in this thesis

Database	Key Variables

2016 Canadian Census - Commute Characteristics	Commute Geography:
	● Origin CT
	● Destination CT
	Mode of Travel:
	● Total – main mode of commuting
	● Car truck or van
	    ○ Driver, alone
	    ○ 2 or more persons shared the ride to work
	    ○ Driver, with 1 or more passengers
	    ○ Passenger, 2 or more persons in the car
	● Sustainable transportation
	    ○ Public transit
	    ○ Active transport
	        ▪ Walked
	        ▪ Bicycle
	● Other method (includes those working from home)

2016 Canadian Census - Socio-Demographic Characteristics	• Median total income of economic families in 2015 ($) Dwelling Type: ● Single Detached House ● Apartment in a building that has five or more storeys ● Semi-detached house ● Row house ● Apartment or flat in a duplex ● Apartment in a building that has fewer than five storeys ● Other single-attached house

Table 3. 3 Summary of the shapefiles used in this thesis

Shapefile	Source	Description
Ontario Street Network (CanMap)	DMTI Spatial[1]	This shapefile contains characteristics of the street network in Ontario, including information such as speed limit, one-way, and road length.
2016 Canadian Census Tract Boundary	Statistics Canada[2]	This shapefile contains the 2016 Canadian Census Tract Boundaries.
2016 Canadian Province Boundary	Statistics Canada[3]	This shapefile contains the 2016 Canadian Province Boundaries. The Ontario polygon was isolated and used to identify CTs in the province.
Great Lakes and Watersheds	U.S. Geological[4] Survey	This shapefile contains the Great Lakes and Watersheds.
US State Boundary	United States Census Bureau[5]	This shapefile contains the US State Boundary at the 500k resolution.

3.3 Variables

3.3.1 Dependent Variable

Although the Census database captured the average commute time of each CT, a few issues prevented me from using this variable in this thesis. Firstly, its aggregated nature cannot reflect mode choice, which is one of the most important independent variables explored in this thesis. Moreover, each mode has different movement speed, such that travelling by car is faster than walking. However, movement speed of certain modes can fluctuate greatly because of peak-hour congestion, thus, this variable cannot accurately represent the objective length between home and work location, defeating the nature of this research.

As such, I chose to use **commute distance** as the dependent variable, computed using the Network Analyst tool on ArcMap. I adopted a simplified model which used the Street Network dataset to calculate the distance between the centroid of the home CT and the centroid of the work CT. The computed distance reflects the average travel distance of all trips between any points within the two CTs, which is justified as the Census dataset does not provide coordinates of individual commute origins or destinations.

---

[1] https://uwaterloo.ca/library/geospatial/collections/canadian-geospatial-data-resources/canada/dmti-canmap-route-logistics
[2] https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm
[3] ibid
[4] https://www.sciencebase.gov/catalog/item/530f8a0ee4b0e7e46bd300dd
[5] https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.2016.html

Existing literature commonly uses commute time as the variable and lacks a standardized definition for “long commute”. In this thesis, I defined long commutes as trips over 70 kilometers (km), such that it corresponds roughly with 60 minutes of commute time, a threshold defined by Statistics Canada as a long commute for cars (Government of Canada, 2019). To put this into perspective, the median Canadian car commute in 2016 was 8.7 km; a trip from Newmarket — a town in Toronto that is just south of Barrie — to the Financial District — the heart of Toronto’s finance and banking firms located near Lake Ontario — is around 70km and 60 minutes by car off peak hours, as shown by Google Maps. This is certainly a conservative definition as peak hour congestion can significantly increase the commute time, but this threshold can capture the most extreme commutes in this region and can provide a fundamental understanding of the commute behaviour in the GGH.

3.3.2 Independent Variables

Income was identified as an important influence for both firm and residential location choice literature. High income jobs in the knowledge sector are typically clustered to benefit from the agglomeration economy, while lower income jobs in the service industry do not show such pattern at a regional level. Households with higher income have greater choices regarding their housing location, while lower income households are more restrained by housing prices and the price of commute. In this thesis, income is represented by “median total income of economic families in 2015 ($)”.

Dwelling type reflects a CT’s typology and density. The Canadian Census divides dwelling type into the following categories: “single-detached house”, “apartment in a building that has five or more storeys”, “semi-detached house”, “row house”, “apartment or flat in a duplex”, “apartment in a building that has fewer than five storeys”, and “other single-attached house”. This thesis used the proportion of single-detached house as an indicator of dwelling type as it is a dominant typology in the GGH.

Lastly, I used **mode** to understand GGH’s commute pattern, which was expressed as the proportion of car commutes in certain spatial and statistical analysis. Longer trips in terms of distance are more likely to be undertaken by automobiles for convenience and faster movement speed, while shorter trips are easier to be done by public and active transport, given the necessary infrastructure.

3.4 Analysis Methods

3.4.1 Data Manipulation and Transformation

The databases and shapefiles were manipulated using Excel, Python, ESRI ArcMap, and R to prepare for analysis. To begin, I computed the commute distance variable using the CT boundary shapefile. I used the "feature-to-point" tool on ArcMap to find the centroid of each CT polygon, which were used as both points of origins and destinations in an OD Cost Matrix analysis to measure the least-cost paths along the street network. This table was then exported to be joined with other databases using their shared unique value, the CT identification (CTUID), creating a master database.

Any trips that originate or end outside of the study area were removed from the database using a Python algorithm. I created new fields to code binary values that determined whether each trip was a long commute (distance >= 70km; 1 = long commute), and whether a commute crosses CMA boundary (origin CMA = destination CMA; 1 = cross-boundary commute). The master database was then summarized by the origin CT and the origin CMA using Excel pivot tables, creating the database to be used for all spatial and statistical analysis.

Finally, employment to population ratio for each CMA was calculated by dividing outgoing commuters by incoming commuters to answer the second research question. The coefficient of 1 means that the CT has the same amount of worker as jobs, a coefficient larger than 1 means that there are more workers than jobs, and a coefficient less than 1 means there are more jobs than workers.

3.4.2 Spatial Analysis

I used the manipulated database that contained data summary by the origin CT for spatial analysis in ArcGIS. Choropleth maps were used to visualize the variables, including average commute distance per CT, median household income per CT, proportion of single detached housing per CT, and proportion of car commuters per CT. I also used choropleth maps to understand the GGH's commute pattern by illustrating the destination CTs for commutes originating from each CMA. I adopted Jenks Natural Breaks Classification for choropleth maps that illustrate proportion and quantile classification for absolute numbers.

3.4.3 Statistical Analysis

Any trips with missing data were excluded during the statistical analysis. I conducted several different statistical analyses at $p<0.01$ using the STATA software to understand the relationship between the independent and dependent variables. I began the analysis with correlation matrices to determine the variables' linear correlation with one another. Values of the correlation coefficient vary from -1 to +1, with +1 indicating a perfect positive correlation, -1 indicating a perfect negative correlation, and 0 indicating the absence of linear association between variables.

I also conducted one-way Analysis of Variance (ANOVA) tests to evaluate if the mean commute distance were different amongst a) CMAs, and b) income groups. Post-hoc Tukey tests were used to learn the specific groups that were significantly different or similar.

Lastly, I undertook multivariate regression analyses to understand how much variability in commute distance were explained by each independent variable. CMA dummy variables were later added to confirm previous spatial results. I also ran a Variance Inflation Factor (VIF) to test for multicollinearity of the independent variables.

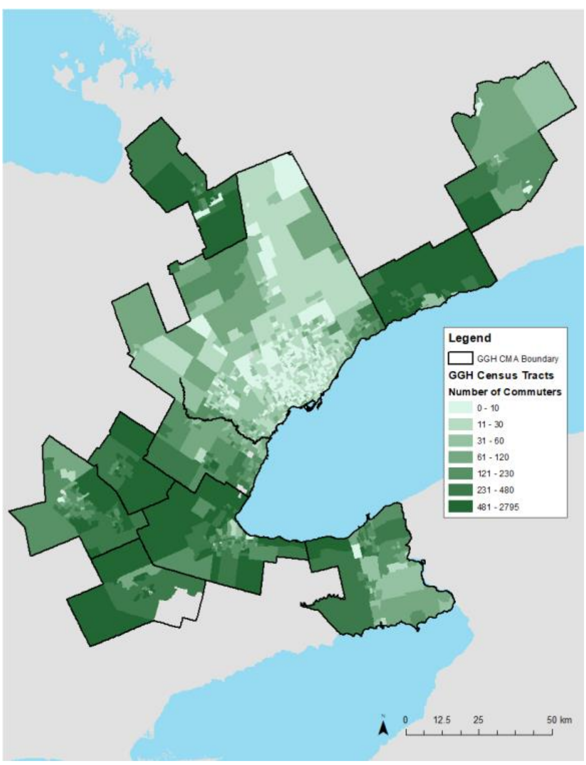
CHAPTER 4: RESULT

I discuss the results in this chapter. I begin by outlining the commute pattern in the GGH, including where people work, the average commute distance, and identifying the geography of long commutes. In section 4.2, I expand on the jobs-housing balance results of the GGH. I end this chapter with each independent variable's influence on commute distance, including median household income, dwelling type, and mode share.

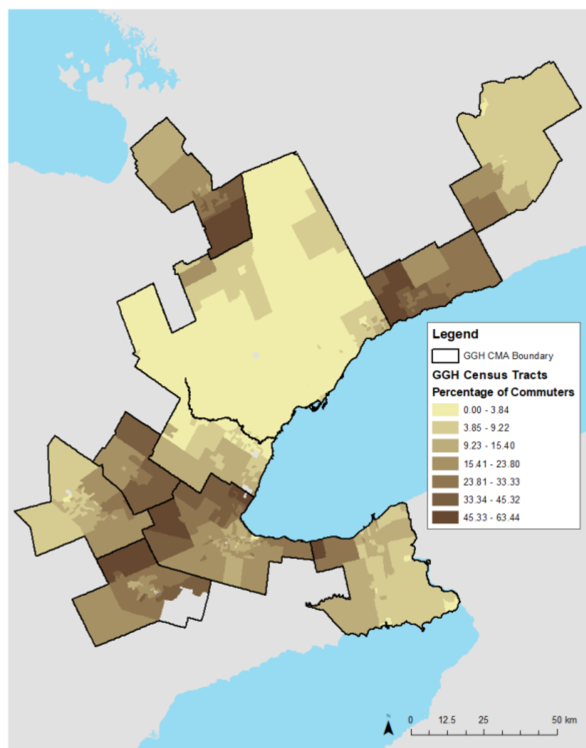
4.1 Commute Pattern

4.1.1 Where are people's working?

Preliminary spatial analysis first investigated the number of commuters working in a different CMA from their home CMA. Map 4.1 and Map 4.2 illustrate the geography of commuters who work in a different CMA. From visual inspection, it seems that CTs near the boarders of Oshawa, Barrie, Hamilton, Guelph, and Brantford tend to have many commuters working in another CMA. This is expected as distance for cross-CMA commutes are shorter near the boarders.



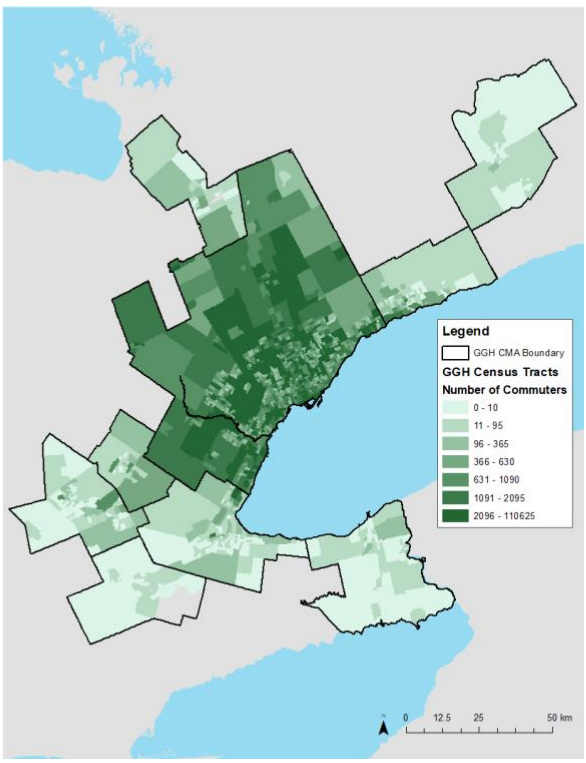
Map 4. 1 Number of Commuters who work in a different CMA



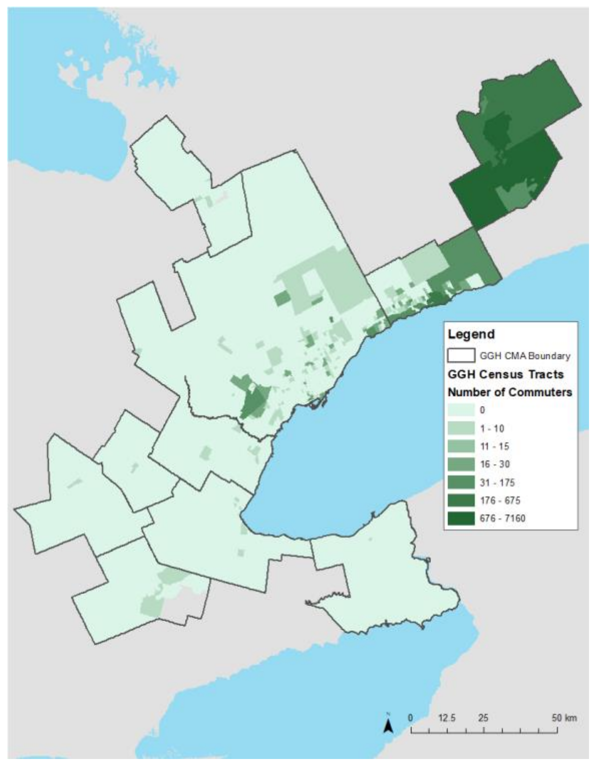
Map 4. 2 Percentage of Commuters who work in a different CMA

Table 4.1 breaks the commute pattern down further by presenting the specific CMA destinations of commuters from each CMA, while Maps 4.3 to 4.11 reveals the specific CT destinations.

Toronto has the least proportion of their residents working in another CMA, at around 2.03% (Map 4.3). This result is expected as it is the largest CMA in the GGH geographically, providing the vast majority job opportunities as the center city of the region. Peterborough has the second highest proportion of its residents that live and work in the same CMA at around 92.44%, while the remaining 7% commuters' destinations are split between Oshawa and Toronto (Figure 4.4).



Map 4. 3 Toronto's Commute Destination

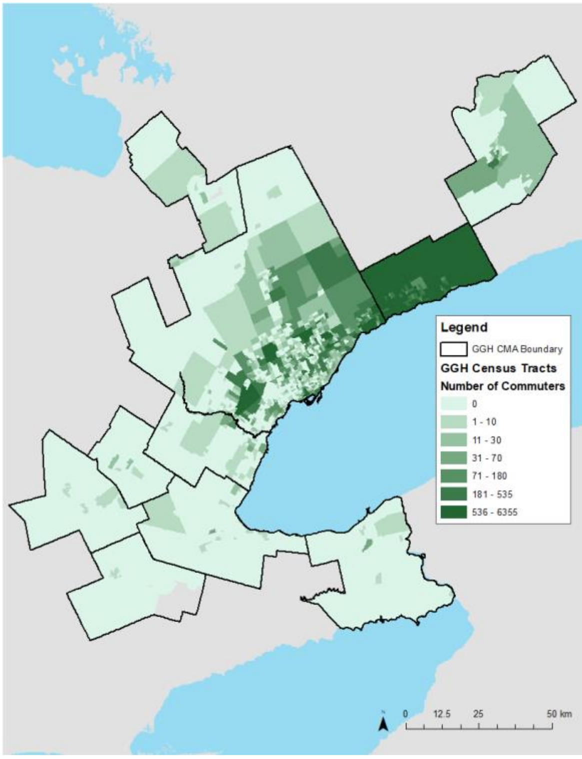


Map 4. 4 Peterborough's Commute Destination

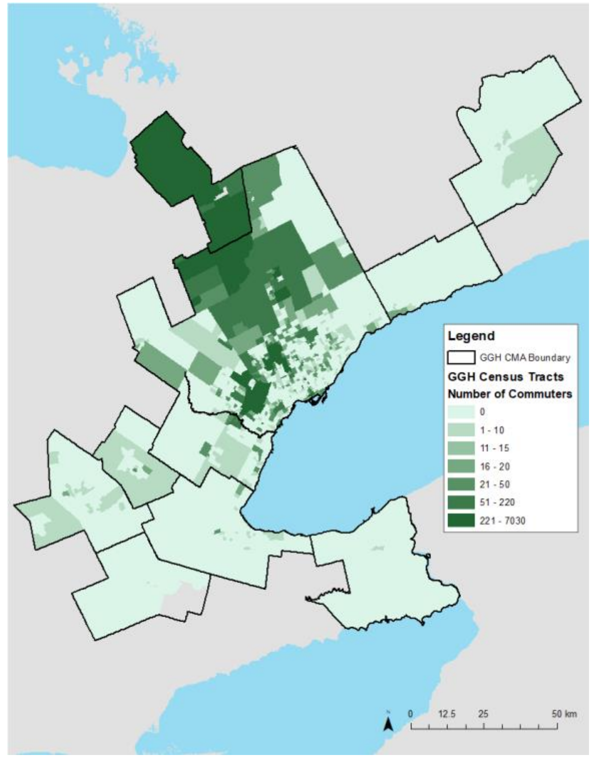
On the other hand, Oshawa has the most amount of its residents working in another CMA (41.71%), where Toronto employs 40.79% of Oshawa residents and mainly located near the two CMA's shared border and across the Southern part of Toronto (Map 4.5). Barrie has the second highest proportion of its residents working in another CMA (30.19%), and Map 4.6 shows that many of them work near the northern part of Toronto, but there are also many that travel to the Southernmost side of Toronto.

Table 4. 1 Commute destinations of CMAs (origin as row, destination as column)

Origin\Destination	Peterborough	Oshawa	Toronto	Hamilton	St. Catharines - Niagara	Kitchener-Cambridge-Waterloo	Brantford	Guelph	Barrie
Peterborough	**92.44**	3.69	3.58	0.13	0.02	0.02	0.02	0.02	0.07
Oshawa	0.60	**58.28**	40.79	0.13	0.06	0.05	0.01	0.04	0.04
Toronto	0.03	0.51	**97.97**	0.84	0.07	0.25	0.02	0.17	0.15
Hamilton	0.01	0.04	20.32	**75.02**	1.58	1.22	1.04	0.75	0.02
St. Catharines - Niagara	0.00	0.06	3.37	6.67	**89.41**	0.26	0.09	0.12	0.02
Kitchener-Cambridge-Waterloo	0.00	0.02	4.96	1.27	0.06	**86.57**	0.89	6.22	0.01
Brantford	0.02	0.02	4.39	11.73	0.15	8.56	**74.15**	0.99	0.00
Guelph	0.03	0.01	11.83	1.99	0.11	10.41	0.14	**75.46**	0.01
Barrie	0.05	0.06	29.51	0.23	0.01	0.18	0.01	0.12	**69.81**



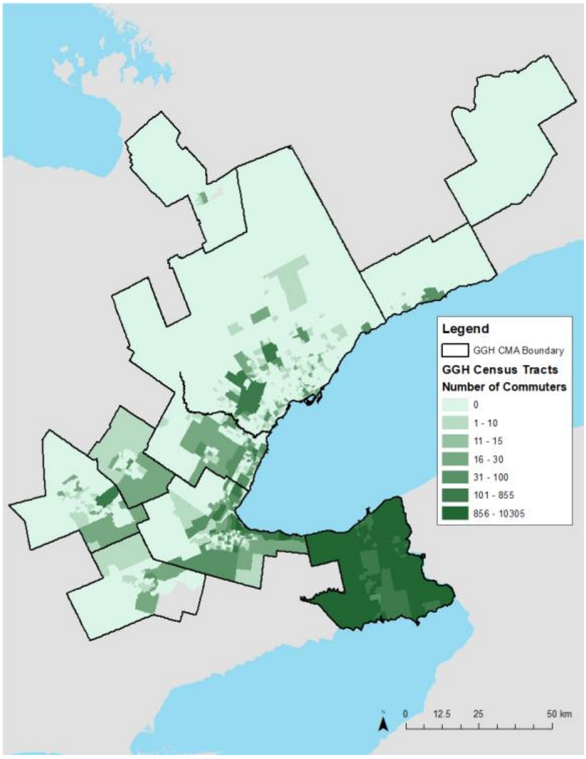
Map 4. 5 Oshawa's Commute Destination



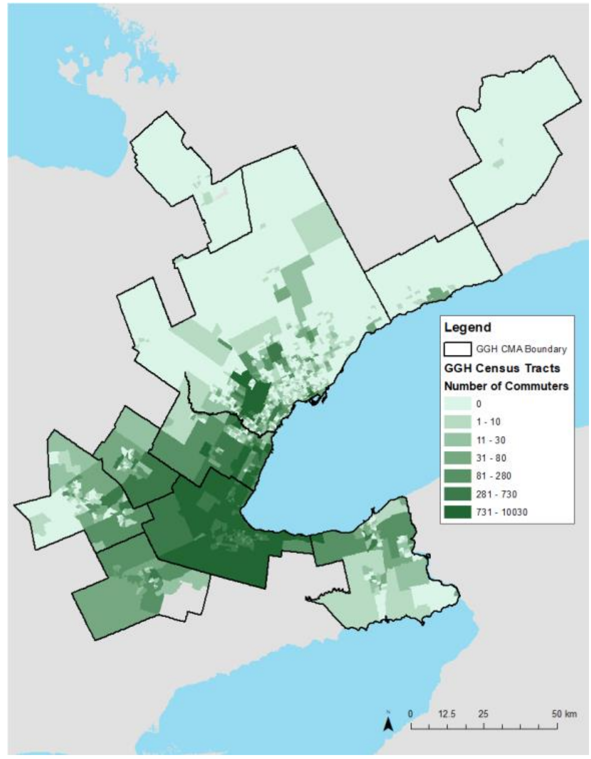
Map 4. 6 Barrie's Commute Destination

While CMAs to the North and East of Toronto exhibit cross-CMA commutes that largely complete in Toronto, the CMAs west of Toronto facilitate more cross-CMA commutes with each other. Furthest from Toronto is St. Catharines – Niagara. Located near the Canadian-U.S. border, it has the third highest proportion of its residents working within the CMA (89.41%). Hamilton is the most popular destination for commuters outside its own CMA (6.67%), while it has the least proportion of commuters travelling to Toronto in this entire region (3.37%). Map 4.7 shows that these workers travel to work in CTs along Lake Ontario.

Hamilton is the second largest CMA in the GGH region with many employment opportunities. However, Map 4.8 shows that 20.32% of Hamilton residents commute to Toronto for work, mainly to the west and south side of the Toronto CMA. Indeed, Hamilton is highly interconnected with Toronto as the two cities form the Greater Toronto and Hamilton Area (GTHA). Many Hamilton commuters also work in neighbouring CMAs including Guelph, Kitchener-Cambridge-Waterloo, and Brantford.



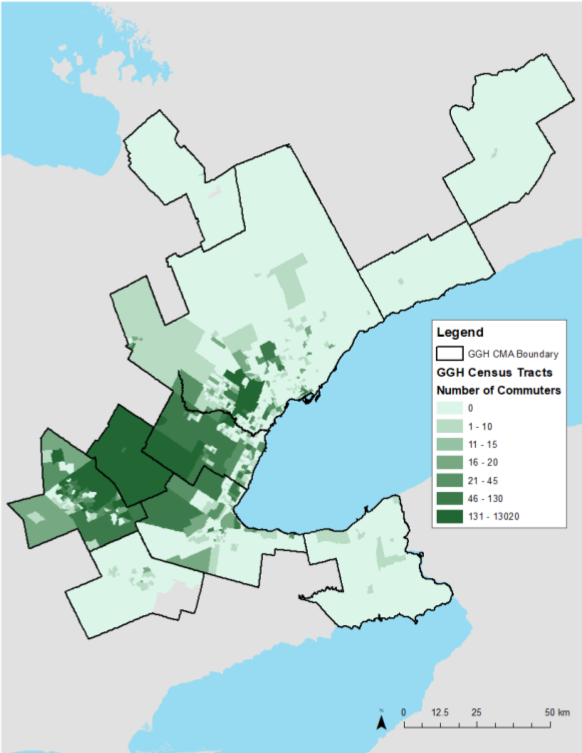
Map 4. 7 St. Catharines – Niagara's Commute Destination



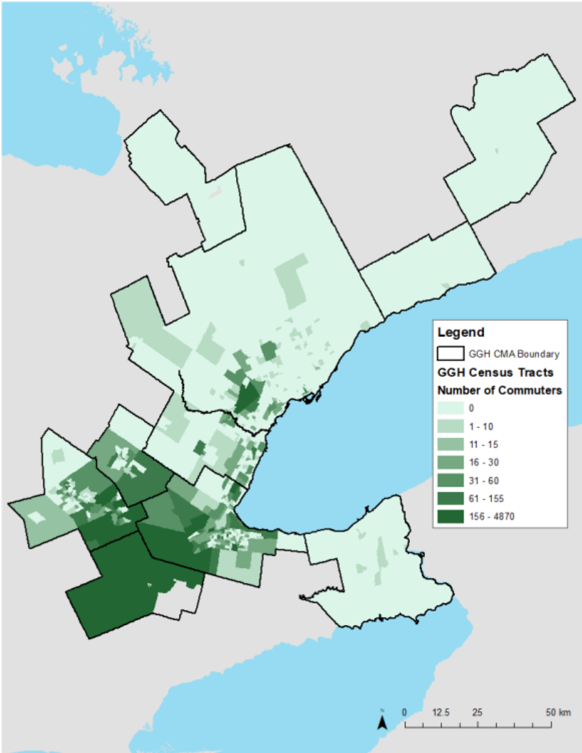
Map 4. 8 Hamilton's Commute Destination

Guelph (Map 4.9) and Brantford (Map 4.10) share similar commute pattern, each have around 75% of intra-CMA commute, and Kitchener-Cambridge-Waterloo is the third most popular CMA destination for both Guelph and Brantford residents (8.56% from Brantford and 10.41% from Guelph). Brantford's second most popular commute destination is Hamilton (11.73%) while Guelph's is Toronto (11.83%).

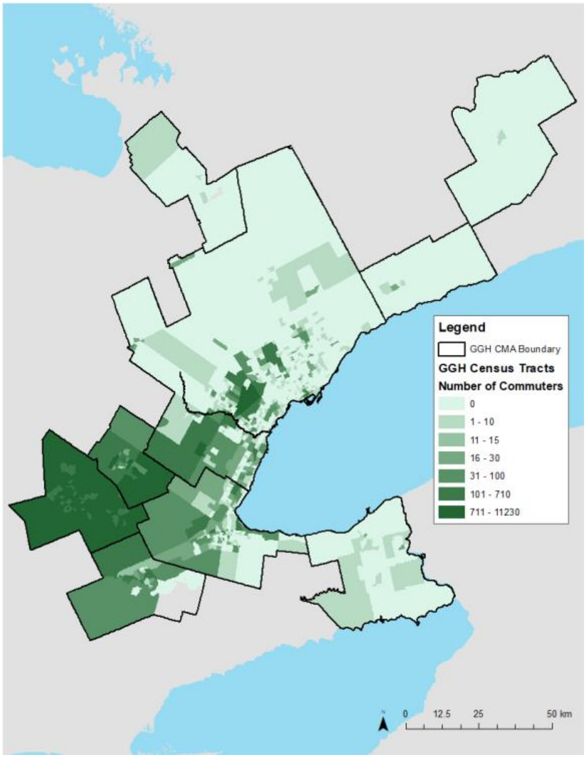
Lastly, Kitchener-Cambridge-Waterloo commuters has the third highest proportion of intra-CMA commute in the region (86.57%), while the remaining work in Guelph (6.22%) and Toronto (4.96%). Map 4.11 illustrate the CT destinations of Kitchener-Cambridge-Waterloo workers.



Map 4. 9 Guelph's Commute Destinations



Map 4. 10 Brantford's Commute Destinations



Map 4. 11 Kitchener-Cambridge-Waterloo's Commute Destination

4.1.2 Average Commute Distance

The histogram presented in Figure 4.1 showcase that the commute distance in the GGH has a positively skewed distribution to the right side. Most trips in the GGH are around zero to nine km (around 1,750,000 commuters), the number of commuters exponentially decreases at every 10km distance intervals.

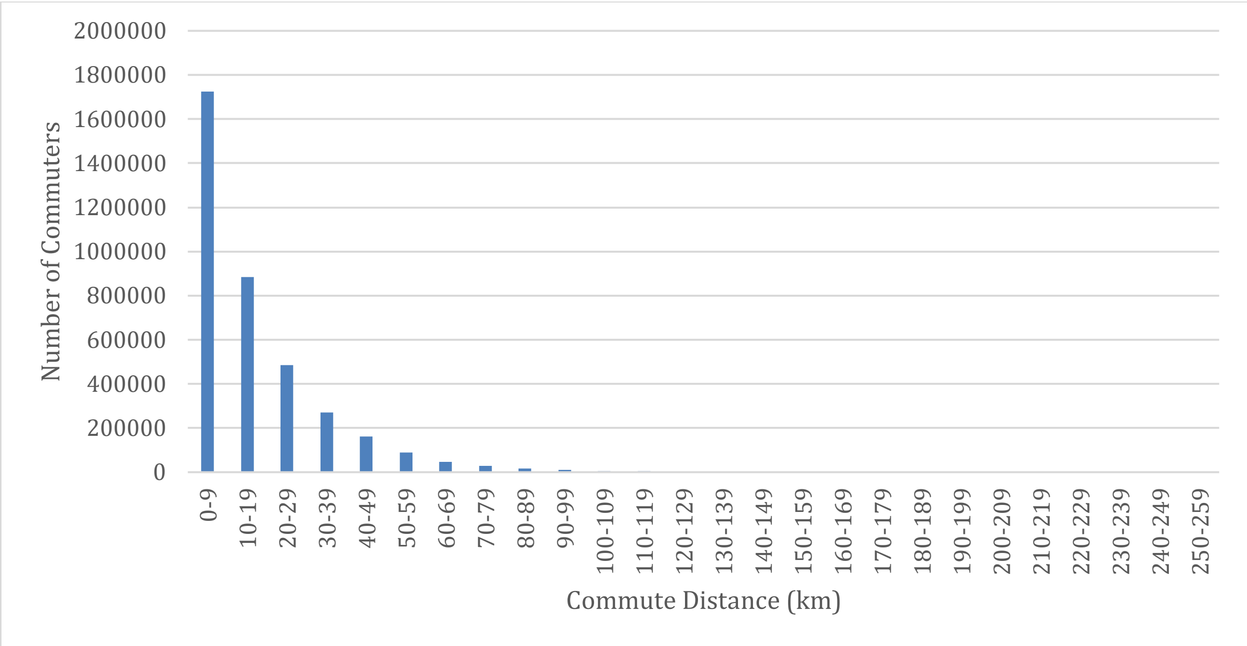
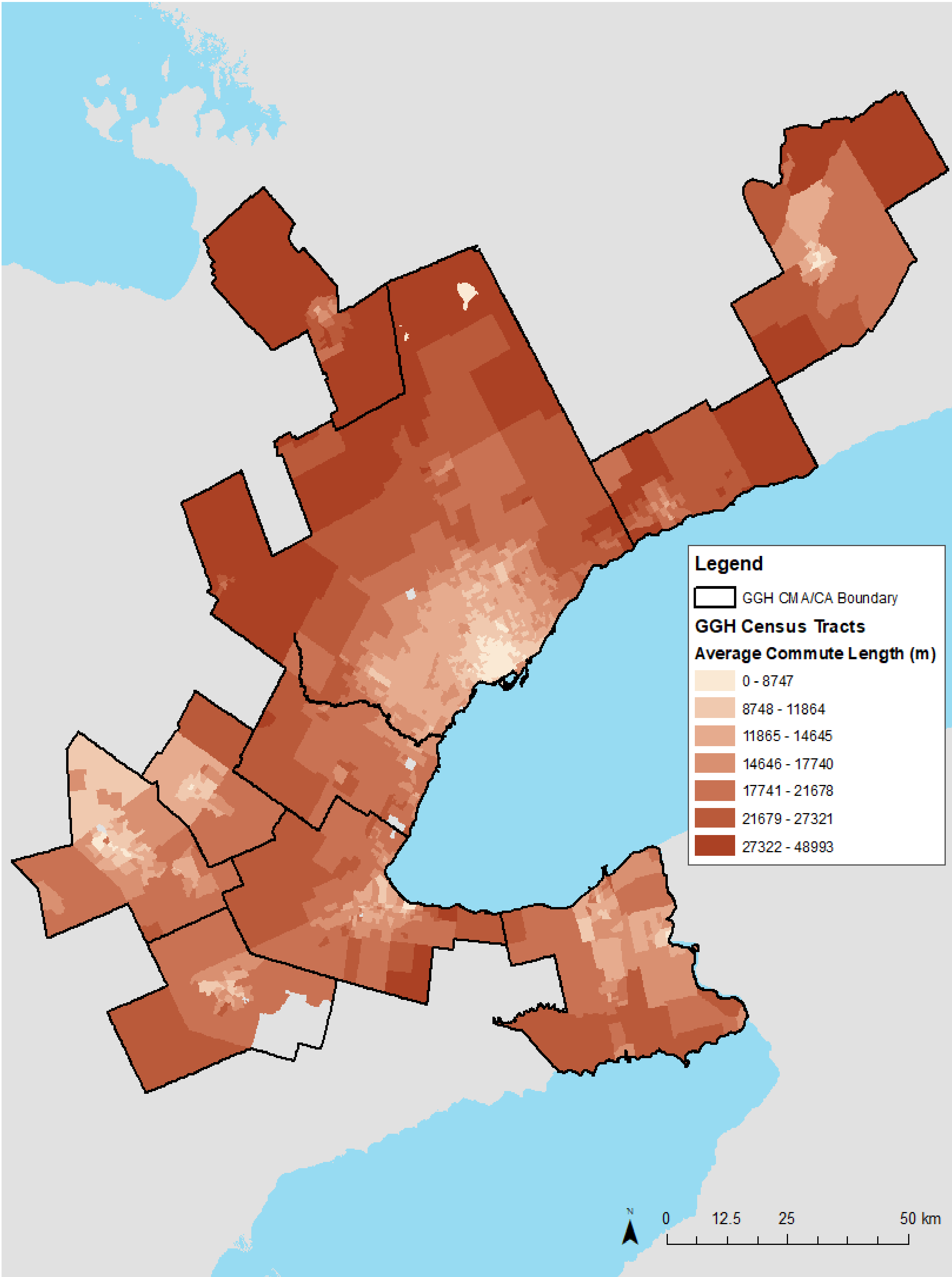


Figure 4. 1 Histogram of commute distance in km

A spatial pattern can be identified through Map 4.12: CTs located in the city centers have shorter average commute time, and commute time increases as distance from the core increases. I ran an ANOVA test to understand whether the average commute distance is significantly different amongst CMAs (Table 4.2). Those who live in Barrie and Oshawa have the longest commutes at 25.23 km and 23.29 km, respectively. On the other hand, commuters from Kitchener-Cambridge-Waterloo have the shortest average commute distance at around 13.68 km. The ANOVA reports a statistically significant result at $F<0.01$, meaning that there exists a significant difference of average commute distance amongst CMAs. The post-hoc Tukey test found that Barrie and Oshawa have no significant difference with each other, but they have significantly longer commute distance than every other CMA (Table 4.3). Another important observation is that Kitchener-Cambridge-Waterloo, Brantford, Guelph, Toronto, and Peterborough have no statistically significant differences with each other.



Map 4. 12 Average Commute Distance by CT (km)

Table 4. 2 ANOVA test for average commute length by CMA

	Summary of Average Commute Distance (km)		
CMA	**Mean**	**Standard Deviation**	**Count**
Barrie	25.23	6.85	42
Brantford	16.29	4.26	29
Guelph	13.76	3.89	30
Hamilton	17.38	3.58	189
Kitchener-Cambridge-Waterloo	13.68	3.54	107
Oshawa	23.29	4.47	84
Peterborough	15.49	7.17	30
St. Catharines-Niagara	16.51	3.76	94
Toronto	15.28	5.49	1,146
Overall	**16.09**	**5.62**	**1,751**

Source	SS	Df	MS	F	Prob>F
Between groups	9.7396e+09	8	1.2175e+09	46.54	0.0000
Within groups	4.5569e+10	1742	26158958.2		
Total	**5.5309e+10**	**1750**	**31604878**		

Table 4. 3 Post-hoc Tukey test for average commute distance by home CMA (* indicates significance at p<0.01)

Origin\Destination	Peterborough	Oshawa	Toronto	Hamilton	St. Catharines - Niagara	Kitchener-Cambridge-Waterloo	Brantford	Guelph	Barrie
Peterborough		*	-	-	-	-	-	-	*
Oshawa			*	*	*	*	*	*	-
Toronto				*	-	-	-	-	*
Hamilton					-	*	-	*	*
St. Catharines - Niagara						*	-	-	*
Kitchener-Cambridge-Waterloo							-	-	*
Brantford								-	*
Guelph									*
Barrie									

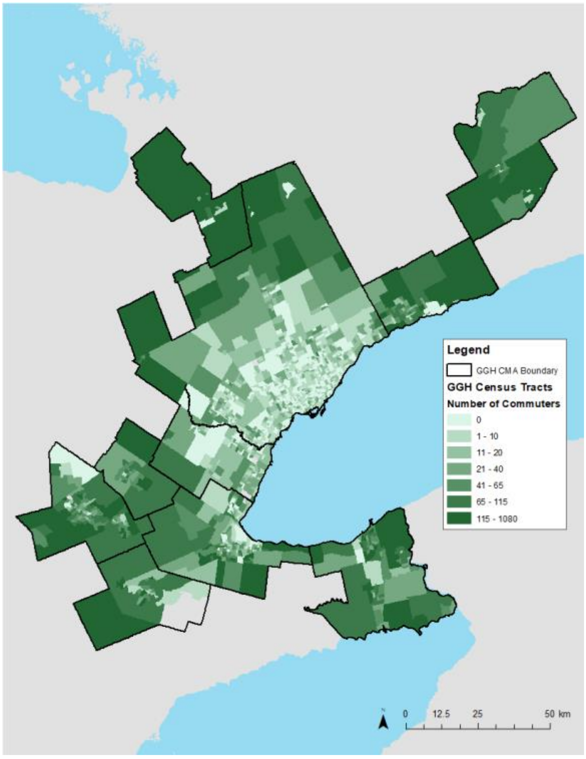
4.1.3 Long Commute

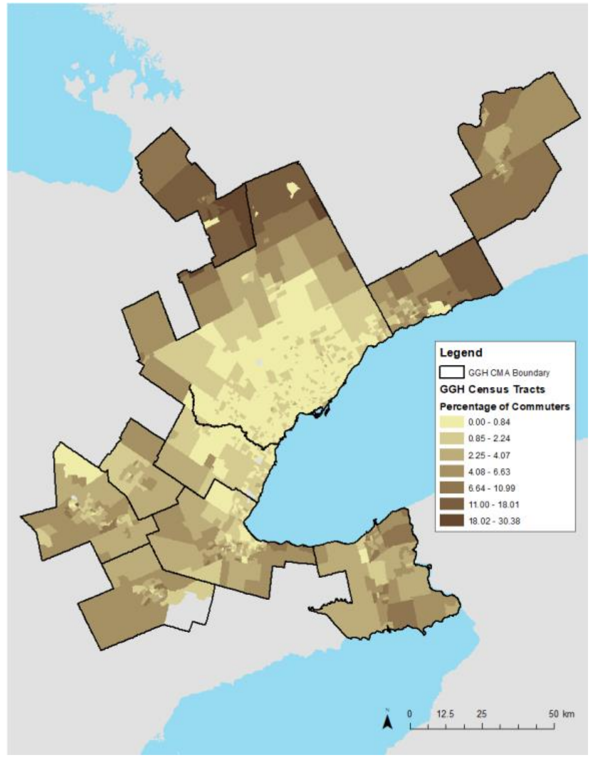
Table 4.4 illustrates the proportion and number of long commuters (commute distance >= 70 km) in each CMA. Barrie has the highest proportion of long commuters (15.38%) while every other CMA has around or less than 5%. Toronto has the least proportion of commuters travelling over 70km at only 0.8%, but the highest absolute number (20,675 people). Surprisingly, Peterborough has the second highest proportion of its residents who has long commutes, even though earlier analysis show that the majority of the people work within the same CMA. Thus, a high proportion of inter-CMA commute that originate from Peterborough are likely to be long commutes.

Table 4. 4 Percentage & absolute number of long commutes from each CMA

CMA Name	Percentage	Absolute Number
Barrie	15.38	12,480
Brantford	4.03	2,190
Guelph	2.98	2,145
Hamilton	2.93	9,465
Kitchener-Cambridge-Waterloo	3.42	8,120
Oshawa	5.03	8,030
Peterborough	5.52	2,505
St. Catharines-Niagara	4.60	7,505
Toronto	0.80	20,675
Overall	**1.96**	**73,115**

Map 4.13 and Map 4.14 showcase the number and percentage of long commuters by CT. Similar to the average commute results (Map 4.12), most of those with long commutes reside in peripheral areas that are far from the city center.





Map 4. 13 Number of Long Commuters **Map 4. 14 Percentage of Long Commuters**

4.2 Jobs-Housing Balance

Jobs-housing balance has been established as a direct cause of long commutes in existing literature. An area with an unproportionate number of jobs versus housing could result in residents seeking employment opportunities in another area, resulting in a longer commute distance. In this thesis, I conducted a simplified employment to population ratio within the boundary of each CMA to establish an elementary understanding. The number of employment opportunities are the number of commute trips that end in a CMA, while the population is the number of commute trips that originate from a CMA. The ratio was calculated through dividing the number of employment opportunities from the population. Thus, a ratio of 1 means that there is a perfect balance between workers and jobs, a ratio less than 1 means that there are more workers than jobs, and a ratio over 1 means that there are more jobs than workers. Of course, it is impossible for every CMA to have an equal number of jobs and residents, thus, I argue that CMAs with a ratio between 0.90 to 1.10 has a relative even balance.

The ratio of each CMA in the GGH is presented in Table 4.5. Only two CMAs have more jobs than workers, which are Guelph (0.94) and Toronto (0.95). In addition to these CMAs, Kitchener-Cambridge-Waterloo, Peterborough, and St. Catharines-Niagara also have a relatively even balance of employment opportunities to population. This result echoes the previous sections as Oshawa and Barrie have the highest ratio at 1.47 and 1.34, respectively. Overall, the result reinforces the narrative that areas with an unbalanced distribution of jobs and housing, or more housing than jobs, have longer commutes.

Table 4. 5 Employment to population ratio in each CMA

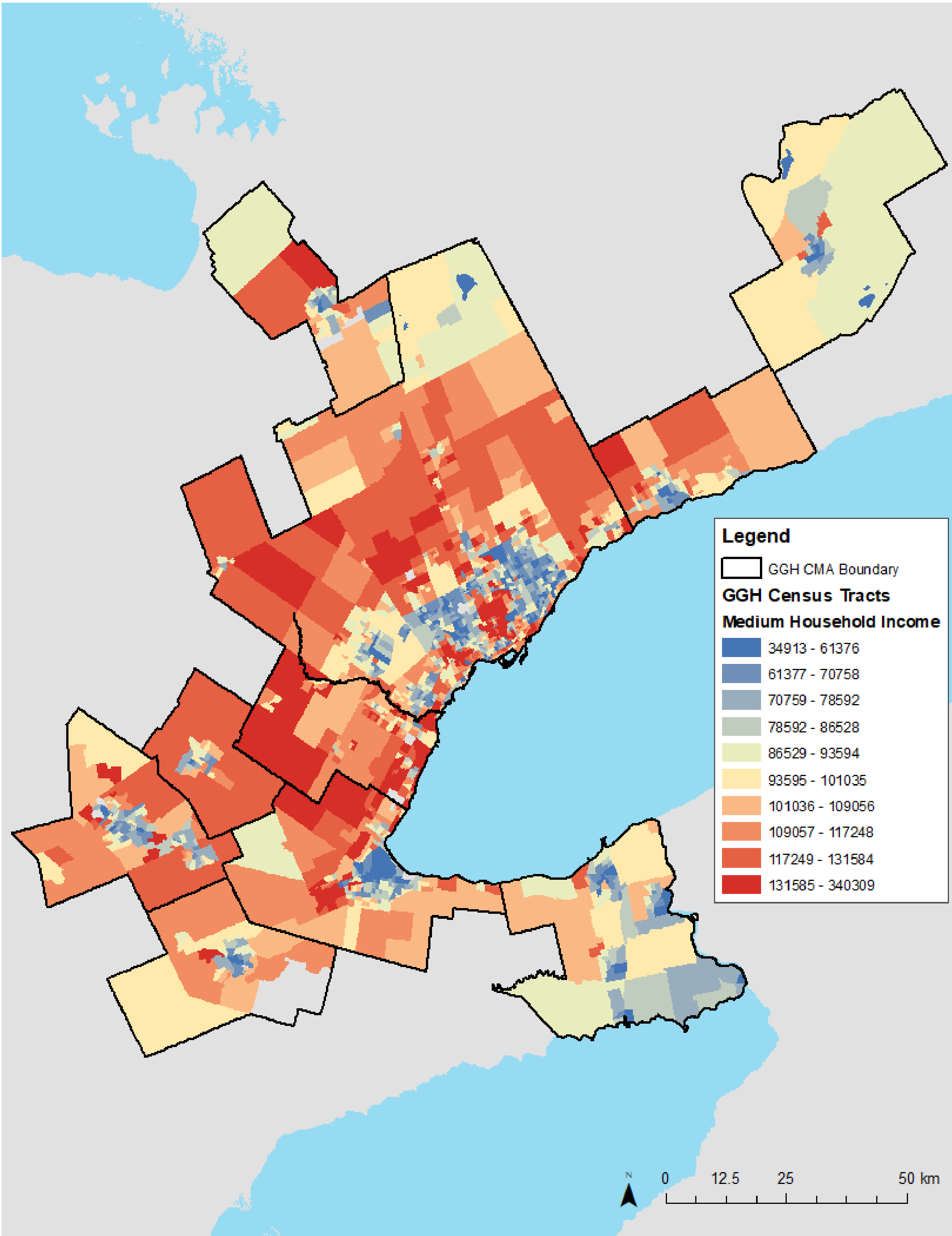
CMA Name	Population	Jobs	Ratio
Barrie	81,130	60,700	1.34
Brantford	54,285	465,65	1.17
Guelph	71,955	76,870	0.94
Hamilton	322,955	286,280	1.13
Kitchener-Cambridge-Waterloo	237,490	228,830	1.04
Oshawa	159,590	108,265	1.47
Peterborough	45,365	43,690	1.04
St. Catharines-Niagara	163,290	153,295	1.07
Toronto	2,600,530	2,732,135	0.95

4.3 Socio-Demographic Influences of Commute Distance

4.3.1 Median Household Income

Results from the correlation matrix show that median household income has a statistically significant positive correlation with average commute distance. However, the correlation is weak as its coefficient is 0.1192. To further understand the relationship between the two variables, median household income was divided into decile groups to facilitate the one-way ANOVA, with "1" coded as CTs with the lowest 10%, and "10" coded as the highest 10% (Appendix B). Map 4.15 shows that each CMA has distinct clustering pattern for income level. Upon visual inspection, I have identified the lower-income deciles to be clustered near the city center of each CMA, with the exception of Toronto, as the pattern is more scattered across the east and west side.

The ANOVA result confirms that that differences amongst income groups is statistically significant at $F<0.01$ (Table 4.6). The income group with the highest average commute distance is the 8th group (median household income between \$109,269 and \$117,325), commuting an average of 18.71 km. Meanwhile, the 1st group (median household income between \$34,912 and \$61,376) has the shortest commute, travelling around 12.95 km. A general parabola shape can be found, where the people with the lowest and the highest income have shorter commutes, and those in the middle decile groups have the longest commute (groups 6, 7, 8, 9).



Map 4. 15 Geography of Average Household Income in 2015 ($) per CT

Table 4. 6 ANOVA test for Average Commute Distance and Income Decile

10 Quantiles of Medium Household Income	Summary of Average Commute Distance (km)	
	Mean	**Standard Deviation**
1	12.95	3.16
2	13.90	3.21
3	14.20	3.65
4	14.47	4.20
5	17.40	5.72
6	18.06	6.52
7	17.97	6.13
8	18.71	5.94
9	17.84	6.01
10	15.47	6.30
Overall	**16.10**	**53.60**

Source	SS	Df	MS	F	Prob>F
Between groups	7.0464e+09	9	782934789	28.47	0.0000
Within groups	4.7732e+10	1736	27495593.1		
Total	**5.4779e+10**	**1745**	**31391841.1**		

The post-hoc Tuckey revealed that income group 1, 2, 3, and 4 are not statistically different with each other (Table 4.8). Moreover, income group 5, 6, 7, 8, 9, are not statistically different with each other. Lastly, income group 10 has significant difference with 1, 5, 6, 7, and 8, but not with 2, 3, 4, and 5. This confirms that income 1 to 4 and 10 has a shorter commute than income level 5-8.

Table 4. 7 Summary of Significance Level for Tukey Test (* indicates significant at p<0.01)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1		-	-	-	*	*	*	*	*	*
2			-	-	*	*	*	*	*	-
3				-	*	*	*	*	*	-
4					*	*	*	*	*	-
5						-	-	-	-	-
6							-	-	-	*
7								-	-	*
8									-	*
9										*
10										

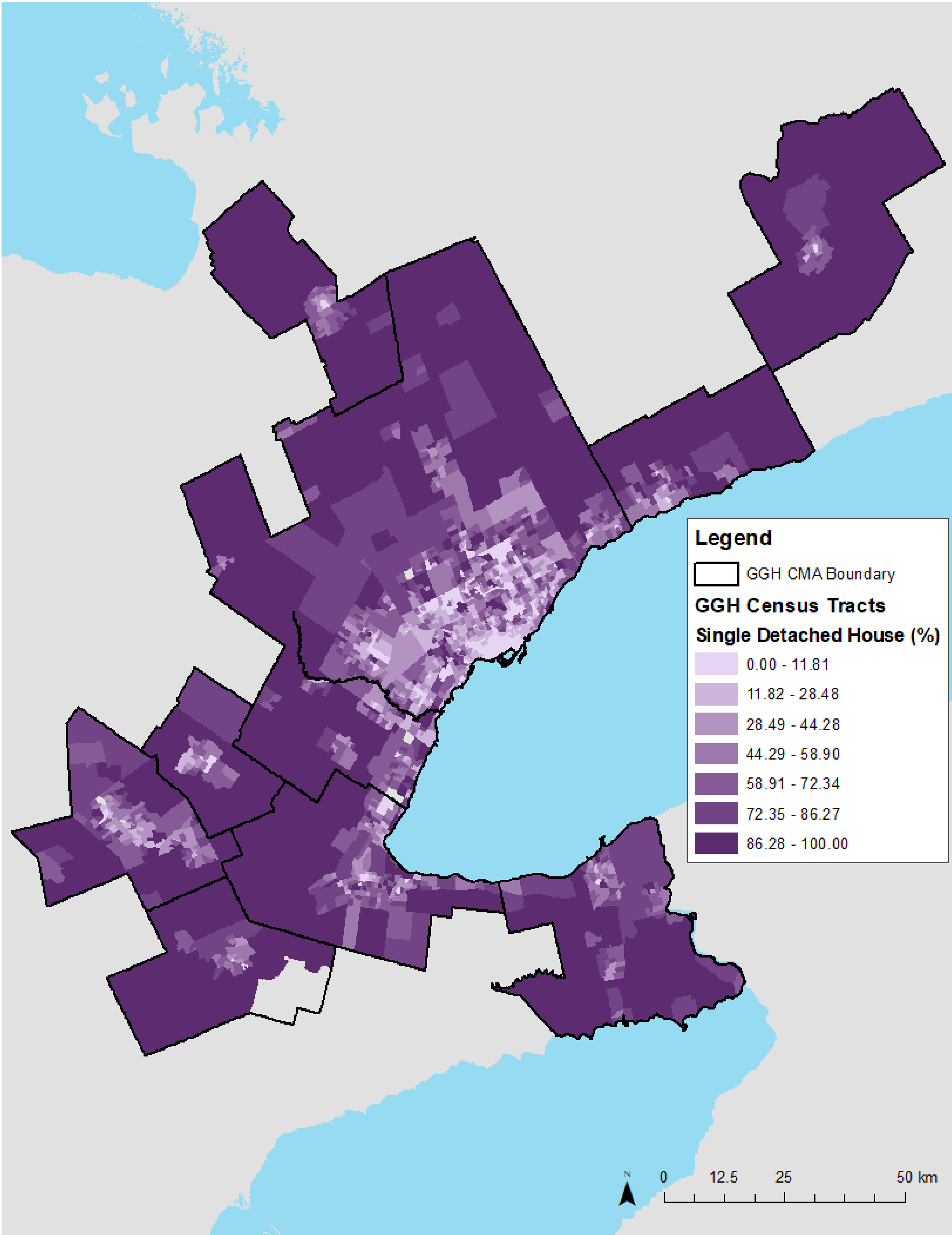
4.3.2 Dwelling Type

Each CT's dwelling type, measured as percentage of single detached housing, is exhibited in Map 4.16. Excluding Toronto, higher proportion of single detached housing seems to be more dominant in the outskirt of each CMA and lower within the centers. The pattern in Toronto demonstrates more random distribution.

I conducted a correlation matrix between average commute distance and dwelling type (Table 4.8). Seven different dwelling variables entered into the matrix as the independent variables. All variables except for "semi-detached house" are significant at p<0.01. Two variables were positively correlated to average commute distance: "single detached house" has a strong correlation (0.54) while "row house" has a weak correlation (0.19). All four remaining dwelling type variables are negatively correlated with average commute distance, where "apartment in building that has five or more storeys" and "apartment in a building that has less than five storeys" has medium correlation (-0.42 and -0.43 respectively), while "apartment or flat in a duplex" and "other single-attached house" has weak correlation (-0.11 and -0.10, respectively). This result is expected as high percentage of single detached houses indicate low density residential developments, which are also characterized by separated land use.

Table 4. 8 Correlation Matrix of Average Commute Distance and Housing Type (* indicates significance at p<0.01)

	Single Detached House	**Apartment in a building that has five or more storeys**	**Semi-detached house**	**Row house**	**Apartment or flat in a duplex**	**Apartment in a building that has fewer than five storeys**	**Other single-attached house**
Average Commute Distance	0.5431* 0.0000	-0.4231* 0.0000	-0.0358 0.1341	0.1912* 0.0000	-0.1144* 0.0000	-0.4308* 0.0000	-0.0992* 0.0000



Map 4. 16 Geography of single detached houses (%) per CT

4.3.3 Mode Share

Mode has a bidirectional relationship with commute distance. On one hand, those who have long commutes are more likely to use automobiles, while on the other hand, those who own automobiles have the option to live further from work as they can have faster movement speed.

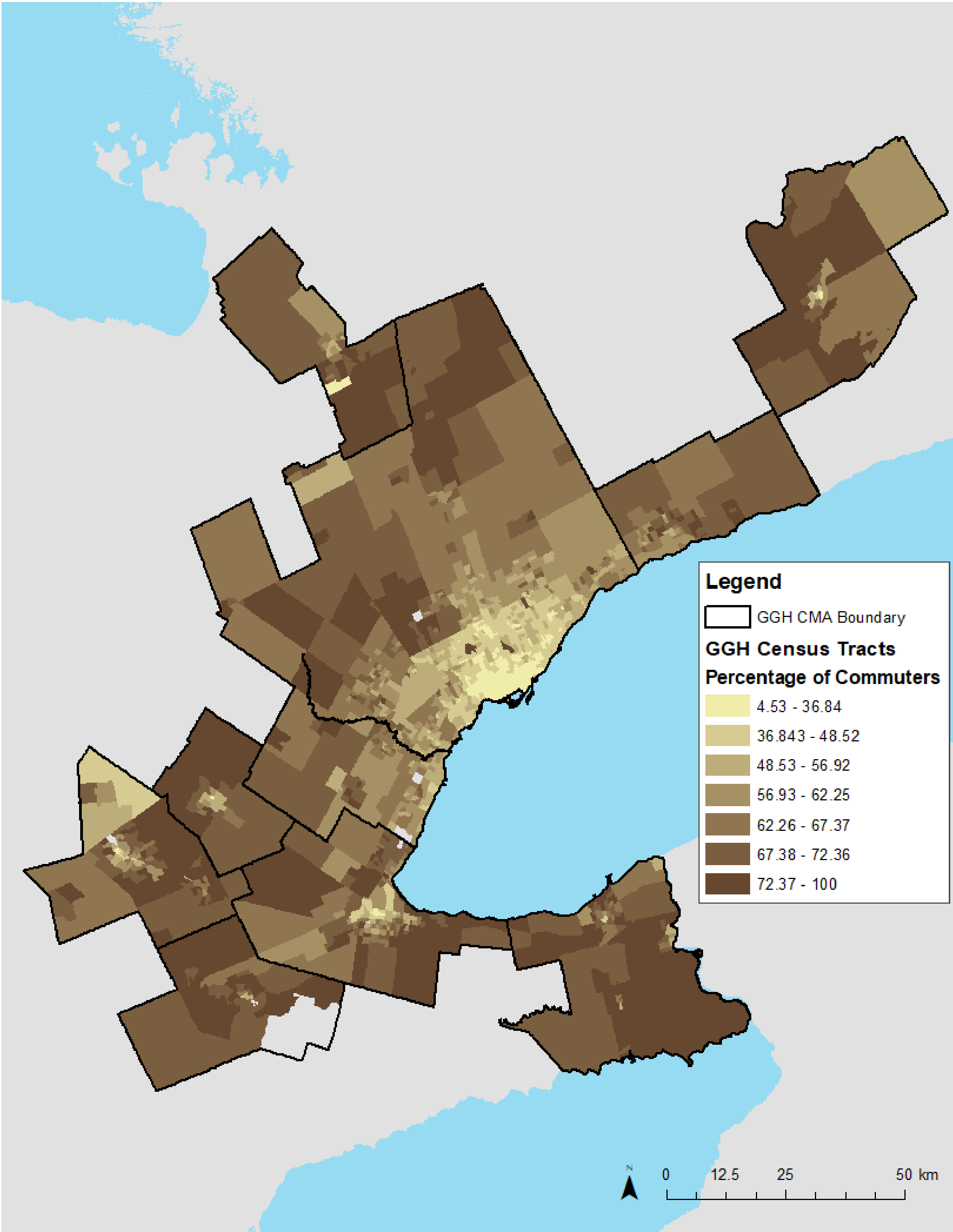
GGH's commute mode share by CMA is presented in Table 4.9 and the percentage of commuters who drive alone by CT is visualized in Map 4.17. Car, truck, or van remains the most common mode in the GGH for the daily commute, making up around 66.58% of total trips, while 18.33% of commute are by public transportation, 6.14% by active transport, and 8.99% by other modes (including working from home). Moreover, around 56.05% of all commuters drive alone, which is the most unsustainable mode.

Out of all CMAs, Brantford has the largest proportion of commuters who drive (84.31%), while Toronto has the least proportion of commuters who drive (60.73%). This finding reflects the geography of existing sustainable transport infrastructure. From the map, it is clear that the CTs in the City of Toronto, and the downtowns of various CMAs have the smallest proportion of commuters who drive alone. Overall, GGH commuters are highly car dependent.

Table 4. 9 Commute mode share (%) by home CMA (* indicates significance at p<0.01)

Origin CMA Name	Car, Truck, or Van*	Driver, Alone	Public Transport **	Active Transport ***	Other****
Barrie	81.31	68.77	4.59	4.25	9.17
Brantford	84.31	72.10	3.10	4.85	8.05
Guelph	78.02	66.10	6.25	6.80	8.92
Hamilton	77.29	66.23	9.64	5.12	8.49
Kitchener-Cambridge-Waterloo	81.01	69.12	5.84	5.37	7.98
Oshawa	78.89	67.65	9.67	3.41	8.13
Peterborough	77.68	65.39	3.92	8.52	10.47
St. Catharines-Niagara	84.15	72.84	2.58	5.33	8.25
Toronto	60.73	50.66	23.396	6.58	9.23
Overall	66.58	56.05	18.33	6.14	8.99

*Includes: driver, alone; 2 or more persons shared the ride to work; driver, with 1 or more passengers; passenger, 2 or more persons in the car
**Includes: subway or elevated rail (Toronto Subway); light rail, streetcar or commuter train (Toronto streetcars, Toronto GO Train); passenger ferry
***Includes: walking; bicycle
****Includes: work from home



Map 4. 17 Proportion of commuters that drive alone by CT

The result of the correlation matrix between average commute distance and modal share confirms the previous statement (Table 4.10). All variables are significant at $p<0.01$. Average commute distance is positively correlated with auto commute types, while negative correlated with active transport, public transport, and other modes. The correlation with all modal variables is strong as their coefficients surpasses ±0.5. "Driver, alone" has the strongest positive correlation with average commute distance, while "other modes", which include working from home, has the strongest negative correlation.

Table 4. 10 Correlation between average commute distance and modal variables (* indicates significance at p<0.01)

Variables	**Auto**	**Driver, Alone**	**Active Transport**	**Public Transport**	**Other Modes**
Average Commute Distance (m)	0.6205* 0.0000	0.6238* 0.0000	-0.5248* 0.0000	-0.5078* 0.0000	-0.0652* 0.0064

4.3.4 Multivariate Regression Analysis

The last statistical analysis was a multivariate regression analysis. I entered a model with average commute length as the dependent variable (y), and the independent variables (x) are median household income, percentage of auto mode share, and percentage of single detached house. The regression yielded an r-squared value of 0.4303, which means that the model explained 43.03% of the variation in average commute length across GGH's CTs (Table 4.11). This is a satisfying explanation considering only three independent variables are in this model. Surprisingly, median household income is not significant in the model at $p>0.01$. Meanwhile, 1% increase in auto mode accounts for 133.31 meters more average commute, and 1% increase in single detached homes in the CT accounts for 53.7 meters more on top of the constant variable, 5280.57 meters.

A VIF test for multicollinearity is conducted to understand whether the three independent variables are related to each other and found that multicollinearity is not an issue to this model.

Table 4. 11 Multivariate regression (distance in meters)

Average commute distance	**Coef.**	**S.E.**	**t**	**P >|t|**	**95% Conf. Interval**	
Income	-0.01	0.00	-2.14	0.03	-0.02	-0.00
Single detached house (%)	133.31	7.43	17.95	0.00	118.74	147.88
Auto mode (%)	53.73	5.45	9.86	0.00	43.05	64.42
Constant	5,280.57	532.98	9.91	0.00	4,235.22	6,325.91

As the previous ANOVA confirmed that average commute length is statistically different across CMAs, thus, I added the CMAs as dummy variables and used Toronto as the reference variable because it is the center city. All variables are significant at p>0.01 in this model (Table 4.12). However, this result is unexpected as it does not match the outcome of the prior ANOVA. Specifically, controlling for all three independent variables, every CMA except for Barrie and Oshawa commutes relatively less than Toronto in this regression. Whereas Hamilton, St. Catharines-Niagara, Brantford, and Peterborough actually have longer average commute distance than Toronto.

Table 4. 12 Multivariate regression with CMA dummy variables (Toronto as the reference variable, distance in meters)

Average commute distance	Coef.	S.E.	t	P >|t|	95% Conf. Interval	
Median household income ($)	-0.02	0.00	-4.61	0.00	-0.02	-0.01
Single detached house (%)	153.88	6.82	22.57	0.00	140.50	167.25
Auto mode (%)	54.44	4.87	11.18	0.00	44.89	63.98
Barrie	5354.33	605.80	8.84	0.00	4166.15	6542.51
Brantford	-3896.92	706.86	-5.51	0.00	-5283.31	-2510.53
Guelph	-4535.25	687.22	-6.60	0.00	-5883.12	-3187.37
Hamilton	-1164.81	304.00	-3.83	0.00	-1761.05	-568.58
Kitchener-Cambridge-Waterloo	-5126.91	388.59	-13.19	0.00	-5889.06	-4364.76
Oshawa	3807.34	430.56	8.84	0.00	2962.87	4651.80
Peterborough	-4367.48	699.57	-6.24	0.00	-5739.76	-2995.20
St. Catharines-Niagara	-4033.41	425.68	-9.48	0.00	-4868.32	-3198.51
Constant	5188.63	467.86	11.09	0.00	4271.01	6106.25

I hypothesize that this outcome may be due to the fact that cost of living and land-value is different in each CMA. Thus, I performed stratified regressions for each CMA to further analyze how the model performs differently across the region. The result is presented in Table 4.13. The r-squared value is the highest for Peterborough (the model explains 67% of all variation) while it is especially low for St. Catharines-Niagara (the model explains 15% of all variation).

Across all CMAs, median household income exerts very weak influence on commute distance, the variable's influence on commute distance ranges from -0.03 meters to 0.13 meters. The variable is only significant in Hamilton, Oshawa, and Toronto. Dwelling type is also only significant in three CMAs: Barrie, Peterborough, and Toronto. Its influence on commute distance is positive except for Hamilton and Oshawa. Mode is statistically significant at five CMAs:

Hamilton, Kitchener-Cambridge-Waterloo, Peterborough, and Toronto. Its influence on commute distance is positive except for Peterborough, where every percent increase of auto mode share in a CT subtracts 194.84 meters from the average commute distance.

Toronto is the only CMA where all three independent variables are statistically significant. On the other hand, neither of the variables are significant in Brantford, Guelph, and St. Catharines-Niagara.

Table 4. 13 Summary of Regression Analysis per CMA (* indicates significance at $p<0.01$; distance in meters)

CMA Name	Median Household Income ($)	Single Detached House (%)	Auto Mode (%)	Constant	R^2
Barrie	-0.01	174.18*	306.91	-10,456.53	0.52
Brantford	0.01	108.44	24.54	6,531.87	0.39
Guelph	0.10	10.46	63.12	-1,581.21	0.47
Hamilton	0.07*	-11.95	72.34*	6,453.63*	0.41
Kitchener-Cambridge-Waterloo	0.01	36.24	116.95*	1,951.16	0.35
Oshawa	0.13*	-15.91	132.61*	492.82	0.47
Peterborough	0.03	288.20*	-194.84*	7344.06	0.67
St. Catharines-Niagara	0.01	61.02	38.83	7961.05	0.15
Toronto	-0.03*	54.71*	161.65*	5749.28*	0.57

CHAPTER 5: DISCUSSION

In this final chapter, I summarize the results from the Chapter 4, compare them with current context and existing literature, and provide my opinion about possible policy intervention to reduce the number of long commutes in this region.

The purpose of this research is to investigate the long commutes in the Greater Golden Horseshoe, which is the fastest growing region in Canada. Three research questions are answered by this thesis. In section 5.1, I answer the first two research questions: what is the commute pattern in the GGH and where are the long commutes, and is there jobs-housing balance in the GGH? Sustainable growth requires the city to be able to supply job opportunities to its residents, as well as supply affordable and suitable housing to its workers. When a city fails to supply either one of the two, people have to sacrifice commute time to obtain both, influencing the long commute issue seen in the GGH. In section 5.2, I answer the third research question, how do socio-demographic factors, including income, dwelling typology, and mode influence commute distance? I conclude this thesis in section 5.3, where I discuss the limitation of this research and future research opportunities.

5.1 Answering RQ1 and RQ2:

The first research question explores the commute pattern in the GGH, and specifically, understanding the origin and destinations of commutes, the average commute distance and locating the long commutes in the region. It is interconnected with the second research question, determining whether there is a jobs-housing balance in the GGH. Thus, section 5.1 answers both of these questions.

On a CT level, the average commute distance is shorter if one lives closer to the downtown cores across CMA and is longer the further away from the cores. This finding is very reasonable as firms tend to cluster and form employment centers in downtown cores due to the benefits of agglomeration economies. Examples are Toronto’s Financial District, Downtown Hamilton, Downtown Barrie. These downtown cores have higher land value and denser residential development. Workers who live close to these employment centers are often employed there, thus having shorter commutes.

There are three top employment CMAs in this region: Toronto, Hamilton, and Kitchener-Cambridge-Waterloo. Toronto as the center of the GGH is the biggest CMA, both in terms of geographical area and population. Having around 2.6 million commuters starting their trips and

2.7 million commuters ending their trips in this CMA, Toronto provides employment for 97.97% of its own residents (2,547,670 people), while the remaining 2.03% commuters travel to the neighbouring Hamilton CMA (21,815 people or 0.84%) and Oshawa CMA (13,225 people or 0.51%) for work. Toronto provides work opportunity for an additional 16.24% of residents living in all other CMAs (184,465 people). It is especially popular for commuters from Oshawa (65,090 people or 40.79%), Barrie (23,945 people or 29.51%), and Hamilton (65,640 people or 20.32%s).

Hamilton as the second largest CMA in the region also receives a large number of cross-CMA commuters. It provides employment for 75.02% of its own residents (242,310 people) and 1.3% for all other CMAs (43,970 people). It is a popular destination for commuters from Brantford (6,365 people or 11.73%) and St. Catharines-Niagara (10,895 people or 6.67%).

Kitchener-Cambridge-Waterloo has grown an agglomeration of big-tech companies over the years (Region of Waterloo, 2018). It is the third largest employment CMA, supporting 86.57% of its residents (146,005 people) and 0.66% of residents from all other CMAs (23,240 people). A large proportion of commuters from Guelph (10.41% or 7,290 people) and Brantford (8.56% or 4,645 people) work in Kitchener-Cambridge Waterloo.

The results confirms that the CMAs in the GGH are very interconnected with each other and especially with the center CMA Toronto. However, the employment-to-population ratio shows that most CMAs provide sufficient work opportunities to satisfy its residents in terms of absolute number of jobs and workers (regardless of employment industry). The exceptions to this statement are Oshawa and Barrie, where less than 60% and 70% of their residents work within their respective CMA. Oshawa and Barrie also have the longest average commute distance in the GGH: Barrie's average commute distance is around 26km, while Oshawa's is around 24km. Moreover, Barrie has the largest proportion of long commuters relative to its commuting population, with around 15% travelling over 70 km, which is much higher than all other CMAs as they have less than 6% of long commuters.

In conclusion, the average commute distance for CMAs in the GGH region is around 16 km, much higher than the Canadian national average of 8.7 km (Government of Canada, 2019). From the geographical commute pattern, it can be understood that most commuters work within their home CMA or travel to a close neighbouring CMA. However, there are also many cases of people commuting extreme distances to work in Toronto, where around 72% of all commutes longer than 70km end. Understanding this phenomenon from a different perspective is that Toronto fails to provide its workers adequate housing options, thus, they trade-off commute distance to live in places such as Oshawa and Barrie where there might be more suitable housing either for their preferences or for necessity (i.e., housing price). Further research should investigate the factors influencing their location choice to understand why they choose to live in their respective CMA, and why they choose to take long commutes to work in Toronto. Policymakers should address the concerns of these residents by providing more working opportunities closer to these CMAs, such as in Aurora, Markham, or Scarborough to achieve jobs-housing balance in these areas.

5.2 Answering RQ3

5.2.1 Median Household Income and Average Commute Distance

Income level and average commute distance have a parabolic relationship in the GGH. The results of the analysis reveal that CTs with median household income between $34,912 and $78,665 or earning over $131,657 have, on average, have the shortest commute in the GGH, while CTs with a median household income between $86,638 and $117,325 have the longest commute. I theorize that this observation is associate with the spatial distribution of job type. Service jobs have lower than average wages (around $19.98 in Ontario in 2016) and are less likely to show regional agglomeration pattern and more likely to be spatially dispersed because service businesses' location choice caters to population demand (Government of Canada, 2021). Thus, lower-income families do not need to travel far for their work, resulting in shorter commutes.

On the other hand, firms in the quaternary sector benefits more from the agglomeration economy, thus, these jobs are more likely to be spatially clustered (Henderson et al., 2001). Examples of such agglomerations are located in prime real estate areas in Ontario, such as the Financial District in downtown Toronto. Housing options near these employment centers are often smaller and more expensive because of land value. Employees in these positions are also paid higher than average wages, such that the average hourly wage for employees in finance, insurance, real estate, rental and leasing (FIRE) is around $31.22 in 2016 (Government of Canada, 2021). Following this logic, the top income decile has the means to afford residential location closer to their employment. Meanwhile, middle classes who are employed in these quaternary employment clusters make relatively well wages, but not enough compared to the top decile to live close to their jobs. As a result, they may choose to move further away for more affordable housing options, but their earnings justify the longer commute. Granted, this is a simplified hypothesis to the analysis result, and it requires further research in the GGH to fully understand the residential location choice.

Another finding is that median household income affects the average commute distance differently across CMAs when accounting for the influence of other independent variables (percentage of single-detached houses and percentage of auto mode share in a CT). The income variable is only statistically significant in Hamilton, Oshawa, and Toronto with minimal influence. Every dollar in income increases the average commute distance by 0.07 meters in Hamilton, 0.13 meters in Oshawa, and 0.26 meters in Toronto. This is hypothesized to be influenced by the different costs of living in each CMA and the geographic location of affordable housing. For example, Map 4.17 indicates that the city center has a cluster of lower-income CTs in Hamilton and Oshawa, whereas the pattern is more random in Toronto. Again, further research on the geographic distribution of median housing prices is needed to provide empirical evidence to this hypothesis. Future research should further investigate the geographic pattern of job opportunities by type, and housing by affordability and suitability to determine whether there is a spatial mismatch. Moreover, policy interventions should be used to ensure that affordable and suitable housing exists within a certain commute radius of such agglomerations.

5.2.2 Dwelling Type and Average Commute Distance

Dwelling type dictates the typology and density of the CT. The higher cost of land often encourages the developers to build vertically and more densely, while peripheries with lower land prices allow for more single-detached housing. Of course, there are also single detached houses in downtown areas, such as Rosedale in Toronto, but it is rare and significantly more expensive, costing buyers a few millions.

Areas with predominantly single-detached houses have fewer opportunities for mixing land-use, which depletes opportunity to shorten the distance between workers and their jobs. Therefore, I hypothesized that since single-detached houses is more likely to be found in the peripheries in Ontario, the CTs with a higher proportion of single-detached houses would have longer commutes, as opposed to higher-density condominiums, apartments, and row houses that are found more often near employment centers. The analysis verified my hypothesis, as the dwelling type variable shows a significant positive correlation with average commute distance and is significant in the multivariate regression analysis as well.

Intensification and diversification are the most important pillars to influence travel behaviour and reduce commute distances. Higher-density developments and mixed land-use place residents closer to employment centers. The opportunity for shorter trips created by compact developments is typically completed by more sustainable transport modes. Dwelling type is also important in the residential location. Households have different needs: larger households require the space that single-detached houses provide, while low-income households require affordable housing. Policymakers should place diverse housing stock near employment centers, which will cater to households with different socioeconomic statuses and reduce the commute distance in areas where existing dwelling type is disproportional to worker's needs.

5.2.3 Mode and Average Commute Distance

Analysis reveals that the GGH has a high rate of car dependency for their commute trips: 66.58% of all commuting trips are completed using auto modes of travel, and 56.05% of all commuting trips are completed by single-occupant cars. This pattern is different across CMAs. Brantford has the highest rate of commuters driving (88.31%), Toronto has the highest rate of commuters using public transit (23.40%), while Peterborough has the highest rate of commuters using active transport (8.42%). Mode choice is influenced by transport infrastructure in place; thus, this result is expected as Toronto is the only CMA with subway system.

Statistical analysis confirms that auto commutes are longer compared to other modes, which confirms existing literature as cars can travel further distances in a shorter time and is more convenient for trip-chaining. However, its influence is quite different across the CMAs. The variable is statistically significant in Hamilton, Kitchener-Cambridge-Waterloo, Oshawa, Peterborough, and Toronto. One percent increase in auto mode share equals to around 72 meters increase for average commute distance in CTs in Hamilton, 117 meters for Kitchener-Cambridge-Waterloo, 132 meters in Oshawa, 161 meters in Toronto, and decreases the average commute distance by 195 meters in Peterborough. The result for Peterborough does not match with the hypothesis and should be investigated in future research. To conclude, if the GGH is aiming to reduce car dependency, it is important to not only provide infrastructure for alternative modes such as public transport, but also to reduce the commute distance, as shorter trips can more easily be replaced by sustainable modes.

5.3 Limitation and Future Research

The results of this thesis provide a fundamental understanding of the GGH's commute pattern and the influences of possible determinants of long commutes. However, I recognize that there are limitations to my methods and hope that future scholarship can address the issues listed in this section.

To begin, the database used in this thesis does not reflect the mixed-modal commutes because the 2016 Canadian Census Survey asks the respondents to answer about the mode that is used for the majority of the trip. This simplifies the complex commute journey into single modal trips. Many commutes in this region are park-and-ride trips: people living outside of the Toronto CMA would drive from their home to the GO-Train station to ride into Toronto, then switching to the TTC to travel within Toronto. Moreover, this database does not reflect people with multiple jobs and multiple commutes, influencing the results of CTs with lower median household income.

Secondly, the unit of analysis has a large influence towards the results. This thesis uses Census Metropolitan Area as one of the main units of analysis since it is also used in the 2016 Canadian Census. However, CMAs in the GGH have various land area and population, such that the Toronto CMA is significantly larger than any other CMA. For analyses such as proportion of inter-CMA commutes, CMAs with larger units would naturally capture more compared to a smaller CMA — if the unit is the GGH boundary, the data would show that 100% of commutes originate and end within the same boundary; versus if the unit is the City of Markham (located in the north side of the Toronto CMA), there would likely be a larger proportion of commuters that work in a different boundary. Moreover, because of the Toronto CMA's large boundary, it has the lowest proportion of long commutes relative to the total number of commuters (0.68%), however, it has the highest absolute number of long commuters (17,580) in the region. The current units of analysis may have contributed to creating a favorable narrative for the Toronto CMA, as well as highlighting inter-CMA commutes while providing limited information on intra-CMA commutes. Future research should use distance-based unit of analysis for more consistent results.

Thirdly, regarding the dependent variable, this thesis uses an OD Cost Matrix and road network within the GGH to solve the commute distance rather than using the self-reported commute time as dependent variable. As such, the calculated commute distance does not account for routes that could be taken outside of the network dataset. Moreover, the variation of calculated commute distance versus real commute distance are bigger in CTs with bigger areas because this thesis uses centroids of CT as points of origin and destinations. My definition of "long commute" as commutes over 70 km is approximated and not rigorously derived as existing literature more commonly uses commute time to determine long commutes.

Furthermore, a more sophisticated statistical model should be used in future research. The variables used in this thesis are determined following the literature review. The current model used three main independent variables — income, dwelling type, and mode. The aggregated nature of the databases limits the inclusion of several other characteristics such as education, family size, ethnicity, immigration status which have been identified by literature as statistically significant. Although the current model accounted for the majority of variation at the regional level, it cannot at the CMA level. The model is especially weak for St. Catharines-Niagara, where the model only explains 15% of the variability. Future research should use exploratory methods to incorporate more variables at a CMA level to construct a better narrative on what factors influences commute distance for each CMA. In terms of spatial analysis, all clustering patterns are visually identified by me. Future research should use spatial statistic methods such as Moran's I for empirical evidence.

Finally, this thesis is a cross-sectional study of the commute pattern in 2016. Future research should conduct longitudinal study to understand how the commute pattern has changed over time. More importantly, the global pandemic caused by Covid-19 may have drastically changed the GGH's commute pattern as more Canadians started to work from home following social distancing measures to curtail the spread of the virus. With better telecommunication technologies and more tolerance towards flexible work arrangements, perhaps the extremely long commutes will decline. The upcoming 2021 Canadian Census will be able to provide the data for such research.

REFERENCES

2016 Census of Population questions, long form (National Household Survey). (n.d.). Retrieved March 12, 2021, from https://www12.statcan.gc.ca/nhs-enm/2016/ref/questionnaires/questions-eng.cfm

Acheampong, R. A. (2018). Towards incorporating location choice into integrated land use and transport planning and policy: A multi-scale analysis of residential and job location choice behaviour. *Land Use Policy*, *78*, 397–409. https://doi.org/10.1016/j.landusepol.2018.07.007

Alonso, W. (2013). Location and Land Use: Toward a General Theory of Land Rent. In *Location and Land Use*. Harvard University Press. http://www.degruyter.com/hup/view/title/320970

Axisa, J. J., Scott, D. M., & Bruce Newbold, K. (2012). Factors influencing commute distance: A case study of Toronto's commuter shed. *Journal of Transport Geography*, *24*, 123–129. https://doi.org/10.1016/j.jtrangeo.2011.10.005

Buliung, R. N., & Kanaroglou, P. S. (2002). Commute minimization in the Greater Toronto Area: Applying a modified excess commute. *Journal of Transport Geography*, *10*(3), 177–186. https://doi.org/10.1016/S0966-6923(02)00010-8

Carlton, D. W. (1983). The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables. *The Review of Economics and Statistics*, *65*(3), 440. https://doi.org/10.2307/1924189

Cervero, R. (1989). Jobs-Housing Balancing and Regional Mobility. *Journal of the American Planning Association*, *55*(2), 136–150. https://doi.org/10.1080/01944368908976014

Florida, R. (2002). The Economic Geography of Talent. *Annals of the Association of American Geographers*, *92*(4), 743–755. https://doi.org/10.1111/1467-8306.00314

Government of Canada, S. C. (2017a, February 8). *About the data, Census Profile, 2016 Census*. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/about-apropos/about-apropos.cfm?Lang=E&wbdisable=true

Government of Canada, S. C. (2017b, February 8). *Focus on Geography Series, 2016 Census—Census metropolitan area of Toronto*. https://www12.statcan.gc.ca/census-recensement/2016/as-sa/fogs-spg/Facts-cma-eng.cfm?LANG=Eng&GK=CMA&GC=535&TOPIC=1&

Government of Canada, S. C. (2018, September 11). *Census data collection—Sampling and Weighting Technical Report, Census of Population, 2016*. https://www12.statcan.gc.ca/census-recensement/2016/ref/98-306/ch1-eng.cfm

Government of Canada, S. C. (2019, February 25). *Study: Long commutes to work by car*. https://www150.statcan.gc.ca/n1/daily-quotidien/190225/dq190225a-eng.htm

Government of Canada, S. C. (2021, January 25). *Employee wages by industry, annual*. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410006401

Government of Ontario. (2020). *A Place to Grow Growth Plan for the Greater Golden Horseshoe Office Consolidation 2020*.

Grengs, J. (2010). Job accessibility and the modal mismatch in Detroit. *Journal of Transport Geography*, *18*(1), 42–54. https://doi.org/10.1016/j.jtrangeo.2009.01.012

Hamilton, B. W., & Röell, A. (1982). Wasteful Commuting. *Journal of Political Economy*, *90*(5), 1035–1053. http://www.jstor.org/stable/1837132

Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, *25*(2), 73–76. https://doi.org/10.1080/01944365908978307

Henderson, J. V., Shalizi, Z., & Venables, A. J. (2001). Geography and development. *Journal of Economic Geography*, 25.

Hunt, J. D. (2010). Stated Preference Examination of Factors Influencing Residential Attraction. In F. Pagliara, J. Preston, & D. Simmonds (Eds.), *Residential Location Choice: Models and Applications* (pp. 21–59). Springer. https://doi.org/10.1007/978-3-642-12788-5_2

Kain, J. F. (1968). Housing Segregation, Negro Employment, and Metropolitan Decentralization. *The Quarterly Journal of Economics*, *82*(2), 175. https://doi.org/10.2307/1885893

Lowry, I. S. (1963). *A Model of Metropolis*. https://www.rand.org/pubs/research_memoranda/RM4035.html

Lyons, G., & Chatterjee, K. (2008). A Human Perspective on the Daily Commute: Costs, Benefits and Trade-offs. *Transport Reviews*, *28*(2), 181–198. https://doi.org/10.1080/01441640701559484

Marshall, A. (1920). *Principles of Economics*. Macmillan and Co.

Meyer, S. P. (2020). Finance, Insurance and Real Estate Firms and The Nature of Agglomeration Advantage Across Canada and Within Metropolitan Toronto. *Canadian Journal of Urban Research*, 34.

Moos, M., Revington, N., & Wilkin, T. (2018). Is there suitable housing near work? The impact of housing suitability on commute distances in Montreal, Toronto, and Vancouver. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *11*(4), 436–459. https://doi.org/10.1080/17549175.2018.1484793

O'Kelly, M. E., & Niedzielski, M. A. (2009). Are Long Commute Distances Inefficient and Disorderly? *Environment and Planning A: Economy and Space*, *41*(11), 2741–2759. https://doi.org/10.1068/a41327

Pagliara, F., Preston, J., & Simmonds, D. (Eds.). (2010). *Residential Location Choice*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12788-5

Peng, Z.-R. (1997). The Jobs-Housing Balance and Urban Commuting. *Urban Studies*, *34*(8), 1215–1235. https://doi.org/10.1080/0042098975600

Redmond, L. S., & Mokhtarian, P. L. (2001). *The positive utility of the commute: Modeling ideal commute time and relative desired commute amount*. 27.

Region of Waterloo. (2018, January 23). *Tech and Digital Media*. https://regionofwaterloo.icreate7.esolutionsgroup.ca/en/doing-business/techanddigitalmedia.aspx

Siedschlag, I., Zhang, X., & Smith, D. (2013). What determines the location choice of multinational firms in the information and communication technologies sector? *Economics of Innovation and New Technology*, *22*(6), 581–600. https://doi.org/10.1080/10438599.2013.783266

Statistics Canada. (n.d.-a). *Census tract: Detailed definition*. Retrieved March 10, 2021, from https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/ct-sr/def-eng.htm

Statistics Canada. (n.d.-b). *CMA and CA: Detailed definition*. Retrieved March 10, 2021, from https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/cma-rmr/def-eng.htm

van Ommeren, J. N., & Gutiérrez-i-Puigarnau, E. (2011). Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics*, *41*(1), 1–8. https://doi.org/10.1016/j.regsciurbeco.2010.07.005

Yaropud, T. (2019). *Results from the 2016 Census: Long commutes to work by car*. 14.

Yigitcanlar, T., Baum, S., & Horton, S. (2007). Attracting and retaining knowledge workers in knowledge cities. *Journal of Knowledge Management*, *11*(5), 6–17. https://doi.org/10.1108/13673270710819762

APPENDIX A: 2016 Canadian Census Survey Questions

Table: Questions in the 2016 Census of Population, long form

#	Question	Answers
3	What is the address of this dwelling?	1: Number (and suffix, if applicable) (e.g., 302, 151 B, 16 1/2) 2: Street name, street type (e.g., DR = Drive), direction (e.g., N = North) 3: Apartment/unit 4: City, municipality, town, village, Indian reserve 5: Province/territory 6: Postal code
42	At what address did this person usually work most of the time?	1: Worked at home (including farms). Go to question 45. 2: Worked outside Canada. Go to question 45. 3: No fixed workplace address. Continue with the next question. 4: Worked at the address specified below. Street address (see example).City, town, village, township, municipality or Indian reserve. Province/territory. Postal code.
43 a)	How did this person usually get to work? *if this person used more than one method of travel to work, mark the one used for most of the travel distance	1: Car, truck or van — as a driver. Go to question 43 b). 2: Car, truck or van — as a passenger. Go to question 43 b). 3: Bus. Go to question 44 a). 4: Subway or elevated rail. Go to question 44 a). 5: Light rail, streetcar or commuter train. Go to question 44 a). 6: Passenger ferry. Go to question 44 a). 7: Walked to work. Go to question 44 a). 8: Bicycle. Go to question 44 a). 9: Motorcycle, scooter or moped. Go to question 44 a). 10: Other method. Go to question 44 a).
43 b)	How many people, including this person, usually shared the ride to work in this car, truck or van?	1: Drove alone 2: 2 people 3: 3 or more people
44 a)	What time did this person usually leave home to go to work?	1: Time in hours and minutes (a.m. or p.m.)

44 b)	How many minutes did it usually take this person to get from home to work?	1: Number of minutes

(Source: *2016 Census of Population Questions, Long Form (National Household Survey)*, n.d.)

APPENDIX B: Median Household Income Deciles

This table exhibits the summary of the ten income groups formed by the decile function in STATA: xtile inc_dec = av_income, nq(10).

Decile Groups of Average Household Income

Median Household Income Deciles	**Summary of Average Income**				
	Min	**Max**	**Mean**	**Std. Dev.**	**Count**
1	34,912	61,376	54,064.54	5,812.70	175
2	61,440	70,758	66,176.49	2,634.04	175
3	70,810	78,592	75,073.72	2,228.45	177
4	78,665	86,528	82,836.67	2,243.19	172
5	86,638	93,594	90,260.13	2,021.32	174
6	93,619	101,035	97,284.97	2,207.98	175
7	101,047	109,056	105,016.4	2,346.48	175
8	109,269	117,248	112,969.17	2,335.88	175
9	117,325	131,584	123,552.11	3,845.74	174
10	131,657	340,309	162,516.74	38,229.86	174
Overall	**34,912**	**340,309**	**96,925.39**	**32,201.84**	**1746**