Understanding Factors Associated With Commuter Rail Ridership
A Demand Elasticity Study of the GO Transit Rail Network

by

Aaron Shantz

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Environmental Studies
in
Planning

Waterloo, Ontario, Canada, 2021

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that this thesis may be made electronically available to the public.

Abstract

Mode share in major North American cities is currently dominated by private automobile use. Planners have theorized that transitioning commuter rail systems to regional rail networks is a viable method to increase ridership and stabilize mode share. This process is currently underway in Ontario, Canada, as the amount and frequency of service is being increased throughout the GO Transit rail network via the GO Expansion Program.

However, previous studies have shown that transit demand does not solely respond to service quantity expansions. Variables related to the built environment, regional economy, network characteristics, and socioeconomic status of the customer base can influence transit demand to varying degrees. Further, the literature states that the travel behavior of commuter rail users is unique, as access mode, distance, socioeconomic status, and the utility derived from varying trip types can differ compared to local transit users. These findings suggest that supplementary policies might be needed to reduce automobile reliance and stimulate demand for regional transit.

Many transit researchers have conducted demand elasticity studies to identify what factors are significantly associated with transit ridership. However, no researcher has conducted this type of analysis specific to the GO Transit rail system. The purpose of this thesis is to fill this gap. Through literature review, variables significantly associated with transit demand were first identified. Station-level datasets were then compiled at monthly intervals from January 2016 to December 2019. During this process, station catchment areas estimated using PRESTO smartcard data were used to extract data related to land use, socioeconomic, and demographic indicators. Additional factors related to station access, service quantity, and availability of substitute transport modes were also compiled. A random effect linear panel data estimator was then applied to obtain demand elasticity estimates.

Of the variables included in the analysis, this study finds that several variables such as service quantity, population density, fuel price, and unemployment rate are significantly associated with transit demand, regardless of trip type examined. Ridership was also responsive to employment density and seasonal variation, although differing signs were shown depending on trip type examined. Surprisingly, demand was relatively unresponsive to enhanced station access options, including park and ride capacity and the quality of feeder bus connections. The results suggest that policies in addition to the service quantity improvements as outlined in the GO Expansion Program should be considered to further increase system demand. Those aimed towards heightened densities and land use diversities around rail stations, increasing the cost of private automobile operation, and the implementation of competitive fare price strategies are outlined.

Notably, desktop research revealed that policies related to these factors have been previously explored by provincial stakeholders. However, only service improvements as proposed within the GO Expansion Program have been committed too. Knowing that demand is responsive to these factors could increase the level of political willingness needed to implement these policies to further increase ridership and subsequently balance mode share. These findings could also be used by Metrolinx to justify the allocation of resources needed to update or implement policies within the study area. Overall, this study highlights that many factors, including those related to the built environment, network characteristics, and the price / availability of substitute transport options are significantly associated with commuter rail demand. Therefore, integrated planning policies should be considered by transit agencies undergoing similar network transitions to ensure that ridership is increased to the greatest extent possible.

Acknowledgements

First, I would like to thank my academic supervisors Jeff Casello and Clarence Woudsma who have patiently supported me during my time at the University of Waterloo. I have good memories discussing transit issues, formulating a research topic, and developing this body of work from an idea into a fully completed thesis. I will especially miss our good-natured hockey talks, which somehow managed to stay level-headed despite the rivalry between our hometown teams.

I'd also like to thank Metrolinx for giving me the opportunity to first work for them in the summer of 2019 as a Student GO Planner. While employed, I was exposed to a study which evaluated change in ridership with respect to fare price changes on the Union-Pearson Express system. Without this knowledge, I would not have identified a demand elasticity study as a potential thesis topic.

I further owe Metrolinx for the financial support and mentorship provided via the 2020 Rob MacIsaac Fellowship Program. I'd like to thank Amy Peebles and Natalie Petra, who recognized the potential of this research and provided professional oversight while working in this position. Various other members of the Service Planning team also played a key role in this project, including Randy Bui, Ian McNeil, Anthony Smith, and others who I might be forgetting. Your help in obtaining various datasets including fare price, ridership, customer origin, and parking statistics allowed me to advance the capacity of this research to a point that I did not think was possible at the onset of this work.

This thesis would have been much more involved without help from Eddy Ionescu, who let me use a script he developed to compile service quantity data. His assistance was greatly appreciated, as his willingness to help saved countless hours of work that would have otherwise been needed. I value the talks we continue to have about transit in Ontario, and hope that we're able to work together again in the future.

Thank-you to my fellow WPTI researchers, who were so welcoming since my first day at the University of Waterloo. I have made many friends who I will miss upon graduation, including Yixin (Candice) Chen, Edward Donato, and Xiaomeng (Ming) Xu. I especially enjoyed the many late-night trips we took to Tim Hortons while working on-campus. Further, the fun times we had while travelling to and participating in the Hands-On Sustainable Mobility conference in Karlsruhe, Germany are a highlight of this degree. I will also never forget the friends I've made who were part of my MA & MES in Planning cohort, including Alison Curtis, Emily Stanley, Hanna Holman, Rex Lin, and Theresa Lynn. The memories we made both in and outside of the classroom will last a lifetime.

A large portion of this degree I owe to my family, including my parents (Susan Shantz and Philip Shantz) my sister Laura Shantz, and my uncle Ross Steckley. They provided continued mental support and encouragement throughout the duration of this program, especially while working at home during the

COVID-19 pandemic. My Mom and Dad also deserve special recognition for letting me live at home for the duration of my studies, which made this undertaking financially possible. I would also like to thank my grandparents (John Meyer, Nora Meyer, Kay Shantz, and Wilbur Shantz), who were a source of inspiration throughout this journey. I hope this achievement has made all of you proud.

This can also be said of my girlfriend Nicole Jankowski, who did the same despite pursuing her own studies at the University of Waterloo's School of Optometry. You have done more for me during these last three years than you could possibly know. I look forward to continuing our journey together once this chapter in our life ends.

Finally, I'd like to thank my friends including Brendan Burton, Sean Canham, Kevin Nogueira, John Pepall, Dan Rickert, Steve Rickert, and those involved in our poker and video game nights. Your continued friendship and our humorous times together have helped make this thesis possible.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Setting the Stage

Transportation in North America, specifically infrastructure and facilities that service large urban centers, has been a topic of debate over the last several decades. Historically, investment in this sector has focused on accommodating automobile use via the construction of roads, expressways, bridges, and other network elements (Hanson, 1992; Moore et al., 2007; Newman & Kenworthy, 1996). As a result, over 90% of travel in some major North American metropolitan regions is completed via private automobile, a mode that has been investigated thoroughly regarding its impact on sustainable outcomes. As urban communities continue to expand, planners are struggling to accommodate increased levels of congestion, greenhouse gas emissions, and other negative externalities being realized due to an uneven mode share. Therefore, questions have arisen as to how municipalities and regions can shift automobile users to modes that are more sustainable, efficient, and effective.

International examples have suggested that the provision of heavy rail infrastructure is an effective solution in facilitating inter-regional transport demand. These systems may operate as both intracity or regional rail services, with the latter facilitating the movement of people from city suburbs into the centre of neighboring metropolitan areas, while operating with 5-20 minute headways on all lines throughout the day (Vuchic, 2007). They are also shown to have numerous advantages when compared to private automobile in terms of speed, capacity, safety, environmental friendliness, energy savings and urban space consumption (Caroline & Yves, 2012). Additionally, rail systems can offer fast and reliable service during peak commuting periods, an aspect that is often not accomplished during automobile travel due to high levels of expressway congestion (Allen & Levinson, 2014; Vuchic, 2007). Various studies have suggested that when these conditions are accomplished, transit demand is stimulated, thus resulting in a mode share that is more evenly distributed. For example, a study of 48 European cities revealed that those with extensive rail coverage were positively correlated with frequent use of the area's transit system when compared to those with minimal offerings (Ingvardson & Nielsen, 2018). A more direct analysis of rail ridership in Karlshrue, Germany, discovered that regional rail ridership increased by 400% once investments in service quantity and quality were implemented (Chisholm, 2002). Therefore, international examples suggest that the provision of such infrastructure, combined with adequate service offerings, can encourage travelers to primarily use public transit when engaging in inter-regional travel.

1.2. The Current State of Rail Transit in North America

Rail service is provided in a variety of large metropolitan cities throughout North America, including Chicago, Boston, New York, Toronto, and Vancouver. However, these systems operate as commuter rail systems, which are characterized by irregular headways and infrequent service offerings, especially during weekends and off-peak periods (Vuchic, 2007). Furthermore, service is typically only provided in a single direction during peak periods (ex. either in to or out of the metropolitan area), meaning that transit service is unavailable to commuters who work in areas outside of the downtown core. As a result, large proportions of residents in these areas continue to choose private automobile as their main mode of transport when engaging in regional travel, as the convenience and flexibility associated with commuter rail use is fairly limited. For example, when the City of Toronto is examined, private automobile continuously accounts for approximately 64.5% of trips completed within the area. Further, public transit has not accounted for a mode share greater than 12% since the beginning of the 21st century (Ashby, 2018; University of Toronto, 2003, 2009, 2014). These figures are even greater when inter-regional public transit figures are examined, as only 1% of trips are completed using GO Transit, the region's commuter rail network. Among other things, large amounts of congestion and inflated travel times along key arterial routes are continuously realized in North American metropolitan areas, as private automobile use is relied on for inter-city accessibility.

1.3. The Role of Regional Rail

One solution that can be implemented to alleviate these negative externalities involves upgrading commuter rail systems to regional rail networks. This strategy is logical as a variety of low-cost, low technology measures can be implemented to upgrade service levels and reduce unit costs of operation (Allen, 1998; Schumann & Phraner, 1994). Embracing this philosophy, GO Transit planners and provincial officials in Ontario, Canada, introduced a plan aimed at improving rail ridership via an increase in flexible and convenient service offerings across all network segments located throughout southwestern Ontario (Government of Ontario, 2018a). Now titled the GO Expansion Program, the bulk of the program involves transforming the current network from one that is primarily focused on satisfying commuter travel behavior to one that provides all-day, two-way service in to and out of the City of Toronto, with 15-minute headways promised during peak travel times. The plan theorizes that 121.3 million additional annual riders will be generated by the time the network transition is completed, representing a 211% increase in ridership compared to ridership figures observed in 2017 (Government of Ontario, 2018b). Furthermore, a cost-benefit analysis of the program estimated that the network transition should generate $42.2 billion dollars in economic benefits, as negative externalities currently realized such as congestion, large travel times for transit users, and environmental emissions will be reduced (Government of Ontario, 2018b).

1.4. Understanding the Determinants of Transit Demand

However, a network transition such as the one proposed in the GO Expansion Program is not a simple task, as the expansion of service coverage and trip quantity does not guarantee that mode shift will occur. Economic theory states that when faced with a variety of purchase decisions, consumers seek to maximize their utility by selecting the good or service that results in the highest level of overall satisfaction (Mankiw, 2013). Since transport demand typically stems from a person's need to temporarily relocate to another area to engage in various activities such as work, social activities, or shopping, those engaging in travel should instead be considered as disutility minimizers, as the mode associated with the lowest total cost incurred by the user is typically selected (Athira et al., 2016; Ben-Akiva & Morikawa, 2002; Casello & Hellinga, 2008; O'Fallon et al., 2004; Yang et al., 2018).

Internal variables including the price, quantity, and quality of transit service provided are various aspects that can influence the amount of utility (and therefore disutility) associated with transit use (Balcombe et al., 2004; Holmgren, 2007; Schimek, 2015; Taylor et al., 2009). External variables such as traveler characteristics, physical and economic characteristics of surrounding urban areas, and the availability of alternative transport modes can further influence mode choice decisions. Research has also illustrated that in the North American context, station access for regional transit systems is primarily facilitated via private automobile, meaning that the capacity of park and ride facilities could significantly influence transit demand (Government of Ontario, 2016; Levinson et al., 2012). Therefore, demand may be more sensitive to variance among station access indicators, rather than fare prices and service quantities, if station access is restricted at some stations compared to others. Fortunately, analytical models can be used to understand the significance and magnitude of influence that various internal and external variables have on transit demand. Undertaking such an analysis allows transportation planners to better understand ridership figures, predict future demand, and implement more informed policy decisions to further encourage mode shift.

1.5. Research Purpose and Questions

The purpose of this research is to further understand what variables significantly influence commuter rail ridership. Using the GO Transit rail network as a case study, the following research questions were formulated:

- What internal and/or external variables are most determinantal to GO Transit rail ridership?
- Do findings differ depending on trip type examined?
- How do these results compare to relationships identified in previous demand elasticity studies?
- Do station accessibility indicators specific to the North American context, such as park and ride capacity and the quality of feeder bus connections, influence demand?
- In addition to the GO Expansion Program, could additional plans or policies be explored to further encourage mode shift and transit demand in the study area?

- What lessons can be transferred to other regional transit agencies looking to grow rail ridership?

1.6. Thesis Structure

To answer these questions, Chapter 2 begins with an overview of the case study area. Trends and patterns relating to demographics, mode choice, and transportation behavior study area are provided, while the history and current structure of the GO Transit rail network is summarized. Plans and policies pertinent to land use and transportation planning in southern Ontario are also highlighted.

Chapter 3 summarizes literature relevant to transit demand studies. This includes an overview of econometric and regression modelling techniques and its application to transit analysis. Previous findings from ridership elasticity studies are also outlined with the purpose of highlighting various internal and external factors that have displayed significant relationships with transit demand. The chapter concludes with a summary of data collection methods, specifically those used to estimate station catchment areas, extract external variable datasets, and measure station accessibility indicators.

Chapter 4 outlines the methodology employed to answer the research questions. Methods used to delineate station catchment areas, measure station accessibility indicators, and analyze the impact that various external and internal variable datasets had on GO Transit rail ridership are therefore justified in this chapter.

Chapter 5 explores data analysis methods. First, the chapter outlines how linear extrapolation, the use of Geographic Information Systems (GIS), and various statistical software programs were used to compile the dependent variable and independent variable datasets that were analyzed. The steps used to select a final subset of independent variables are also outlined. Finally, the modelling framework and process is presented.

Chapter 6 presents the findings generated from the modelling outputs and discusses the relationship and significance of each variable in relation to transit demand. A discussion regarding model performance and explanatory capacity is also highlighted in this chapter.

Chapter 7 further discusses the model outputs and compares and contrasts findings relative to those identified in previous studies. Policy implications in relation to the GO Expansion Program are further discussed. This chapter concludes with a discussion regarding the limitations of the study and suggestions for future research.

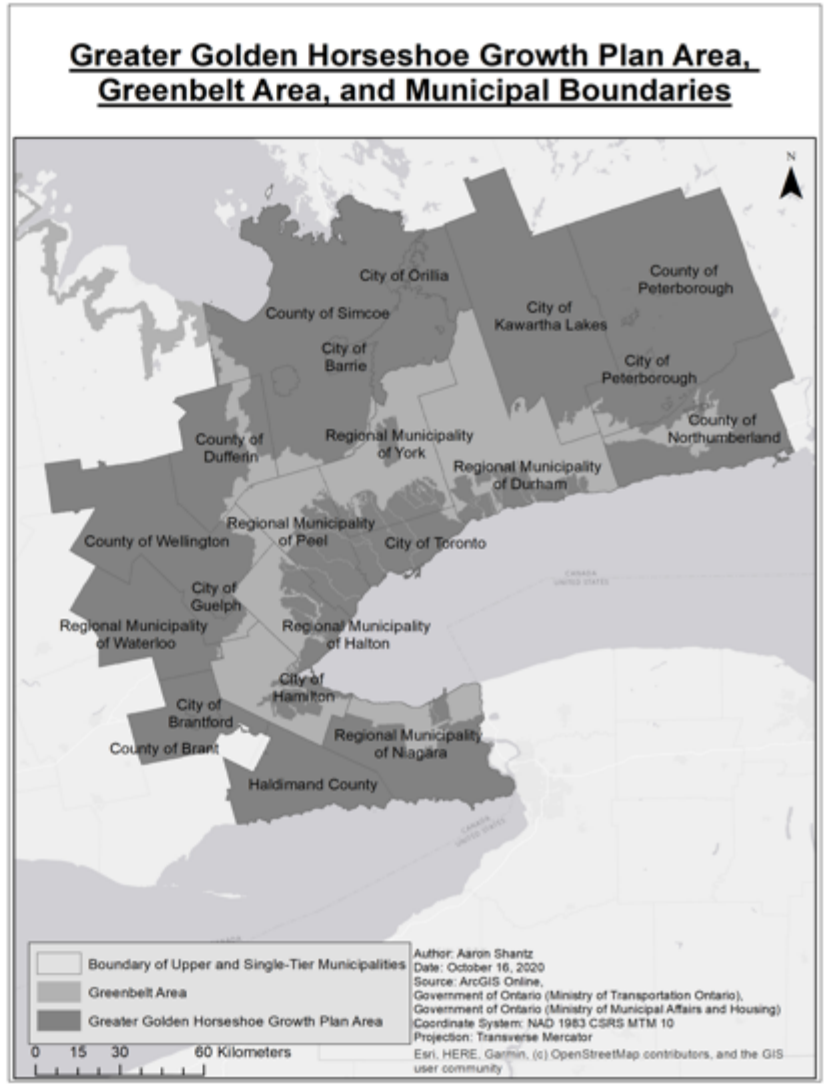
2. Overview of Study Area

2.1. The Greater Golden Horseshoe

2.1.1. Spatial Context

The Greater Golden Horseshoe (GGH) is a geographical region located in southern part of Ontario, Canada. Conceptualized by the provincial government in 2006, the region begins on the western shore of the Niagara River and wraps around the western end of Lake Ontario, thereby encompassing the Greater Toronto and Hamilton Area (GTHA). As shown in Figure 1, the GGH further extends along the northwestern shore of Lake Ontario before terminating at Oshawa. Further inland, municipalities including Brantford, the Regional Municipality of Waterloo, Guelph, the City of Barrie, Peterborough, and Kawartha Lakes are contained within the region. More specifically, the GGH consists of 110 separate municipal jurisdictions in total, 21 of which are single or upper-tier municipalities, and 90 of which are lower-tier municipalities (Allen, R. Campsie, 2013). Various pieces of provincial legislation, including the Greenbelt Act, the Niagara Escapement Plan, and the Oak Ridges Moraine Conservation Act restrict development in areas that have been identified as environmentally sensitive throughout the region.

Figure 1 - *Spatial Extent of the Greater Golden Horseshoe and Corresponding Single/Upper-Tier Municipalities*



Commonly known as the Greenbelt, this area encompasses 22% of the region's 32,000 km^2 land mass (Government of Ontario, 2017a).

2.1.2. Demographic and Transportation Trends

The GGH is one of the most populous areas in Canada. Currently, the region is home to 9 million residents, and is forecasted to grow to a population of 13.5 million residents by 2041 (Government of Ontario, 2020a). This accounts for roughly 67% of Ontario's population, and more than a quarter of the national population according to the 2016 Census of Population (Statistics Canada, 2017). When individual municipalities within the region are examined, the City of Toronto contains the largest number of residents with a population of 2,731,571 (Statistics Canada, 2017). As a result, the City of Toronto is not only the region's most populous municipality, but the largest in Canada. At an international level, the City of Toronto ranks as the 4th largest city in North America. The GGH is also significant in terms of economic output. Currently, businesses in the region generate 25% of Canada's Gross Domestic Product, while highlighted sectors include finance, insurance, real estate, industrial, and technology-based firms (Government of Ontario, 2020a).

Currently, the majority of residents within the GGH are dependent on private automobile as their main mode of transport. According to the University of Toronto's Transportation Tomorrow Survey, only 43% of trips completed in the region are completed using sustainable modes, resulting in an average of 2.6 automobile trips generated per household on a daily basis (Ashby, 2018). The survey also highlights that while the population of the region has grown by 40% since the beginning of the 21st century, mode share has remained largely stagnant (Ashby, 2018). Automobile use continuously accounts for approximately 64.5% of trips completed in the region, while public transit has not accounted for a mode share greater than 12% in any given year. As shown in Table 1, these figures are even greater when inter-regional public transit figures are examined, as approximately 99% of trips are completed using modes other than regional rail (Ashby, 2018; University of Toronto, 2003, 2009, 2014).

As a result, the GGH is plagued by large and highly variable travel times. Real-time traffic data compiled by TomTom, a worldwide Global Positioning System company, found that a commuter in the City of Toronto is expected to spend 33% more time in traffic during peak periods compared to those experienced in off-peak periods (TomTom International BV, 2020b). Based on this finding, they theorize that a typical commuter logs 142 hours of lost time per year as a result of congestion (TomTom International BV, 2020a). Compared to other large cities that were also included in the study, Table 2 illustrates that the City of Toronto ranks as the 6th most congested city in North America.

Table 1 - *Historical Mode Share Conditions in the Greater Golden Horseshoe*

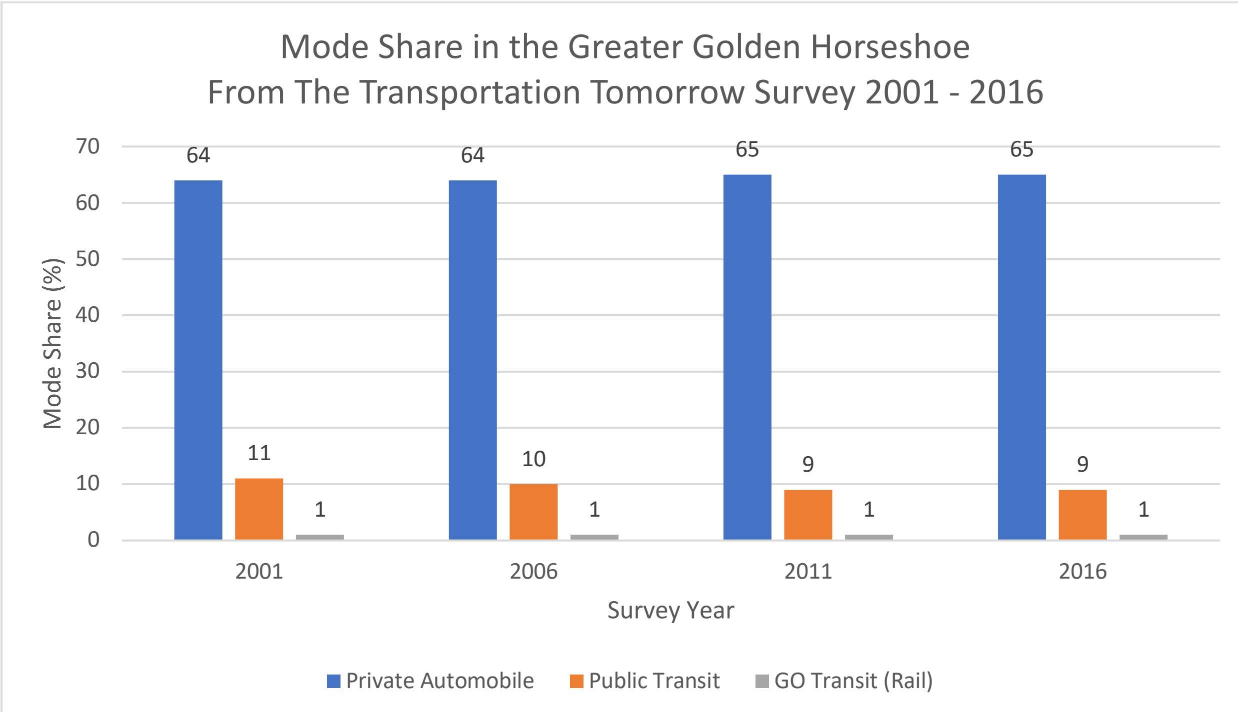
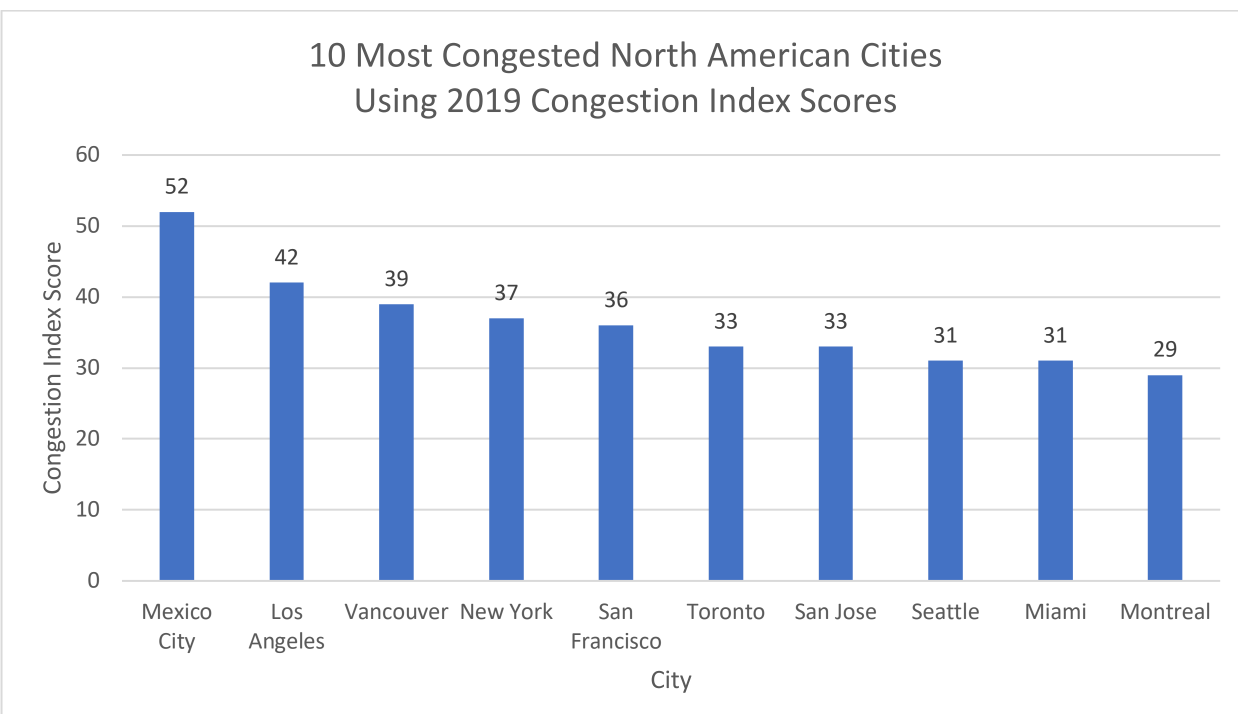


Table 2 - *Congestion in the City of Toronto Relative to Other North American Cities*



2.1.3. History of Regional Public Transit in the GGH

Regional public transit has been provided to municipalities throughout the GGH long before the region was conceptualized. GO Transit, abbreviated for Government of Ontario Transit, was founded by the provincial government in 1967 to provide regional public transit to commuters working in the City of Toronto (Government of Ontario, 2017b). Commuter rail service was first provided along Lake Ontario's shoreline between Pickering and Hamilton, on what is now known as the Lakeshore West and Lakeshore East corridors. Initially launched as a pilot project, passenger volumes grew rapidly and quickly outpaced projections. As such, the size and extent of the network was expanded throughout the 20th century to satisfy demand.

At the end of the 20th century, the majority of service provided by GO Transit was facilitated via service agreements with municipalities and freight companies, as GO Transit did not own any rights of way on which they operated (Lysyk, 2016). Additionally, little policy was in place to give the organization any legislative mandate to expand service throughout the region, resulting in stagnant service expansion. To address this, Premier Dalton McGuinty and the Liberal Government of Ontario drafted the Greater Toronto Transportation Authority Act in the spring of 2006, which involved the establishment of a corporation responsible for regional transportation planning throughout the GGH (Greater Toronto Transportation Authority Act, 2006). After the bill received royal assent in the provincial government on June 22, 2006, the province became the sole stakeholder responsible for regional transportation planning in the GGH. The official objectives of the corporation were:

1. To provide leadership in the co-ordination, planning, financing, and development of an integrated and multimodal regional transport network,
2. To act as a central procurement agency for purchase of public transportation assets on behalf of Ontario municipalities,
3. To be responsible for the operation of the GO Transit system and the provision of other transit services throughout the operating area (Greater Toronto Transportation Authority Act, 2006).

The establishment of the Greater Toronto Transportation Authority officially allowed a single organization to plan for and organize an integrated regional transportation system. Therefore, the introduction of the Greater Toronto Transportation Authority was seen as a step in the right direction for regional transportation planning in the area, as a consistent regional transportation plan could be drafted and acted on with provincial authority. During this time, large advancements in regional transportation planning across the region were realized. This was shown as "The Big Move", the region's first regional transportation plan, was released in 2009 (Government of Ontario, 2008). This established a coordinated direction and vision about how regional transit in the area should evolve until 2031, and also identified "quick wins" that could be established by the corporation to enhance transit connections and service in the area. Additionally, several complementary pieces of legislature, such as the Greater Toronto and Hamilton Area Transit Implementation Act of 2009, combined the Greater Toronto Transportation Authority and GO Transit into one corporation titled "Metrolinx" (Greater Toronto and Hamilton Area Transit Implementation Act, 2009). This further streamlined the provision of regional public transportation, as planning and service implementation now laid under one corporation and governing legislation. Furthermore, these policies enabled the province to direct funds and prioritize investment in regional transit infrastructure and service as they saw fit.

2.2. Current GO Transit Network

2.2.1. Network Layout

The current GO Transit rail system provides commuter rail service to 68 stations throughout the GGH via five radial lines and a single diametrical line. All corridors feed into Union Station in the City of Toronto, which acts as the main hub of the network. The location and extent of each corridor is illustrated in Figure 2. Appendix A provides a more detailed overview of each corridor currently in operation.

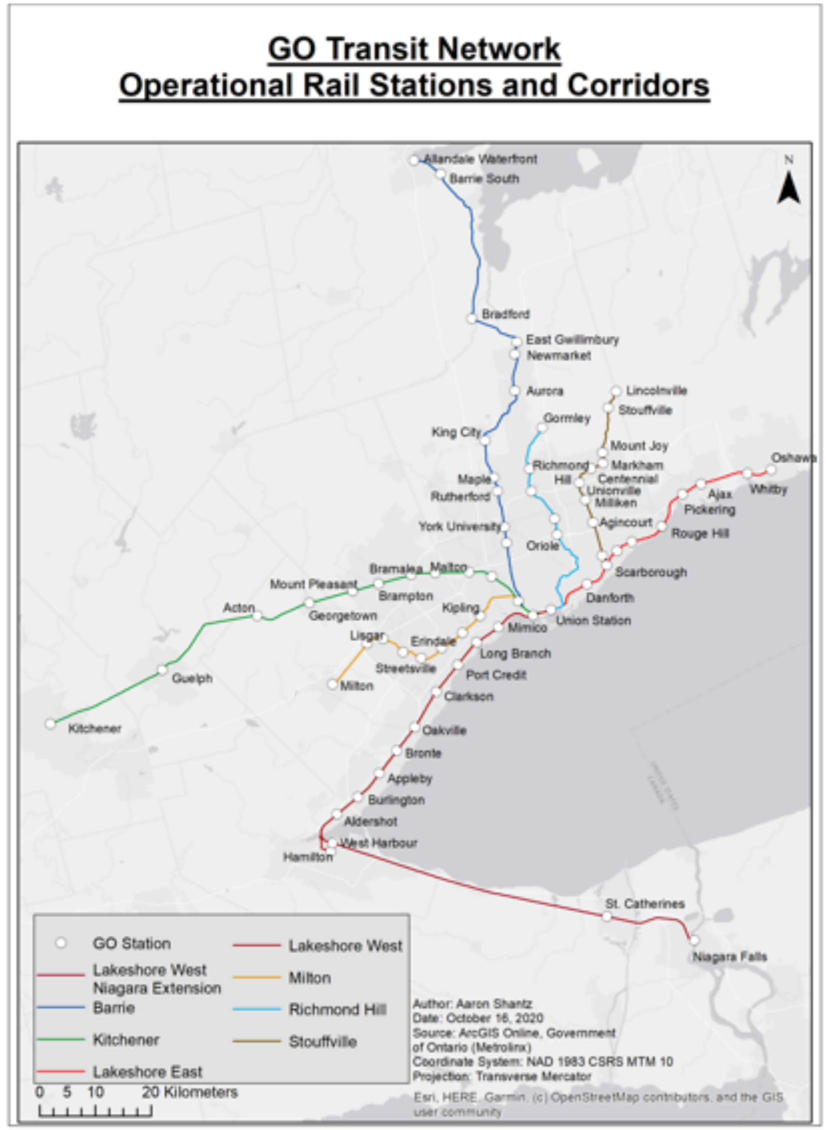
2.2.2. Trackage Ownership

While GO Transit maintains full ownership of its rolling stock, various infrastructure components throughout the network are owned by private corporations and/or other crown corporations. When service first began, GO Transit formulated agreements with major freight rail companies Canadian National (CN) Railway and Canadian Pacific (CP) Railway to use their freight corridors for public transit purposes (Collenette, 2016). As a result, GO Transit paid a usage fee to CN and CP in exchange for track usage rights, therefore avoiding large upfront costs for trackage and corridor construction (Collenette, 2016). Beginning in 1999, GO Transit recognized that service expansion would be difficult based on these track usage agreements, as freight traffic and freight priority prevented GO Transit from increasing service levels (Lysyk, 2016). As such, GO Transit began to negotiate track and land purchases with both CN and CP to allow them to dictate operating agreements and prioritize public transit over freight transport throughout the network. As of 2016, approximately 80% of trackage throughout the network has been purchased by GO Transit, while the remaining portions are still shared with CN and CP under similar track usage agreements. Appendix A provides further information on the state of the network with respect to track ownership rights.

2.2.3. Fare Structure and Collection Methods

GO Transit uses a zonal fare structure to assign ticket prices. Unlike a flat fare structure where all customers pay the same fare price regardless of distance travelled, zonal fare structures assign a ticket price depending on the origin and destination of the traveler. This allows the agency to assign a fare that is correlated with the distance travelled by the customer, resulting in a more accurate and equitable fare policy. 55 fare zones are present within the GO Transit rail network, with fare prices assigned depending on how many fare zones the customer travels between during their trip (Smith, 2020). A variety of fare payment systems are used to collect fares, although the majority of fares are collected via the PRESTO smartcard system. Otherwise known as a proof of payment system, PRESTO allows users to load stored values onto a smartcard, which is then collected from automated fare collection machines located within stations and platforms when they board at their origin and alight at their destination. The system then deducts the appropriate fare based on the distance travelled by the user. This also allows GO Transit to track reliable data regarding system ridership, fare price data, and the distribution of boardings and alightings throughout the network.

Figure 2 - *GO Transit Rail Network*



Approximately 90% of GO Transit rail riders pay their fares using PRESTO cards (Smith, 2020). This system also offers riders the ability to register their PRESTO card online so that stored values can be refunded to the user in the event that it is lost or stolen. Through this process, important demographic data of the user is also recorded and made available to GO Transit, including the postal code of the rider's residential address. PRESTO was first introduced as a pilot project in 2009, and slowly expanded to include networkwide coverage by mid 2012 (McCarter, 2012). Due to various operating glitches during program launch and initial use, PRESTO smartcard data is only considered to be reliable from 2014 onwards (Smith, 2020; McCarter, 2012).

2.3. Planning Policy Framework

2.3.1. The Planning Act

The Planning Act legislates land use planning in Ontario (Planning Act, 1990). Mainly, it describes how land uses may be controlled and by whom they are controlled, and specifies the role of the province and individual municipalities with respect to land use planning matters. The Planning Act also provides a basis for considering provincial interests, preparing planning policies to guide future development, and provides a variety of implementation tools that can be used to facilitate planning. At the provincial level, the Planning Act specifies that the role of the province is to promote provincial interests, including the support of public transit and sustainable infrastructure development.

2.3.2. The Provincial Policy Statement

The Provincial Policy Statement (PPS) helps to explain and interpret the guidelines set forth in the Planning Act (Government of Ontario, 2020b). In other words, the Planning Act provides a framework for land use planning In Ontario, while the PPS provides an overall policy direction that should be pursued to address planning matters of provincial interest. Three major policy sections are outlined in the PPS, including "Building Strong Healthy Communities", "Wise Management of Resources", and "Protecting Public Health and Safety". Policies addressed within these sections formulate the basis for land use planning in Ontario. The provision and development of public transit service is addressed under "Building Strong Healthy Communities" and outlines how it can be used to facilitate the growth of communities in a sustainable and efficient manner. Specifically, the PPS states that:

1. Public transportation should be provided to facilitate the movement of people, and that dense and mixed-use developments should be promoted to reduce the number of trips made by private automobile while encouraging the use of public transit,

2. Providing for an efficient, cost-effective, and reliable multimodal transportation system that is integrated with other systems and jurisdictions is a key component in ensuring long-term prosperity in the region,
3. Intensification along public transportation corridors, integrated with transit-supportive development, should be encouraged to shorten commute journeys and decrease congestion (Government of Ontario, 2020b).

Therefore, the PPS is clear in establishing the connection between public transit provision and sustainable land use planning, as it can be used as a tool to move people efficiently, influence settlement patterns, generate economic output, and reduce greenhouse gas emissions in the study area.

2.3.3. A Place to Grow – Growth Plan for the Greater Golden Horseshoe

A Place to Grow, Growth Plan for the Greater Golden Horseshoe (hereby referred to as "the Growth Plan"), is a regional planning framework meant to guide government investment and land use planning activities in the region (Government of Ontario, 2020a). The Growth Plan recognizes that population densities, employment figures, and settlement patterns in the GGH are unique compared to other areas throughout the province. Essentially, it builds on the PPS to establish unique land use planning objectives that are specific to the regional context.

The Growth Plan states that the built environment is currently not optimal to accommodate current and projected population growth throughout the GGH. Notably, the plan highlights that recent development within the region is sprawled and fragmented, meaning that public transit is ineffective in serving these populations. As a result, the majority of residents within the area are reliant on private automobile as their main mode of transport. Secondly, the Growth Plan highlights that the economic foundations of the region are changing, and that the regional transportation system has not adapted to facilitate this transition. Notably, the current transportation network favors the movement of freight and goods, rather than the movement of people and ideas. Finally, a summary of various negative externalities being realized as a result of the current mode share is provided. Large levels of greenhouse gas emissions and congestion have been shown, resulting in impacts to air quality, water quality, and loss of economic output due to increases in travel times. Therefore, the Growth Plan highlights that the objectives as outlined in the PPS have not been accomplished throughout the GGH. As a result, the potential economic benefits realized from, dense, growing urban settlements could be marginalized if interventions are not implemented to influence where and how growth throughout the GGH occurs.

Fortunately, the Growth Plan recognizes that an integrated and high quality regional public transit system can be used as a tool to alleviate these challenges. Section 3.2.2 of the Growth Plan states that the transportation system within the GGH will be planned and managed to:

- "Offer a balance of transportation choices that reduces reliance upon the automobile and promotes transit",
- "Be sustainable and reduce greenhouse gas emissions by encouraging the most financially and environmentally appropriate mode for trip- making",
- "Offer multimodal access to jobs, housing, schools, cultural, and recreational opportunities, and goods and services" (Ontario, 2020a, pg. 32).

The Growth Plan further states that "Public transit will be the first priority for transportation infrastructure planning and major transportation investments", with the purpose of maximizing the efficiency and viability of existing and planned service levels, increasing the capacity of existing transit systems, and increasing the mode share of transit in the area (Ontario, 2020a, pg. 32). Therefore, specific plans and policies directed at increasing the vitality of regional public transit in the region are required to ensure that the objectives outlined in the PPS and the Growth Plan are accomplished. To date, the 2041 Regional Transportation Plan and the GO Expansion Program are the most significant policies released by Metrolinx with the purpose of supporting these objectives.

2.3.4. 2041 Regional Transportation Plan and GO Expansion

The 2041 Regional Transportation Plan outlines how Metrolinx and GO Transit plan to utilize regional public transit as a tool to satisfy the objectives as outlined in the PPS and the Growth Plan (Government of Ontario, 2018a). Building on the initial objectives as outlined in The Big Move, the vision of the 2041 Regional Transportation Plan is to implement a sustainable transportation system that is aligned with appropriate land uses, while supporting healthy and complete communities. Consistent with the key objectives identified in the PPS and the Growth Plan, the goal of such a system is to provide convenient and reliable connections, support a high quality of life, stimulate a prosperous and competitive economy, and protect the environment.

The Regional Express Rail program is a major initiative outlined in the 2041 Regional Transportation Plan as a means of accomplishing these goals. Now titled the GO Expansion Program, the program aims to transform the current GO Transit rail network from a commuter focused system to one that is a comprehensive regional rail system. GO Expansion is expected to increase the quantity and quality of service provided to customers, increase the amount of residential and employment areas that are accessible to GO Transit rail service, encourage intensification of both residential and employment developments located within close proximity to GO Transit rail service, and provide system access to all subgroups of the population. In order to do this, the GO Expansion Program has proposed the following deliverables:

1. Expand service by over 1,000 new trips per day,
2. Implement two-way, all-day service throughout the region,
3. Increase the provision and coordination of station access options, including feeder bus connections and parking capacity,
4. Increase frequencies to ensure 15 minute or better headways along priority transit corridors,
5. Implement a transit-oriented development framework that engages both private developers and public stakeholders,
6. Ensure that the cost of transit use is competitive with other transport modes (Government of Ontario, 2018a).

Metrolinx believes that implementing these deliverables will have a significant impact in mitigating the land use planning challenges currently being realized in the GGH. Mainly, the expected increase in service quantity is expected to increase ridership by 211% by 2031, with daily ridership projected to be 630,000 trips per day (Government of Ontario, 2018a). Metrolinx anticipates shifting 145,000 car trips per day to rail, therefore influencing mode share and decreasing time lost due to congestion by a total of 6.5 million hours per year.

GO Expansion is further expected to make rail transit a viable and competitive option for residents and workers throughout the region. 42% of the region's population is expected to live within five kilometers of a GO Transit rail station providing two-way, all-day service, whereas 34% of the region's population will be able to reach Union Station in 45 minutes via train (Government of Ontario, 2018a). This level of accessibility, coupled with increased service provisions, is expected to decrease average commute times by ten minutes per trip compared to the current state.

The number of connections to major employment centers throughout the region is expected to increase as a result of this network transition. Employment centers outside of downtown Toronto will also be accessible throughout the day, as all-day two-way connections are expected to be provided to emerging employment generators such as Kitchener, Barrie, and Oshawa (Government of Ontario, 2020a). Additionally, 42% of all jobs within the Greater Toronto and Hamilton Area are expected to be within a 45-minute rail trip from Union Station, resulting in increased flows of people and ideas throughout the region (Government of Ontario, 2018a). Additionally, GO Expansion is expected to influence settlement patterns throughout the region. 40% of all homes and 45% of all jobs within the region are planned to have access to 15-minute, two-way all-day service on priority transit corridors, thereby making these areas attractive for commuters and employers who do not live or work in downtown Toronto (Government of Ontario, 2018a). Increased densities and diverse land uses are also recommended for these areas via the Growth Plan, thereby further concentrating growth towards areas of the built environment that are transit supportive. Concentrating development in areas supplemented by frequent transit service is expected to reduce auto-centric behavior, as transit could therefore be used to satisfy both work-related and discretionary transport demands.

Finally, Metrolinx has stated that in order to remain cost competitive with other modes, substantial fare price increases will not be implemented to help fund the project (Government of Ontario, 2018a). Instead, fare prices will only be implemented that are consistent with inflation, while other fare and transfer agreements with municipal service providers will be investigated to decrease the overall cost of transit use. These fare policies should thereby encourage ridership by remaining competitive with other modes, while remaining affordable for marginalized subsections of the population.

2.4. Conclusion

In summary, various plans and studies have recognized that significant growth within the Greater Golden Horseshoe is expected. However, the current state of the area's regional transportation network could prevent sustainable and healthy development from occurring. Negative externalities stemming from an auto centric mode share, including congestion, increased travel times, and lost economic productivity are expected to continue if policy interventions are not implemented.

Provincial planning policies relevant to the Greater Golden Horseshoe have stated that increased provision of regional transit service should be prioritized to reduce these effects. Notably, the GO Expansion Program states that transitioning the current GO Transit rail network to a regional rail network will be an effective solution. Substantial service increases that are planned as a part of transition are theorized to be a significant driver of increased transit demand in the study area.

3. Literature Review

Previous research has revealed that transit demand can be influenced by a variety of internal and external variables. Typically, econometric analysis and the application of linear regression techniques are used to understand the relationship between transit demand and change in these factors. As mentioned above, the study area considered for this analysis is governed by a distinctive land use and transportation planning framework, as settlement patterns, demographics, and public transit systems in the region are unique compared to other areas throughout Ontario. Therefore, relationships identified within the context of this study might differ from those previously identified in various academic papers or professional studies.

This chapter begins with an overview of concepts and theory related to econometric analysis and its application to transit demand studies. Secondly, findings from previous transport behavior, demand, and/or mode share studies are summarized with the purpose of creating a list of variables that could be determinantal to transit demand in the study area. Where possible, attempts were made to focus on the inclusion of studies completed in the North American context, which also included commuter / regional rail ridership in their analysis. Finally, quantitative methods used to identify and extract various datasets are reviewed, with a unique focus on methods used to obtain feeder bus connection quality datasets and those used to delineate station catchment boundaries. The paragraph concludes by identifying relevant gaps in the literature to be filled by this body of work.

3.1. Regression Analysis

3.1.1. Foundational Concepts

Econometric modeling uses mathematical equations to describe various relationships. Demand modeling is a type of econometric modeling used to describe how the demand of a good or service changes based on economic factors impacting the customer. Many studies have revealed that a variety of economic factors such as fare price, service quantity, economic status, underlying socioeconomic conditions, and demographic characteristics can impact the quantity of transit demanded. If sufficient data is available, the change in transit demand attributed to these variables can be measured.

Regression analysis is used to measure these relationships. Essentially, it is a mathematical equation that measures the impact that a single independent variable, or multiple independent variables, have on a dependent variable. Results generated from a regression model can be used to understand:

- The influence of the independent variable(s) on the dependent variable,
- The statistical significance of the relationship,

- Which independent variable is most important in influencing the dependent variable, given the inclusion of multiple independent variables in the model,
- How the dependent variable should change if fluctuation in an independent variable occurs (Wooldridge, 2012).

3.1.2. Simple Linear Regression

Linear regression is the most common method of regression analysis. A simple linear regression attempts to explain the value of the dependent value based on the independent variables included in the model, and those not included (ex. the error term). A simple linear regression takes the following form:

$$y = \beta_0 + \beta_1 x + \mu \qquad \textit{Eq. 1}$$

Where;

- y = the dependent variable,
- x = the independent variable,
- β_0 = the intercept parameter, otherwise known as the constant term,
- β_1 = the slope parameter (i.e., the effect) that the independent variable has on the dependent variable,
- β_1 = the slope parameters (ex. the effect) that "x" has on "y" holding other factors in "μ" constant,
- μ = the error term (represents factors not captured in the model that effect "y").

3.1.3. Ordinary Least Squares Regression

Ordinary Least Squares (OLS) is a common simple linear regression method used in econometric modeling. When using the OLS method, the sum of squared residuals is used to compute the parameters of the model, with the goal of estimating a fitted line that minimizes the distance between observed values and fitted values (Wooldridge, 2012). This method is commonly used as it is the best linear approximation of the true relationship between the dependent and independent variable(s), and allows researchers to estimate unbiased and consistent statistical properties of relationships (James et al., 2013). Per Wooldridge (2012), the equation is first rearranged to solve for the intercept and slope parameters, where:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \qquad \textit{Eq. 2}$$

and:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad \textit{Eq. 3}$$

Using Equation 3, a fitted value for each observation in the sample is estimated, while the residual for a given observation is the difference between the actual "y" value and the fitted value. The sum of squared residuals is then made as small as possible given the following form:

$$\sum_{i=1}^{n} \hat{u}_i^{\,2} = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^{\,2} \qquad \textit{Eq. 4}$$

The OLS regression line can then be estimated, where:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \qquad \textit{Eq. 5}$$

3.1.4. Multiple Linear Regression

The OLS method can also be applied when a dependent variable is regressed on multiple factors. Additional independent variables may be added to the model to minimize the error term, while increasing the amount of information made available to the researcher about the relationship (Wooldridge, 2012). The form of the regression line looks similar to Equation 1, where:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 +_{\dots} + \hat{\beta}_k x_k \qquad \textit{Eq. 6}$$

And:

- κ = the number of independent variables included in the model.

3.1.5. Application to Transit Demand Analysis

In transit demand elasticity studies, the slope estimate ($\hat{\beta}_k$) of each independent variable is of primary interest to the researcher. The sign of the slope explains the direction of the relationship with the dependent variable; a positive sign indicates a positive correlation, while a negative sign indicates a negative correlation. The slope estimate also predicts the anticipated change in the dependent variable should the independent variable increase by one unit. Per Taylor et al. (2009), analysis based on this model is the most efficient method of calculating ridership elasticities, as the slope estimates can be easily used to calculate demand elasticities. Additionally, a variety of variables can be included in the model to estimate their relationship and impact on transit demand, including transit fares, travel times, service supply, service attributes, passenger characteristics, prices of alternative transport modes, urban characteristics, and regional characteristics (Taylor et al., 2009).

3.1.6. Advanced Methods

While simple linear regression methods are adequate in assessing one-dimension datasets, more advanced techniques are often used for datasets that have a cross-sectional and/or time-series component. These datasets, otherwise known as panel or longitudinal data, observe the behavior of multiple entities across time (Torres-Reyna, 2007). Compared to a one-dimensional dataset, such as a study that analyzes station-level ridership at a given point in time, panel data is usually preferred as it provides more information, contains more variability, and demonstrates less collinearity amongst independent variables. When transit demand studies are considered, panel data also allows the researcher to utilize readily available internal variable datasets, and combine them with more detailed station level observations such as land use information, parking availability, and the provision of feeder bus connections, so that ridership figures can be explained more efficiently (Guerra & Cervero, 2011).

When analyzing panel datasets, the simple OLS method is rarely used as subsequent observations inherently influence each other. Additionally, the impact of unobserved factors also has a consistent impact on model performance (Torres-Reyna, 2007). However, various methods that build upon the OLS method are commonly used to analyze panel datasets which account for these issues, including pooled OLS, fixed effect and random effect estimators (Wooldridge, 2012). A variety of statistical tests, such as the Lagrange Multiplier and the Hausman Test, are then applied to the model outputs to determine the method that best suits the dataset (Guerra & Cervero, 2011; Lee & Lee, 2013; R. Liu, 2018; Stover & Christine Bae, 2011).

3.2. Factors Associated With Transit Demand

A review of ridership elasticity studies generated by academics, transit authorities, and research centers presented in the Canadian Urban Transit Association's Ridership Trends Study found that a variety of built environment, socioeconomic, transit service, and external factors can be determinantal to ridership demand (E. J. Miller et al., 2018). Of these, population density, employment density, fare price, service quantity, fuel price, vehicle ownership, unemployment rate, and income were found to have a consistent and statistically significant correlation with transit ridership. This literature review identified several additional variables that could be determinantal to commuter rail demand, including seasonality, distance to the central business district, age, households with children, park and ride capacity, and the presence of feeder bus connections. A summary of findings pertinent to each factor is presented below.

3.2.1. Fare Price

Previous studies found fare price to be significantly correlated with transport demand. Since transport demand analysis originally focused on assessing the impact of fare price changes on ridership figures, a large body of literature in relation to this variable exists (Curtin, 1968; Mayworm et al., 1980; Webster & Bly, 1981). A metadata analysis of transport demand studies identified a common range of demand elasticities with respect to fare price (Balcombe et al., 2004). Notably, their findings suggest that fare price elasticities differ depending on the geographical context examined. When all modes were considered, an overall demand elasticity of -0.44 was identified for studies conducted in the United Kingdom, while a lesser elasticity of -0.35 was obtained when Australian and North American studies were analyzed. The authors theorize that this difference could be a result of differing urban morphologies, or the presence of high fare prices and poor service quality in the United Kingdom compared to other regions included in the study.

Transport demand studies conducted solely in the North American context have attempted to further quantify this relationship (Boisjoly et al., 2018). An analysis of transport ridership trends in 25 North American cities identified a statistically significant fare price elasticity of -0.219, while a similar elasticity of -0.207 was found once service quantity metrics were disaggregated by mode.

Differing figures have been shown when trip type, time period, and size of the metropolitan area were controlled for. Using data obtained from 265 urban areas throughout the United States, Taylor et al. (2009) found a statistically significant fare price elasticity of -0.42, whereas a larger elasticity of -0.51 was estimated once per capita transport demand was estimated. The results suggest that large metropolitan areas are more sensitive to fare price changes compared to small urban centers, most likely due to a greater proportion of commuters that use transit willingly in large cities.

Demand elasticities were estimated using data obtained from 103 Canadian transport agencies over a 14 time-series (Diab et al., 2020). Total personal expenditure on public transport was used as a fare price indicator, which displayed a statistically significant elasticity of -0.143. Separate models were also produced that controlled for the size of the transit agency, where a lower elasticity of -0.147 was found for large transit agencies compared to a larger elasticity of -0.162 which was found for agencies with less than 1.2 million yearly trips. When contrasted with results produced by Taylor et al. (2009), the results suggest even within the North American context, fare price changes affect users in Canadian cities differently compared to those in American urban areas.

Furthermore, an analysis of American transit agencies attempted to control for both time-period and size of metropolitan area examined (Schimek, 2015). An overall fare elasticity of -0.32 was found in the short-run, comparable to that estimated by Balcombe et al. (2004). Separate models were also computed after controlling for the size of urban areas examined, and produced results separated into short-run and long-run estimates. Fare price elasticities in small urban areas were revealed to be -0.38 in the short-run and -0.73 in the long-run, whereas lesser elasticities of -0.2 and -0.48 were discovered when ridership in large urban areas was analyzed. Notably, the difference in magnitude between short-run and long-run elasticities was shown to be comparable in both models, further indicating that users display a greater sensitivity to fare price changes over time. Furthermore, the results suggest that fare price changes have a greater impact on those living in small urban centers compared to those living in large metropolitan areas. Like Diab et al. (2020), the author notes that this is likely due to differences in the socioeconomic status of the respective customer bases, as a large proportion of customers in large metropolitan areas are wealthy commuters, compared to those in small urban centers who are low-income captive riders.

Various studies have specifically focused on analyzing the determinants of commuter rail ridership and have suggested that fare price elasticities might differ compared to other modes examined. A study of 59 rail transit projects in the United States found that ridership is expected to decrease by 4.55% if fare price is increased by 10% (Guerra & Cervero, 2011). Notably, ridership figures gathered for their analysis were aggregated, meaning that demand elasticities were not calculated for individual rail modes. Regardless, the authors state that demand for commuter rail service is most likely inelastic to fare changes compared to other modes, as the majority of users are workers. Therefore, users are unaffected by fare price changes as the utility associated with the trip outweighs the additional disutility generated by the fare price change. As a result, the authors suggest that external variables such as population or employment density might be more significant in influencing commuter rail ridership. In addition to overall transit demand, Balcombe et al. (2004) identified fare elasticities specific to commuter rail systems. A fare elasticity of -0.58 was identified when studies conducted in the United Kingdom were considered, whereas a lesser elasticity of -0.37 was obtained when Australian and North American studies were analyzed. Unlike Guerra & Cervero (2011), the authors theorize that rail might have a larger elasticity compared to other local modes such as metro and bus, as longer average trip lengths associated with commuter rail can result in automobile being a direct competitor.

Using data collected from several major North American cities over a 10-year period, demand elasticities were estimated for a variety of transit modes including commuter rail (Iseki & Ali, 2014). Notably, a fare price elasticity of -0.353 was considerably larger compared to those calculated for other rail modes examined, including light rail and metro. The authors theorize that commuter rail users could be more sensitive to fare price increases as prices are often already high due to the use of distance or zonal based fare schemes. Similar results were generated in an analysis of transit demand in Chicago (Nowak & Savage, 2013). A comparison of rail ridership in the city and surrounding metropolitan area identified a -0.42 demand elasticity for suburban rail, whereas an insignificant fare price elasticity of 0.038 was estimated for the city's metro and light rail systems. The authors state that this difference could be attributed to a fare price increase that occurred on the city bus system during the study period, thus resulting in increased city rail ridership. The results indicate that the provision, availability, and pricing of alternative transport systems should be considered when interpreting demand elasticity estimates.

A review of New Jersey Transit's commuter rail system identified both short and long-run fare elasticities in relation to network ridership (C. Chen et al., 2011). Using data collected at monthly intervals between January 1996 and February 2009, a fare price elasticity of -0.4 was found in the short-run, while a fare elasticity of -0.8 was found in the long-run. Similar to Schimek (2015), their results suggest that commuter rail demand is twice as sensitive to fare changes over time. The authors are quick to note that their elasticity estimates might be overestimated, as the metropolitan area surrounding New Jersey is one of few in North America where public transit is a direct substitute to private automobile use. Therefore, the availability of substitute transit systems, such as local bus and light rail transit routes, could entice users to switch to other public transit modes if commuter rail fares are increased.

Paulley et al. (2006) further explored findings compiled by Balcombe et al (2004) to explore the impact that trip type might have on commuter rail fare elasticity estimates. Notably, the results show that demand elasticities differ considerably once trip type is accounted for. For example, an estimated short-run off-peak period fare elasticity of -0.79 was found for suburban rail systems located within the United Kingdom, whereas an elasticity of -0.34 was estimated for peak period riders in the same geographical context. Consistent with findings from Guerra and Cervero (2011), the results indicate that commuters are less sensitive to fare price changes compared to discretionary travelers, and that these impacts are realized to a greater extent when commuter rail demand is examined.

Regardless, a variety of studies have found fare price to be insignificant in explaining ridership. Focusing on transit ridership in the State of Washington, Stover and Bae (2011) assessed ridership data obtained from several counties throughout the state. Fare price was found to be statistically significant in only four areas examined, whereas a model that considered both cross-sectional and temporal differences demonstrated a statistically insignificant elasticity of -0.0388. Additionally, two counties demonstrated a positive coefficient, when a negative sign was expected. An examination of travel survey data in Germany found that fuel price and household vehicle ownership, rather than fare price, were the main determinants of ridership (Frondel & Vance, 2011). Unlike the literature reviewed above, fare price demonstrated a statistically insignificant relationship with ridership. The authors theorize that the high level of service quantity and quality associated with their transit systems, such as the InterCity Express rail system, could render fare price insignificant if people are willing to pay for these services. In the Canadian context, data obtained from 85 urban transit agencies throughout the country also found demand to be relatively inelastic to changes in fare price (Kohn, 2000). While demand elasticities were not calculated, the author discovered that despite fare price increases and service decreases, fare box revenue increased throughout the time-series analyzed, thus indicating an inelastic relationship. The author theorizes that commuter related transport could be the driving force behind this trend, as the increased cost of a ride is marginal compared to the cost of operating a vehicle and paying for parking at employment destinations.

Furthermore, a variety of studies that analyzed commuter rail demand found fare price to be insignificant in explaining ridership figures. For example, a study of rail ridership in Canada excluded fare price as a variable from their final models altogether due to insignificance, whereas an examination of rail ridership in California only found fare price to be a significant factor in two of four systems examined (Durning & Townsend, 2015; R. Liu, 2018). Studies completed in recent years have attempted to quantify the main determinants of commuter rail ridership, but excluded fare price as a candidate variable altogether (Brown et al., 2014; S. H. Chen & Zegras, 2016; C. Liu et al., 2016; Rahman et al., 2019). This could suggest that data availability in relation to this variable is hard to obtain and/or interpret due to the use of zonal or distance-based fare schemes, or it could indicate that researchers are more interested in assessing the impact that other internal and external variables might have on commuter rail ridership.

To summarize, a negative correlation between fare price and ridership is suggested by the literature, although these affects might not be significant in explaining ridership on commuter rail systems.

3.2.2. Service Quantity

Service quantity has been revealed to be extremely significant in explaining transit demand, and is commonly included in demand elasticity studies as it is one of few internal variables that is easy to quantity and assess (Schimek, 2015). When analyzed, aspects such as varying geographies, trip types, and modal classification are shown to have similar influences on demand elasticities as those identified in relation to fare price. The following paragraph summarizes these findings.

Taylor et al. (2009) tested both total revenue vehicle hours and service frequency (ex. total vehicle revenue hours divided by route miles) in their assessment of transit demand in urbanized American cities. Both variables were included in their final model as they demonstrated a statistically significant positive relationship with ridership. An elasticity of 1.23 was shown for vehicle revenue hours, and an elasticity of 0.48 was found for service frequency. The authors state that these findings, coupled with demand elasticities identified in relation to fare price, indicate that adjustments to internal variables could be implemented to double ridership figures in North America.

An early examination of transit use in Canada found service quantity to be the most influential variable in explaining ridership demand (Kohn, 2000). After considering a variety of variables including population, city size, and transit usage rates, only fare price and vehicle revenue kilometers remained in their final model. Of these, service quantity was found to be most significant in explaining ridership, as a much larger t-statistic was shown for this variable. Further to Taylor et al. (2009), the author suggests that changes only need to be made to these two internal variables to increase ridership.

A study of transit ridership in Canada determined ridership to be largely elastic to service supply changes (Diab et al., 2020). Vehicle revenue hours demonstrated the largest elasticity in their model outputs, as an elasticity of 1.009 was shown. Separate models were also estimated for transit agencies with more than 1.2 million trips in 2016, and those with less than 1.2 million trips to investigate if results differed depending on the size and extent of the transit agency. However, similar elasticities of 1.044 and 0.827 were found (Boisjoly et al., 2018). The results indicate that regardless of city size, ridership in North America can be grown via service expansion efforts.

Service supply was shown to have a varying influence on transit demand in the State of Washington (Stover & Christine Bae, 2011). Service elasticities ranging in value from 0.34 to 1.39 were identified, as separate models were developed for each network included in the study. However, a panel data estimator identified an average demand elasticity with respect to service quantity of 0.7. The results suggest that the impact of service provision on transit demand can differ greatly as evidenced by the range of elasticities computed.

Demand elasticities with respect to service quantity have been shown to fluctuate depending on the time period analyzed. Schimek (2015) was able to separate service elasticity estimates into short-run and long-run estimates, using vehicle revenue miles as a service quantity indicator; elasticities of 0.41 and 0.79 were found respectively. Li et al. (2020) identified similar findings in their study of transit ridership in Canada, as a short-run demand elasticity of 0.227 was found, compared to a long-run demand elasticity of 1.31. Much like findings in relation to fare price, the results indicate that users respond more drastically to service quantity changes over time. Schimek (2015) further controlled for the city size in an attempt to see if short and long run elasticities differed depending on the population of the study area. Notably, elasticities were found to be much larger in populous urban areas compared to smaller cities. A long-run service elasticity of 1.12 was noted in large urban areas, compared to a 0.67 elasticity identified for areas with less than 1 million residents. A similar trend, although not as drastic, was noted when short-run elasticities were examined; a 0.45 elasticity was found in large urban areas compared to an estimate of 0.35 identified for less populous urban areas. The author suggests that the variance between estimates may be a result of socio-economic characteristics correlated with city size, as large urban centers are more likely to have high concentrations of wealthy commuters who choose to use transit in an effort to avoid vehicle operation costs, whereas a greater ratio of riders in small, urban areas consists of captive, low income riders. Therefore, users in larger urban areas are more likely to shift to public transit if the level of convenience is comparable to that of driving, thereby making them more elastic to change in this variable.

Mode specific research has suggested that rail users are twice as sensitive to service changes compared to alternative transit users (Balcombe et al., 2004). Using vehicle revenue kilometers as a service quantity indicator, a short-run demand elasticity of 0.75 was computed, compared to an elasticity of 0.38 identified for bus systems included in the analysis. The results indicate that commuter rail users are twice as sensitive to service changes compared to bus users, but authors caution that results should be interpreted carefully as only three rail studies were included in their analysis, compared to 27 bus studies. They suggest that further research on rail ridership demand elasticities is needed for this finding to be corroborated.

Ridership demand in Chicago was regressed on a variety of independent variables, including average daily revenue vehicle miles (Nowak et al., 2013). Mode specific demand models found that commuter rail demand is expected to increase by 5.69% if a 10% increase in service quantity is implemented. A nearly identical service elasticity of 0.536 was identified in a separate examination of rail systems in the United States (Guerra & Cervero, 2011). Comparatively, demand elasticities with respect to service quantity generated for other modes ranged from 0.173 to 0.298 (Nowak & Savage, 2013). The results further suggest that commuter rail users within North America are much more sensitive to service supply compared to bus or local rail users.

Furthermore, a study of transit ridership in urban areas throughout the United States generated differing results (Iseki & Ali, 2014). Commuter rail, bus, and heavy rail ridership demonstrated comparable demand elasticities with respect to service quantity, as values ranged between 0.263 and 0.299. The authors note that of the entities included in their study, data regarding commuter rail ridership obtained from South Florida Regional Transportation Authority demonstrated a negative correlation with service quantity, which could have resulted in an underestimation of the commuter rail service quantity elasticity.

Multiple studies undertaken in the New Jersey area identified varying demand elasticities with respect to service quantity. Chen et al., (2011) identified a short-run service elasticity of 0.13, considerably lower than previous estimates. The authors theorize that since their study solely analyzed commuter rail ridership, the level of service currently associated with the system is so extensive that additional service offerings result in marginal ridership increases. In comparison, a short-run elasticity of 0.973 was identified by Yanamx-Tuzel et al. (2010), although all modes offered by the New Jersey Transit authority including commuter rail were aggregated in their analysis. The results suggest the importance of modal separation when transit demand elasticities are calculated, as mode specific conditions and characteristics could heavily influence demand elasticity estimates. This postulation is further highlighted by Boisjoly et al. (2013), as very different demand elasticities of 0.0093 and 0.465 with respect to service quantity were found once service quantity statistics were disaggregated by mode in their modelling outputs.

Additional studies have further suggested the level of baseline service offered can influence demand elasticities with respect to service quantity. A cross-sectional study of four major rail systems in

California found that an aggregate demand elasticity with respect to service quantity of 0.643, but relatively low elasticity values of 0.19 and -0.067 in cities where high levels of baseline service are offered (R. Liu, 2018). An assessment of public transit ridership in the San Francisco Bay Area generated similar results, as service quantity was not found to be a significant factor in explaining transit demand (Wasserman, 2019). The results suggest that simply expanding service on rail systems might not increase ridership figures, especially for those associated with service supply levels that are already consistent and convenient. Instead, integrated and multi-faceted planning approaches that decrease the disutility of system use in other ways might be required to stimulate transit demand.

The findings above illustrate that commuter rail ridership is most likely sensitive to change in service quantity, as more convenient and flexible trip offerings decrease the amount of disutility associated with system use. However, aspects including size of metropolitan area examined, time period analyzed, and level of baseline service can all have significant impacts on model outputs.

3.2.3. Distance to Central Business District

Various studies have revealed that transit use is correlated with distance from the study area's central business district (CBD). However, the direction of the relationship differs depending on the mode examined. A survey of homemakers in New York City was completed to see if internal or external characteristics were most significant in influencing transport behavior (C. Chen & McKnight, 2007). After analyzing travel diaries completed by over 11,000 households, the researchers noticed a significance difference in mode share depending on geographical location of the respondent. For example, those living in outer lying metropolitan areas completed only 1% of trips using public transit, whereas those living in central city areas completed upwards of 17% of trips via public transit. An analysis of transport behavior in Montreal used distance to downtown as an independent variable when estimating the likelihood of public transit use during the a.m. peak period (Grimsrud & El-Geneidy, 2013). Based on survey data obtained in 1998, 2003, and 2008 the authors found that distance to downtown was the only variable that demonstrated a consistent level of significance throughout all three study periods. A negative sign was consistently displayed, meaning that commuter rail, metro, and municipal bus ridership decreased as distance from the CBD increased. The authors theorize that the abundance of transit in the CBD, compared to the lack of accessibility options in outer lying suburban locations, was a contributing factor.

These findings were further reinforced in a station-level analysis of commuter rail ridership in the Boston area (S. H. Chen & Zegras, 2016). Although distance to CBD was only included in models that analyzed off-peak and weekend ridership, they found demand elasticities of -0.303 and -0.260 respectively. An analysis of ridership behavior in the Chicago area illustrated similar findings (Lascano Kežić & Durango-Cohen, 2018). They identified a significant increase in ridership at rail stations located within a

10 kilometer radius of the city's CBD, whereas ridership remained stagnant at stations located outside of this boundary. Both studies theorized that urban renewal, population density increases, and low commute times associated with living in close proximity to work was the main reason for this observation.

In contrast to Kežić et al. (2018), a study of station-level rail ridership in Chicago found that transit demand is significantly influenced by distance from the station to the CBD, although this relationship is not as significant as it once was (C. Miller & Savage, 2017). Using ridership data obtained at the station-level in 2004, 2006, 2009, and 2013, separate ridership models were estimated for each time period. Distance from downtown continuously displayed a negative coefficient in all model outputs, although smaller coefficients were displayed in more recent observation periods. For examples, elasticities of -0.024 and -0.027 were found in 2004 and 2006, while elasticities of -0.013 and -0.009 were found in 2013. Although small in magnitude, the trend suggests that demand for transit in suburban areas has increased considerably over the past 15 years.

An analysis of GO Transit rail users in Toronto, Ontario, was undertaken to better understand station access behavior (Engel-Yan et al., 2014). The authors theorized that several variables, including station location relative to the CBD, could influence average station access distance and associated ridership figures. The authors found that suburban stations demonstrated much larger station catchment areas compared to those located in urban areas, therefore resulting in larger customer bases. They theorized that the provision of commuter rail service might be inefficient when offered in close proximity to the CBD, as the availability of walking, cycling, and alternative transit choices may be preferred by those residing within the city. An examination of transit use in Toronto, Ontario, reinforced this observation (Mahmoud et al., 2014). The authors found that users are most likely to choose metro, rather than commuter rail, if both types of service are available to the user. Like Engel-Yan et al. (2014), the authors theorize that this could be a result of increased costs and transfer penalties associated with commuter rail systems, whereas metro systems are more attractive in terms of real costs and convenience. Therefore, the availability of substitutes could result in decreased commuter rail demand in high density urban environments, where alternative transit modes are typically offered.

To summarize, the literature suggests that demand for transit demonstrates a negative correlation with distance to the CBD, although the opposite relationship should be expected when commuter rail demand is analyzed.

3.2.4. Station Accessibility Indicators

More station access options typically results in increased transit demand. A variety of variables, including station-level parking supply, were tested for their significance in influencing station access distance, the extent of station catchment areas, and subsequent commuter rail demand in the Greater

Toronto and Hamilton Area (Engel-Yan et al., 2014). They found that all three factors were heavily influenced by the number of parking spaces available, as the majority of riders access the system via private automobile. Further, their results indicated that users disperse between adjacent stations depending on parking availability during the a.m. peak period, thereby suggesting that parking capacity is a limiting factor on station-level ridership demand. These findings were reinforced by Mahmoud et al. (2014), who discovered that station choice in the Greater Toronto and Hamilton Area is largely dependent on availability of park and ride infrastructure and local transit service connections. Their results suggest that in order to stimulate commuter rail demand, increasing the number of park and ride spaces at the station-level may be necessary.

A direct demand model of rail ridership in the Washington D.C and Maryland area used dummy variables to indicate the presence of station-level park and ride lots and feeder bus service (C. Liu et al., 2016). When only commuter rail ridership was modelled, the author found that ridership increased significantly in the presence of feeder bus connections, while all other candidate variables were insignificant. The influence of parking supply on demand could not be tested as the researchers found that park and ride lots were present at all commuter rail stations included in their analysis.

In contrast, an examination of commuter rail ridership in Orlando, Florida, found that the provision of park and ride lots had a drastic influence on the number of boardings and alightings. However, data with respect to feeder bus connections was unavailable for analysis (Rahman et al., 2019). The results suggest that station accessibility indicators are significant in influencing commuter rail ridership, although the joint influence of both factors on ridership warrants further examination.

A study of light-rail transit ridership in the United States suggested that the provision of feeder bus service might have a greater impact on growing ridership compared to additional park and ride spaces (Kuby et al., 2004). They found that the addition of a bus connection resulted in 123 more weekday boardings, indicating that riders not only respond to the presence of multi-modal connections, but also the frequency of such offerings. The number of park and ride spaces at each station was also included in their analysis; the authors identified an elasticity of 0.77. Both variables demonstrated a p-value < 0.001, indicating that station access played a large role in influencing boardings throughout the study area. Similar results were found in an analysis of rail rapid transit ridership in Canada, as feeder bus connections and the presence of park and ride lots were included in their direct demand model (Durning & Townsend, 2015). The availability of bus connections was represented using a dummy variable, while a continuous variable representing the number of park and ride spaces at each station was analyzed. Station-level ridership was found to increase by 40.88% with the provision of feeder bus connections, while ridership increased by 16.2% for every additional park and ride space offered. Both authors highlight that coordinating and facilitating a greater amount of alternative transit connections, rather than providing additional parking facilities, may be more effective in growing rail ridership.

The influence of station access indicators on demand has been found to be more influential to some users once trip type is controlled for. Chen and Zegras (2016) used dummy variables to indicate the presence of feeder bus connections in their station-level demand elasticity analysis of Boston's commuter rail network. Differing elasticities were found depending on trip type; an elasticity of 0.44 was found during a.m. peak period, while an elasticity of 0.252 was found during the p.m. peak period. Furthermore, Wasserman (2019) found parking supply to be a significant determinant to ridership in the a.m. peak period, whereas an insignificant relation was identified when ridership in the p.m. peak was examined. Their findings indicate that commuters may be more sensitive to the provision of multi-modal connections, specifically during the a.m. peak period, compared to other users.

The literature reviewed suggests that park and ride capacity and the provision of feeder bus connections has a positive influence on ridership, although further research is warranted.

3.2.5. Population

Various studies have attempted to quantify the relationship between population density and public transit use. Several researchers have found that people are more likely to use transit in dense areas, but are less likely to travel long distances to satisfy their trip purposes. A review of transport behavior in the United Kingdom found that when the density of an urban space is less than five persons per hectare, only 0.11 rail journeys per person per week were generated, compared to a rate of 0.63 journeys per person per week in areas with a population density greater than 50 people (Balcombe et al., 2004). However, distance travelled per person per week via all public transit modes increased with density, but overall distance travelled demonstrated an inverse relationship. Transport behavior studies in both the European and North American context reiterated these findings. A study of homemakers in New York City found those living in dense urban areas demonstrated a transit mode share of approximately 16.5%, compared to a transit mode share of 1% for those living in suburban locations (C. Chen & McKnight, 2007). A study of German transport behavior also found that transit use differed significantly urban and suburban respondents (Frondel & Vance, 2011). As suggested by Frondel & Vance (2012), this is likely a result of increased transit service coverage and density typically shown in urban communities, which therefore shortens transit access distance and travel times compared to transit systems that operate in rural areas. Furthermore, the authors suggest that the distance between residential locations and places of importance are reduced when population density increases, which further incentivizes transit use as the disutility of using a private automobile to complete a short distance trip is much greater compared to transit modes.

Previous studies have shown that total population has demonstrated a significant positive influence on transit demand. Boisjoly et al. (2018) examined ridership trends in 25 cities throughout Canada and the United States, although only areas with a population greater than 1.5 million were included in his analysis. The total population of surrounding metropolitan areas was used as an independent variable in their analysis. A demand elasticity of 0.339 was found, indicating that population has a consistent impact on ridership in large urban areas. Population was also found to be statistically significant in an analysis of light-rail transit ridership in the United States, as the results indicated that for population increases of 100 residents, an additional 9.2 boardings would be realized (Kuby et al., 2004). The authors note that policies aimed at increasing population size and densities should be investigated to encourage ridership growth. An analysis of station level rail ridership in San Francisco also found population to be determinantal in explaining ridership, although significance differed depending on trip type examined (Wasserman, 2019). Total population within a 0.5-mile radius of each station was tabulated and included in their model. Smaller, but still significant elasticities of 0.126 and 0.105 were found when weekday a.m. and p.m. peak ridership was examined. Notably, population was found to be insignificant in explaining weekend demand. The results suggest that population may be more determinantal in explaining commuter related travel, compared to discretionary demand.

Population density has been used as a factor in other transit demand studies rather than total population, although similar results have been shown. Guerra et al. (2011) found a statistically significant demand elasticity of 0.37 when population density was included in their station-level regression model. They state that while their findings were statistically significant, the majority of station areas included in their study were located in suburban locations. Therefore, demand with respect to population density could be understated, as a substantial amount of variation in their population density dataset did not occur. The authors suggest that the impact of population density on ridership might be more significant if estimates obtained from a wide range of environments were included in their model. More specifically, population and dwelling density were included as independent variables in a demand elasticity study of 342 rail stations throughout Canada (Durning & Townsend, 2015). Dwelling density was eliminated from their analysis as it displayed multi-collinearity with other independent variables, whereas population density was shown to demonstrate a statistically significant elasticity of 0.326. The author suggests that increasing density around existing stations, coupled with increased density targets around new station locations, could be instrumental in growing ridership figures. On a smaller scale, Miller et al. (2017) included population density as an independent variable in their rail demand study of the Chicago Transit Authority. Like Wasserman (2019) separate demand models were estimated for different trip types. Their overall ridership model found population density to be statistically significant, as did their model that only assessed Saturday demand. The authors theorize that Saturday ridership might be particularly impacted as high-density neighborhoods throughout the study area are often correlated with low levels of car ownership, thereby meaning that transit is needed if residents want to take leisurely weekend trips. Additionally, the authors theorize that these locations could be large leisure trip generators, as an abundance of recreational and social activities that occur in these areas draw people in from other areas of the city, but do not have a large amount of parking supply. Therefore, public transit is needed both for outgoing residents and incoming tourists. These findings suggest that the inclusion of other variables in the model that might capture these relationships, such as employment density or parking supply, could further distinguish the role of population density in influencing ridership demand.

Both total population size and population density have been included as factors in several transit demand studies. An analysis of transit use in Atlanta, Georgie found demand elasticities of 1.284 and 0.5, indicating that consumer response to change in population might be overestimated if aggregate population is used as an indicator (Brown et al., 2014). Chen and Zegras (2016) also included both population and population density in their station-level examination of rail ridership in Boston. Only a single variable representing population was selected for inclusion in each model based on which variable had the greatest impact on model performance. When a.m. peak period ridership was modelled, population density was shown to be most significant in determining ridership; a demand elasticity of 0.576 was identified. Total population was also found to be statistically significant when off-peak and p.m. peak period ridership was analyzed, as elasticities of 0.34 and 0.33 were shown.

More advanced studies have also shown that population density can be positively correlated with other important factors, such as service supply. Therefore, two-stage methods are sometimes used to estimate appropriate service supply indicators, while population is represented using alternative metrics that do not display multi-collinearity with other independent variables. Taylor et al. (2009) initially identified a positive relationship with population density and service supply. They theorize that this is the case as increased levels of service supply are expected to be delivered by transit agencies that service large populations, thereby increasing ridership as service supply increases. To prevent collinearity in their models, they first used population density as a factor to explain total transit ridership, and then used total population as an instrumental variable to estimate service supply. Consistent with the previous literature, an elasticity of 0.42 was found. Per capita ridership was then modelled, where population density was instead used as the instrumental variable to estimate service supply, and geographic land are of the surrounding metropolitan area was used to represent population. A smaller yet still significant demand elasticity with respect to population of 0.19 was found.

A similar approach was also used in a study of Canadian transit demand, as vehicle revenue hours was estimated using population (Diab et al., 2020). Different dwelling types were instead used as a proxy measure of population density, as the author theorized that the presence of single-family homes are often correlated with sprawled urban areas, whereas apartments and row houses are more prevalent in dense environments. When overall ridership was modelled, apartments and row houses demonstrated a significant and positive correlation with ridership, while an elasticity of -0.342 was shown in relation to single-family dwellings. However, the sign of the demand elasticity with respect to apartments shifted once transit agency size was controlled for. For example, when ridership for large agencies was modelled, a demand elasticity of -0.404 was found, but a positive elasticity of 0.535 was identified when small agency ridership was modelled. The authors suggest that the increased presence condos in large Canadian cities might discourage the use of transit in these areas, as increased wealth and automobile ownership is also typical of condo owners. The results suggest that if collinearity between service quantity variables and population factors is significant, proxy measures of population can be used as a viable source of information. However, population should be represented using magnitude or density measures if quantifying this relationship is of significant importance to the researcher, as information garnered via proxy measures could be difficult to translate into practice or policy recommendations.

To summarize, the results suggest that a variety of metrics can be used to estimate the impact of population size and density on transit demand, however a statistically significant positive relationship should be expected.

3.2.6. Employment

Findings have suggested that number of jobs, in addition to number of residents, can influence transit demand to a large degree. A variety of metrics, including total number of employees, employment density, and number of commercial or retail locations have been used in demand elasticity studies to evaluate the impact that economic output has on ridership. A cross-sectional analysis of transit ridership in Houston and San Diego found that ridership was heavily influenced by the distribution, location, and density of jobs, employment centers, and workplaces (Kain & Liu, 1999). After analyzing ridership and socioeconomic statistics obtained between 1980 and 1990, an employment elasticity of 0.25 was identified in both cities. Regardless, the authors estimated that employment increases had a greater impact on ridership growth experienced in San Diego compared to that experienced in Houston, as employment growth in San Diego was concentrated in the CBD of the city. The authors suggest that the implementation of zoning by-laws was key in facilitating the development realized in San Diego, whereas development in Houston is allowed to happen in a sporadic and haphazard fashion. The authors theorize that demand for transit can be maximized if land use policies that guide employment intensification in key areas are implemented in conjunction with adequate transit service between key residential and employment locations.

Number of jobs within a 0.5 mile radius of each station was used as a factor to explain rail demand in the United States (Guerra & Cervero, 2011). After applying a fixed effect panel data estimator, a statistically significant demand elasticity of 0.597 was found, the largest elasticity produced by the analysis. An analysis of light rail transit ridership in the United States also found that total employment was significant in explaining transit demand, and estimated that an additional 100 jobs should result in an additional 2.3 boardings (Kuby et al., 2004). However, the authors note that their findings should be interpreted with caution, as different metrics were used to compile employment data across entities. Regardless, the results suggest that station-level transit demand is sensitive to the number of employment opportunities in surrounding areas.

Total employment opportunities in the surrounding metropolitical area has been used as a factor in the absence of station-level employment statistics. Using this metric, Schimek (2015) found that employment was statistically significant in both the short and long run, as demand elasticities of 0.21 and 0.41 were found respectively. Notably, a study of rail ridership in Chicago used the same metric, but found that demand was not responsive to change in total employment (Lascano Kežić & Durango-Cohen, 2018). However, the authors note that their findings could be minimized by the fact that little variation in employment occurred over the time-series analyzed, suggesting that the results computed by Schimek (2015) are more reliable.

Consistent with Schimek (2015), additional studies have revealed that demand elasticities with respect to employment can differ depending on time period examined. A review of New Jersey Transit's commuter rail system identified a short-run elasticity of 0.00, whereas a long-term elasticity of 0.59 was also estimated (C. Chen et al., 2011). The authors highlight that a short-run demand elasticity of 0.00 is unusual, but note that there is often a lag between change in employment and change in settlement and transport behavior, which could explain this finding.

Wasserman (2019) included jobs at destination and jobs at origin as employment variables in his station level analysis of San Francisco's rail network. His results suggest once trip type is controlled for, employment has varying impacts on transit demand. For example, jobs at origin demonstrated an elasticity of 0.121 in the a.m. peak period, while an elasticity of 0.347 was displayed in the p.m. peak period. Nearly opposite elasticities were found when the impact of jobs at destination was evaluated, as an elasticity of 0.364 in the a.m. peak period was identified compared to an elasticity of 0.173 during evening rush hour. The author states that these results are expected, as the majority of jobs in San Francisco are concentrated in the CBD, resulting in heightened transport flows between this location and residential areas throughout the city. The results indicate that the identification and provision of service between key trip generators can be effectively used to encourage mode shift.

Both aggregate employment and employment density were included as independent variables in rail ridership studies conducted in Boston, Massachusetts, and Atlanta, Georgia (Brown et al., 2014; S. H. Chen & Zegras, 2016). Brown et al. (2014) theorized that employment figures at the riders' destination would have a significant impact on rail use, whereas Chen & Zegras (2016) also theorized that employment would be a primary driver of ridership. Interestingly, Brown et al. (2014) found a demand elasticity with respect to total employment of 1.231, but discovered a negative correlation between employment density and ridership. Elasticities generated by Chen & Zegras (2016) displayed the expected sign regardless of the employment indicator used, although elasticities differed depending on trip type examined. An elasticity of -0.137 was identified during the a.m. peak period, while an elasticity of 0.507 in the p.m. peak was shown. Their results conform to the expectation that the majority of trips in the a.m. peak period are home-based, whereas the opposite is true during the p.m. peak period, thus indicating that the majority of users are commuters. Furthermore, the results generated by Brown et al. (2014) could indicate that the majority of jobs are located at the fringe of the city, therefore indicating that transport and commuter related flows in Atlanta are not concentrated towards the CBD during peak periods.

Various studies have determined that employment type can influence transit demand to differing degrees. Zhang & Wang (2014) included statistics relating to the amount of retail and storage area in their model of metro ridership in New York City. The authors theorized that the presence of retail areas would sustain traffic throughout the day, thus resulting in increased transit demand. Their demand model found that the presence of such establishments had a significant impact on ridership; station level demand was estimated to be 3.19% higher if the amount of retail area increased by 10%, whereas ridership decreased by 0.5% if the amount of storage area increased by 10%. Nearly identical estimates were calculated in an analysis of rail ridership throughout Canada (Durning & Townsend, 2015). After concluding that job density displayed multi-collinearity with other independent variables, commercial site density was selected as a proxy indicator to measure the impact of economic make-up on ridership. The authors found a significant demand elasticity of 0.327. Chen & Zegras (2016) further identified that density of retail outlets can influence transit demand, as an elasticity of 0.303 was computed, however findings were only estimated with respect to off-peak ridership. An analysis of the Orlando SunRail commuter rail system also found that the presence of commercial and financial centers within a 1500 metre radius of stations had a significant impact on the number of station-level boardings (Rahman et al., 2019). Weekday boardings were significantly higher if commercial and financial centers were located in close proximity to the station, whereas weekday alightings were also significantly higher if commercial centers were nearby. The results suggest that the presence of mixed-use developments in surrounding station areas is key in influencing both commuter related and discretionary demand for rail systems, as ridership is stimulated in all time periods.

The results indicate employment has a statistically significant positive impact on ridership, while employment types such as retail and commercial uses can influence demand regardless of trip type examined.

3.2.7. Unemployment Rate

Unemployment rate has been shown to demonstrate differing relationships with transit demand. Surprisingly, limited findings with respect to this variable were presented in the reviewed literature. An analysis of ridership in Chicago, Illinois, found unemployment rate to be determinantal in influencing demand, especially when commuter rail ridership was examined (Nowak & Savage, 2013). Unemployment rate demonstrated an elasticity of -0.149, whereas elasticities ranging in value from -0.055 to 0.005 were found when heavy rail, city bus, and suburban bus systems were examined. The results indicate that commuter rail systems are more sensitive to economic downturns and corresponding shifts in household socioeconomic statistics. The authors theorize this should be expected, as regional transit systems are more heavily utilized by commuters than casual users.

Differing findings were presented in a study of rail ridership throughout the state of Washington (Stover & Christine Bae, 2011). Separate regression models demonstrated a significant positive relationship in six of seven transit systems examined. Additionally, a panel data estimator identified a significant elasticity of 0.29. Unlike Nowak & Savage (2013), the results suggest that as unemployment increases, transit demand will increase. The authors theorized that sustained economic downturn could transition choice riders to captive users if they cannot afford costs associated with vehicle ownership, and that decreasing commuter flows during recessionary periods could be offset by those switching to public transit in order to save money on transport.

These findings were reinforced in a study of transit demand in 10 major urbanized areas throughout the United States (Iseki & Ali, 2014). When the authors regressed ridership on a variety of variables, unemployment rate was found to demonstrate a statistically significant positive relationship with ridership in all modes examined except commuter rail. For example, percentage-wise elasticities ranging in value from 0.028 to 0.0343 were found, whereas unemployment rate was excluded from the commuter rail model altogether due to insignificance. Unemployment rate was also found to be insignificant in explaining transit demand in a variety of studies that assessed transit demand in areas where demand was largely comprised of commuters (Boisjoly et al., 2018; Diab et al., 2020; Taylor et al., 2009).

The results suggest that unemployment rate could demonstrate an insignificant relationship with demand if the customer base is primarily comprised of commuters.

3.2.8. Seasonality

Various studies have shown that transit demand can demonstrate seasonal fluctuations. Seasonal dummy variables were included to assess the impact of climate stability on ridership in nine major American cities (Lane, 2010). Using winter as a baseline, the authors found that seasonal factors were significant in explaining demand for those cities located in temperate climates, as demand was significantly reduced during winter months. Similar results were found in an analysis of light rail transit ridership throughout the United States, as cities with stable climates were associated with up to 300 additional boardings per station compared to the average (Kuby et al., 2004). In contrast, ridership in cities associated with temperate climates saw ridership reduced by the same amount during cooler seasons. Further, a demand study of metro riders in Chicago found that demand for transit during the a.m. peak period was significantly lower in winter and spring seasons compared to the summer baseline. The authors suggest that the abundance of adverse weather conditions and cold temperatures during winter months is a rational explanation, as public transit users are often exposed to environmental conditions when waiting for vehicles. Therefore, demand often shifts to private modes as the utility of an enclosed, sheltered, and warm vehicle for the majority of the trip outweighs the disutility of vehicle operation costs.

Various studies have found that the disutility of transit use can also increase in the presence of adverse weather conditions, regardless of season. A demand analysis of rail users in Orlando, Florida, found that daily use declinded in the presence of rain and wind (Rahman et al., 2019). An analysis of metro riders in New York City illustrated similar findings, although impacts differed depending on time of day examined (Singhal et al., 2014). For example, a.m. peak period ridership was most significantly impacted by rain and abnormal temperatures, while the presence of rain and snow was correlated with a decrease in demand during the midday off-peak period. Ridership during the p.m. peak period also declined significantly if snow was present, whereas rain demonstrated insignificant impacts. Additional models were developed by Singhal et al. (2014) to determine if weather related variables had differing impacts on station-level ridership depending on the type of platform infrastructure used. They found that in the presence of various weather events including rain, heavy rain, wind speed, and warm days, ridership at elevated stations was significantly impacted compared to ridership at underground stations. Notably, heavy snow was found to increase ridership at both station types, while regular snow displayed a negative correlation. The authors theorize that the presence of snow may deter riders due to safety concerns related to station access, while heavy snow could result in private auto users shifting to public transit due to unsafe driving conditions. Regardless, their results suggest that elevated stations may benefit from station infrastructure improvements that include weather protection features.

To summarize, the literature indicates that cold temperatures, snow, and adverse weather conditions related to seasonal climate trends can negatively impact transit ridership for systems located in temperate climates.

3.2.9. Income

Previous research has revealed that the effect of income on transit demand is significant. Balcombe et al. (2004) states that the sign and magnitude of demand elasticities with respect to income will vary depending on the income level of the user. They note that all things being equal, an increase in income typically results in an increase in car ownership, therefore leading to a reduction in the demand for transit. Otherwise known as the income effect, a study of transport behavior in Montreal, Quebec, revealed a similar relationship as the authors identified a negative relationship between income and the likelihood of using transit (Grimsrud & El-Geneidy, 2013). Further, the authors estimated that an increase in income of $1000 decreases the likelihood of using public transit by 1%. Holmgren (2007) completed a metadata analysis of transit demand studies in an attempt to identify a common income elasticity. Their analysis was further disaggregated by geography and time period examined, but an elasticity of -0.62 was shown in all estimates. The authors theorize that as disposable income increases, demand for public transit decreases as other modes become more affordable and accessible to the user (Balcombe et al., 2004; Grimsrud & El-Geneidy, 2013). Therefore, the literature suggests that public transit is an inferior service in all contexts.

Median household income was included in a mode share study conducted in the Calgary, Alberta (Pasha et al., 2016). 185 community areas were assessed in an attempt to determine the main factors influencing transit use and associated mode share percentages throughout the city. Income was classified into four levels representing basic income, low income, middle class, and wealthy households. Consistent with Grimsrud et al. (2013), transit use was positively associated with those earning less than $40,000 per year, whereas the opposite was found for those earning more than $125,000 per year. The authors suggest that these findings should be used to develop more equitable transit policies and programs.

Schimek (2015) identified a similar relationship, as their overall demand model identified a short-run elasticity with respect to per capita income of -0.36. Unlike Holmgren (2007), effects were shown to be more drastic over long time periods, as a long-run elasticity of -0.69 was identified. Notably, they found that demand responded differently when urban area size was controlled for, as large urban areas demonstrated short and long-run elasticities of -0.37 and -0.91, whereas elasticities of -0.26 and -0.5 were found in small urban areas. The authors theorize demand could be more elastic in large urban areas as these spaces typically have a large share of commuters with more disposable income. In contrast, the customer base in small urban areas is typically comprised of low-income captive riders. Therefore, the income effect is more prevalent in large urban areas as a greater proportion of residents with alternative transport options live in these spaces.

Regardless, studies specific to commuter-related transit systems found an opposite relationship. Transit ridership in major American cities demonstrated an income elasticity of 0.92, indicating that demand is stimulated when income increases (Taylor et al., 2009). Balcome et al. (2004) further identified that income has a positive impact on rail demand, as a range of income elasticities varying between 0.11 and 2.07 were found. Notably, the largest income elasticity computed in their study was for those travelling between suburban locations and the CBD. Chen and Zegras' (2016) found that income was positively correlated with commuter rail ridership in Boston, but only during the a.m. peak period; an elasticity of 0.398 was found. Consistent with Chen et al. (2016), an income elasticity of 0.27 was identified in an examination of station level ridership in New York City (Zhang & Wang, 2014). Taylor et al. (2009) suggests that these findings could be explained by the presence of high paying jobs typically located in the central business district of metropolitan areas, therefore resulting in the use of regional transit infrastructure to facilitate commuter related travel patterns. Previous studies have also theorized that since most rail users already have access to a private automobile, additional income rarely generates an additional rider (Paulley et al., 2006). Instead, additional work responsibilities associated with increases to income could additional journeys to and from work, therefore increasing transit demand in the process.

To summarize, the literature suggests that transit demand demonstrates a negative correlation with income, but only in areas where the customer base consists of a variety of users. The opposite correlation should be expected of systems widely comprised of wealthy commuters, as variables other than income have a greater impact at mode choice decisions.

3.2.10. Vehicle Ownership

Vehicle ownership has shown to be correlated with transit demand, but findings are scarce relative to other variables such as fare price and service levels. Various studies have suggested that findings are limited as vehicle ownership is often correlated with other factors such as income, thus warranting it's removal from the demand model (Balcombe et al., 2004; S. H. Chen & Zegras, 2016; Grimsrud & El-Geneidy, 2013; Iseki & Ali, 2014). Regardless, a negative correlation with transit demand has been illustrated, although further research is commonly recommended. A mode share review of various countries throughout the European Union notes a distinctive difference between vehicles owned per capita and mode share (Balcombe et al., 2004). Of 15 countries reviewed, the number of vehicles owned per capita often demonstrated a negative correlation with public transit mode share. Regardless, results were not consistent, as some countries with a high mode share also had a high level of vehicle ownership, while the opposite was true for others. The authors theorize that regional contexts play a role in these relationships, therefore making it hard to identify a "common" car ownership elasticity.

Holmgren (2007) further states that demand research with respect to vehicle ownership is relatively new. In their metadata analysis of ridership determinants in Europe, America, and Australia, a variety of elasticities ranging from -0.21 to -2.75 are identified. Unlike Balcombe et al. (2004), he notes no variation based on geography or period of analysis, and recommends an overall demand elasticity of -1.48. Regardless, the author states that only eight studies included in his review accounted for car ownership in their regression models, and therefore recommends that results should be interpreted with caution.

Research that included a large cross-section of entities found that the proportion of carless households was statistically significant in influencing transit demand (Boisjoly et al., 2018). In their model that aggregated bus and rail service levels, a demand elasticity of 0.447 was identified, the second largest generated. An additional model which disaggregated bus and rail service quantity into separate categories found a lesser but still statistically significant elasticity of 0.253. Based on the results, the authors suggest that policies aimed at reducing vehicle ownership could be most effective at increasing transit demand.

A study of transit demand in Boston found similar results, as average household ownership demonstrated a statistically significant negative relationship with station-level ridership (S. H. Chen & Zegras, 2016). An elasticity of -0.469 was found for their daily weekday ridership model, consistent with findings from Boisjoly et al. (2018). However, the authors note that factor was removed due to multi-collinearity when other sources of information, such as distance to CBD and station accessibility indicators, were included in the model. The results provide further explanation as to why findings in relation to this variable are scarce when demand elasticity studies are undertaken.

Finally, the likelihood of using transit in Montreal was estimated using a number of independent variables including number of cars per license (Grimsrud & El-Geneidy, 2013). This variable was selected as it demonstrated less correlation with demographic factors compared to household vehicle counts. Vehicle ownership was found to display the greatest impact on transit use, as an elasticity of -2.98 was found. Notably, once observation period was controlled for, the impact of this variable on transit use decreased significantly, suggesting that a growing number of people are keeping automobiles but might not be using them for commuter related transport. Their findings suggest that the impact of vehicle ownership on transit use could have differing impacts depending on trip type examined.

To summarize, vehicle ownership is expected to have a negative correlation on ridership, although findings with respect to differing trip types is somewhat unclear.

3.2.11. Gender

Previous research has shown that gender can influence demand for transit. Like income, these findings are relatively limited in the space of demand elasticity studies, but travel surveys and qualitative studies have suggested this correlation. An analysis of origin-destination survey data collected in Montreal was used to explain station access distance during the a.m. peak period (Vijayakumar et al., 2011). Notably, gender was found to be one of the most influential factors, as they estimated that station access distance of males is 12.5% larger compared to females. The results suggest areas that contain a higher percentage of male residents should result in larger ridership figures, as the station catchment area and associated customer base is greater compared to environments with a larger proportion of females. In contrast, a similar analysis of origin-destination surveys in Montreal found that female respondents were more likely to use public transit for commuter related trips (Grimsrud & El-Geneidy, 2013). Unlike Vijayakumar et al. (2011), survey data obtained in 1998, 2003, and 2008 was analyzed in an attempt to explain transit use. Notably, transit use was found to be higher amongst female respondents, although the significance of this relationship declined over time. For example, observations obtained in 1998 found this relationship to be statistically significant in all age groups examined, whereas a statistically significant relationship was only present in the 50-54-year-old age group in 2008. The results suggest that transit demand could be influenced by gender, although impacts might decline as service and system infrastructure is developed to higher standards. Further investigation is recommended by the authors.

Like Grimsrud & El-Geneidy (2013), a ridership demand study in Calgary found that communities with a large proportion of male residents were more likely to take transit (Pasha et al., 2016). Further, communities associated with female lone parents did not use transit often in their study. An analysis of data obtained from the Utah Household Travel Survey further found that females have a significantly larger cost associated with transit use compared to males, therefore resulting in lower trip generation rates (Farber et al., 2014). Both studies indicate that the level of disutility associated with transit use is greater for females compared to males, although rationale is not provided. Notably, the demand study produced by Farber et al. (2014) included both bus, light rail, and commuter rail ridership, indicating that these effects could be expected regardless of mode examined.

Various qualitative studies have theorized that gender differences in transport behavior are often a result of safety concerns. In Los Angeles, California, travel surveys were distributed to residents located within six station catchment areas before and after the corridor became operational, to see if mode share or trip generation rates changed significantly. Respondents were asked to record trip activity, including information regarding socioeconomic characteristics, over a seven-day period. Additional questions ranked on a 7-Point Likert scale were asked regarding environmental beliefs and safety concerns related to transit use. The analysis revealed that change in transport behavior varied depending on the gender of the respondent. For example, a ridership was negatively correlated with gender once a a female respondent dummy variable was included in the model. Additionally, the respondent's measure of safety was also significantly correlated with female respondents and transit use. A qualitative analysis of female transit users in Irvine, California, generated similar findings (Hsu, 2008). The authors hypothesized that in the face of sexual harassment vulnerability, female riders might alter their travel patterns or switch modes altogether if they feel unsafe. Using a panel of 18 middle-aged female respondents, questions regarding their perception, personal experience, position regarding proposed and existing policies aimed to address these issues were asked. A qualitative review of responses found that safety and security might be an important consideration for females when making decisions regarding transport behavior. For example, respondents were found to make changes to travel patterns, routes, or boarding/alighting destination if they were concerned about the occurrence of sexual harassment. The results suggest that females are less likely to use transit compared to males due to safety and security concerns disproportionally realized by the female gender.

A study of transport behavior in Chicago suggested an alternative explanation (C. Miller & Savage, 2017). When year over year ridership trends were examined, the author also found that station catchment areas containing large proportions of males were significantly correlated with increases in station level ridership. Furthermore, results were significant in all day types examined, including weekdays, Saturdays, and Sundays / holidays. Unlike the results summarized above, the authors theorize that changes to the urban environment, rather than attitudinal concerns regarding safety and security, can explain these trends. For example, the authors outline that neighborhoods with unusually large proportions of males emerged in areas that were gentrified during the study period, and that the first "new wave" of settlers in these areas were men. In contrast, females were found to live more frequently in neighborhoods where development remained stable. They further note that neighborhoods associated with gentrified developments also increased trip generation rates during off-peak hours, as restaurants and nightlife associated with these areas attracts traffic in all time periods. Therefore, they theorize that trends associated with gentrification, settlement, and change to the urban form are likely to explain gender differences in transit demand.

To summarize, the literature suggests that gender differences regarding safety and security can negatively impact female's likelihood of using transit, but contextual information regarding development patterns and changes to the urban form can also explain observed trends.

3.2.12. Age

Various studies have suggested that younger age cohorts are more likely to use public transit. The direct impact of age on mode choice behavior was analyzed in a study of origin-destination surveys by Grimsrud et al. (2013). After computing separate mode share models that controlled for age of the respondent, they determined that transit mode share was larger for young persons compared to older aged populations. The authors theorize that this could be attributed to an increase in environment and sustainability related material in the school curriculum, thus resulting in younger age groups choosing travel modes that are more sustainable. They also theorize that attitudes and behaviors of younger adults are changing, meaning that negative perceptions associated with public transit are not as abundant compared to older populations. The author recommends that discounted fares for young professionals could be implemented to further increase ridership, whereas discount fare cards received by university and high school students should be sustained.

Demand elasticity studies that controlled for age revealed similar findings, as a study of rail users in the United Kingdom found that children were less affected by fare price changes compared to older populations (Balcombe et al., 2004). A fare price elasticity of -0.47 was calculated, whereas adults and the elderly/disabled demonstrated fare elasticities of -0.59 and -0.77 respectively. The authors note that children are more likely to be captive users, while older aged populations are likely to be discretionary users. Therefore, the authors state that differing elasticities should be expected as income, vehicle ownership, and trip purpose can vary between age cohorts.

Percent of population in college was included as an independent variable in an examination of transit ridership trends throughout the United States (Taylor et al., 2009). The authors identified a significant relationship with respect to total ridership, as an elasticity of 0.228 was estimated. Similar results were identified in an analysis of ridership trends in Canada, as the proportion of postsecondary students demonstrated a statistically significant elasticity of 0.117 (Diab et al., 2020). Taylor et al. (200) notes that the inclusion of this variable rendered other variables, such as poverty, insignificant as areas examined with the largest poverty rates were those located within college towns. Like Balcombe et al. (2004), the results indicate that younger populations are more likely to take transit compared to older populations, most likely a result of low income and unavailability of other mode choices. A statistically significant, yet substantially smaller demand elasticity of 0.02 was identified in a demand elasticity study of 67 urbanized areas throughout the United States, but the authors theorize their results could be minimized by the fact that annual data was used to record this variable, whereas the majority of other variables included in the dataset were made available at monthly intervals (Lee & Lee, 2013).

Diab et al. (2020) also included percent of population that are children and percent of population that are senior to see if transit demand was stimulated by those not in the workforce. Both variables were excluded from their overall ridership model due to insignificance. However, percentage of population that are senior was included in their model that specifically examined ridership in large transit agencies, where a statistically significant elasticity of 0.123 was found. The results indicate that older aged adults in large urban areas are more likely to use transit, most likely to access recreational or discretionary uses within the city.

The literature suggests that younger people are expected to use transit if alternative travel options are not readily available, while older people could be more likely to use transit in off-peak periods due to discretionary related travel patterns.

3.2.13. Households with Children

The literature illustrates that transit demand is negatively influenced by the number of dependants, specifically children, at the household level. A mode share study in Montreal found a statistically significant negative correlation between transit use and the presence of children aged 5 years or younger. A negative elasticity of -0.22 was found when mode choice across all time periods and age groups was considered (Grimsrud & El-Geneidy, 2013). Similar findings were identified in a study of transit ridership in Utah, as transit use demonstrated an inverse relationship with the number of children and retirees at the household level (Farber et al., 2014). El-Geneidy et al. (2013) theorized that adult members are faced with more responsibility when young children are introduced to the household. Therefore, the disutility of private auto ownership is offset by the convenience and travel time benefits generated, thus making it more favorable compared to transit. Findings generated by a case study of the Orlando SunRail commuter rail system reinforced this observation, as the authors found that station-level ridership was reduced significantly in the presence of education centers (Rahman et al., 2019). Assuming that a greater proportion of families and children live in close proximity to schools, observed transit figures could be explained by a large proportion of the customer base using private auto to facilitate school-related transport patterns in the a.m. and p.m. peak periods. El-Geneidy et al. (2013 further disaggregated their analysis by age group, and found greater significance amongst individuals whose age ranged from late 20s to early 30s. Their results suggest that transit demand is temporarily reduced for adults once children are born, but impacts are reduced as children eventually grow older and become more transit dependant.

A study of transit ridership in Calgary tested many factors in an attempt to quantify the impact of children on mode choice behaviour (Pasha et al., 2016). Of these, number of children less than 14, a family size of three, share of lone parent families with female lone parent, and share of couple families without children were all found to be statistically significant in explaining mode share. A negative coefficient associated with number of children less than 14, coupled with a positive coefficient of share of couple families without children, indicated that communities with large proportions of young children were more likely to be auto dependant, thus reducing transit usage. The authors theorize that inconvenience associated with travelling on public transit with kids could be a factor, while kids might be prevented from using transit themselves due to parental concerns regarding safety and security. However, the authors also found a statistically significant positive correlation with family size and transit mode share, suggesting that large families with older children are more likely to use transit due to diversified trip destinations. Consistent with El-Geneidy et al. (2013), the results suggest that presence of pre-teen children is a limiting factor on household transit use.

The literature indicates that private modes are favoured when children are introduced to the household, likely due to convenience and travel time benefits that cannot be matched by public transit modes. Therefore, the presence of children at the household level is expected to have a negative influence on ridership.

3.2.14. Fuel Price

A variety of studies has shown that the price of gas is a determinantal variable that can influence transit demand. Frondel (2011) examined the main determinants of transit ridership in Germany, paying specific attention to the elasticity of ridership with respect to fuel price. Using household travel surveys obtained from the German Mobility Panel, he found that increased fuel prices were a major factor in explaining transit demand, as an elasticity of 0.262 was identified. Various studies have reinforced these observations, but have noted that demand elasticities are influenced by the availability of alternative modes. For example, a survey of residents in Austin, Texas, found that only 17.7% of private auto users would switch to transit if fuel prices were increased (Bomberg & Kockelman, 2012). In contrast, an analysis of rail demand throughout the United States found that demand elasticities with respect to fuel price were typically largest in cities with expansive transit systems, such as Cleveland, Ohio, and Seattle, Washington. Average price per gallon of fuel was included as a factor in an analysis of transit ridership throughout the state of Washington, where an aggregate elasticity of 0.172 was identified when all entities were modelled (Stover & Christine Bae, 2011). However, insignificance was identified in four of the eleven areas examined when individual regression models were computed. Findings were further disaggregated when urban area size was controlled for, as a long-run elasticity of 0.22 was calculated for large urban areas, compared to a 0.13 long-run elasticity shown in small urban spaces. The results indicate that adequate transit service offerings need to be in place for consumer response to occur, as mode shift cannot occur if alternatives are not available.

Several studies have noted that short run consumer response to fuel price increases could be minimal. A study of New Jersey Transit ridership identified a short-run elasticity of 0.11, compared to a long-run elasticity of 0.19. A comparable but insignificant short-run elasticity was identified in a separate study of New Jersey Transit ridership, but the authors note that their time-series analyzed might not have been long enough to capture changes in consumer response (Yanmaz-Tuzel & Ozbay, 2010). Similar conclusions were reached in a study of city and regional transit systems in Chicago (Nowak & Savage, 2013). When Metra commuter rail ridership was assessed, a ridership elasticity with respect to gas price of 0.002 was calculated. The authors theorize that their results might be lower than previous estimates, as their 12-month data collection period would not have captured long run effects. These findings indicate that users are relatively insensitive to short-term fluctuations in gas price, and that maintained increases are instead needed to encourage mode shift.

Demand elasticities with respect to fuel price have further shown to differ depending on spatial context. A meta-analysis of demand elasticities studies identified a short-run elasticity of 0.4 for those conducted in Europe, compared to a short-run elasticity of 0.82 for those completed in America and Australia. Similar trends were noted when long-run elasticities were summarized. The findings suggest that European users are less responsive to fuel price increases compared to American and Australian examples. Stover and Bae (2011) theorize that this can occur in the presence of paid parking schemes, congestion, and toll roads, as the cost of driving could be increased to a threshold where fuel price changes would have marginal impacts. Guerra and Cervero (2011) also note that demand elasticities with respect to fuel price should be interpreted with caution, as fuel price is only one cost associated with private automobile operation. Therefore, change in other factors such as parking costs and toll road charges could further influence elasticity estimates if not controlled for. Balcombe et al. (2004) further states that the initial cost of car ownership is high in Europe therefore preventing market access for a large proportion of consumers. As a result, consumers in these markets are less responsive to change in automobile operation costs, as other variables associated with automobile ownership and access prevents mode shift from occurring. These findings highlight the importance of controlling for factors specific to the regional context when conducting demand elasticity studies.

Currie and Phung (2007) suggests that mode examined can heavily influence elasticity estimates. They calculated an aggregate ridership elasticity with respect to gas price of 0.104 for all modes, but a negative and insignificant ridership elasticity of -0.093 was calculated when commuter rail systems were analyzed in a separate model. In contrast, elasticity values ranging from 0.27 to 0.28 were calculated in an assessment of Philadelphia's regional rail system (Maley & Weinberger, 2009). The authors theorize that their findings could be influenced by the presence of more choice riders, whereas users of other transit systems in the area, such as city bus, are typically captive riders. Both authors note that other variables such as fare and service data were not included in their analysis, meaning that results should be interpreted with caution.

Demand elasticities have also been shown to differ depending on the baseline price of gasoline. For example, when commuter rail ridership was examined, a demand elasticity with respect to fuel price of 0.61 was identified. However, an elasticity of 0.527 was shown when a price threshold of three dollars per gallon was reached. Similar results were identified in Bomberg's (2012) study, who estimated that ridership increased to a greater extent when gas price increased and stayed at four dollar per gallon. The results suggest transit demand could be stimulated if high gas prices are sustained.

To summarize, the literature suggests that a wide range of factors, including price and availability of alternatives, mode examined, and baseline price can influence demand elasticities with respect to fuel price. However, a consistent and significant negative correlation with ridership should be expected if long-run impacts are captured.

3.3. Data Collection Methods

3.3.1. Delineation of Station Catchment Areas

According to the literature, station catchment areas are used to extract external variable datasets included in demand elasticity studies. A station catchment area is defined as a spatial boundary that encompasses the area where the majority of non-transferring passengers originate from (Andersen & Landex, 2008). In other words, the station catchment area can be viewed as the customer base for a given transit station or network. When using household statistics to formulate external variable datasets, station catchment areas should be as accurate as possible to ensure that data being captured is reflective of the customer base.

Station catchment areas are delineated using a variety of methods. Most commonly, station catchment areas are determined as a function of the maximum distance a transit user is willing to reach. In North America, this is commonly done by implementing a circular Euclidian buffer around a given station, ranging in distance from 400-800 metres (Andersen & Landex, 2008; Durning & Townsend, 2015; El-Geneidy et al., 2014; Guerra et al., 2012). The size of the catchment area reflects the assumption that a passenger walking at a speed of 1.3 metres per second can reach the station within ten minutes (Guerra et al., 2012). External variable datasets lying within the buffer are then extracted for further analysis. Network buffers are also commonly used to delineate station catchment areas. Unlike Euclidian buffers, aspects of the urban environment that might impede station access are incorporated into the estimate (Andersen & Landex, 2008). Therefore, distribution and layout of road networks, pathways, buildings, rivers, and other natural / built features are incorporated to generate a more realistic station catchment area that is not uniform in size or shape. Both Euclidian and network buffers can be estimated in terms of access time, rather than access distance, if the researcher chooses to do so.

A study of 21 transit systems across the United States was conducted to determine if demand model accuracy differed on the Euclidian distance used to extract external variable datasets (Guerra et al., 2012). Data was extracted from 1,449 station catchment areas in 21 American cities, using buffers ranging in size from 0.25 miles to 1.5 miles. Model results did not improve depending on the size of the Euclidian buffer used, indicating that a “correct” station catchment area size is far from clear, and likely varies with the spatial context of the study. The authors state that based on these findings, researchers should use datasets that are easily calculatable and readily available when extracting external variable data. However, commuter rail stations were not included in their analysis.

The use of Euclidian and network buffer thresholds can be improved by weighing extracted data according to station access distance, as transit ridership is negatively correlated with this variable. Keiier and Rietveld (2000) found that people living 500-1000 metres from rail stations use transit 20% less compared to people living within 500 metres from the station. Another study in south Florida used onboard surveys to determine that the majority of trips originated within 1,800 feet of a transit stop, while few trips originated more than 2,700 feet from the access station (Zhao et al., 2003). Using this data, the author was able to model a distance-decay curve which indicated that transit use beyond a 0.5 Euclidian buffer is 3% of that within a 300-foot Euclidian buffer. Therefore, the authors recommend that distance decay-functions should be applied to external variable datasets, as their results indicate that data extracted from areas in close proximity to the station is more characteristic of the customer base.

A direct demand analysis of transit riders in Madrid, Spain incorporated a distance decay function while extracting external variable data for use in their study (Gutiérrez et al., 2011). Consistent with Zhao et al. (2003), they stated that external variables extracted from station catchment areas should not be weighted equally, as households in close proximity to the station have a larger impact in explaining demand. Several demand models were tested using a variety of station catchment area delineation methods. They found that using a network distance threshold, coupled with a distance-decay function, improved model fit by 5.2% compared to the model using a Euclidian distance threshold with no distance decay function. Their findings indicate that ridership models that use sophisticated methods to estimate station catchment areas produce more accurate demand models.

Regardless, previous studies have suggested that station access distance varies depending on a number of factors. A study of origin-destination surveys in Montreal identified a mean walking access distance of 1259 metres for commuter rail, 873 metres for subway, and 484 metres to 897 metres for bus, depending on the type of service provided (El-Geneidy et al., 2014). A separate analysis incorporated demographic and socioeconomic data into their estimation of station access distance, also obtained via an onboard survey (Vijayakumar et al., 2011). The authors found that gender, age, park and ride capacity, overall trip length, and service quantity / quality variables were statistically significant in determining station access distance. For example, the station access distance of males was found to be 12.5% longer compared to females, while an additional 100 park and ride spaces at the station level resulted in an increased access distance of 0.38%. Further to this, a GIS analysis of transit riders found that station access distance was largely a function of station access mode (Wang et al., 2016). Customer origin points were collected for over four thousand public transit users after a survey was administered in Beijing, China. GIS tools were applied to this dataset to determine the most likely station access distance and associated access mode used by the sample. The authors found that walkers had a mean access distance of 430 metres, compared to 9115 metres by those accessing the system via private automobile. Their results indicate that the station catchment area of high order transit systems is much larger, especially when commuter rail systems are considered since the majority of customers use private auto to access the system. Therefore, station catchment boundaries should be delineated using techniques other than Euclidian or network-based buffers to ensure that areas are digitized that reflect the actual customer base.

An analysis of station access quality throughout the GO Transit rail network echoes the concerns brought forth in the previous studies (Engel-Yan et al., 2014). The authors state that using a 400 to 800-meter arbitrary buffer is not applicable for commuter rail systems, as the majority of passengers access the system via automobile. Therefore, buffers estimated using alternative methods are needed to ensure that the majority of riders and associated external variable datasets are captured. Instead of using arbitrary buffers that are larger in size, Engel et al. (2014) outlines a methodology that uses customer origin data to formulate station catchment areas to ensure that the size and shape of the buffer is representative of customer origin locations. In their analysis, the authors gathered customer origin point data via station level surveys, and mapped customer origin points using GIS tools. After the removal of outliers (ex. customer origins located more than 10km from the station, and an additional 10% of observations located farthest from the centre mean of observations) a convex hull polygon was digitized around these points. The authors highlight notable differences between using observed customer origin locations compared to arbitrary buffers. For example, they note that for some stations located in industrial areas, the estimated catchment area did not even include the station. They also note that for most estimates, the station was not centered in the catchment area, with the boundary typically stretched in the opposite direction of the CBD. Their results reinforce that the creation of station catchment areas is more accurate when estimated using observed customer origin locations, rather than methods that generalize the station access behavior of the customer base.

3.3.2. Feeder Bus Connection Quality Indicators

As noted in Section 3.2.4, previous studies have identified that transit demand is dependent on the availability of station access options, as customers need a means of accessing the station. If station accessibility is not easily available or convenient, riders will be pushed to other modes that are more accessible, resulting in stagnant ridership growth. Various studies have suggested that station access is a key component of the GO Expansion Program. The 2016 Station Access Plan notes that approximately 62% of GO Transit riders access the station using park and ride infrastructure, but approximately 85% of parking lots are at or are near capacity (Government of Ontario, 2016). Therefore, Metrolinx theorizes that additional park and ride spaces need to be constructed to account for increased demand, or alternative station access choices should be explored to accommodate new customers.

A study of station access options in the study area found that while feeder bus connections are provided to GO Transit stations, accessibility and quality of connecting services could be improved (Engel-Yan et al., 2014). For example, the authors note that most feeder bus routes do not serve areas where customers live, and that direct service is rarely provided to GO Transit stations. Therefore, customers often chose private auto to facilitate station access, as large travel times due to service layouts results in a large amount of disutility experienced by the user. Their findings suggest that the accessibility, service quantity, and service quality of feeder bus systems should be considered in demand elasticity studies, as a combination of these aspects, rather than just service provision, are determinantal in growing regional rail ridership.

Of the sources reviewed, ridership elasticity studies have traditionally assessed the impact of feeder bus connections on ridership through the use of a dummy variable to indicate the presence of service, or by recording the number of station level connections / feeder bus lines. Therefore, this metric ignores the characteristics highlighted in the previous paragraph as being determinantal to use. A new methodology for recording feeder bus connection data that incorporates both system accessibility, service quality, and service quantity is recommended which incorporates these aspects into the indicator used to measure feeder bus connection quality.

3.4. Justification for Study

The literature revealed that a variety of factors can influence transit demand. Table 3 summarizes how factors related to station accessibility, metropolitan economy, demographics, the price and availability of alternative services, and transit system characteristics are expected to influence commuter rail demand. This table could be referenced by regional transit planners or transportation researchers when researching demand elasticities in relation to commuter rail ridership.

Interestingly, demand elasticity estimates were found to vary depending on a number of aspects. Most notably, the sign of various relationships differed depending on the transit mode examined. For example, commuter rail ridership demonstrated an opposite relationship with distance to CBD, unemployment rate, and income compared to studies that analyzed overall transit demand. The significance and magnitude of various relationships was also impacted by modal classification, as commuter rail demand was shown to be less responsive to change in fare price and vehicle ownership, compared to other modes such as bus and light rail transit. Alternatively, service quantity was found to have a greater impact on commuter rail demand compared to alternative transit modes examined.

Demand elasticities were also found to differ depending on the geographical context of the study. Specifically, demand elasticities with respect to fuel price were typically found to have twice the impact on ridership in North American cities compared to European examples. Differences in climatic conditions were also shown to have a significant impact on transit demand.

Studies that disaggregated ridership figures by time of day further illustrated that trip type can influence the size and significance of various demand elasticities. Most notably, population density was found to have a significant impact on ridership during the a.m. peak period, while employment density was found to have a greater impact on p.m. peak ridership. Station accessibility indicators were also found to have differing impacts once trip type was controlled for.

Finally, the literature suggests that transit demand responds more significantly to various factors over the long-run compared to the short-run. For example, considerable lags were noted when fare price, service quantity, employment, and fuel price demand elasticities were calculated over both short and long-run time periods. Therefore, previous demand elasticity estimates could be understated if a lack of data was available for the researcher to analyze.

The literature suggests that a general understanding of public transit and subsequent commuter rail demand can be gathered from previous studies. However, a demand elasticity study has not been conducted in the context of the Greater Golden Horseshoe that:

1. Calculated demand elasticity estimates specific to commuter rail ridership,
2. Evaluated how commuter rail demand responds relative to trip type,
3. Considered the impact that various station accessibility indicators may have on demand,
4. Utilized a longitudinal dataset to incorporate both cross-sectional and temporal information into the analysis.

Further, the Canadian Urban Transit Association recommends that when specific research questions relating to transit demand are proposed, studies that analyze a specific mode, geography, and trip type are most effective as they include and account for factors and variables specific to the regional context being examined (E. J. Miller et al., 2018). Therefore, a demand elasticity study specific to the GO Transit rail system is needed to determine the main determinants of ridership, and to see if service quantity is the most significant variable that ridership responds to. The results will indicate if the service expansions proposed within the GO Expansion Program are the most effective means of growing ridership and encouraging mode shift in the Greater Golden Horseshoe, or if additional policy directions should be considered by Metrolinx to ensure that mode share and ridership targets proposed within the GO Expansion Program are achieved.

Table 3 - *Hypothesized Relationships between Ridership and Independent Variables*

Expected Relationship With Dependent Variable		
Independent Variable	**Expected Relationship**	**Rationale**
Population Density	+	More population results in larger customer base.
Employment Density	+	Larger levels of employment generate commuter trips.
Gender - Female	-	Safety / convenience of travel more important to females compared to males, results in higher disutility for public transport.
Income	+ / -	(-) Greater value of time, availability of substitute transport modes increases. (+) More CBD bound trips generated due to increased employment responsibilities.
Unemployment Rate	+ / -	(+) More unemployment results in more captive riders. (-) More unemployment reduces number of commuter related trips.
Age	+ / -	(+) Older aged adults less likely to engage in commuter related travel, but more likely to generate discretionary related trips during off-peak periods. (-) Young professionals / students more likely to use cost effective transport modes.
Households With Children	-	Number of dependents increases cost and decreases utility of public transport, thereby making other transport options more convenient.
Vehicle Availability / Ownership	+ / -	(+) Vehicle ownership complementary access mode for majority of commuter rail users. (-) Vehicle availability results in less captive riders.
Fuel Price	+	Increase in fuel price disincentivizes alternative transport modes.
Service Quantity	+	More service increases convenience, flexibility associated with transit.
Fare Price	-	Larger fares disincentivizes transit use.
Distance to Central Business District – Near	-	Greater selection of cost competitive and more convenient transit modes.
Distance to Central Business District – Far	+	Less mode choice options, GO Transit only transit mode available for long distance travel.
Number of Parking Spaces	+	More parking spaces can accommodate more park and ride customers.
Feeder Bus Connection Quality (transit access time)	-	Longer transit access times decrease system accessibility.
Winter	-	Rain, snow, cold temperatures discourage transit use.
Spring	+	Increase in discretionary trips related to tourism.
Summer	+	Increase in discretionary trips related to tourism.
Fall	-	Rain, snow, cold temperatures discourage transit use.

3.5. Chapter Summary

The first section of this literature review summarized methods frequently used by transit researchers to understand demand. Most commonly, econometric analysis is undertaken which explains how ridership is influenced by change in a variety of internal and external variables. The literature identified that this analysis is most commonly completed using an ordinary least squares estimator, although more advanced methods are used if both cross-sectional and temporal information is available to the researcher.

The next section reviewed academic literature to identify variables that have been shown to influence transit demand. Variables such as service quantity, fare price, population and employment density continuously demonstrate a consistent relationship with transit demand, while findings in relation to other demographic, socioeconomic, and station accessibility indicators are less certain. Notably, the influence, significance, and importance of these relationships were shown to vary depending on mode examined, regional context of the study, trip type, and data availability. The results suggest that a demand elasticity study specific to the study area is needed in order to accurately answer the research questions proposed in Section 1.5. A complete list of all studies summarized in this section of the literature review can be found in Appendix B. This resource could be used by transit researchers interested in identifying demand elasticity research that has previously been conducted.

The literature review further summarized data collection methods commonly used in transit demand studies. Specifically, methods used to delineate station catchment variables were summarized, as the accuracy of external variable datasets and corresponding model performance is dependent on the method selected by the researcher. Commonly, station catchment boundaries are digitized using Euclidian or network-based buffers, but inaccurate datasets could be captured if the station access distance of the customer base is extensive. Since this is the case for commuter rail systems, alternative methods such as the use of customer origin data to delineate station catchment boundaries is recommended. Methods used to represent station access indicators, including feeder bus connections, were also summarized. Notably, the literature states that this variable is commonly represented using a dummy variable, and other aspects that could be determinantal to use and subsequent ridership are understated. The use of an indicator that incorporates the quality and accessibility of feeder bus routes is therefore recommended.

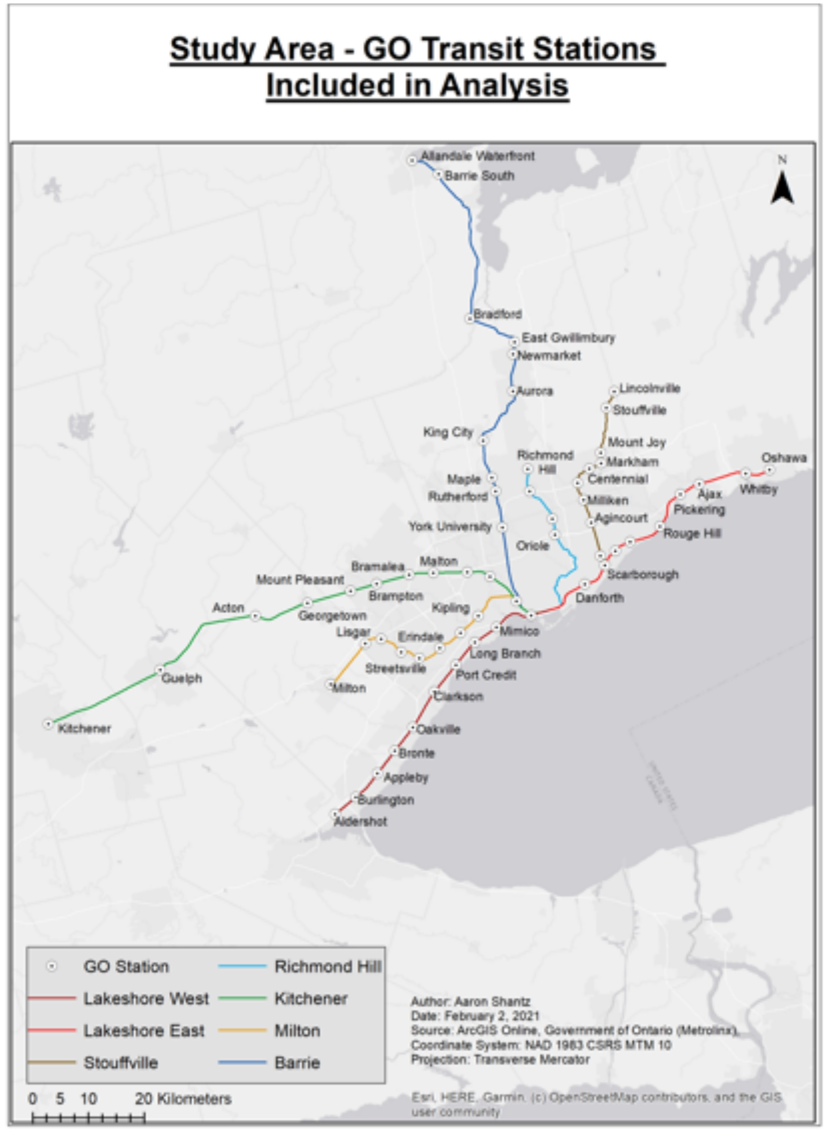
Various gaps in the literature were then identified. Notably, it was found that a demand elasticity study is needed in relation to GO Transit rail ridership to understand what variables are most determinantal to commuter rail demand. The following chapter outlines the research approach adopted to answer these questions.

4. Methodology

4.1. Station Selection

Data was collected at the station-level from January 2016 to December 2019. All stations along the Lakeshore West, Lakeshore East, Milton, Kitchener, Barrie, Richmond Hill, and Stouffville corridors were initially considered for inclusion in the study. Stations along the Niagara Falls corridor were excluded as regular weekday service was not offered during this time period. To ensure that the effects of network expansion were not captured in the demand elasticity models, Gormley GO Station and Downsview Park GO Station were excluded as they only became operational in December 2016 and January 2018 respectively. Union Station was also excluded from the analysis as its inclusion would have produced skewed model outputs, as ridership is concentrated at this station in all time periods except for the a.m. peak period. Hamilton GO Centre and West Harbour GO Station were also eliminated from the analysis, as minimal service was provided to these stations in contrast to two-way, all-day service that was provided to all remaining stations along the Lakeshore West corridor for the duration of the time-series. As a result, Aldershot GO Station was selected as the terminus station for this corridor, as the number of trips originating and terminating

Figure 3 - *Study Area - GO Transit Rail Stations Included in Analysis*



at this location was more characteristic of a termini station compared to Hamilton GO Centre or West Harbour GO Station. After these adjustments were made, a total of 61 stations were included in the analysis as shown in Figure 3.

Danforth GO Station is the only station in the network other than Union Station that services multiple corridors, as vehicles travelling on both the Lakeshore East and Stouffville corridors pass through Danforth GO Station. Ridership, fare price, and level of service data at this station is therefore aggregated to reflect both customer bases.

4.2. Time Parameters and Trip Types Analyzed

Data was recorded at monthly intervals, resulting in a total of 48 observation periods. Data collection was further disaggregated by time of boarding so that trip type could be controlled for. Therefore, separate datasets and models were created for the a.m. peak, midday off-peak, p.m. peak, and evening off-peak time periods. The time parameters as indicated in Table 4 were used to delineate these datasets. An observation was only included in the analysis if outbound service was offered at the given station at a given observation period. All stations in the study area offer outbound service during the a.m. peak period; therefore, all 61 stations were included in the a.m. peak period analysis, resulting in a balanced panel. Fewer stations were included in the midday off-peak and evening off-peak models, as two-way, all-day service is currently only provided to stations located along the Lakeshore West and Lakeshore East corridors. A reduced number of stations were also included in the p.m. peak analysis, as termini stations were excluded if outbound service was not offered during this time period. Since the dependent variable used to record ridership was number of boardings, ridership figures could not be obtained if a departing in-service trip was not avaliable. For this reason, the majority of termini stations outside of the a.m. peak period saw reduced levels of service if two-way service was not provided.

4.3. Modelling Framework

4.3.1. Panel Data Analysis

As outlined in Section 3.1.6, panel data estimators build on the simple OLS model as additional terms which account for serial correlation and unobserved factors are introduced into the model. This methodology was selected as the dataset being analyzed has both a temporal and cross-sectional component, as observations at the station-level are being analyzed over a 48-month time-series.

Table 4 - Time Parameters Used to Define Trip Types

Trip Type Time Parameters		
Trip Type	**Start Time**	**End Time**
A.M. Peak	5:00	9:30
Midday Off-Peak	9:31	14:59
P.M. Peak	15:00	19:00
Evening Off-Peak	19:01	26:00

Therefore, the panel data methodology results in the estimation of more reliable and efficient demand elasticities compared to those computed using a simple OLS methodology.

4.3.2. Model Assumptions

When linear regression analysis is conducted utilizing the OLS framework, several conditions should be met to ensure that model outputs are efficient, unbiased, and accurate (Wooldridge, 2012). Otherwise known as the Gauss-Markov assumptions, these include:

- Linearity,
- Random Sampling,
- Non-Collinearity,
- Exogeneity,
- Homoscedasticity.

If these assumptions are not satisfied, inaccurate model outputs could be produced. For example, when multi-collinearity amongst two independent variables is present, this indicates that both factors are correlated with each other in addition to the dependent variable. In other words, this introduces redundancy into the model, as the same relationship is being captured by two independent variable datasets. When this occurs, the significance of the true relationship can be understated, as some the variance in the dependent variable is captured by a redundant supply of information. This could lead the researcher to conclude that both factors are insignificant in explaining the dependent variable, thus warranting exclusion from the model, while the relationship could in fact be significant if represented by a single source of information. In addition to multi-collinearity, Table 5 further outlines issues that could be encountered by the researcher if the Gauss-Markov assumptions are violated. The use of preliminary analysis and various test statistics are often used by researchers to ensure that the Gauss-Markov assumptions are met before the regression analysis procedure is completed.

Table 5 - *Outline of Gauss-Markov Assumptions for Linear Regression Models, from Wooldridge (2012)*

Assumption	Algebraic Expression	Description	Issues If Violated
Linearity	$\gamma = \beta_o + \beta_1 x + u$	Assumes the dependent variable is a linear function of the independent variable(s).	Coefficients could be biased and inefficient.
Random Sampling	$\{(x_i, y_i): i = 1, 2, \ldots, n\}$	Assumes that the sample is obtained from a random subset of the population.	Results cannot be interpreted / applied to general population.

Non-Collinearity	$\{x_i, i = 1, \ldots, n\}$	Assumes that independent variables are not correlated with one another.	Coefficients could be overestimated and display opposite sign. Model will be less efficient, but unbiased.
Exogeneity	$E(u \mid x) = 0$	Assumes that the error has an expected value of zero given any value of the independent variable (s). Therefore, the error term cannot be correlated with the independent variable(s).	Coefficients could be overestimated. Model could be biased if omitted variables, measurement errors, etc. introduce error into the model.
Homoskedasticity	$Var(u \mid x) = \sigma^2$	Assumes that the variance of the error term is the same given any value of the independent variable (s).	Coefficients could be overestimated. Model will be less efficient, but unbiased.

Wooldridge (2012) further states that when panel datasets are analyzed, the structure of the data results in the inherent violation of various Gauss-Markov assumptions. For example, the assumption of random sampling is fundamentally violated, as data points are continuously obtained from the same entities over a defined time-series. Further, this leads to the violation of homoskedasticity, as values obtained in an observation period naturally depend on conditions observed in a prior observation period. Therefore, serial autocorrelation and spatial dependencies often arise.

When analyzing panel datasets, it is common practice to first ensure that multi-collinearity is not present to ensure that coefficients and relationships are accurate. This is commonly accomplished using cross-correlation tables and variance inflation factor scores, which highlight independent variables are highly related with each other. If a high level of correlation is identified between two factors, one is removed from the model to ensure that the relationship is only captured by one source of information.

Panel data estimators are then used to estimate model outputs, which build on the simple OLS framework but include additional terms that can account for exogeneity of the error terms and spatial / temporal dependencies. These estimators are outlined in the following sections.

4.3.3. Pooled Ordinary Least Squares

Pooled OLS is the most simplistic approach used when modeling panel data, as a simple OLS regression is run on all observations included in the dataset. This approach is commonly selected when an unbalanced panel is being analyzed, or when sample entities vary greatly throughout the time-series (Wooldridge, 2012). Notably, assumptions regarding individual heterogeneity are ignored using this approach, meaning that spatial and temporal effects which could impact the true nature of the relationship between the independent variable(s) and dependent variables are not accounted for. As a result, many pooled OLS outputs appear to be good, as statistically significant coefficients, expected relationships, and large R-squared values are computed (Gill-Carcia, JR. Puron-Cid, 2014). However, since spatial and temporal relationships are ignored, auto-correlation within the data results in the overestimation of model parameters (Gill-Carcia, JR. Puron-Cid, 2014). Therefore, pooled OLS is rarely used in final model outputs, and is instead used as a baseline to introduce more advanced modeling frameworks.

4.3.4. Fixed Effect

A fixed effect model is used when the researcher is interested in analyzing the impact of independent variables that vary over time. The model assumes that each entity has its own individual characteristics that could influence the dependent variable, and therefore need to be controlled (Torres-Reyna, 2007; Wooldridge, 2012). Time-invariant characteristics and the associated effect of these variables are removed, but are captured in the unknown intercept term for each entity to control for these effects. As a result, the model assumes that the entity's error term and its independent variables are correlated. Due to this process, a fixed effect estimator does not consider cross-sectional relationships, and is only concerned with analyzing change in the dependent variable attributed to temporal change in the independent variables. The model takes the following form:

$$\gamma_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \dots \beta_k x_{itk} + a_i + u_{it} \qquad \textit{Eq. 7}$$

- γ_{it} = a given dependent variable (where i = entity and t = time),
- x_{itk} = a given independent variable,
- β_k = the coefficient for a given independent variable,
- a_i = the unknown intercept, or unobserved fixed effects, for each entity,
- u_{it} = the idiosyncratic error term.

4.3.5. Random Effect

A random effect model can be used instead of a fixed effect model if the idiosyncratic error terms are assumed to be uncorrelated with the associated independent variables (Wooldridge, 2012). Since this assumption holds true, time-invariant variables can be included in this estimator (Torres-Reyna, 2007). However, unlike a fixed effect estimator that captures entity-specific characteristics in the error term, this assumption does not hold true for a random effect estimator. Therefore, entity specific characteristics that could influence the dependent variable (such as geographical location, sampling time, etc.) need to be controlled for to ensure that the effects being analyzed are truly random. If this is not completed, outputs generated using this approach may produce unrealistic results. The random effect model takes the following form:

$$\gamma_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots \beta_k x_{itk} + a + u_{it} + \varepsilon_{it} \qquad \textit{Eq. 8}$$

- γ_{it} = a given dependent variable where (i = entity and t = time),
- x_{itk} = a given independent variable,
- β_k = the coefficient for a given independent variable,
- a = the intercept of the model,
- u_{it} = the combined time-series and cross-sectional error,
- ε_{it} = the individual specific cross-sectional error.

4.3.6. Selection of Appropriate Estimator

When selecting the appropriate estimator, it is common practice for the researcher to analyze the panel dataset using both pooled OLS, fixed effect, and random effect models. A variety of statistical tests are then applied to the model outputs to determine the method that best suits the dataset (Guerra & Cervero, 2011; Lee & Lee, 2013; R. Liu, 2018; Stover & Christine Bae, 2011).

Of these, a Breusch-Pagen Lagrange Multiplier is commonly used to compare the pooled OLS estimator to both the fixed and random effect estimators. The null hypothesis is that there are no panel effects; if the test produces a significant result, this indicates that time and/or entity effects are present. Therefore, the use of a fixed or random effect estimator will produce more reliable and unbiased results and should be further explored by the researcher.

A Hausman test can also be used to compare the efficiency of the fixed effect estimator to the random effect estimator. Per Wooldridge (2012), a fixed effect estimator should be prioritized when the entity's error term is correlated with the same entity's independent variables, as the model allows for this correlation to occur. A significant Hausman test statistic indicates that this relationship is present, thereby indicating that the fixed effect estimator should compute results that are more reliable and unbiased compared to the random effect estimator.

4.3.7. Controlling for Heteroskedasticity

It is also common practice for the researcher to further diagnose the model for the presence of spatial and/or temporal dependencies, as the presence of such relationships would result in inefficient model outputs. As mentioned in Section 4.3.2, it is common that spherical errors such as heteroskedasticity and autocorrelation are present within panel datasets due to the sampling nature. Various test statistics can be applied to detect these effects, including the Durbin-Watson test for serial correlation in panel data models and the Breusch-Pagen test for heteroskedasticity. If detected, non-constant variance estimates can be used to compute regression outputs that are robust to heteroskedasticity and autoregressive effects. Regression outputs can be estimated using a variety of non-constant variance estimators, however White robust standard errors are commonly used in econometric analysis and transit demand elasticity studies that have utilized a panel data approach (Lee & Lee, 2013; Torres-Reyna, 2007, 2010).

5. Methods

5.1. Data Collection

5.1.1. Dependent Variable

As outlined in Section 2.2.3, specific and accurate ridership counts are recorded by Metrolinx via the PRESTO smartcard system. Therefore, the number of boardings as indicated by the PRESO system was used to formulate the dependent variable dataset for this study. Filters were applied so that monthly boarding counts could be obtained at the station-level, separated by trip type. Only weekday ridership counts were included in the analysis. Station-level monthly ridership is therefore expressed by the following equation:

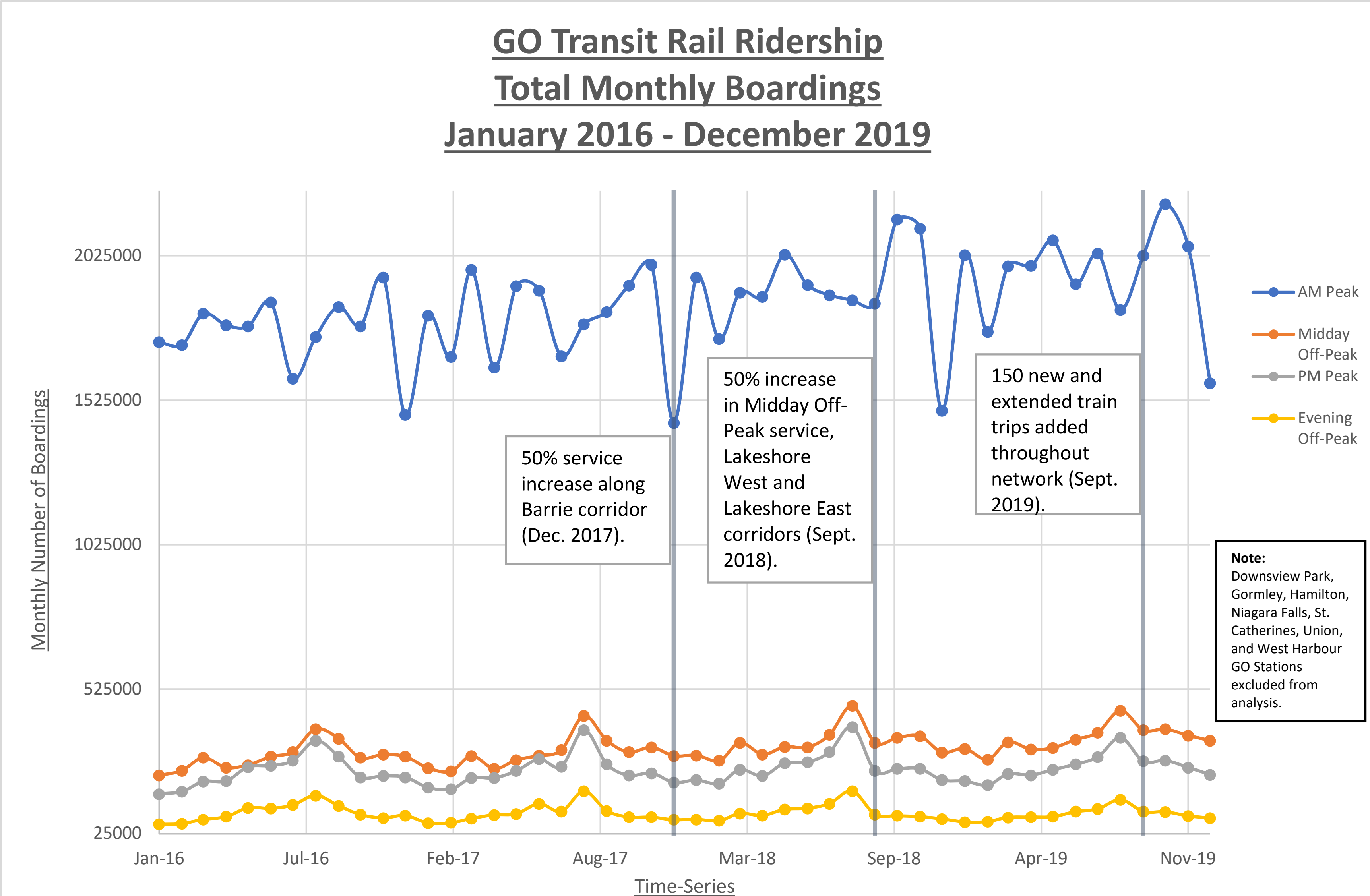
$$Monthly\ Ridership_{kl} = \sum_{i=1}^{n} Boardings_{i_{kl}} \quad \forall\ stations, l\ and\ months, k \qquad Eq.\ 9$$

Where:

- k = a given month,
- l = a given station,
- i = a given weekday,
- $Boardings$ = number of boardings as recorded by PRESTO,
- $\forall$ = for all.

Figure 4 illustrates how GO Transit rail ridership at the network level changes throughout the time-series. The graph illustrates that ridership during the a.m. peak period is drastically larger than other time periods examined, although this was expected as the majority of trips during the p.m. peak originate at Union Station and were thus excluded from the analysis. Notably, ridership during the midday off-peak, p.m. peak, and evening off-peak periods is shown to demonstrate seasonal dependencies, whereas the

Figure 4 - *Distribution of GO Transit Rail Ridership, Jan. 2016 - Dec. 2019*



number of boardings during the a.m. peak period fluctuates more drastically on a monthly basis. Drastic service increases, such as those that occurred in December 2017, September 2018, and September 2019 are highlighted. Notably, the number of boardings was seen to increase in the month(s) following the service expansion.

It was further theorized that variation in the number of business days per month could influence monthly ridership counts. To account for these differences, ridership figures were normalized by the number of business days in a given month. The dependent variable therefore took the form of average daily boardings per month:

$$Average\ Daily\ Ridership_{kl} = \frac{Monthly\ Ridership_{kl}}{N_k} \forall\ stations, l\ and\ months, k \qquad Eq.\ 10$$

Where:

- k = a given month,
- l = a given station,
- N = the total number of business days,
- ∀ = for all.

5.1.2. Delineation of Station Catchment Areas

As summarized in Section 3.3.1, Euclidian and network-based buffers are commonly used when delineating station catchment areas. This approach was initially considered, as it is an efficient delineation method when estimated using GIS tools. However, a Metrolinx study found that approximately 81.5% of GO Transit rail users access the station via private automobile in some capacity (Government of Ontario, 2016). Additionally, findings generated by Engel et al. (2014) indicate that station catchment areas previously estimated around GO Transit stations are not uniform in size and shape, most likely attributed to the access mode share and modal classification of the system (Grimsrud & El-Geneidy, 2013; Vijayakumar et al., 2011; Wang et al., 2016). Therefore, to ensure that more accurate station catchment areas could be created, customer origin data was used to delineate station catchment areas throughout the study area.

A methodology outlined by Engel et al. (2014) used customer origin data obtained from a survey of GO Transit rail passengers to digitize station catchment boundaries. For this study, customer origin data as indicated by PRESTO records were selected and downloaded for analysis. The use of PRESTO data was prioritized for a number of reasons. As outlined in Section 2.2.3, PRESTO references customer origin data to the postal code of the user's residence. Additional information such as the number of boardings associated with each customer origin location is also recorded. Therefore, very specific and accurate station catchment boundaries weighted by the intensity of demand could be estimated via this data source. Secondly, PRESTO customer origin data was available for the duration of the study period, while survey data only represents station access behavior for a specific point in time. Therefore, PRESTO data allowed for a more realistic understanding of station catchment behavior for the duration of the time series, thus resulting in the delineation of accurate station catchment boundaries. Finally, approximately 90% of customers use the PRESTO fare payment system, meaning that catchment areas mapped using this data source are an accurate generalization of station access behavior for the entirety of the customer base.

The first step when delineating station catchment areas is to define the urban environment, various rights of way, and station locations throughout the network (Andersen & Landex, 2008). A baselayer containing data on roadways and topological features within the study area was loaded into ArcMap via ArcGIS Online, while station location data was downloaded from the Metrolinx Open Data Inventory.

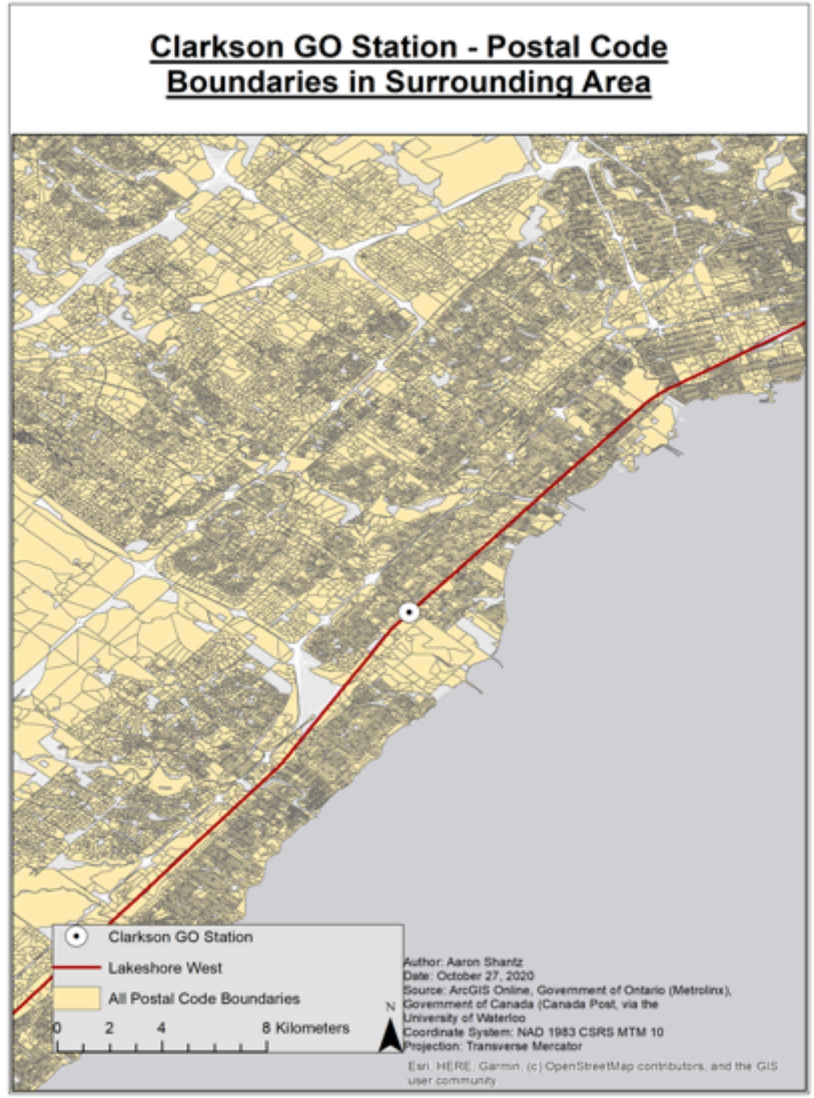
Customer origin data was then obtained from the PRESTO server. A temporal filter was applied to remove observations not related to the study period. Each observation specified the postal code, number of boardings, and access station of the customer. Observations were then disaggregated by access station of the customer to ensure that unique station catchment areas could be created for each station within the study area. Observations were then uploaded to ArcGIS so that station catchment areas could be digitized. Current GIS software does not allow for postal code data to be georeferenced if corresponding spatial information, such as latitude and longitude coordinates, are not specified. Instead, observations need to be linked to a baselayer where postal code locations have already been georeferenced. As shown in Figure 5, a shapefile containing all possible postal code addresses and associated locations in the study area was downloaded from the University of Waterloo's Geospatial Centre so that customer origin data could be georeferenced to the appropriate location. In some cases, it was found that a single postal code address was represented by multiple polygons. This occurred if postal code addresses were separated by a natural or urban feature, or if one postal code address was used to represent multiple parcels of land within the study area. Polygons sharing the same postal code address were dissolved into a single polygon to avoid double counting customer origin data.

Customer origin data was then georeferenced to the postal code boundary shapefile. Initial outputs revealed that customer origin locations were widely dispersed throughout southern Ontario. Following the methodology outlined in Engel et al. (2014), all customer origin records located further than 10km of the access station were eliminated. This was done to ensure that only home-based trips were included in the analysis. A heatmap as shown in Figure 6 was then created using the remaining customer origin observations. The heatmap was weighted according to the number of boardings to ensure that areas with a larger concentration of riders were reflected more heavily in the analysis. A polygon(s) was then digitized around the computed heatmap, which was then saved and exported as the corresponding station catchment area. The output generated for Clarkson GO Station as shown in Figure 6 reveals that some station catchment areas were not continuous. This can be explained by the presence of natural or built features, such as the Greenbelt or freeways, that cause disconnects in the urban environment surrounding GO Transit rail stations. Additionally, some GO Transit rail stations are located in employment and industrial areas where residential locations are not permitted, resulting in an absence of customer origin data and a subsequent disconnect in the station catchment boundary. This process was completed for each station in the study area.

5.1.3. Extrapolation of Census Data

Few countries collect household demographic and socioeconomic statistics more than twice per decade. Previous ridership elasticity studies have used linear extrapolation to estimate monthly values from annual or quinquennial census-based data sources (Chiang et al., 2011; Lee & Lee, 2013). Linear extrapolation provides a more finite amount of information to the research, but also ensures that change in demographic and socioeconomic variables is captured and accounted for in the modelling process. Observations from the

Figure 5 - *Delineation of Station Catchment Area at Clarkson GO Station, Postal Code Boundary Shapefile*



2016, 2011, and 2006 Canadian Census of Population were used to complete the projections, as a minimum of three census periods are required to complete accurate census data projections (Lewis, 2018). In some cases, values were also obtained from the 2011 National Household Survey as different sampling techniques and variable classifications were used during the 2011 census period. Data was collected at the dissemination area scale as this is the most disaggregated census dataset available to the public, which was downloaded from the Computing in the Humanities and Social Sciences (CHASS) Data Server via the University of Toronto. Characteristics representing population, gender, income, unemployment rate, age, and number of households with children were obtained. A descriptive list of the characteristics downloaded to represent these factors is provided in Appendix C.

Data was initially downloaded and processed for all dissemination areas throughout Ontario. Redundant observations were then removed from the dataset to decrease file size and processing times. Only dissemination areas located with the Greater Golden Horseshoe were considered for further analysis.

Figure 6 - *Final Output, Delineation of Station Catchment Area at Clarkson GO Station*

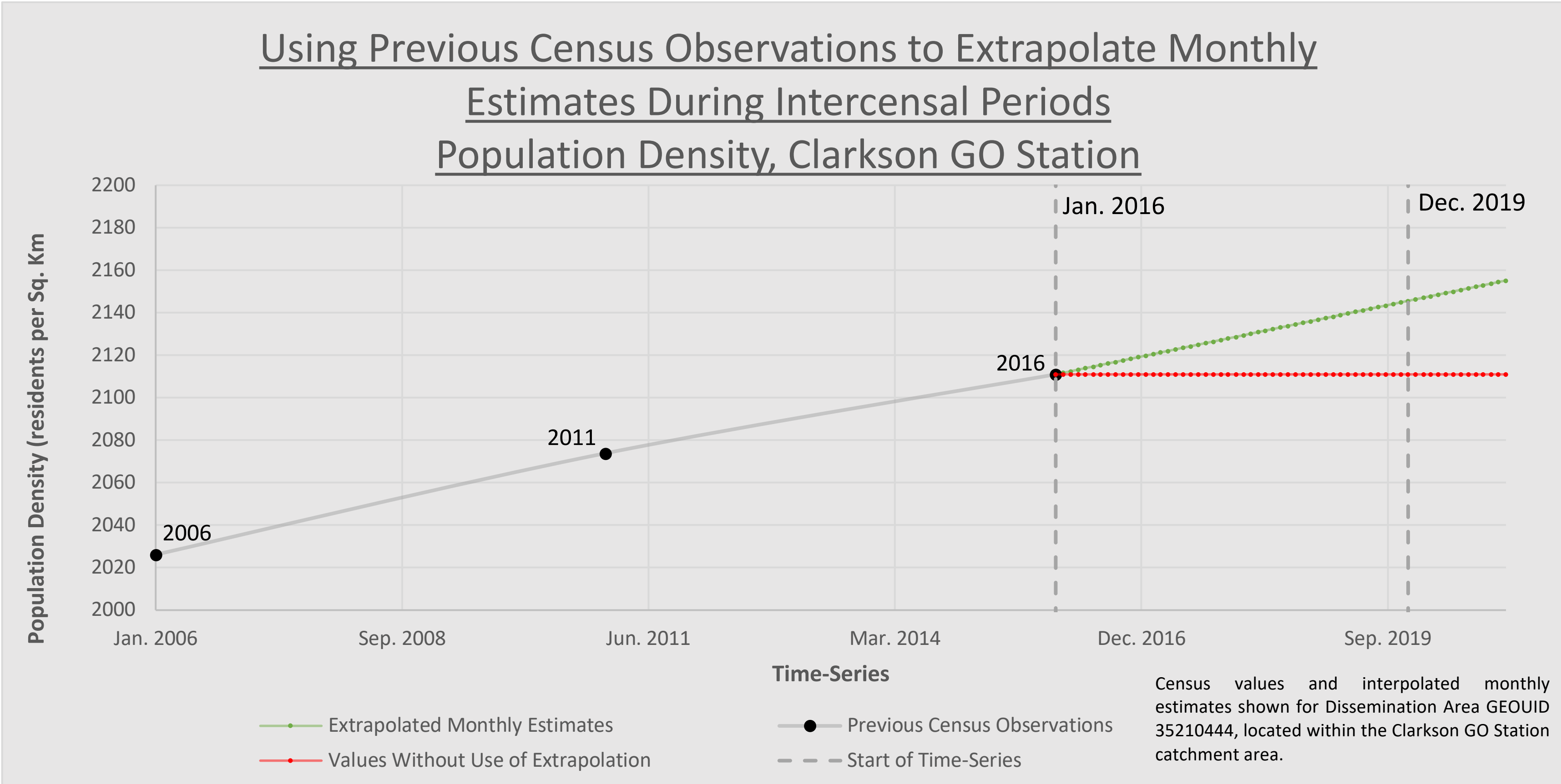
Several datasets were cleaned and adjusted to ensure that all candidate variables closely represented those identified in the literature as being determinantal to transit demand. For example, Section 3.2.12 highlights that households with children can be determinantal to transit demand, but the Canadian Census of Population provides separate counts based on the material status of the household. These datasets were therefore aggregated to provide a single count tabulating all households with children in each census period.

Varying measures of central tendency were also used to track the age of the population throughout the study period. Median age was reported at the dissemination area scale in the 2011 census, whereas average age was reported at this scale in the 2016 census. However, both average and median age were reported at the provincial scale in the 2016 census. These two values were compared, and a difference of 0.73% was estimated between the two measures. As described in Appendix D all average age values identified in the 2016 census at the dissemination area scale were adjusted by this value to reduce any inaccuracies in the dataset. Additionally, a central measurement of age was not reported in the 2006 census altogether. Values provided from the 2011 census were carried over to 2006 to account for this void.

Observations were then merged in Microsoft Excel so that values corresponding with each dissemination area could be visualized in the same spreadsheet. The merge was completed using the Geographic Unique Identifier (GEOUID) of the dissemination area. Observations were then extrapolated in Microsoft Excel using the "Trend" function. This process was completed for each Dissemination Area within the study area. Figure 7 illustrates how the process was conducted, and further demonstrates how extrapolated figures accurately reflect changing socioeconomic and demographic conditions compared to the use of time-invariant values. Once the projection was completed, each dissemination area contained a projected monthly estimate for each of the factors obtained from the CHASS server.

After completion, it was found that a small proportion of dissemination areas within the Greater Golden Horseshoe did not contain observations for all three census periods, or that null/zero values were recorded for certain characteristics. As a result, extrapolated values could not be estimated for these entities. While dissemination area boundaries remain relatively stable over time, their delineation is correlated with population size, which is targeted from 400 to 700 persons (Canada, 2018). Therefore, population increases likely resulted in various dissemination areas being resized, thus resulting in a larger count of dissemination areas in 2016 compared to 2011 and 2006. To ensure data consistency, values obtained from 2016 census products were used if missing or null values were found in previous census periods. This correction was required for approximately 20% of dissemination areas within the study area.

Figure 7 - *Extrapolated Socioeconomic and Demographic Values*



5.1.4. Extracting Census Data Using Overlay Analysis

Station-level external variable datasets were then identified using overlay analysis in ArcGIS. A spreadsheet containing the projected external variable data was first uploaded to ArcGIS. A shapefile representing the size and extent of all dissemination areas in Ontario, consistent with the 2016 Census, was also uploaded to ArcGIS via Statistics Canada.

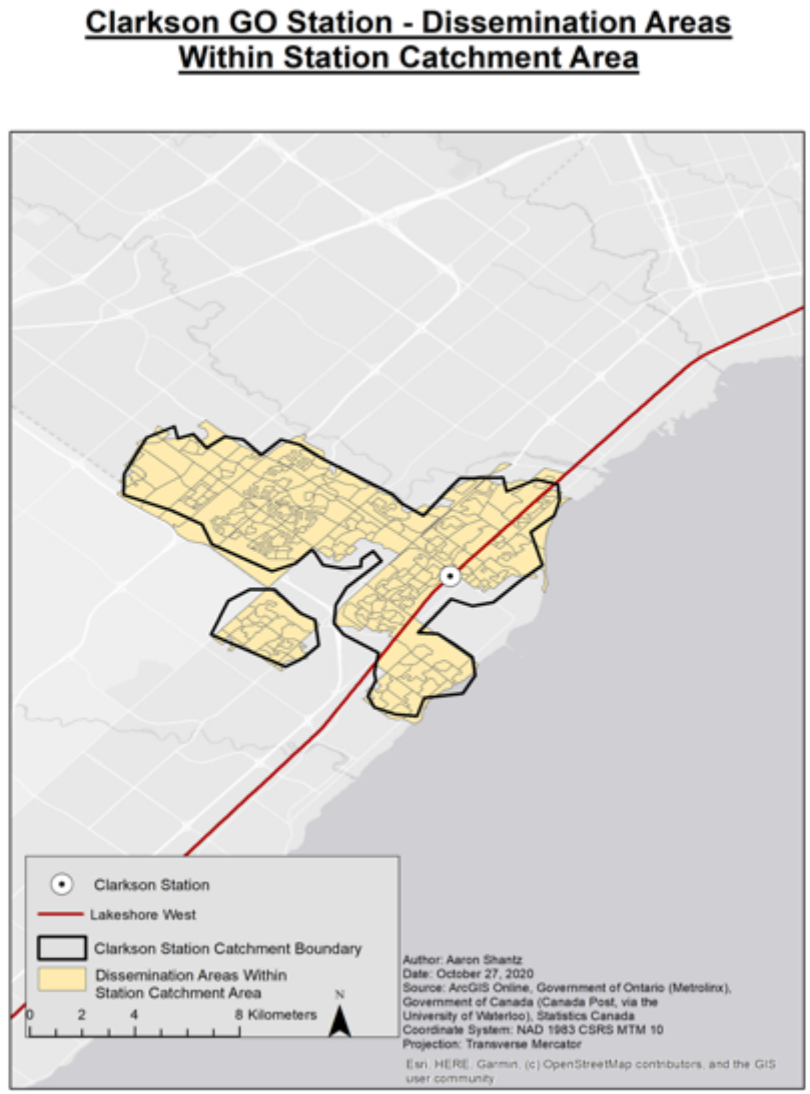
Figure 8 shows how station catchment boundaries previously delineated in Section 5.1.2 were used to identify dissemination areas and corresponding external variable datasets correlated with customer origin location. Overlay analysis was used to ensure that external variable datasets obtained at the station-level were representative of proven customer origin locations and the intensity of boardings. As shown in Figure 9, only dissemination areas whose centroid was located within the estimated station catchment area boundary were selected. Using the "Calculate Geometry" tool, the total area of all dissemination areas selected was also calculated. A spreadsheet containing the projected datasets within the estimated station catchment boundary and the area of the associated dissemination areas was extracted using the "Export" tool for further processing in Microsoft Excel.

Figure 8 - *Overlay Analysis to Extract Socioeconomic and Demographic Datasets*

5.1.5. Final Form of Demographic and Socioeconomic Variables

Characteristics were then adjusted to ensure that the final form of each external factor was consistent with those identified in the literature. Factors representing population density, density of households with children, percent of population female, median household income, median age, and unemployment rate were formulated. Since data was obtained from multiple dissemination areas within each station catchment area, figures were aggregated to provide a single monthly value for each entity included in the analysis. Some factors were normalized by the size of the station catchment area so that density measures could be computed. Other variables that were already expressed using a measure of central tendency, such as median age, were normalized by the number of dissemination areas within the identified station catchment area to provide a single averaged value. The expression of each external variable estimated using census products is outlined below:

Figure 9 - *Dissemination Areas Selected to Formulate Socioeconomic and Demographic Datasets*



$$Population\ Density_{kl} = \frac{\sum_{i=1}^{n} Population_{i_{kl}}}{\sum_{i=1}^{n} Area_{i_l}} \forall\ stations, l\ and\ months, k \quad \text{Eq. 11}$$

$$Density\ of\ Households\ With\ Children_{kl} = \frac{\sum_{i=1}^{n} Households\ With\ Children_{i_{kl}}}{\sum_{i=1}^{n} Area_{i_l}} \forall\ stations, l\ and\ months, k \quad \text{Eq. 12}$$

$$Percent\ Population\ Female_{kl} = \left(\frac{\sum_{i=1}^{n} Female\ Population_{i_{kl}}}{\sum_{i=1}^{n} Population_{i_{kl}}}\right) x\ 100 \forall\ stations, l\ and\ months, k \quad \text{Eq. 13}$$

$$Median\ Household\ Income_{kl} = \frac{\sum_{i=1}^{n} Median\ Household\ Income_{i_{kl}}}{N_l} \forall\ stations, l\ and\ months, k \quad \text{Eq. 14}$$

$$Median\ Age_{kl} = \frac{\sum_{i=1}^{n} Median\ Age_{i_{kl}}}{N_l} \forall\ stations, l\ and\ months, k \quad \text{Eq. 15}$$

$$Unemployment\ Rate_{kl} = \frac{\sum_{i=1}^{n} Unemployment\ Rate_{i_{kl}}}{N_l} \forall\ stations, l\ and\ months, k \quad \text{Eq. 16}$$

Where:

- k = a given month,
- l = a given station's catchment boundary,
- i = a given dissemination area within a station's catchment boundary,
- $Area$ = area of the dissemination area in Sq. Km.,
- N = the number of dissemination areas within a station's catchment boundary,
- $\forall$ = for all.

5.1.6. Employment Density

Data obtained from the Census of Population only records employment statistics at the household level. Therefore, any employment data obtained from this source is a function of the population living in the catchment area, and does not convey information about regional economic output. However, the 2016 Census of Population asks respondents to indicate their census tract of residence and their census tract of workplace. Therefore, count data indicating census tract of workplace is an approximate estimate of the level of employment within a given census tract.

Data was downloaded from the 2016 Census into Microsoft Excel. Data was originally given as origin-destination pairs, separated by census tract of origin and census tract of destination. Pivot tables were used to sum the total number of incoming commuters in each given census tract. Notably, only a single observation for each census tract was provided, and archived data from previous Census' could not be obtained. Therefore, projected values could not be computed.

This dataset was uploaded into ArcMap so that data from census tracts located within the previously identified station catchment areas could be downloaded for further analysis. Count data for each census tract was uploaded and linked with a shapefile illustrating all census tract boundaries in southern Ontario. Data was extracted using the same process as outlined in the Section 5.1.4. A separate dataset was download and exported for each station in the study area. Each dataset was processed so that a value representing employment density could be estimated. This was done by summing the number of incoming commuters in a given month in a given station area, and normalizing it by the size of all census tracts within the station catchment boundary. Employment density thereby took the following form:

$$Employment\ Density_l = \frac{\sum_{i=1}^{n} Incomming\ Commuters_{i_l}}{\sum_{i=1}^{n} Area_{i_l}} \forall\ stations, l \quad \text{Eq. 17}$$

Where:

- l = a given station's catchment boundary,
- i = a given census tract within the station's catchment boundary,
- $Area$ = area of the census tract in Sq. Km,
- ∀ = for all instances of.

5.1.7. Vehicle Ownership

Vehicle ownership statistics were downloaded from the 2016, 2011, and 2006 Transportation Tomorrow Survey, also obtained from the CHASS Data Centre at the University of Toronto. Count data tabulating the number of vehicles owned per household at continuous levels (0, 1, 2…99 vehicles owned) was provided for all upper-tier municipalities located throughout the Greater Golden Horseshoe. To obtain monthly estimates for each value, these figures were also extrapolated using the same process as outlined in Section 5.1.3. Vehicle ownership in each upper-tier municipality was then calculated by multiplying the estimated count data by the corresponding household level of car ownership:

$$VehicleOwnership_{kl} = \sum \begin{matrix} 1(NumberofHouseholdsWithOneVehicleOwned_{kl}) \ldots \\ 99(NumberofHouseholdsWithNinetyNineVehiclesOwned_{kl}) \end{matrix} \forall\ upper - tier\ municipalities, l\ and\ months, k \qquad \text{Eq. 18}$$

Where:

- k = a given month,
- l = a given upper-tier municipality,
- ∀ = for all instances of.

Since data was only provided for upper/single-tier municipalities located within the Greater Golden Horseshoe, overlay analysis could not be used to delineate station-level values. Therefore, stations were assigned the value corresponding to the municipality in which they are located. Appendix E further describes this process and outlines what values were allocated to each station in the study area.

5.1.8. Fuel Price

Fuel price data was obtained from the Ontario Fuel Price Survey via the Ontario Data Catalogue. The survey estimates the average price of one liter of unleaded fuel at monthly intervals for various regions throughout the province. Values for the Toronto West, Toronto East, and Southern Ontario regions were downloaded for the duration of the time-series.

Notably, geographic boundaries delineating these regions are not specified. Stations were assigned a value based on their geographical location relative to the City of Toronto. Toronto East and Toronto West values for a given month were averaged and were assigned to stations located within the City of Toronto. Remaining stations located along the Milton and Lakeshore West corridors were assigned the Toronto West value, while remaining stations located along the Lakeshore East corridor were assigned the Toronto East value. All other stations along the Kitchener, Barrie, Stouffville, and Richmond Hill corridors were assigned the Southern Ontario Value. This process is further specified in Appendix E.

5.1.9. Final Form of Internal Variable Datasets

A variety of internal variables including service quantity, fare price, distance to CBD, and the availability of station access alternatives were included in the analysis. The majority of variables included were not readily available from internal databases and were therefore downloaded and processed from external sources. This paragraph outlines this process and the final form that each internal variable took in the analysis.

5.1.9.1. Service Quantity

Service data was originally compiled using archived service vetting reports supplied by Metrolinx. However, it was discovered that service quantity was being recorded at the corridor level, using a metric of number of trips per corridor. Unfortunately, this metric ignores the impact of express and local trips on service distribution, as all stations along a corridor were considered to receive the same level of service regardless of service type. Further investigation of internal databases revealed that service data was only recorded using this method, meaning that industry standard metrics, such as vehicle revenue hours were unavailable for analysis. To resolve this problem, it was determined that number of trips per station, rather than corridor level counts, should be used to represent service supply as this metric delineates between stations that have a greater supply of service compared to others.

A Python script was used to extract trip data from archived GO Transit General Transit Feed Specification (GTFS) files obtained from "transitfeeds.com" (Ionescu, 2020). GTFS files were used as they contain archived service schedules, meaning that number of trips per station could be estimated for the duration of the time-series. Individual estimates could have been obtained for each month in the time-series, however internal Metrolinx service reports indicated that service changed only 17 times over the time-series. Therefore, trip data gathered from a single GTFS file could be assigned to multiple observation periods, as the same level of service was offered over consecutive observation periods. A list of GTFS feeds used in the analysis is outlined in Appendix F.

Once processed, trip data at the station level including the arrival time, departure time, direction of travel, and station destination of each vehicle was obtained. Pivot tables were used to extract station-level trip counts for each trip type. Trips arriving but terminating at a given station were excluded from the count, as no boardings would result from these trips. Unlike the number of boardings dataset, final counts were not normalized by the number of business days in the observation period as data was already obtained at the daily level. Service quantity therefore took the following form:

$$Service\ Quantity_{kl} = \sum Outbound\ Trips_{i_l}\ \forall\ stations, l\ and\ months, k \qquad \textit{Eq. 19}$$

Where:

- k = a given month,
- l = a given station,
- i = a given business day,
- $Outbound\ Trips$ = a trip arriving at a given station and continuing service to other stations along the corridor,

- $\forall$ = for all.

5.1.9.2. Fare Price

Consistent fare price values could not be obtained as fare price depends on the number of zones through which a customer travels. Additionally, archived reports specifying the average cost of travelling from one zone to another could not be obtained. Therefore, fare price values were estimated using data obtained from the PRESTO system.

PRESTO estimates fare price at a given station by averaging the total fare revenue obtained by the total number of boardings. As a result, station-level fare price is a function of the number of boardings and the average distance travelled. Pivot tables were used to separate estimates by trip type, while monthly estimates were provided for the duration of the time-series. The final values used can be expressed by the following equation:

$$Fare\ Price_{kl} = \frac{\sum Fare\ Revenue_{kl}}{Monthly\ Ridership_{kl}} \forall\ stations, l\ and\ months, k \qquad \text{Eq. 20}$$

Where:

- k = a given month,
- l = a given station,
- $\forall$ = for all.

5.1.9.3. Park and Ride Capacity

Park and ride statistics were obtained from the Metrolinx Open Data Portal. Total number of parking spaces at the station level were provided for the duration of the time-series. All types of park and ride spaces, including handicap, priority, and electric vehicle spaces were included. Only spaces owned and operated by Metrolinx were included in the analysis. The variable took the following form:

$$Park\ and\ Ride\ Capacity_{kl} = \sum Parking\ Spaces_{kl} \forall\ stations, l\ and\ months, k \qquad \text{Eq. 21}$$

Where:

- k = a given month,
- l = a given station,
- $\forall$ = for all.

5.1.9.4. Distance to Central Business District

Straight-line distance from a station to the largest CBD within the operating area has shown to influence transit demand. However, preliminary analysis showed that this metric was highly correlated with fare price, as stations located further away from the CBD were associated with larger fare price due to longer distances being travelled by the customer base. Instead, dummy variables were used to indicate if a station was located within the City of Toronto municipal boundary, as this area contains the largest CBD in the study area.

Stations located within the City of Toronto were identified using geospatial analysis. A shapefile containing all station locations was loaded into the software, along with a shapefile illustrating the City of Toronto municipal boundary, obtained from the City of Toronto's Open Data Portal. A data query was then used to select stations that were within the boundary. Stations within the boundary were assigned a value of "1", while stations located in other municipalities were assigned a value of "0".

5.1.9.5. Feeder Bus Connection Quality

Feeder bus connection quality can be recorded by determining the transit travel time from areas that have a concentrated number of riders to the associated access station. Otherwise known as the transit access time, this measure allows service characteristics, such as service quality and service quantity, to influence the travel time estimate. Additionally, estimating the travel time from areas of proven customer origin allows system accessibility to be incorporated into the measure, as access times will increase if minimal or dispersed service is provided within these areas.

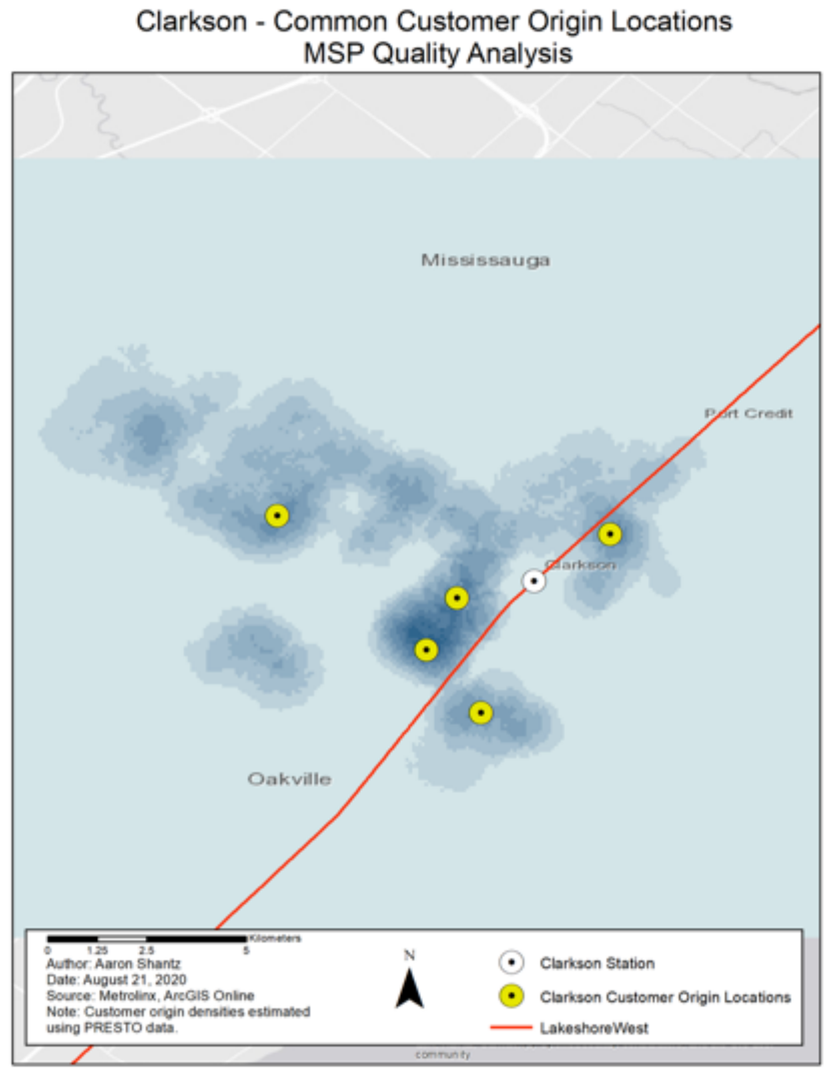
Transit access times can be calculated using the Network Analyst tool in ArcGIS Pro. The software allows the user to upload a point of customer origin, a GTFS file outlining the extent and quality of transit service in the area, a point of customer destination, and a road network within the study area. Once a departure time is specified, the software can calculate how long it will take the customer to travel from their origin to their destination, using the uploaded transit network as their main mode of transport. Notably, this analysis can be completed using archived GTFS files, meaning that unique values could be calculated for the duration of the time-series.

The first step in such a process is determining areas with concentrations of GO Transit riders. Section 5.1.2 outlines how heatmaps illustrating concentrations of GO Transit riders were previously created when station catchment boundaries were delineated. The same process was followed here, but customer origin observations within an 800-meter radius of the station were also removed. This was completed as customer origin locations at some stations were concentrated within an 800-meter radius of the station, therefore skewing the results.

Figure 10 illustrates how points were digitized over areas where customer origin was shown to be most concentrated. Five customer origin points were created for each station in the study area. Archived GTFS files of local bus networks operating within the study area were then downloaded from "transitfeeds.com". GTFS feeds could have been downloaded at monthly intervals, but the majority of local transit authorities only make service changes at the beginning of each season. Therefore, GTFS files were downloaded and processed at times coinciding with Metrolinx's seasonal board period changes as indicated in Appendix F. If a GTFS file was not available for a specified board period, data from the next closest board period was used to ensure data consistency.

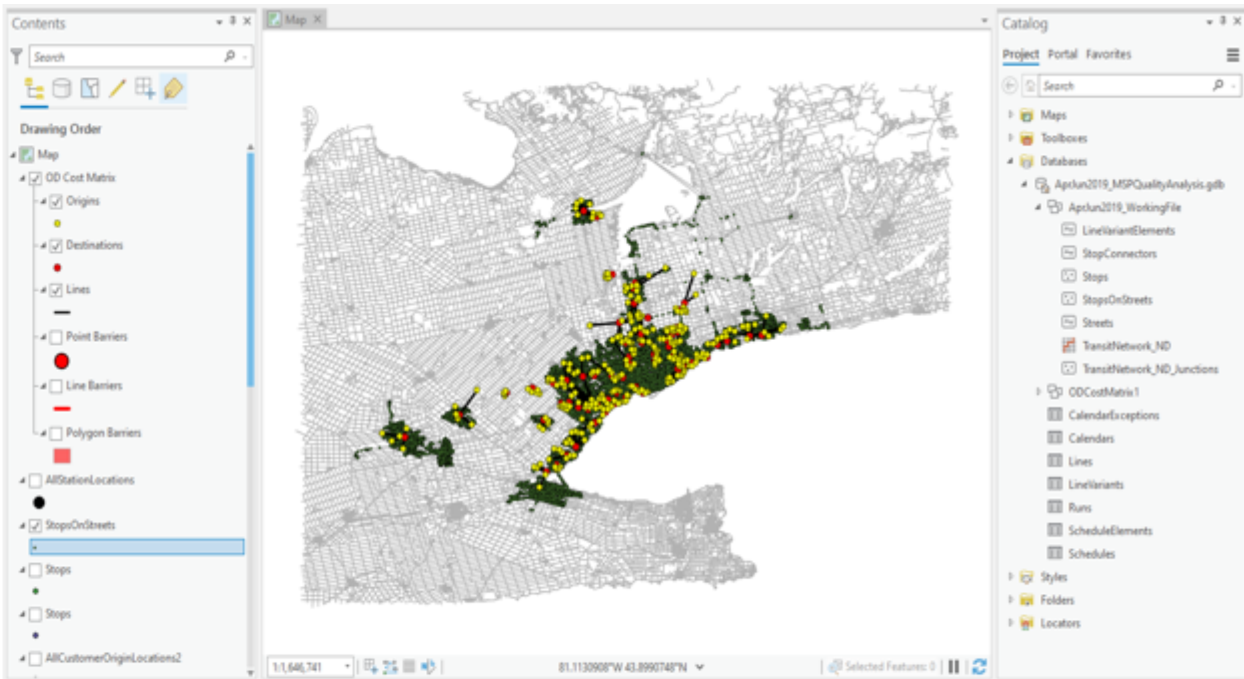
As shown in Figure 11, a virtual transit network in each board period was built so that transit stops, routes, scheduling information, and associated road network data could be incorporated into the travel time estimate. This was completed using the steps outlined in the World Bank's Introduction to the General Transit Feed Specification (GTFS) and Informal Transit System Mapping tutorial, and the ArcGIS Network Analyst Tutorial (ESRI, 2019, 2020; World Bank Group, 2020). Different theoretical departure times as specified in Table 6 were inputted to ensure that different travel time estimates could be obtained for the various trip types included in the analysis.

Figure 10 - *Common Customer Origin Points Surrounding Clarkson GO Station Used in Feeder Bus Connection Quality Analysis*



After completion, a spreadsheet specifying each origin-destination pair and the associated transit travel time between each pair was computed. Separate estimates were calculated for each observation period and time period. These files were exported to Microsoft Excel for further processing.

Figure 11 - *Virtual Transit Network of All Municipal Service Providers in the Greater Golden Horseshoe, Used to Obtain Feeder Bus Connection Quality (ex. Travel Time) Estimates*



Data was then processed in Microsoft Excel to control for outliers and account for origin-destination pairs that were incorrectly specified by the software. For example, abnormally large transit travel times were calculated for some origin-destination pairs, most likely a result of minimal or no feeder bus service being provided in these areas. In these situations, the software assumed that the customer would access the station on foot, thus resulting in extremely large travel times. To account for this, an upper limit of 60 minutes was used. Additionally, the Network Analyst tool sometimes connected a customer origin point to the wrong access station. This occurred if the software determined that another GO Transit rail station, rather than the station associated with the customer origin location, was shown to have a lesser transit access travel time. If this occurred, a transit access time of 60 minutes was also assigned, as

Table 6 - Time Parameters Used to Define Trip Types for Feeder Bus Connection Quality Analysis

Trip Type Time Parameters	
Trip Type	**Theoretical Departure Time**
A.M. Peak	7:00
Midday Off-Peak	12:30
P.M. Peak	16:00
Evening Off-Peak	19:00

the results indicate that feeder bus service in the specific area does not adequately provide service to the associated access station.

Transit access time was then estimated by calculating the average transit access time between each origin-destination pair at a given station for a specific observation period:

$$FeederBusConnectionQuality_{kl} = \frac{\sum_{i=1}^{5} Travel\ Time_{i_{kl}}}{5} \forall\ stations, l\ and\ months, k \qquad \text{Eq. 22}$$

- k = a given board period,
- l = a given station,
- i = a given origin/destination pair,
- $\forall$ = for all.

5.2. Data Analysis

As outlined in Section 5.1, monthly station-level observations for one dependent variable and 16 independent variables were obtained over a 48-month time-series from January 2016 to December 2019. As noted in the descriptive analysis of the dependent variable, it was theorized that ridership could be influenced by seasonal effects. Therefore, dummy variables indicating the season of observation were also included in the analysis. In total, one dependent and 19 independent variables were compiled in Microsoft Excel for processing. Separate spreadsheets that compiled observations by trip type were further created. A summary of all variables included in the analysis is outlined in Table 7.

Table 7 – *Final List of Variables Compiled for Analysis*

Final List of Variables Compiled For Analysis					
Variable	**Indicator**	**Type of Data**	**Data Source / Spatial Scale (if applicable)**	**Are Values Time-Variant?**	**Unique Values Obtained for Different Trip Types?**
Ridership	Average daily number of boardings	Continuous	PRESTO server	Yes	Yes
Population Density	Population density	Continuous	Census products, Dissemination Area	Yes, unique values obtained via extrapolation	No

Final List of Variables Compiled For Analysis					
Variable	**Indicator**	**Type of Data**	**Data Source / Spatial Scale (if applicable)**	**Are Values Time-Variant?**	**Unique Values Obtained for Different Trip Types?**
Employment Density	Density of incoming commuters	Continuous	Census products, Census Tract	No	No
Gender – Female	Percentage of population female	Continuous	Census products, Dissemination Area	Yes, unique values obtained via extrapolation	No
Income	Median household income	Continuous	Census products, Dissemination Area	Yes, unique values obtained via extrapolation	No
Unemployment Rate	Unemployment rate	Continuous	Census products, Dissemination Area	Yes, unique values obtained via extrapolation	No
Age	Median age	Continuous	Census products, Dissemination Area	Yes, unique values obtained via extrapolation	No
Households With Children	Density of households with children.	Continuous	Census products, Dissemination Area	Yes, unique values obtained via extrapolation	No
Vehicle Availability / Ownership	Total amount of private vehicles owned	Continuous	Transportation Tomorrow Survey, Upper/Single-Tier Municipality	Yes, unique values obtained via extrapolation	No
Fuel Price	Price of one liter of unleaded fuel	Continuous	Ontario Fuel Price Survey, Region relative to City of Toronto	Yes, unique monthly observations available for duration of time-series	No
Service Quantity	Number of outbound trips per station	Continuous	Generalized Transit Feed Specification files	Yes, unique monthly observations available for duration of time-series	Yes
Fare Price	Average fare price	Continuous	PRESTO server	Yes, unique monthly observations available for duration of time-series	Yes
Distance to Central Business District – Near	Station located within City of Toronto	Nominal	Metrolinx Open Data Portal	Yes, unique monthly observations available for duration of time-series	No

Final List of Variables Compiled For Analysis					
Variable	**Indicator**	**Type of Data**	**Data Source / Spatial Scale (if applicable)**	**Are Values Time-Variant?**	**Unique Values Obtained for Different Trip Types?**
Distance to Central Business District – Far	Station not located within City of Toronto	Nominal	Metrolinx Open Data Portal	Yes, unique monthly observations available for duration of time-series	No
Number of Parking Spaces	Total number of parking spaces	Continuous	Metrolinx Open Data Portal	Yes, unique monthly observations available for duration of time-series.	No
Feeder Bus Connection Quality	Average transit access time	Continuous	Generalized Transit Feed Specification files	Yes, unique monthly observations available for duration of time-series	Yes
Winter	Observation occurred in January, February, or March	Nominal		No	No
Spring	Observation occurred in April, May, or June	Nominal		No	No
Summer	Observation occurred in July, August, or September	Nominal		No	No
Fall	Observation occurred in October, November, or December	Nominal		No	No

Summary statistics for the above datasets were then calculated using the "summary" function in R-Studio. These results are presented in Appendix H.

5.2.1. Initial Formatting of Datasets

Data was then organized using a long panel data format, where each row represents one time point per entity. Therefore, the entity (station) and time (month and year) were specified in the first two columns, with corresponding dependent / independent variable observations specified in the following columns.

Observations were only included if service was offered at a given station during a given time period. Since service is offered at all GO Transit rail stations during a.m. peak period, this resulted in a balanced panel for this model. Other time periods, such as midday off-peak, were not balanced as service was gradually added to some stations throughout the duration of the time series.

5.2.2. Adjusting for Inflation

To account for inflation, variables including fare price, income, and fuel price were converted into January 2016 Canadian dollars. This was done to ensure that change in these values over time was a result of changing urban and socioeconomic conditions, rather than change in the value of Canadian currency. Using Canada's Consumer Price Index (monthly, not seasonally adjusted), a deflator value referenced to the January 2016 Consumer Price Index was calculated for each observation period throughout the time series (Government of Canada, 2020). Real values in each observation were then estimated as shown in Appendix I.

5.2.3. Natural Logarithm Transformations

All continuous variables were then transformed by their natural logarithm. Log-log transformations are commonly used in ridership elasticity studies for several reasons. First, the coefficients estimated in regression model outputs can be directly interpreted as ridership elasticities, because they represent the percent change in demand when an independent variable is increased by 1% (Holmgren, 2007; Li et al., 2020; Stover & Christine Bae, 2011; Taylor et al., 2009). Secondly, the use of logarithm transformations can normalize the skewness of datasets, which has been shown to increase the fit of models and thereby improve model performance (Durning & Townsend, 2015; Taylor et al., 2009). The transformation was completed using the "LM" function in Microsoft Excel.

5.2.4. Data Cleaning

A number of observations were found to contain fare price and ridership counts with a value of less than one. This could have been a result of internal processing errors, as employees could have been testing the PRESTO system or associated proof of payment infrastructure, resulting in invalid observations. Since the natural log of any number less than one is a negative number, this prevented the regression analysis from occurring as the software package used to compute the regression analysis is constrained to positive values. Observations containing a value of less than one in either category were therefore removed to prevent this error from occurring. In total, 35 observations were removed from the p.m. peak dataset, 136 observations were removed from the evening off-peak dataset, while no observations were removed from the a.m. peak and midday off-peak datasets.

5.2.5. Examination of Multi-Collinearity

Spreadsheets were then loaded into R-Studio so that independent variables could be examined for multi-collinearity. As summarized in Section 4.3.2, redundant supplies of information need to be identified and eliminated to ensure that accurate model parameters are computed. Before a regression model is estimated, the amount of correlation between independent variables can be estimated by calculating the correlation coefficient of each pair of independent variables included in the regression. Variance Inflation Factor (VIF) scores are also used to illustrate the variance, or error, in a regression coefficient that is increased due to multi-collinearity with other independent variables. However, little consensus exists as to how these scores should be applied to address multi-collinearity issues.

Wooldridge (2012) states that pairs of independent variables that demonstrate perfect correlation should be eliminated from the model, but a definitive correlation coefficient value that can be used as a cut-off to eliminate other independent variables that are problematic to the regression analysis is not definitive. While a correlation coefficient value of 0.7 is commonly used as a cut-off, a variety of thresholds ranging from 0.4 to 0.85 have been identified depending on the level of restrictiveness the author wishes to implement. (Dormann et al., 2013). Wooldridge (2013) further states that all things being equal, a model with less correlation between independent variables will perform better than one that suffers from multi-collinearity, but the researcher risks losing information if too many variables are eliminated. For this study, a fairly unrestrictive correlation coefficient cut-off of 0.85 was used to ensure minimal loss of information from the analysis.

VIF scores were also used to identify independent variables whose correlation coefficients might be overestimated due to multi-collinearity. A cut-off of 10 was used to identify such variables. The "corrplot" package in R-Studio was used to estimate correlation coefficients for all independent variables in each time period (Wei & Simko, 2017). VIF scores for each independent variable were also estimated in R-Studio using the "faraway" package (Faraway, 2016). A full summary of the results is provided in Appendix J.

As shown in Table 8, independent variables with coefficients and VIF scores above the established thresholds are highlighted in red and were removed from each model. In all time periods, households with children were removed as it demonstrated significant correlation with population density. Households with children, rather than population density, were selected for removal as population density has been shown to be more significant in influencing ridership and is more relevant to the research question.

Once removed, all correlation coefficients and VIF scores were below the predetermined thresholds. Of note, a perfect correlation was shown to exist between distance to CBD – near and distance to CBD – far. However, this was expected as these variables represent a two-category dummy variable. Therefore, both variables remained in the regression and were not eliminated during the evaluation process. Updated correlation plots after removal of households with children are further summarized in Appendix J.

Table 8 – *Initial Unrestricted Model Correlation Plots, Highly Correlated Independent Variables*

A.M. Peak Correlation Plot	Population Density	Distance to CBD – Far
Households With Children	0.98	-
Distance to CBD – Near	-	-1*

Midday Off-Peak Correlation Plot	Population Density	Distance to CBD – Far
Households With Children	0.95	-
Distance to CBD – Near	-	-1*

P.M. Peak Correlation Plot	Population Density	Distance to CBD – Far
Households With Children	0.97	-
Distance to CBD – Near	-	-1*

Evening Off-Peak Correlation Plot	Population Density	Distance to CBD – Far
Households With Children	0.97	-
Distance to CBD – Near	-	-1*

5.2.6. Initial Unrestricted Models

Panel data frames were then created in R-Studio using the "pdata.frame" function. A separate panel data frame was created for each time period. Using the "plm" package, ridership in each time period was regressed on the final list of independent variables using both pooled OLS, fixed effect, and random effect estimators (Croissant & Miinno, 2008). Independent variables including employment density, distance to CBD – near, and distance to CBD – far were excluded from the fixed effect analysis as they are time-invariant. In all models, the distance to CBD - near dummy variable was compared to the baseline of distance to CBD – far, and the winter, spring, and summer dummy variables were compared to the fall baseline. Initial results from the unrestricted model outputs are summarized in Table 9 to Table 12.

Table 9 – *A.M. Peak Unrestricted Model Outputs*

A.M. Peak Unrestricted Model Outputs												
	-	-	-	-	-	-	-	-	-	-	-	-
	Pooled OLS Model				**Fixed Effect Model**				**Random Effect Model**			
	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value
(Intercept)	-10.193	1.230	-8.284	< 0.001	-	-	-	-	-6.763	4.943	-1.368	0.171
Service Quantity	0.661	0.018	37.745	< 0.001	0.174	0.029	5.919	< 0.001	0.208	0.028	7.429	< 0.001
Fare Price	0.487	0.060	8.139	< 0.001	-0.337	0.063	-5.326	< 0.001	-0.283	0.062	-4.589	< 0.001
Feeder Bus Connection Quality	-0.202	0.024	-8.400	< 0.001	0.006	0.020	0.295	0.768	-0.008	0.020	-0.411	0.681
Population Density	0.413	0.021	19.281	< 0.001	0.049	0.439	0.111	0.911	0.397	0.125	3.174	0.002
Gender - Female	-0.017	0.111	-0.151	0.880	-0.061	0.045	-1.362	0.173	-0.057	0.046	-1.221	0.222
Unemployment Rate	0.720	0.053	13.520	< 0.001	-0.086	0.265	-0.325	0.745	0.568	0.182	3.128	0.002
Income	1.015	0.058	17.601	< 0.001	-0.698	0.293	-2.385	0.017	0.151	0.231	0.654	0.513
Age	-1.300	0.202	-6.435	< 0.001	3.306	1.674	1.975	0.048	1.201	1.047	1.147	0.252
Employment Density	0.056	0.015	3.658	< 0.001	-	-	-	-	-0.134	0.097	-1.390	0.165
Fuel Price	0.077	0.090	0.859	0.391	0.217	0.041	5.258	< 0.001	0.199	0.041	4.902	< 0.001
Vehicle Ownership	0.148	0.021	7.067	< 0.001	0.650	0.255	2.546	0.011	0.265	0.109	2.432	0.015
Park and Ride Capacity	0.238	0.005	48.743	< 0.001	0.019	0.008	2.374	0.018	0.043	0.008	5.593	< 0.001
Distance to CBD - Near	-0.522	0.037	-14.198	< 0.001	-	-	-	-	-1.626	0.177	-9.187	< 0.001
Winter	0.024	0.021	1.120	0.263	0.018	0.009	2.105	0.035	0.022	0.009	2.440	0.015
Spring	0.057	0.022	2.630	0.009	0.014	0.009	1.498	0.134	0.022	0.009	2.428	0.015
Summer	0.019	0.021	0.899	0.369	-0.014	0.009	-1.564	0.118	-0.007	0.009	-0.837	0.403
	Number of Observations: 2928. **Total Sum of Squares:** 3110.5. **Residual Sum of Squares:** 473.62. **R-Squared:** 0.84773. **Adj. R-Squared:** 0.8469. **F-statistic:** 1012.91 on 16 and 2911 DF, p-value < 0.001..				**Number of Observations:** 2928. **Total Sum of Squares:** 84.317. **Residual Sum of Squares:** 73.809. **R-Squared:** 0.12463. **Adj. R-Squared:** 0.10192. **F-statistic:** 29.0135 on 14 and 2853 DF, p-value < 0.001.				**Number of Observations:** 2928. **Total Sum of Squares:** 94.99. **Residual Sum of Squares:** 79.542. **R-Squared:** 0.16263. **Adj. R-Squared:** 0.15802. **Wald Chi-Squared:** 565.342 on 16 DF, p-value < 0.001.			

Table 10 - *Midday Off-Peak Unrestricted Model Outputs*

Midday Off-Peak Unrestricted Model Outputs												
	Pooled OLS Model				**Fixed Effect Model**				**Random Effect Model**			
	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value
(Intercept)	-1.891	2.261	-0.836	0.403	-	-	-	-	-57.082	8.400	-6.796	< 0.001
Service Quantity	0.978	0.017	57.003	< 0.001	0.535	0.016	33.054	< 0.001	0.565	0.016	35.244	< 0.001
Fare Price	1.261	0.086	14.741	< 0.001	-0.288	0.069	-4.199	< 0.001	-0.232	0.069	-3.382	< 0.001
Feeder Bus Connection Quality	-0.176	0.046	-3.849	< 0.001	-0.116	0.033	-3.560	< 0.001	-0.091	0.033	-2.726	0.006
Population Density	0.523	0.051	10.213	< 0.001	0.062	1.289	0.048	0.962	1.442	0.249	5.779	< 0.001
Gender - Female	-0.087	0.141	-0.618	0.537	0.011	0.057	0.199	0.842	0.010	0.059	0.167	0.868
Unemployment Rate	-0.078	0.101	-0.765	0.444	-0.733	0.515	-1.422	0.155	0.427	0.319	1.338	0.181
Income	0.199	0.101	1.976	0.048	3.919	0.746	5.253	< 0.001	1.593	0.427	3.727	< 0.001
Age	-0.420	0.411	-1.023	0.306	27.475	2.415	11.378	< 0.001	11.824	1.710	6.913	< 0.001
Employment Density	0.001	0.036	0.038	0.969	-	-	-	-	-0.504	0.187	-2.695	0.007
Fuel Price	0.392	0.151	2.594	0.010	0.514	0.067	7.683	< 0.001	0.495	0.067	7.421	< 0.001
Vehicle Ownership	-0.195	0.070	-2.804	0.005	0.153	0.517	0.297	0.767	-0.859	0.222	-3.878	< 0.001
Park and Ride Capacity	0.031	0.011	2.700	0.007	-0.010	0.015	-0.666	0.506	0.002	0.015	0.141	0.888
Distance to CBD - Near	-0.991	0.074	-13.441	< 0.001	-	-	-	-	-1.689	0.303	-5.571	< 0.001
Winter	-0.138	0.035	-3.925	< 0.001	-0.127	0.015	-8.758	< 0.001	-0.136	0.015	-9.253	< 0.001
Spring	-0.136	0.036	-3.787	< 0.001	-0.104	0.015	-6.751	< 0.001	-0.121	0.015	-7.875	< 0.001
Summer	0.092	0.035	2.632	0.009	0.103	0.014	7.226	< 0.001	0.094	0.014	6.475	< 0.001
	Number of Observations: 1735. **Total Sum of Squares:** 2553.2. **Residual Sum of Squares:** 448.71. **R-Squared:** 0.82425. **Adj. R-Squared:** 0.82262. **F-statistic:** 503.595 on 16 and 1718 DF, p-value < 0.001.				**Number of Observations:** 1735. **Total Sum of Squares:** 189.17. **Residual Sum of Squares:** 69.147. **R-Squared:** 0.63448. **Adj. R-Squared:** 0.6216. **F-statistic:** 207.675 on 14 and 1675 DF, p-value < 0.001.				**Number of Observations:** 1735. **Total Sum of Squares:** 198.78. **Residual Sum of Squares:** 75.675. **R-Squared:** 0.61974. **Adj. R-Squared:** 0.6162. **Wald Chi-Squared:** 2856.94 on 16 DF, p-value < 0.001.			

Table 11 *– P.M. Peak Unrestricted Model Outputs*

P.M. Peak Unrestricted Model Outputs												
	Unrestricted Pooled OLS Model				**Unrestricted Fixed Effect Model**				**Unrestricted Random Effect Model**			
	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value
(Intercept)	-4.862	1.831	-2.655	0.008	-	-	-	-	-32.576	6.868	-4.743	< 0.001
Service Quantity	1.627	0.023	69.965	< 0.001	0.351	0.037	9.437	< 0.001	0.553	0.034	16.140	< 0.001
Fare Price	1.482	0.044	33.498	< 0.001	0.088	0.032	2.750	0.006	0.132	0.033	3.989	< 0.001
Feeder Bus Connection Quality	-0.214	0.036	-6.006	< 0.001	0.094	0.026	3.687	< 0.001	0.091	0.026	3.443	< 0.001
Population Density	0.395	0.031	12.640	< 0.001	-2.328	0.543	-4.290	< 0.001	0.390	0.175	2.226	0.026
Gender - Female	0.002	0.160	0.014	0.989	-0.001	0.054	-0.020	0.984	0.013	0.057	0.230	0.818
Unemployment Rate	0.590	0.088	6.697	< 0.001	1.919	0.356	5.399	< 0.001	2.662	0.244	10.932	< 0.001
Income	-0.113	0.084	-1.351	0.177	2.475	0.383	6.466	< 0.001	1.591	0.315	5.044	< 0.001
Age	0.475	0.334	1.420	0.156	11.006	2.044	5.384	< 0.001	2.935	1.453	2.020	0.043
Employment Density	0.165	0.023	7.016	< 0.001	-	-	-	-	0.259	0.136	1.900	0.057
Fuel Price	0.289	0.136	2.126	0.034	0.119	0.051	2.316	0.021	0.169	0.052	3.231	0.001
Vehicle Ownership	-0.174	0.033	-5.353	< 0.001	1.653	0.355	4.652	< 0.001	-0.352	0.167	-2.110	0.035
Park and Ride Capacity	-0.037	0.009	-4.014	< 0.001	0.006	0.010	0.583	0.560	-0.002	0.010	-0.162	0.871
Distance to CBD - Near	-0.044	0.059	-0.743	0.457		-	-		0.333	0.248	1.341	0.180
Winter	-0.042	0.032	-1.288	0.198	-0.097	0.011	-8.781	< 0.001	-0.099	0.011	-8.668	< 0.001
Spring	0.016	0.033	0.493	0.622	0.035	0.011	3.076	0.002	0.026	0.012	2.191	0.028
Summer	0.163	0.032	5.107	< 0.001	0.214	0.011	19.719	< 0.001	0.205	0.011	18.153	< 0.001
	Number of Observations: 2690. **Total Sum of Squares:** 6144.5. **Residual Sum of Squares:** 902.95. **R-Squared:** 0.85305. **Adj. R-Squared:** 0.85217. **F-statistic:** 969.783 on 16 and 2673 DF, p-value < 0.001.				**Number of Observations:** 2690. **Total Sum of Squares:** 169.77. **Residual Sum of Squares:** 97.334. **R-Squared:** 0.42666. **Adj. R-Squared:** 0.41089. **F-statistic:** 139.107 on 14 and 2617 DF, p-value < 0.001.				**Number of Observations:** 2690. **Total Sum of Squares:** 185.6. **Residual Sum of Squares:** 109.37. **R-Squared:** 0.41093. **Adj. R-Squared:** 0.40741. **Wald Chi-Squared:** 1898.47 on 16 DF, p-value < 0.001.			

Table 12 - *Evening Off-Peak Unrestricted Model Outputs*

Evening Off-Peak Unrestricted Model Outputs												
	Unrestricted Pooled OLS Model				**Unrestricted Fixed Effect Model**				**Unrestricted Random Effect Model**			
	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value	Coefficient	SE	t-value	p-value
(Intercept)	-0.549	2.801	-0.196	0.845	-	-	-	-	-45.478	10.957	-4.150	< 0.001
Service Quantity	1.014	0.016	61.803	< 0.001	0.458	0.017	26.396	< 0.001	0.488	0.017	29.407	< 0.001
Fare Price	1.454	0.051	28.750	< 0.001	0.067	0.032	2.068	0.039	0.118	0.033	3.596	< 0.001
Feeder Bus Connection Quality	-0.035	0.051	-0.695	0.487	-0.054	0.040	-1.346	0.178	-0.060	0.040	-1.496	0.135
Population Density	0.453	0.045	10.012	< 0.001	-3.231	0.877	-3.682	< 0.001	0.567	0.261	2.175	0.030
Gender - Female	-0.120	0.233	-0.516	0.606	-0.007	0.086	-0.078	0.938	-0.006	0.088	-0.073	0.942
Unemployment Rate	1.019	0.135	7.557	< 0.001	3.721	0.615	6.054	< 0.001	2.346	0.381	6.151	< 0.001
Income	0.056	0.128	0.435	0.664	6.181	0.662	9.339	< 0.001	3.130	0.505	6.198	< 0.001
Age	-0.618	0.495	-1.249	0.212	19.649	3.511	5.596	< 0.001	5.397	2.283	2.364	0.018
Employment Density	0.301	0.035	8.713	< 0.001	-	-	-	-	0.264	0.202	1.306	0.191
Fuel Price	-0.221	0.206	-1.073	0.283	0.298	0.084	3.549	< 0.001	0.204	0.084	2.434	0.015
Vehicle Ownership	-0.350	0.045	-7.769	< 0.001	-2.207	0.577	-3.823	< 0.001	-1.493	0.247	-6.035	< 0.001
Park and Ride Capacity	0.050	0.012	4.130	< 0.001	-0.002	0.015	-0.154	0.877	0.000	0.015	0.026	0.979
Distance to CBD - Near	0.610	0.096	6.379	< 0.001	-	-	-	-	1.065	0.377	2.822	0.005
Winter	-0.084	0.049	-1.722	0.085	-0.131	0.018	-7.357	< 0.001	-0.131	0.018	-7.206	< 0.001
Spring	0.136	0.049	2.765	0.006	0.109	0.019	5.846	< 0.001	0.103	0.019	5.427	< 0.001
Summer	0.315	0.048	6.596	< 0.001	0.313	0.018	17.777	< 0.001	0.307	0.018	17.079	< 0.001
	Number of Observations: 2515. **Total Sum of Squares:** 8160.8. **Residual Sum of Squares:** 1787.3. **R-Squared:** 0.78099. **Adj. R-Squared:** 0.77959. **F-statistic:** 556.741 on 16 and 2498 DF, p-value < 0.001.				**Number of Observations:** 2515. **Total Sum of Squares:** 417.74. **Residual Sum of Squares:** 227.01. **R-Squared:** 0.45656. **Adj. R-Squared:** 0.44077. **F-statistic:** 146.605 on 14 and 2443 DF, p-value < 0.001.				**Number of Observations:** 2515. **Total Sum of Squares:** 440.87. **Residual Sum of Squares**: 244.92. **R-Squared**: 0.44449. **Adj. R-Squared**: 0.44094. **Wald Chi-Squared:** 2015.03 on 16 DF, p-value < 0.001.			

5.2.7. Evaluation of Initial Unrestricted Models

Model diagnostic tools were applied to each set of models to determine which estimator best fit the data. First, a Breusch-Pagen Lagrange Multiplier test was applied using the "plmtest" function from the "plm" package to compare the pooled OLS and random effect estimators (Croissant & Miinno, 2008). For each set of models, the test returned significant results, indicating that the use of a random effect estimator should be further investigated.

Second, a Breusch-Pagen Lagrange Multiplier test was applied using the "pFtest" function from the "plm" package to compare pooled OLS and fixed effect estimators (Croissant & Miinno, 2008). Again, the test returned significant results for each set of models, indicating that the use of pooled OLS is not an efficient estimator for these datasets.

A Hausman test was then conducted to compare the fixed effect estimator to the random effect estimator using the "phtest" function (Croissant & Miinno, 2008). The test returned significant for each set of models, indicating that a fixed effect model is most applicable for these datasets.

5.2.8. Model Selection

Per the Hausman test, a fixed effect estimator was selected, and a stepwise regression procedure was completed for each dataset. However, the results were too degraded to generate informative results and discussion. As the stepwise regression procedure continued, the majority of independent variables included in the unrestricted model were eliminated as they were shown to be insignificant in explaining transit demand, while only a few independent variables remained at the conclusion of the procedure. Additionally, some of the remaining independent variables demonstrated inflated coefficients and counter-intuitive signs that did not align with previous estimates identified in the literature.

Torres-Reyna (2007) states that this can occur when the structure of the dataset being analyzed does not align with the mathematical concept of the fixed effect estimator. Because a fixed effect analysis only considers change over time and does not consider cross-sectional differences, such models are only efficient when independent variable datasets demonstrate considerable temporal variation. Since the external variable datasets used were generated from projections and do not reflect real values, this could have resulted in little fluctuation over the time-series, thus resulting in insignificant outputs. Additionally, some independent variables such as employment density and distance to CBD were excluded from the model altogether as they are time invariant. Eliminating these sources of information could have further skewed model outputs and increased model error, as the reviewed literature indicated that these are significant sources of information that help explain transit demand. As noted in Lee et al. (2013), the use of a fixed effect estimator can be inefficient if time-invariant variables are theorized to be significant in explaining transit demand.

The literature further notes that if the researcher has reason to believe that cross-sectional differences have some impact on the dependent variable being analyzed, a random effect estimator should be selected (Torres-Reyna, 2007). When station-level transit demand is analyzed, cross-sectional differences in external variable datasets such as population and employment density are expected to have an influence on ridership, as catchment areas with more residents and workers inherently results in a larger customer base compared to stations located in sprawled and dispersed environments.

A random estimator was therefore selected to produce final modelling outputs as this method allowed transit demand to be explained by both cross-sectional, temporal, and time-invariant factors, therefore increasing the accuracy and explanatory capacity of the models. The final form of each unrestricted random effect model is outlined below:

$$\begin{aligned}(RidershipAMPeak)_{it} &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} \\ &+ \beta(FeederBusConnectionQuality)_{it} + \beta(PopulationDensity)_{it} \\ &+ \beta(Gender - Female)_{it} + \beta(UnemploymentRate)_{it} + \beta(Income)_{it} \\ &+ \beta(Age)_{it} + \beta(EmploymentDensity)_{it} + \beta(FuelPrice)_{it} \\ &+ \beta(VehicleOwnership)_{it} + \beta(ParkandRideCapacity)_{it} \\ &+ \beta(DistancetoCBD - Near)_{it} + \beta(Winter)_{it} + \beta(Spring)_{it} \\ &+ \beta(Summer)_{it} + a_1 + u_{it} + \varepsilon_{it}\end{aligned} \tag{Eq. 23}$$

$$\begin{aligned}(RidershipMiddayOffPeak)_{it} &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} \\ &+ \beta(FeederBusConnectionQuality)_{it} + \beta(PopulationDensity)_{it} \\ &+ \beta(Gender - Female)_{it} + \beta(UnemploymentRate)_{it} + \beta(Income)_{it} \\ &+ \beta(Age)_{it} + \beta(EmploymentDensity)_{it} + \beta(FuelPrice)_{it} \\ &+ \beta(VehicleOwnership)_{it} + \beta(ParkandRideCapacity)_{it} \\ &+ \beta(DistancetoCBD - Near)_{it} + \beta(Winter)_{it} + \beta(Spring)_{it} \\ &+ \beta(Summer)_{it} + a_1 + u_{it} + \varepsilon_{it}\end{aligned} \tag{Eq. 24}$$

$$\begin{aligned}(RidershipPMPeak)_{it} &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} \\ &+ \beta(FeederBusConnectionQuality)_{it} + \beta(PopulationDensity)_{it} \\ &+ \beta(Gender - Female)_{it} + \beta(UnemploymentRate)_{it} + \beta(Income)_{it} \\ &+ \beta(Age)_{it} + \beta(EmploymentDensity)_{it} + \beta(FuelPrice)_{it} \\ &+ \beta(VehicleOwnership)_{it} + \beta(ParkandRideCapacity)_{it} \\ &+ \beta(DistancetoCBD - Near)_{it} + \beta(Winter)_{it} + \beta(Spring)_{it} \\ &+ \beta(Summer)_{it} + a_1 + u_{it} + \varepsilon_{it}\end{aligned} \tag{Eq. 25}$$

$$
\begin{aligned}
(RidershipEveningOffPeak)_{it} &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} \\
&+ \beta(FeederBusConnectionQuality)_{it} + \beta(PopulationDensity)_{it} \\
&+ \beta(Gender - Female)_{it} + \beta(UnemploymentRate)_{it} + \beta(Income)_{it} \\
&+ \beta(Age)_{it} + \beta(EmploymentDensity)_{it} + \beta(FuelPrice)_{it} \\
&+ \beta(VehicleOwnership)_{it} + \beta(ParkandRideCapacity)_{it} \\
&+ \beta(DistancetoCBD - Near)_{it} + \beta(Winter)_{it} + \beta(Spring)_{it} \\
&+ \beta(Summer)_{it} + a_1 + u_{it} + \varepsilon_{it}
\end{aligned}
\qquad \textit{Eq. 26}
$$

5.2.9. Model Diagnostics

Each random effect unrestricted model was tested for the presence of heteroskedasticity and serial correlation of error terms. A Durbin-Watson test for serial correlation in panel data models was conducted using the "pdwtest" function, while a Breusch-Pagen test for heteroskedasticity was conducted using the "bptest" function obtained from the "lmtest" package (Croissant & Miinno, 2008; Zeileis, 2002). Both tests were completed in R-Studio. Significant results were found in all models, therefore indicating the presence of heteroskedasticity and serial correlation. Model outputs produced during the stepwise regression procedure were thereby estimated with White robust standard errors, clustered at the station level, to ensure that estimated coefficients and the level of significance associated with each individual variable included in the analysis was correctly specified (Guerra & Cervero, 2011; Torres-Reyna, 2007; Zeileis, 2002).

Table 13 - *Unrestricted Random Effect Model Outputs Estimated Using Robust Standard Errors*

A.M. Peak Unrestricted Model Estimated Using Robust Standard Errors

-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-6.763	5.893	-1.148	0.251
Service Quantity	0.208	0.033	6.248	< 0.001
Fare Price	-0.283	0.112	-2.522	0.012
Feeder Bus Connection Quality	-0.008	0.015	-0.546	0.585
Population Density	0.397	0.139	2.861	0.004
Gender - Female	-0.057	0.005	-11.696	< 0.001
Unemployment Rate	0.568	0.273	2.080	0.038
Income	0.151	0.203	0.743	0.458
Age	1.201	1.217	0.987	0.324
Employment Density	-0.134	0.089	-1.510	0.131
Fuel Price	0.199	0.036	5.524	< 0.001
Vehicle Ownership	0.265	0.072	3.651	< 0.001
Park and Ride Capacity	0.043	0.010	4.282	< 0.001
Distance to CBD - Near	-1.626	0.134	-12.140	< 0.001
Winter	0.022	0.009	2.511	0.012
Spring	0.022	0.009	2.435	0.015
Summer	-0.007	0.009	-0.780	0.435

Number of Observations: 2928. **Total Sum of Squares:** 94.99. **Residual Sum of Squares:** 79.542. **R-Squared:** 0.16263. **Adj. R-Squared:** 0.15802. **Wald Chi-Squared Test:** 1441.1 on 16 DF, p-value < 0.001.

-

Midday Off-Peak Unrestricted Model Estimated Using Robust Standard Errors

-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-1.891	2.261	-0.836	0.403
Service Quantity	0.978	0.017	57.003	< 0.001
Fare Price	1.261	0.086	14.741	< 0.001
Feeder Bus Connection Quality	-0.176	0.046	-3.849	< 0.001
Population Density	0.523	0.051	10.213	< 0.001
Gender - Female	-0.087	0.141	-0.618	0.537
Unemployment Rate	-0.078	0.101	-0.765	0.444
Income	0.199	0.101	1.976	0.048
Age	-0.420	0.411	-1.023	0.306
Employment Density	0.001	0.036	0.038	0.969
Fuel Price	0.392	0.151	2.594	0.010
Vehicle Ownership	-0.195	0.070	-2.804	0.005
Park and Ride Capacity	0.031	0.011	2.700	0.007
Distance to CBD - Near	-0.991	0.074	-13.441	< 0.001
Winter	-0.138	0.035	-3.925	< 0.001
Spring	-0.136	0.036	-3.787	< 0.001
Summer	0.092	0.035	2.632	0.009

Number of Observations: 1735. **Total Sum of Squares:** 198.78. **Residual Sum of Squares:** 75.675. **R-Squared:** 0.61974. **Adj. R-Squared:** 0.6162. **Wald Chi-Squared Test:** 2237.1 on 16 DF, p-value < 0.001.

Table 14 - *Unrestricted Random Effect Model Outputs Estimated Using Robust Standard Errors (continued)*

P.M. Peak Unrestricted Model Estimated Using Robust Standard Errors				
-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-32.576	7.182	-4.536	< 0.001
Service Quantity	0.553	0.055	10.154	< 0.001
Fare Price	0.132	0.045	2.910	0.004
Feeder Bus Connection Quality	0.091	0.024	3.763	< 0.001
Population Density	0.390	0.187	2.084	0.037
Gender - Female	0.013	0.006	2.263	0.024
Unemployment Rate	2.662	0.246	10.806	< 0.001
Income	1.591	0.341	4.670	< 0.001
Age	2.935	1.508	1.947	0.052
Employment Density	0.259	0.144	1.801	0.072
Fuel Price	0.169	0.054	3.139	0.002
Vehicle Ownership	-0.352	0.185	-1.902	0.057
Park and Ride Capacity	-0.002	0.013	-0.119	0.906
Distance to CBD - Near	0.333	0.271	1.227	0.220
Winter	-0.099	0.011	-9.203	< 0.001
Spring	0.026	0.011	2.298	0.022
Summer	0.205	0.012	17.347	< 0.001
Number of Observations: 2690. **Total Sum of Squares:** 185.6. **Residual Sum of Squares:** 109.37. **R-Squared:** 0.41093. **Adj. R-Squared:** 0.40741. **Wald Chi-Squared Test:** 1438.6 on 16 DF, p-value < 0.001.				

-

Evening Off-Peak Unrestricted Model Estimated Using Robust Standard Errors				
-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-45.478	13.414	-3.390	< 0.001
Service Quantity	0.488	0.027	17.884	< 0.001
Fare Price	0.118	0.046	2.568	0.010
Feeder Bus Connection Quality	-0.060	0.045	-1.338	0.181
Population Density	0.567	0.243	2.329	0.020
Gender - Female	-0.006	0.019	-0.344	0.731
Unemployment Rate	2.346	0.360	6.513	< 0.001
Income	3.130	0.552	5.672	< 0.001
Age	5.397	2.579	2.092	0.036
Employment Density	0.264	0.206	1.284	0.199
Fuel Price	0.204	0.086	2.366	0.018
Vehicle Ownership	-1.493	0.219	-6.809	< 0.001
Park and Ride Capacity	0.000	0.017	0.023	0.982
Distance to CBD - Near	1.065	0.386	2.758	0.006
Winter	-0.131	0.017	-7.679	< 0.001
Spring	0.103	0.018	5.789	< 0.001
Summer	0.307	0.019	16.296	< 0.001
Number of Observations: 2515. **Total Sum of Squares:** 440.87. **Residual Sum of Squares:** 244.92. **R-Squared:** 0.44449. **Adj. R-Squared:** 0.44094. **Wald Chi-Squared Test:** 38734 on 16 DF, p-value < 0.001.				

5.2.10. Stepwise Regression

Table 13 outlines the results produced from the unrestricted random effect models estimated using robust clustered standard errors. Initial examination of the model outputs show that a variety of variables are statistically insignificant in explaining ridership, as large p-values are associated with several factors. A backwards stepwise regression procedure was therefore used to eliminate insignificant independent variables, with the purpose of improving the explanatory capacity of each model.

The stepwise regression was completed by selecting and eliminating the independent variable associated with the largest p-value in each model. Once a variable was removed, model outputs were recalculated so that updated statistical measures could further inform the stepwise regression process. The stepwise regression process continued until all variables included in each model were statistically significant, as evidenced by a corresponding p-value less than or equal to 0.1. Analytical reasoning was also used to eliminate variables where the sign or magnitude of the coefficient was counterintuitive. Once a variable was removed, model outputs were recalculated until all variables included in the model were shown to be significant. The stepwise regression process undertaken for each model is detailed below.

Age was eliminated from every model as it displayed an abnormally large coefficient, and its exclusion did not dramatically decrease model performance. After removal, feeder bus connection quality, income, and employment density were eliminated from the a.m. peak model due to insignificance. Summer was also shown to be insignificant, but remained in the model as a separate linear hypothesis test showed that all seasonal dummy variables were jointly significant in explaining demand (Fox & Weisberg, 2019). Employment density and park and ride capacity were eliminated from the midday off-peak model due to insignificance. Park and ride capacity were also eliminated from the p.m. peak and evening off-peak model as it was not significantly correlated with demand. All variables in the evening off-peak model were shown to be significant once gender - female and feeder bus connection quality was removed.

After the stepwise regression procedure, the p-value associated with all variables included in the final models were shown to be less than the predetermined significance threshold, indicating statistical significance. Final results generated for each trip type are summarized in the following chapter. A complete copy of the regression outputs used to inform the stepwise regression process can be found in Appendix K.

5.3. Chapter Summary

In sum, this section summarized the steps taken to clean and prepare the datasets, the process used to remove multi-collinearity from each model, how the panel data model estimator was selected, and the stepwise regression procedure used. Findings from the restricted model outputs are therefore presented in the following chapter. A full reproducible code used to complete the analysis as outlined above can be obtained by contacting the author.

6. Results

In this chapter, separate regression outputs are presented for each trip type analyzed. Ridership elasticities pertaining to internal, socioeconomic, demographic, and station access indicators are presented in Table 16 and Table 17. As described in Table 15, results are color coded to indicate the level of significance and relationship demonstrated by each independent variable included in each model. A summary of the significance and magnitude associated with each demand elasticity is further provided, as is a comparison of similar and/or differing results between models. A discussion regarding the explanatory capacity of each model is also outlined.

6.1. A.M. Peak Model

After adjusting for multi-collinearity, eliminating insignificant variables, and controlling for heteroscedasticity and spatial correlation, the restricted a.m. peak ridership model takes the following form:

$$\begin{aligned}(RidershipAMPeak)_{it} &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} + \beta(PopulationDensity)_{it} \\ &+ \beta(Gender - Female)_{it} + \beta(UnemploymentRate)_{it} \\ &+ \beta(FuelPrice)_{it} + \beta(VehicleOwnership)_{it} \\ &+ \beta(ParkandRideCapacity)_{it} + \beta(DistancetoCBD - Near)_{it} \\ &+ \beta(Winter)_{it} + \beta(Spring)_{it} + \beta(Summer)_{it} + a_1 + u_{it} + \varepsilon_{it}\end{aligned} \qquad \textit{Eq. 27}$$

As shown in Table 16, all continuous variables were found to significantly explain ridership. Some variables, including service quantity, gender - female, price of fuel, park and ride capacity, vehicle ownership, and distance to CBD - near were shown to have a greater influence on a.m. peak period ridership as a p-value < 0.001 was associated with these factors. When considered independently, dummy variables representing seasonal effects had differing impacts on ridership. However, a joint significance test revealed that all seasonal variables were significant in explaining the number of boardings. As a result, summer remained in the model regardless of the large p-value associated with this factor. The results indicate that a combination of internal and external variables were significant in explaining ridership during the a.m. peak time period.

Table 15 - Regression Model Color Coding Scheme

Sign With Respect to A Priori Assumption		
A Priori Assumption Realized	A Priori Assumption Not Realized	A Priori Assumption Not Formulated

Table 16 - *Restricted Random Effect Model Outputs Estimated Using Robust Standard Errors*

A.M. Peak Restricted Model Estimated Using Robust Standard Errors				
-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-0.983	1.154	-0.852	0.394
Service Quantity	0.207	0.033	6.261	< 0.001
Fare Price	-0.296	0.114	-2.591	0.010
Population Density	0.284	0.111	2.550	0.011
Gender - Female	-0.056	0.004	-12.625	< 0.001
Unemployment Rate	0.532	0.268	1.990	0.047
Fuel Price	0.203	0.036	5.642	< 0.001
Vehicle Ownership	0.305	0.080	3.816	< 0.001
Park and Ride Capacity	0.041	0.010	4.194	< 0.001
Distance to CBD - Near	-1.543	0.145	-10.611	< 0.001
Winter	0.021	0.009	2.396	0.017
Spring	0.020	0.009	2.213	0.027
Summer*	-0.009	0.009	-0.898	0.369
Number of Observations: 2928. **Total Sum of Squares:** 93.71. **Residual Sum of Squares:** 79.02. **R-Squared:** 0.15676. **Adj. R-Squared:** 0.15329. **Wald Chi-Squared Test:** 1291.4 on 12 DF, p-value < 0.001. *Summer dummy variable not eliminated as a joint significance test found that all seasonal variables were significant in explaining ridership.				

-

Midday Off-Peak Restricted Model Estimated Using Robust Standard Errors				
-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-13.069	5.143	-2.541	0.011
Service Quantity	0.571	0.030	19.334	< 0.001
Fare Price	-0.244	0.095	-2.566	0.010
Feeder Bus Connection Quality	-0.085	0.035	-2.449	0.014
Population Density	0.770	0.147	5.240	< 0.001
Gender - Female	0.022	0.008	2.811	0.005
Unemployment Rate	0.625	0.274	2.278	0.023
Income	1.515	0.383	3.953	< 0.001
Fuel Price	0.532	0.068	7.833	< 0.001
Vehicle Ownership	-0.692	0.173	-3.992	< 0.001
Distance to CBD - Near	-0.882	0.215	-4.100	< 0.001
Winter	-0.139	0.014	-9.847	< 0.001
Spring	-0.126	0.015	-8.515	< 0.001
Summer	0.091	0.015	6.210	< 0.001
Number of Observations: 1735. **Total Sum of Squares:** 200.49. **Residual Sum of Squares:** 78.324. **R-Squared:** 0.60989. **Adj. R-Squared:** 0.60694. **Wald Chi-Squared Test:** 1798.4 on 13 DF, p-value < 0.001.				

Table 17 - *Restricted Random Effect Model Outputs Estimated Using Robust Standard Errors (continued)*

P.M. Peak Restricted Model Estimated Using Robust Standard Errors				
-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-20.945	4.479	-4.676	< 0.001
Service Quantity	0.562	0.054	10.448	< 0.001
Fare Price	0.135	0.045	2.993	0.003
Feeder Bus Connection Quality	0.094	0.024	3.902	< 0.001
Population Density	0.307	0.171	1.790	0.074
Gender - Female	0.016	0.006	2.650	0.008
Unemployment Rate	2.703	0.238	11.344	< 0.001
Income	1.506	0.337	4.467	< 0.001
Employment Density	0.342	0.136	2.503	0.012
Fuel Price	0.176	0.054	3.292	0.001
Vehicle Ownership	-0.339	0.180	-1.887	0.059
Distance to CBD - Near	0.485	0.258	1.878	0.060
Winter	-0.100	0.011	-9.235	< 0.001
Spring	0.024	0.011	2.178	0.030
Summer	0.204	0.012	17.326	< 0.001
Number of Observations: 2690. **Total Sum of Squares:** 186.53. **Residual Sum of Squares:** 110.07. **R-Squared:** 0.41012. **Adj. R-Squared:** 0.40703. **Wald Chi-Squared Test:** 1090.5 on 14 DF, p-value < 0.001.				

-

Evening Off-Peak Restricted Model Estimated Using Robust Standard Errors				
-	-	-	-	-
	Coefficient	SE	t-value	p-value
(Intercept)	-23.063	6.666	-3.460	< 0.001
Service Quantity	0.493	0.027	18.115	< 0.001
Fare Price	0.121	0.046	2.639	0.008
Population Density	0.410	0.210	1.950	0.051
Unemployment Rate	2.388	0.350	6.828	< 0.001
Income	2.824	0.522	5.414	< 0.001
Employment Density	0.403	0.175	2.300	0.022
Fuel Price	0.197	0.086	2.306	0.021
Vehicle Ownership	-1.407	0.213	-6.593	< 0.001
Distance to CBD - Near	1.267	0.362	3.496	< 0.001
Winter	-0.132	0.017	-7.761	< 0.001
Spring	0.099	0.018	5.621	< 0.001
Summer	0.306	0.019	16.198	< 0.001
Number of Observations: 2515. **Total Sum of Squares:** 443.3. **Residual Sum of Squares:** 246.98. **R-Squared:** 0.44288. **Adj. R-Squared:** 0.44021. **Wald Chi-Squared Test:** 759.55 on 12 DF, p-value < 0.001.				

Service quantity demonstrated a coefficient of 0.207, indicating that a 10% increase in service supply could increase ridership by 2.07%. This suggests that users are more sensitive to fare price increases compared to service quantity changes, as a demand elasticity with respect to fare price of -0.296 was found. Station location demonstrated the largest coefficient in the analysis. The base scenario is stations located outside of the city of Toronto, so the negative sign of distance to CBD – near indicates that stations located within the city of Toronto were associated with fewer boardings than those located beyond the municipal boundary.

Variables related to station access, including park and ride capacity and feeder bus connection quality generated differing results. Feeder bus connection quality was removed from the model during the stepwise regression procedure due to insignificance. However, park and ride capacity was statistically significant in explaining ridership as a p-value < 0.001 was found. A coefficient of 0.041 indicates that a 10% increase in parking spaces could yield a 0.41% increase in boardings. These results suggest that park and ride lots are a significant means of connecting the current customer base to the network, whereas feeder bus connection quality does not play a significant role in facilitating ridership.

Population density demonstrated a coefficient of 0.284, indicating that ridership should increase by 2.84% if the concentration of residents is increased by 10%. Notably, employment density was removed during the stepwise regression procedure due to insignificance This conforms to the expectation that stations located in areas with high employment densities have less trip production during the a.m. peak period, as the majority of boardings during this time period are derived from home-based commuter related trips. Unemployment rate significantly explained ridership, while the coefficient indicates that a 10% increase in unemployment could increase ridership by 5.32%. Consistent with Stover & Bae (2011), the results suggest that sustained periods of economic downturn could render vehicle ownership infeasible, therefore resulting in increased ridership. Gender - female demonstrated a significant negative correlation with ridership. As expected, this indicates that station catchment areas with a greater proportion of female residents were less likely to take transit. Variables related to automobile ownership were positively correlated with ridership. This was expected for fuel price, as rising vehicle operation costs can entice travelers to use transit as it is more affordable compared to private vehicle use. However, vehicle ownership also demonstrated a positive relationship, indicating that an increase in the number of vehicles within the study area has a positive impact on ridership. These findings indicate that vehicle availability can complement transit use, as long as the cost of regional transport via private automobile is large. Seasonality was also shown to have a significant impact on the number of boardings. The base scenario is observations that were recorded in October, November, and December. An insignificant coefficient was generated for observations recording during the summer, indicating that ridership during these months did not differ greatly from ridership figures observed in the fall. The positive sign for winter and spring indicates that ridership was significantly larger during the first half of any given year.

6.2. Midday Off-Peak Model

After adjusting for multi-collinearity, eliminating insignificant variables, and controlling for heteroscedasticity and spatial correlation, the restricted midday off-peak model takes the following form:

$$\begin{aligned}(RidershipMiddayOffPeak)_{it} &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} \\ &+ \beta(FeederBusConnectionQuality)_{it} + \beta(PopulationDensity)_{it} \\ &+ \beta(Gender - Female)_{it} + \beta(UnemploymentRate)_{it} + \beta(Income)_{it} \\ &+ \beta(FuelPrice)_{it} + \beta(VehicleOwnership)_{it} \\ &+ \beta(DistancetoCBD - Near)_{it} + \beta(Winter)_{it} + \beta(Spring)_{it} \\ &+ \beta(Summer)_{it} + a_1 + u_{it} + \varepsilon_{it}\end{aligned} \qquad \textit{Eq. 28}$$

Table 16 highlights that all variables except unemployment rate, feeder bus connection quality, and fuel price significantly explained ridership, as p-values > 0.001 were found. Additional factors including unemployment rate, feeder bus connection quality, and fuel price were also found to be correlated with demand, as p-values less than the a priori criterion were identified. Similar to the a.m. peak model, a variety of internal and external variables were shown to influence the number of boardings during the midday off-peak time period, however the direction and significance of several independent variables shifted.

Service quantity demonstrated a coefficient of 0.571, indicating that a 10% increase in service supply could increase ridership by 5.71%. While still significant, fare price demonstrated a lesser coefficient of -0.244. Unlike the a.m. peak period model, these findings indicate that midday riders are more responsive to service changes compared to fare price increases. Station level ridership was again negatively impacted if the station was located in close proximity to downtown Toronto. Compared to the findings from the a.m. peak period model, a smaller but statistically significant coefficient of -0.882 was identified.

Park and ride capacity was removed during the stepwise regression due to insignificance, while feeder bus connection demonstrated a statistically significant relationship with ridership. A coefficient of -0.085 was estimated, indicating that a 10% increase in transit access time decreases ridership by 0.85%. In contrast to the a.m. peak model, these findings indicate that the quality of feeder bus connections is important to midday off-peak riders, whereas the provision of park and ride infrastructure is not influential.

Population density demonstrated a coefficient of 0.77, suggesting that a 10% increase in the concentration of residents could increase ridership by 7.7%. Again, employment density was removed from the model due to insignificance, so coefficients comparing the impact of additional residents vs. additional jobs were not shown. Household income generated the largest coefficient in the model; an estimate of 1.515 was computed, indicating that a 10% increase in household income could increase ridership by 15%. These results suggest that stations located in affluent areas are more likely to generate riders compared to those located in marginalized neighborhoods. The coefficient for unemployment rate was similar to that generated in the a.m. peak period model, suggesting that rising unemployment figures have a positive impact on transit ridership. Large proportions of female residents had a statistically significant positive impact on ridership, thereby displaying the opposite sign generated by the a.m. peak model. These results illustrate that females are more likely to use transit during the midday off-peak period compared to the a.m. peak period. Again, seasonality had a significant impact on ridership. Demand during winter and spring was significantly less compared to the fall baseline scenario, whereas observations obtained during the summer were correlated with a larger number of riders. These observations conform to the expectation that large amounts of midday trips are generated during summer months due to vacations, leisure trips, and other tourism related travel patterns. Fuel price generated a coefficient similar to that generated in the a.m. peak model, as an elasticity of 5.32% was found. Notably, the findings with respect to vehicle ownership suggest that a 10% increase in the number of vehicles within the study area could decrease ridership by 6.9%. These findings are unlike those identified in the a.m. peak model and suggests that private automobile competes with, rather than compliments rail demand outside of the a.m. peak period.

6.3. P.M. Peak Model

After adjusting for multi-collinearity, eliminating insignificant variables, and controlling for heteroscedasticity and spatial correlation, the restricted p.m. peak model takes the following form:

$$\begin{aligned}(RidershipPMPeak)_{it} &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} \\ &+ \beta(FeederBusConnectionQuality)_{it} + \beta(PopulationDensity)_{it} \\ &+ \beta(Gender - Female)_{it} + \beta(UnemploymentRate)_{it} + \beta(Income)_{it} \\ &+ \beta(EmploymentDensity)_{it} + \beta(FuelPrice)_{it} \\ &+ \beta(VehicleOwnership)_{it} + \beta(DistancetoCBD - Near)_{it} \\ &+ \beta(Winter)_{it} + \beta(Spring)_{it} + \beta(Summer)_{it} + a_1 + u_{it} + \varepsilon_{it}\end{aligned} \qquad \textit{Eq. 29}$$

The results summarized in Table 17 illustrate that demand during the p.m. peak period was most sensitive to service quantity, fare price, unemployment rate, household income, and seasonal factors. Fare price and fuel price were also shown to be statistically significant as p-values < 0.01 were associated with these factors. While not as influential, employment density, population density, vehicle ownership, and distance to CBD displayed p-values less than the 0.1 significance threshold, thus warranting inclusion in the final model.

Service quantity demonstrated the expected sign and produced a coefficient of 0.562, indicating that a 10% increase in service could result increase ridership by 5.62%. Fare price displayed the opposite sign, indicating that riders are willing to pay for service when offered. These findings, coupled with those summarized above, indicate that a.m. peak users are more sensitive to the monetary cost associated with transit use compared to other riders. Distance to CBD was further shown to influence ridership, as stations located close to downtown Toronto generated more ridership compared to those located in suburban environments. This was expected as employment centers and districts are disproportionally concentrated within the city of Toronto compared to the rest of the study area. Employment density further generated a statistically significant relationship with demand, therefore validating these findings.

Park and ride capacity was eliminated during the stepwise regression procedure due to insignificance, while feeder bus connection quality tested significant; however, the opposite sign was displayed. These results indicate that boardings during the p.m. peak period could be concentrated at stations associated with poor quality feeder bus connection, or that transit access quality has worsened while ridership has grown over the time-series. Ridership during this time period might be better explained by station access infrastructure that is more abundant in the downtown core, such as walking connections or cycling infrastructure, as ridership is concentrated at these stations.

Population density demonstrated a coefficient of 0.307, indicating that 10% increase in the concentration of residents could increase demand by 3.07%. Comparatively, an elasticity of 3.42% was found for employment density. These findings indicate that ridership during the p.m. peak period is more dependent on the presence of employment centers and business parks, although this finding was expected as p.m. peak ridership is mainly dominated by trips that originate from a person's place of employment.

Household income and unemployment rate generated the two largest coefficients in the model as elasticities of 15.06% and 27.03% were found. The findings with respect to household income are consistent with those identified by the midday off-peak model, while the impact of increasing unemployment is much larger compared to the other time periods analyzed. Again, fuel price demonstrated a significant relationship with ridership, as an elasticity of 1.76% was identified. Ridership was negatively impacted by large levels of vehicle ownership, further indicating that the availability of alternative transport modes impacts demand outside of the a.m. peak period. Using fall as the base scenario, ridership was again impacted by seasonality. Demand during winter months was significantly less, while spring and summer seasons generated an increase in boardings. These results conform with the expectation that holidays and weather have a negative impact on ridership during cooler time periods.

6.4. Evening Off-Peak Model

After adjusting for multi-collinearity, eliminating insignificant variables, and controlling for heteroscedasticity and spatial correlation, the restricted evening off-peak model takes the following form:

$$\begin{aligned}(RidershipEveningOffPeak)_{it} \\ &= \beta(ServiceQuantity)_{it} + \beta(FarePrice)_{it} + \beta(PopulationDensity)_{it} \\ &+ \beta(UnemploymentRate)_{it} + \beta(Income)_{it} \\ &+ \beta(EmploymentDensity)_{it} + \beta(FuelPrice)_{it} \\ &+ \beta(VehicleOwnership)_{it} + \beta(DistancetoCBD - Near)_{it} \\ &+ \beta(Winter)_{it} + \beta(Spring)_{it} + \beta(Summer)_{it} + \ a_1 + u_{it} + \varepsilon_{it}\end{aligned} \qquad \textit{Eq. 30}$$

In this model, a p-value < 0.001 was identified with respect to a variety of factors, therefore indicating statistical significance. Employment density, fuel price, and population density were shown to have a less significant impact on demand, but p-values less than the predetermined cut-off value were still identified.

Service quantity produced a coefficient of 0.493, indicating that a 10% increase in service supply could increase ridership by 4.93%. Consistent with findings from the p.m. peak model, fare price demonstrated a significant positive correlation with ridership. Again, these findings indicate that riders who board during the latter half of the day are willing to pay for service. Distance to CBD played a significant role in explaining ridership, as stations located close to downtown Toronto generated a greater number of riders. This could be a result of concerts, gatherings, or sporting events that take place within Toronto, therefore resulting in more discretionary trips.

Park and ride capacity was excluded from the model during the stepwise regression procedure, indicating that park and ride infrastructure is an insignificant station amenity for evening riders. Feeder bus connection quality was also eliminated from the model due to insignificance. These findings are most likely a result of parking spaces being available during the evening off-peak period, thus reducing the importance of feeder bus connections and overall parking supply.

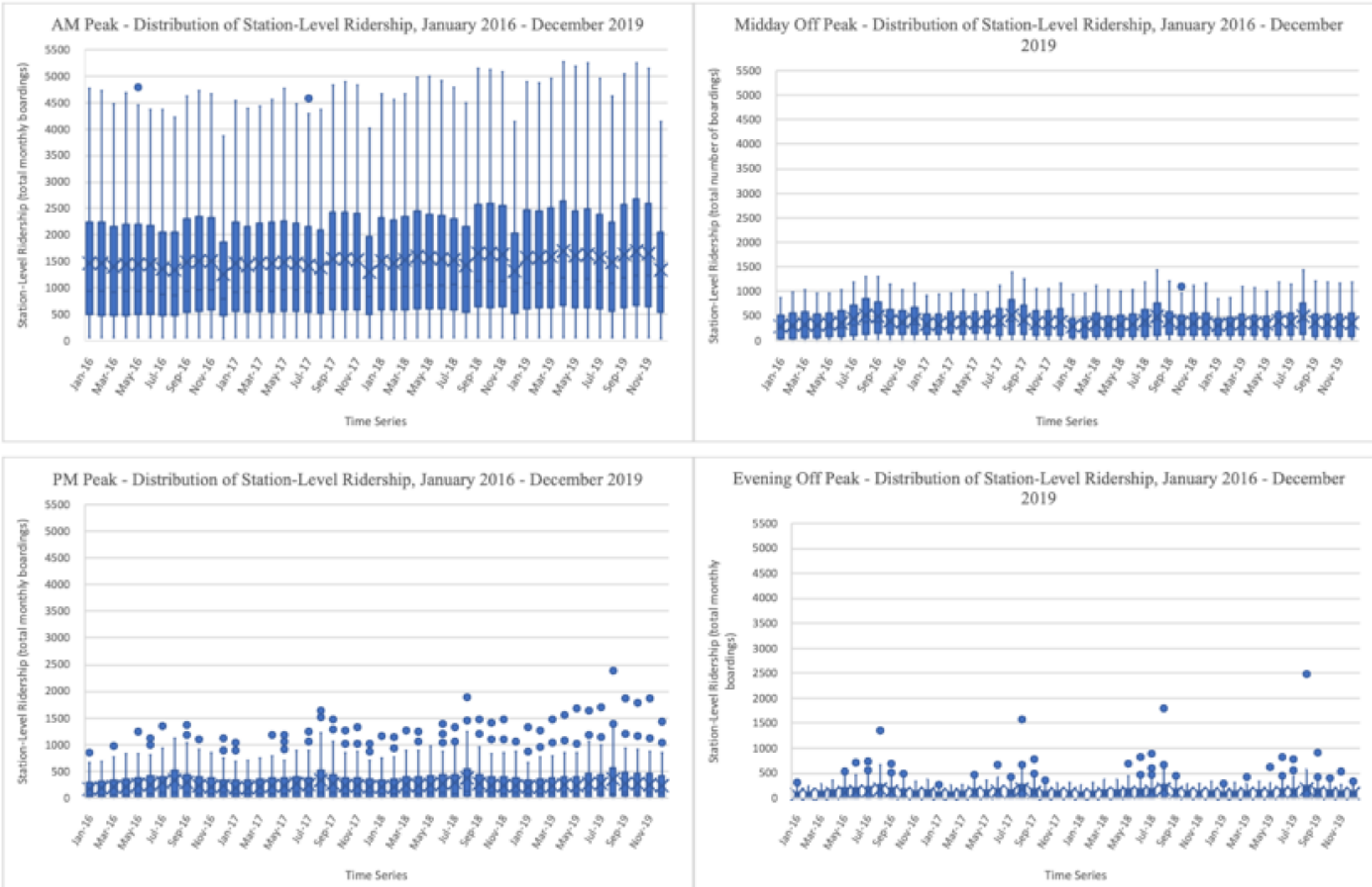
Consistent with findings from the p.m. peak model, population and employment density generated comparable coefficients, as elasticities of 4.1% and 4.03% were identified. This is most likely attributed to trips generated by the presence of service and retail industries, both of which are open later than typical white-collar businesses. Household income and unemployment rate were also found to generate the largest coefficients in the dataset, as elasticities of 23.88% and 28.24% were estimated. Again, these findings suggest that both wealthy and marginalized residents are likely use regional rail services either due to increased travel time valuations or as a means of reducing overall transport costs. Ridership increased in the presence of growing fuel prices, as an elasticity of 0.197 was found. Vehicle ownership also demonstrated a negative relationship with ridership, consistent with findings from midday off-peak and p.m. peak models. Seasonality had a significant impact on ridership, as ridership during winter months was significantly lower compared to the fall baseline. Alternatively, ridership increased significantly during spring and summer months, therefore illustrating that tourism and leisure-based trips are a valid source of ridership outside of the a.m. peak period.

6.5. Model Fit

All models were statistically significant in explaining the observed ridership figures. Wald's Chi-squared tests of overall significance returned significant results for each restricted model, indicating that the dependent variable was significantly affected by the explanatory variables included in each analysis. Notably, the r-squared test statistic found that the explanatory capacity of some models was greater than others. Wooldridge (2013) notes that the r-squared value is a number that summarizes how well the computed regression line fits the data. Essentially, it explains the fraction or percentage of variation in the dependent variable that is explained by the factors included in the analysis. Previous studies have revealed that transport behaviour during the a.m. peak period is predictable and consistent, as the majority of demand generated during this time period is a function of home-based work trips. Therefore, it was theorized that the r-squared value associated with the a.m. peak period model would be the largest relative to the other trip types analyzed. However, the model representing a.m. peak ridership demonstrated an r-squared value of 0.157, indicating that only 15.7% of variation in ridership during this time period was explained by the factors included in the model. In contrast, an r-squared value of 0.61 was shown for the midday off-peak period model, while the r-squared value of both the p.m. peak and evening off-peak models was greater than 0.4.

Descriptive analysis suggests that the spread and range of station-level ridership during the a.m. peak period was a contributing factor. Box and whisker plots presented in Figure 12 illustrate that the distribution of station-level ridership varied greatly during the a.m. peak period compared to the other trip types analyzed. Since datasets with more variation are harder to explain, the lower r-squared value as shown for the a.m. peak ridership model is plausible. The results suggest that ridership, specifically during the a.m. peak period, is heavily influenced by explanatory factors not included in the analysis.

Figure 12 - *Distribution of Station-Level Ridership by Trip Type, December 2016 - January 2019*



Regardless, Miller et al. (2017) states that this should be expected when station-level ridership is analyzed. A variety of variables, such as neighbourhood gentrification, the opening / closing of local businesses, construction, and special events such as festivals, concerts, and success of sports teams can randomly influence station-level ridership. However, these occurrences are too descriptive to be included in demand studies of this scale. Therefore, the r-squared values identified in all models is within an acceptable range as suggested by the author, and is comparable to previous transit demand studies that used linear panel data estimators in their analysis (C. Miller & Savage, 2017; Stover & Christine Bae, 2011).

6.6. Chapter Summary

This chapter summarized results generated from regression models that analyzed GO Transit rail demand between January 2016 and December 2019. Separate models with respect to a.m. peak, midday off-peak, p.m. peak, and evening off-peak ridership were presented. A variety of internal, external, and station accessibility indicators demonstrated a statistically significant relationship with demand. Further, post regression test statistics revealed that these models efficiently explained demand during the time period analyzed.

Demand was shown to respond differently to a variety of factors, including fare price, vehicle ownership, and seasonality depending on trip type examined. Further, demand was unresponsive to station accessibility indicators outside of the a.m. peak period. However, this analysis found that several independent variables including service quantity, population density, fuel price, unemployment rate, and distance to CBD are significant in explaining ridership regardless of trip type examined. This suggests that policies targeted toward these factors will effectively stimulate ridership demand. Therefore, policies in addition to the service quantity improvements proposed in the GO Expansion Program should be explored to further improve ridership and stimulate mode shift throughout the Greater Golden Horseshoe. Policy and practice recommendations that could be considered as a result of these findings are explored in the following chapter.

7. Discussion

The purpose of this thesis was to identify variables associated with station-level ridership throughout the GO Transit rail network. 12 independent variable datasets were used to develop statistically viable regression models that analyzed transit demand across a 48-month time-series. Results from these models were highlighted in the previous chapter.

This thesis also sought to understand how these relationships might differ depending on the trip type examined. Separate regression models which analyzed demand during the a.m. peak, midday off-peak, p.m. peak, and evening off-peak time periods were therefore formulated. The results indicate that some factors are more influential in explaining station-level ridership in some time periods compared to others. The following section contrasts our results with previous studies and theorizes how the identified relationships may have been realized.

This chapter further explores these differences to formulate practice and policy recommendations that could be implemented to further stimulate demand throughout the GO Transit network. Discussion is directed towards variables that were found to be influential across all trip types, as it would be most effective for transit agencies to implement policies that stimulate demand throughout the day. Based on our findings, it is recommended that planners in the Greater Golden Horseshoe focus on policies related to the density and diversity of the built environment, the cost of vehicle operation, and the price of short-distance trips within the city of Toronto to stimulate demand. Desktop research revealed that even in the absence of this study, planners throughout the GGH have suggested that policies related to these factors should be implemented to increase ridership. However, implementation barriers and a lack of political commitment have prevented this from occurring. Therefore, recommendations are also provided which could be applied to overcome these barriers.

This chapter concludes by outlining limitations that were realized during the research process, and states future research directions that could pursued by transit researchers.

7.1. Consistent Findings Between Models

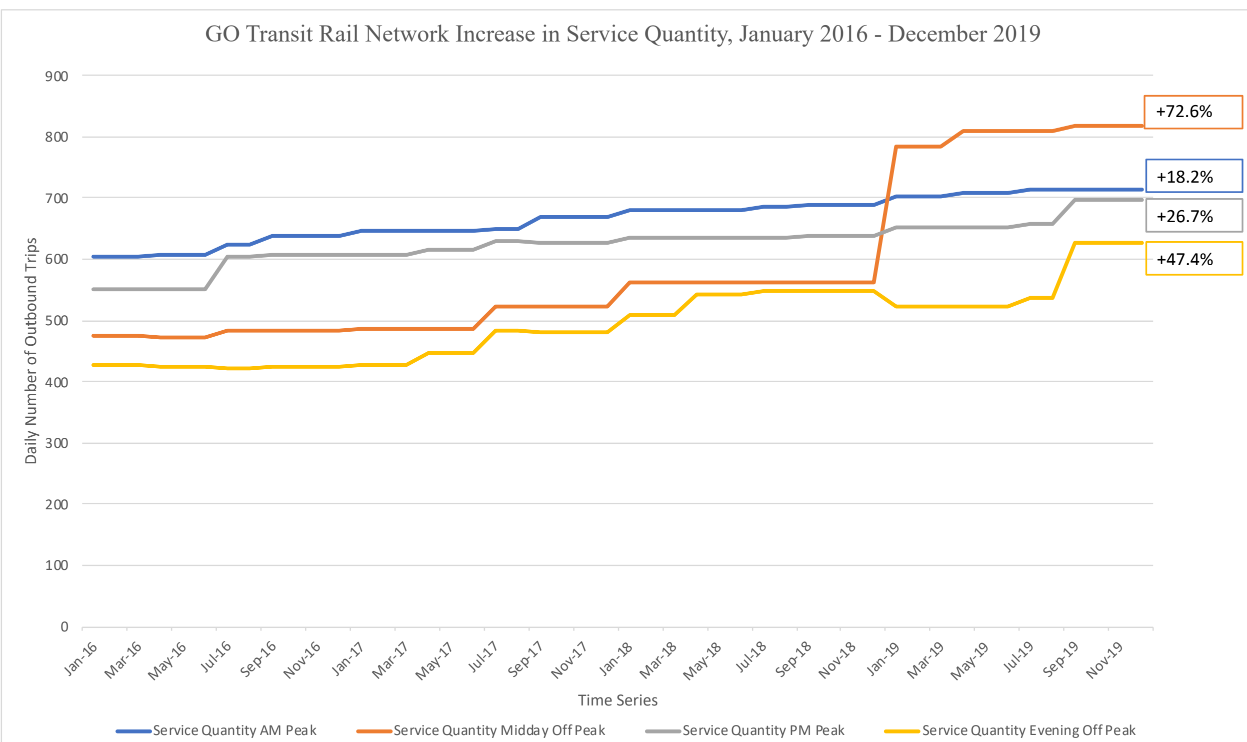
The results illustrate that GO Transit rail ridership is impacted by a variety of variables, but that the sign, significance, and magnitude of influence can differ depending on the type of trip examined. However, the sign and significance of service quantity, average price, of fuel, unemployment rate, and population density remained consistent between models.

7.1.1. Service Quantity

Elasticity estimates for service quantity remained positive regardless of trip type examined, as elasticities ranging from 0.207 to 0.571 were calculated. Consistent with previous findings, the results support the notion that increasing the amount of commuter rail trips will increase demand regardless of trip type or purpose (Balcombe et al., 2004; Guerra & Cervero, 2011; Nowak & Savage, 2013; Taylor et al., 2009). Furthermore, three of the four models found that demand was more responsive to service quantity than fare price. This suggests that customers in the GGH are not overly sensitive to monetary costs associated with transit use, but instead are significantly influenced by the increase in utility generated by decreased headways and more convenient trip options. These findings indicate that in service supply, compared to fare price reductions, should have a more significant impact on increasing all day ridership (Balcombe et al., 2004; Kohn, 2000; Taylor et al., 2009).

Notably, our results found that demand during the a.m. peak period was less responsive to service quantity changes compared to other trip types examined. This could be explained by extensive baseline service that was already offered during the a.m. peak period, as previous studies have shown that consumer response can be limited if little benefit is realized by the addition of an extra trip (C. Chen et al., 2011; R. Liu, 2018; Wasserman, 2019; Yanmaz-Tuzel & Ozbay, 2010). As shown in Figure 13, a lesser elasticity during the a.m. peak period was also expected as the number of trips during this time period increased at a lower rate compared to the other trip types analyzed. These findings suggest that demand should increase as service quantity continues to expand, but that marginal gains should be expected as customers become climatized to adequate service quantity levels.

Figure 13 - *GO Transit Rail Network Service Quality Increases, January 2016 - December 2019*



7.1.2. Population Density

Population density demonstrated a positive relationship with ridership in all models estimated, as demand elasticities ranging from 0.284 to 0.77 were identified. Consistent with previous studies, the results suggest that rail demand can be increased significantly when dense residential developments are constructed within close proximity to the network (Boisjoly et al., 2018; Brown et al., 2014; Durning & Townsend, 2015; Guerra & Cervero, 2011). Further, the consistency of the relationship across models indicates that policies aimed at intensifying residential areas around GO Transit rail stations is an effective method of influencing mode share and transit use for all trip types.

While a statistically significant relationship was continuously noted, elasticity values differed between model outputs. Previous studies have found that population density has a strong influence on transit demand during the a.m. peak period, as the majority of travel during this time period consists of home-based work trips. Therefore, it was expected that ridership during the a.m. peak period would be elastic to increased population densities relative to other models estimated..

Instead, an elasticity of 0.284 was identified, nearly half of that identified by the midday off-peak model. As noted by Miller & Savage (2017), large population densities are expected to stimulate demand in off-peak periods, as the abundance of recreational and social activities can draw people into these areas from other regional locations. Our findings suggest that discretionary demand is more sensitive to increased population densities instead of commuter related trips.

This value could have been minimized by the fact that key trip generators during the a.m. peak period are located in low density suburban environments. For example, stations associated with large ridership figures such as Oakville, Clarkson, and Whitby GO Stations are surrounded by urban areas that consist of sprawled, single family homes. In contrast, stations located in dense urban environments, such as Exhibition, Bloor, and Danforth GO Stations struggle to generate ridership during this time period. The results suggest that ridership could be further increased if policies are implemented that incentivize GO Transit rail use for inter-city travel.

Various articles have suggested that even if large population densities are realized, the presence of children at the household level significantly impacts ridership as more complex trip chains are required (Currie & Delbosc, 2011; Rahman et al., 2019). Therefore, transit use is reduced in favour of private automobile due to the convenience and flexibility associated with this mode. The impact of this factor on ridership could not be assessed, as households with children was eliminated from all models during the stepwise regression process as a significant correlation with population density was detected. For future research, a metric different than the one specified for this study should be used to record the presence of children at the household level, to see if elasticities with respect to population density change once the presence of children is accounted for.

7.1.3. Fuel Price

The price of fuel significantly explained ridership demand, as relatively high demand elasticities were estimated across all models. The significance of these results indicate that push techniques aimed at increasing the disutility of private automobile use is a viable method to encourage mode shift and increase ridership demand (C. Chen et al., 2011; Lane, 2010; Maley & Weinberger, 2009; Taylor et al., 2009). Strategies aimed at increasing the cost of parking and expressway use could play a role, as these costs could be strategically implemented to increase the cost of inter-regional transport without negatively impacting local or work-related transport patterns where private automobile use is necessary. However, collaboration and cooperation from various levels of government could prove difficult, as the optics associated with such policies decreases political willingness regardless of the potential benefits that could be realized (M. J. Bianco et al., 1997).

7.1.4. Unemployment Rate

Surprisingly, unemployment rate demonstrated a relationship commonly seen in studies where local transit ridership, rather than regional transit demand is analyzed. For example, various studies have suggested that local transit ridership increases when the unemployment rate rises, as choice riders transition to public transit due to cost concerns. However, the opposite has been shown for commuter rail ridership as a reduction in the amount of traffic between residential and employment locations is expected to occur when a downturn in economic activity is realized (Iseki & Ali, 2014; Nowak & Savage, 2013; Stover & Christine Bae, 2011). Most likely, the sign and significance of this relationship was skewed as extrapolated values, rather than real values, were used to obtain monthly unemployment rate estimates throughout the study area. Previous demand studies have shown that linear extrapolation is commonly used to project and obtain socioeconomic and demographic values during intercensal periods (Chiang et al., 2011; Lee & Lee, 2013), but this method fails to capture local industrial and economic trends that have a significant impact on unemployment rates (Weden et al., 2015). Therefore, it is likely that the unemployment rate dataset used for this study was misspecified. It is recommended that results in relation to this variable be interpreted with caution, as the use of real household employment statistics are needed to further explain the relationship between unemployment rate and transit demand.

7.2. Differing Findings Between Models

Model comparison revealed that a variety of internal and external variables significantly explained transit demand for some trip types but not others. Of the variables assessed, the sign, significance, and coefficient associated with fare price, seasonality, employment density, and vehicle ownership differed between models.

7.2.1. Fare Price

Fare price demonstrated a negative correlation with ridership demand, but only during the a.m. peak and midday off-peak periods. These findings are unlike those identified in the literature, who found that larger fare prices should discourage transit demand, regardless of trip purpose (Boisjoly et al., 2018; C. Chen et al., 2011; Diab et al., 2020; Guerra & Cervero, 2011; Iseki & Ali, 2014; Schimek, 2015). Furthermore, previous studies have theorized that peak period users are less sensitive to fare price changes, as the utility generated by employment offsets the cost of transit use (Guerra & Cervero, 2011). Several factors could explain why differing relationships were identified in this analysis.

First, previous studies have found that commuter rail demand responds differently to fare price changes compared to other modes (Brown et al., 2014; S. H. Chen & Zegras, 2016; Durning & Townsend, 2015; R. Liu, 2018; Rahman et al., 2019; Stover & Christine Bae, 2011). Kohn (2000) theorized that many vehicle owners chose to use commuter rail as high vehicle operation and parking costs experienced at employment destinations is much larger than fare prices charged by commuter rail agencies. Therefore, additional transport costs that are experienced as a result of a fare price increase is marginal compared to increased costs that would be realized if the rider chose to use private automobile for the entirety of their trip. This can result in fare price elasticities that are low relative to other transit modes. While not directly measured, high parking costs or a lack of parking in the study area's CBD could explain the fare price elasticities identified. The results suggest that sustained demand should be expected until fare prices are comparable to parking costs experienced in the city of Toronto.

Secondly, drastic service quantity increases occurred throughout the time-series while only marginal fare price changes were experienced. Further to findings identified by Stover et al. (2011), fare price elasticities can be understated if few and/or marginal fare price changes occur during the study period. These findings, coupled with the drastic level of service changes that were realized, could have resulted in the underestimation of demand elasticities with respect to fare price for the models estimated.

Third, it is important to note that fare price changes that occurred during the time-series were not uniform in direction. For example, a memo released by Metrolinx stated that fare prices were increased for long distance trips, while those 10 kilometers or less in length were reduced to a flat rate of $3.70 (Woo & Childs, 2019). As a result, the relatively stable elasticities identified outside of the a.m. peak time period suggest that long distance riders lost as a result of the fare increase were balanced out by additional riders engaging in more short-distance trips. Per Farber et al (2014), these findings should be interpreted with caution by the transit agency being reviewed, as declines in average trip length could decrease revenues and encourage the use of unsustainable modes when completing long distance trips. Further research could be conducted to see if average trip length did decrease after the fare adjustment took place, to indicate if rider behaviour did change significantly with respect to fare price.

7.2.2. Seasonality

Seasonality was shown to have a significant impact on demand in all models. Typically, observations that occurred during the winter demonstrated a negative correlation with demand, while ridership increased in summer months. This was expected as previous studies have noted that rail ridership is heavily impacted by seasonal effects for systems located in temporal climates (C. Chen et al., 2011; Lane, 2010). As summarized in Section 3.2.8, inadequate station infrastructure could result in decreased ridership during winter months, as this creates a disutility of having to wait outside to bear cold temperatures and precipitation events (Singhal et al., 2014). His study found that rail users whose origin station was elevated were particularly elastic to the presence of snow and rain, and that cold temperatures also impacted ridership figures (Singhal et al., 2014). He further states that internal characteristics, such as service frequency and total trip time, can reduce the impacts caused by climatic related events, and station upgrades can be implemented to otherwise deter seasonal fluctuations.

Notably, the majority of stations throughout the GO Transit network use elevated platforms, meaning that passengers need to stand outside when boarding and alighting vehicles. Additionally, the majority of stations currently in use by GO Transit were constructed in the 20th century and were not designed to satisfy high volume passenger flows. Further, the largest increase in ridership during the time series was seen at stations where new station infrastructure has been implemented, such as Oakville and Clarkson GO Stations. Therefore, inadequate station infrastructure could be a driver of decreased winter ridership, as the disutility of being exposed to cold weather and precipitation events may encourage choice riders to use private automobile rather than transit. Per Singhal et al. (2014), changes to station infrastructure, such as the number of heated shelters, the amount of indoor seating, and reduced headways during winter months, are all internal changes that could be implemented to decrease the amount of disutility associated with system use during winter months. Investigating the implementation of covered or enclosed stations, as seen in Europe, could further deter impacts.

An exception to this was ridership during the a.m. peak period, as relatively small coefficients suggest that morning commuters are not overly sensitive to seasonal effects. However, this was anticipated as more utility is generated by trips that occur during the a.m. peak period, which therefore result in greater consistency and less seasonal deviation in demand compared to the other time periods examined. Regardless, upgrades to station infrastructure could still be considered as a greater level of customer satisfaction could be realized by these improvements.

7.2.3. Employment Density

Employment density demonstrated a positive correlation with transit demand, but only during the p.m. peak and evening off-peak time periods. As identified in the literature, these findings conform to the expectation that the majority of trips during the later half of the day are work-based trips that consist of commuters returning to residential areas (S. H. Chen & Zegras, 2016; Wasserman, 2019). Notably, population density also demonstrated a significant correlation with transit demand during these time periods, suggesting that stations surrounded by a variety of land uses are associated with larger ridership figures compared to those surrounded by uniform uses. This was expected as the presence of retail stores, commercial outlets, and eating establishments, which are often located in mixed-use areas, generate increased amounts of discretionary traffic due to late operating hours and recreational offerings (S. H. Chen & Zegras, 2016; Durning & Townsend, 2015; Rahman et al., 2019; Zhang & Wang, 2014). The results of this study indicate that increasing the diversity of land uses surrounding GO Transit rail stations could be effective in increasing transit demand during p.m. peak and evening off-peak hours.

7.2.4. Station Location

Station location was significantly associated with ridership, although the sign of this relationship differed between models. Stations near the study areas CBD were associated with fewer boardings during the a.m. and midday off-peak periods, while the opposite relationship was identified in the p.m. and evening off-peak models. The results indicate that trips undertaken during the first half of the day are concentrated in suburban areas, while those in the p.m. and evening off-peak are concentrated within the city of Toronto. This was expected, as the majority of GO Transit rail users are those who reside in suburban locations and commute to downtown Toronto, as employment opportunities are concentrated in this area.

Regardless, the fact that boardings were disproportionally concentrated at suburban and inner-city stations during peak periods suggests that residents within the study area rarely use GO Transit for trips other than commuter related travel. The negative relationship identified during the a.m. peak period further suggests that residents within the city of Toronto likely use other modes to satisfy transport patterns, despite the fact that stations are more accessible compared to those located in rural or suburban areas. Additional policies could be implemented to increase the utility associated with inner-city trips, therefore increasing the amount of station-level ridership within the city of Toronto. Competitive fare pricing strategies could be explored to ensure that the amount of disutility associated with system use does not exceed that associated with local service providers who operate within the study area.

7.2.5. Household Vehicle Ownership

More vehicles were associated with increased transit demand during the a.m. peak period, while the opposite relationship was identified for remaining trip types examined. Consistent with Balcombe (2004), these findings were expected as the utility generated by travel time savings and increased productivity can influence automobile owners to use the system for part of their journey. Most likely, large amounts of congestion in and around the city of Toronto during the a.m. peak period can further explain this relationship, as increased travel times, unreliability, and reduced productivity increase the utility associated with rail use compared to private automobile.

Regardless, the other models indicate that private automobile availability is a competitive good that negatively impacts ridership outside of the a.m. peak period. As suggested by Grimsrud (2013), this indicates that vehicles might be kept and utilized for discretionary related transport, such as shopping, during off-peak periods. Furthermore, improved traffic conditions are commonly realized throughout the study area once the a.m. peak period concludes, meaning that the disutility otherwise associated with private automobile use is reduced. Much like the results identified in relation to fuel price, these findings suggest that methods which disincentivize automobile use during all time periods could encourage the use of public transit.

7.2.6. Park and Ride Capacity

Park and ride capacity was associated with more boardings, but only during the a.m. peak period. A demand elasticity of 0.041was calculated during this time period, further suggesting that ridership is not overly sensitive to increases in parking capacity. The results indicate that the presence of park and ride facilities have enabled ridership figures that are currently realized, but the expansion of such infrastructure is not needed to encourage additional demand. Consistent with previous studies, this suggests that the provision of alternative station access infrastructure such as pull-up-drop-off circles, cycling connections, and direct pedestrian routes could have a more significant impact on ridership (Engel-Yan et al., 2014; Government of Ontario, 2016).

Previous research has also noted that users are likely to access nearby GO Transit rail stations if parking capacity has been reached at their regular access station (Engel-Yan et al., 2014; Mahmoud et al., 2014). These findings suggest that parking utilization may have a greater impact on station-level demand as this could be a driving factor that influences station choice and associated station-level demand. Further research could measure how transport behaviour responds to variation in parking utilization, rather than parking capacity, to see if ridership is elastic to change in this variable.

7.2.7. Feeder Bus Connection Quality

Much like parking capacity, feeder bus connection quality does little to explain station-level ridership throughout the study area. This factor was found to have a significant impact on ridership during the midday off-peak and p.m. peak periods, but the expected sign was only displayed during the midday off-peak period. Unlike the literature reviewed, a marginal coefficient of -0.085 suggests that demand is relatively unaffected when poor quality feeder bus connections are provided (S. H. Chen & Zegras, 2016; Durning & Townsend, 2015; R. Liu, 2018; Rahman et al., 2019; Wasserman, 2019).

The results could be explained by the provision of indirect routes, large headways, or dispersed service coverage that is common of local transit providers that operate in North American suburban centers (Alshalalfah & Shalaby, 2012; C. Chen & McKnight, 2007). Further, the access times computed were shown to be extremely uncompetitive with other access modes such as private auto, as travel times greater than 30 minutes were continuously realized. These results, coupled with those identified in relation to the built environment, suggest that increased land use densities and diversities within station catchment areas is first needed before the impact of feeder bus service on commuter rail ridership can be accurately evaluated.

7.3. Practice and Policy Recommendations

This study is intended not only to determine what variables are most determinantal to GO Transit rail ridership, but to use these findings to further inform the implementation of the GO Expansion Program. The following paragraph outlines recommendations that could be incorporated into the GO Expansion Program to encourage mode shift and stimulate regional transit demand in the study area.

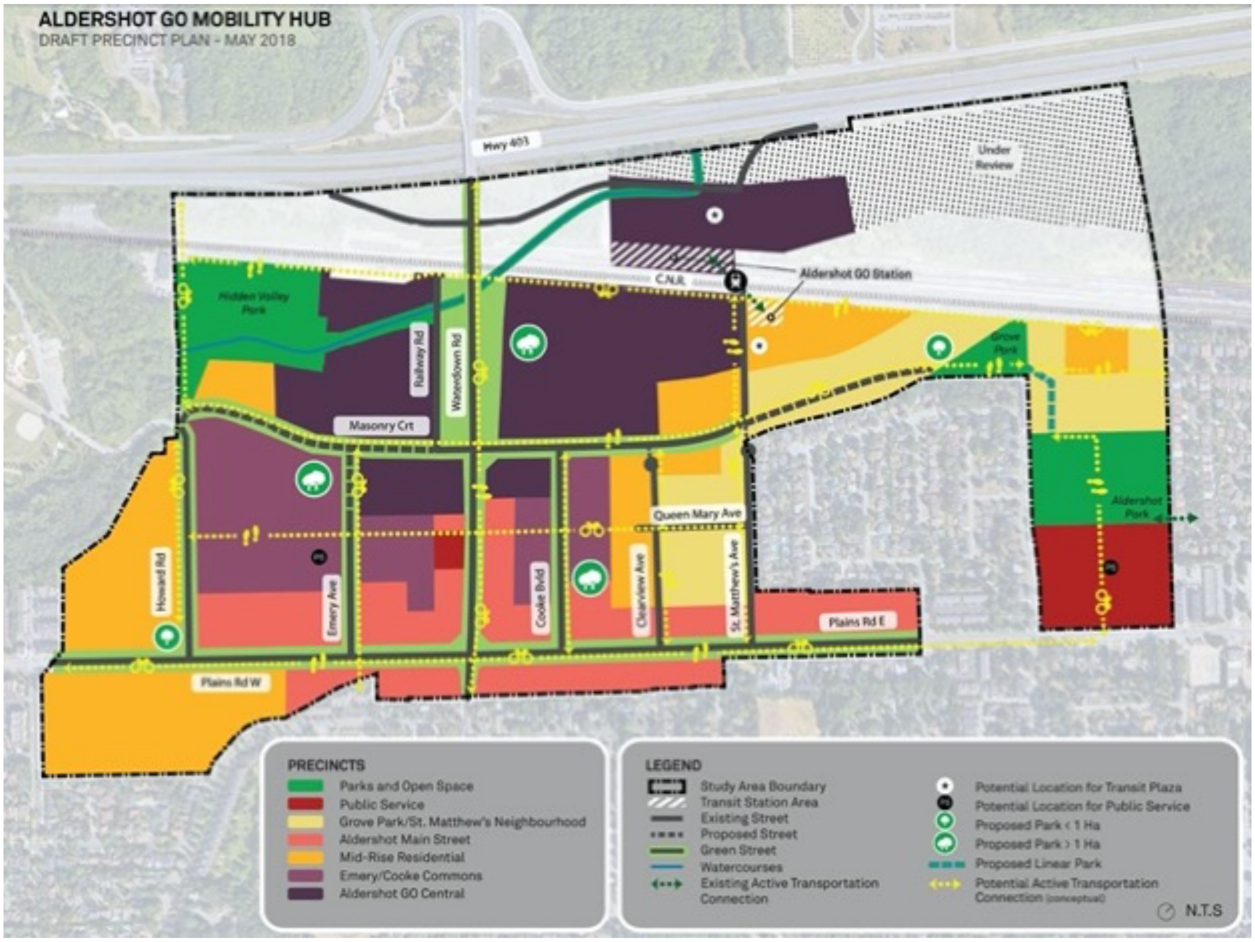
7.3.1. Expand the Role of Mixed-Use Development

The analysis identified that transit demand is positively influenced by large population and employment densities. Therefore, areas that are planned for dense, mixed-use developments should stimulate transit demand in all time periods compared to those where sprawled and singular land uses are permitted. Recognizing this, the GO Expansion Program states that lands surrounding GO Transit rail stations should be planned in a transit-supportive fashion, but a formal development strategy is not outlined. Instead, Metrolinx has stated provincial planning guidelines and municipal planning policies will guide the development of transit-supportive communities within GO Transit station catchment areas. A review of provincial and municipal planning documents was undertaken to see if appropriate planning guidelines and policies have been specified.

As outlined in Section 2.3, the implementation of land use plans and minimum density by-laws in Ontario is a municipal responsibility. However, the province can specify unique land use planning objectives via the Growth Plan to ensure that sustainable growth occurs throughout the GGH. The Growth Plan recognizes the connection between dense, mixed-use spaces and transit ridership, and that all lands within a 500-800 meter radius of GO Transit stations located along priority transit corridors should be designated as a Major Transit Station Area (MTSA) (Government of Ontario, 2020a). Therefore, municipal plans need to specify that these lands should be planned for a variety of land uses, and developments should be able to meet a minimum density target of 150 residents and jobs combined per hectare (Government of Ontario, 2020a). As shown in Figure 16, 37 stations within the study are located on priority transit corridors. Therefore, municipal planning documents currently in-effect for lands surrounding these stations should conform to the development specifications as outlined in the Growth Plan.

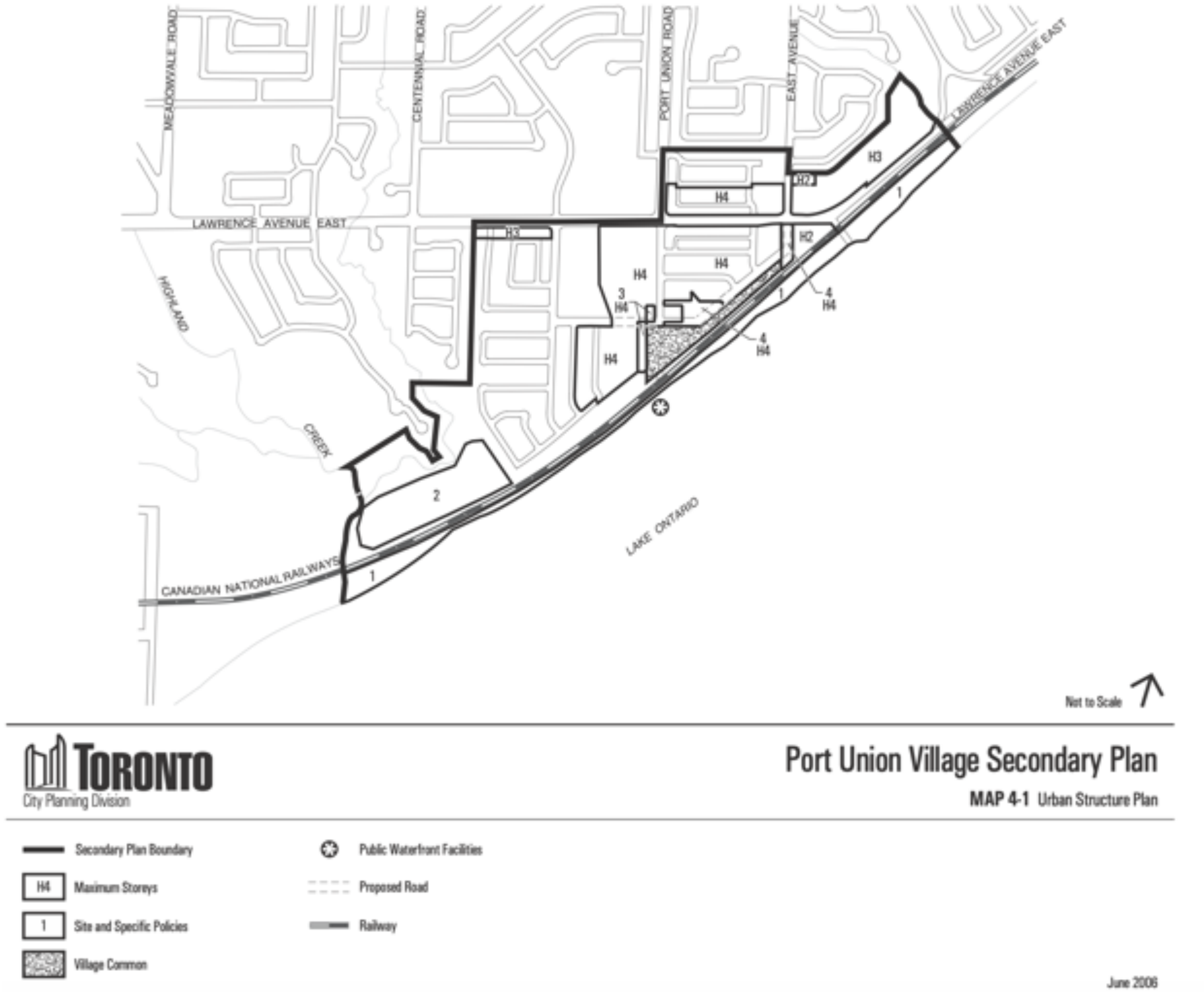
However, a review of current municipal planning documents revealed that appropriate land use policies have not been implemented by the majority of municipalities located throughout the study area. For example, uniform and incompatible land use designations are currently in effect for lands surrounding Aldershot, Appleby, and Burlington GO Stations, despite the drafting of transit-supportive land use plans as illustrated in Figure 14. Further, the Port Union Village Community Secondary Plan, which outlines site specific land use and zoning by-laws for areas surrounding Rouge Hill GO Station in the city of Toronto, mainly contains policies aimed at the preservation of low and medium density residential developments throughout the area (City of Toronto, 2006, 2019; Lintern, 2019). Similar issues were identified when land use plans in the cities of Brampton, Mississauga, and Markham were examined, as low density land use policies are currently in-effect for lands surrounding GO Transit rail stations in these municipalities. These findings suggest that differing municipal priorities, inadequate budgets, and staffing limitations are potential implementation barriers currently being realized by municipal stakeholders that is preventing the implementation of appropriate transit-supportive land use planning guidelines. Regional partnerships should be developed with appropriate municipal stakeholders to aid in the development of appropriate land use plans at the station-level, to ensure that appropriate plans are drafted and implemented as soon as possible. Funding should also be provided to municipal partners to aid in the development of such plans if internal capacity is unavailable.

Figure 14 - *Suggested Land Uses Surrounding Aldershot GO Station*



City of Burlington. (2018). *Aldershot GO Mobility Hub Draft Precinct Plan – May 2018.* City of Burlington. https://www.burlington.ca/en/your-city/resources/Grow-Bold/Mobility-Hubs/Appendix-A---Aldershot-GO-Draft-Precincts.pdf

Figure 15 - *Current Land Use Schedule In-Effect Surrounding Rouge Hill GO Station*



City of Toronto. (2006). *Port Union Village Secondary Plan - Urban Structure Plan*. City of Toronto. https://www.toronto.ca/wp-content/uploads/2017/11/9049-cp-official-plan-SP-4-PortUnionVillage.pdf

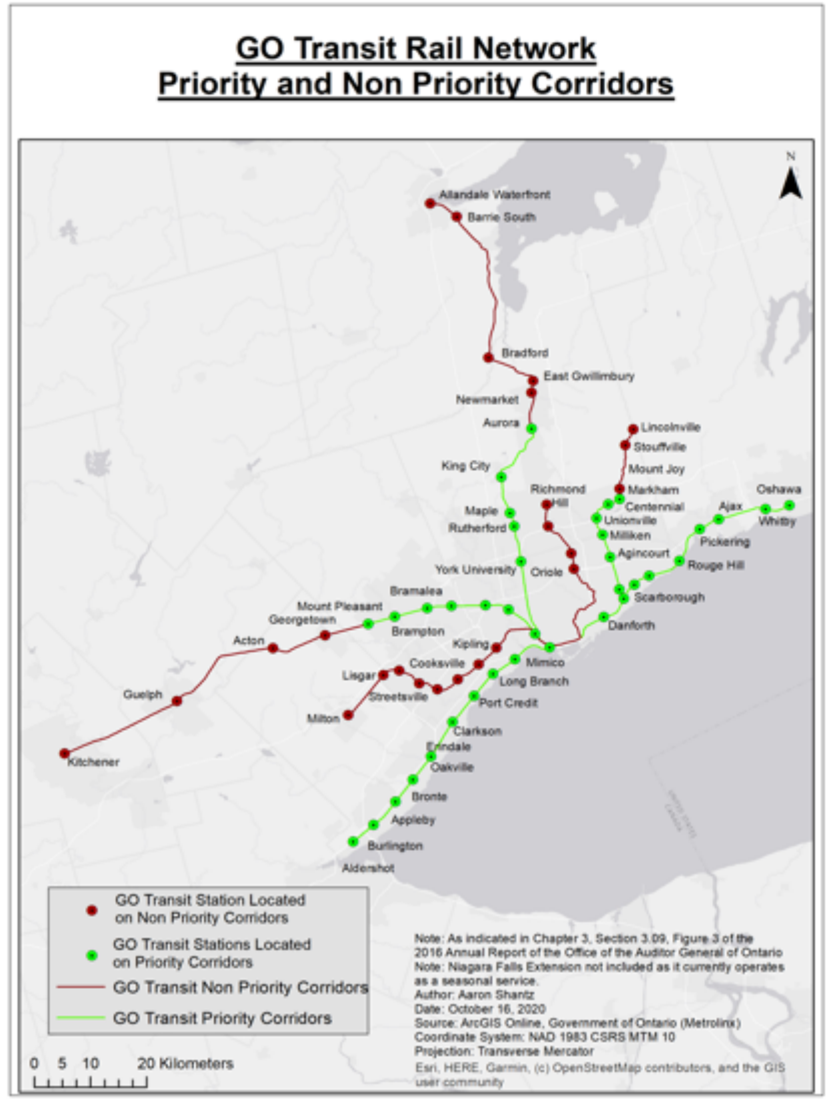
As mentioned above, the Growth Plan specifies that MTSA designations and associated transit-oriented development guidelines are required for lands surrounding stations that are located on priority transit corridors. Therefore, municipalities are not required to implement transit-supportive land use policies for lands surrounding GO Transit rail stations if located on a non-priority transit corridor. The lack of a consistent transit-supportive development framework could be a limiting factor on demand, as 24 of the 61 stations located within the study area are located on non-priority transit corridors. Ridership growth as projected in the GO Expansion Program could be at risk if low density, uniform developments continue to occur throughout the study area.

Additional policy is needed to encourage transit-supportive development at all GO Transit rail stations, regardless of corridor status. Land use policies that build upon the current MTSA framework should be mandated by the province, although local context should be taken into account when specifying minimum density targets and the extent of such policies. For example, the extent of the MTSA designation and associated minimum density requirements could be reduced for stations located in rural environments, such as Georgetown GO Station. A consistent land use framework could ensure that ridership growth is realized at all stations throughout the network, therefore further encouraging mode shift throughout the GGH. If not accomplished, increased ridership figures might not be realized during the early stages of the GO Expansion Program's lifecycle.

7.3.2. Exploration of Integrated Transport Pricing Strategies

This study found that users are sensitive to vehicle operation costs, such as fuel price. These results are consistent with those identified in the literature, and indicate that demand could be stimulated by increasing the price of fuel (C. Chen et al., 2011; Guerra & Cervero, 2011; Maley & Weinberger, 2009; Paulley et al., 2006). Previous studies have further stated that vehicle operation costs include all expenses incurred per vehicle

Figure 16 - *GO Transit Rail Network, Spatial Extent of Priority and Non-Priority Rail Corridors*



kilometer of operation, including fuel, parking costs and toll road charges (Guerra & Cervero, 2011; Voith, 1997). Transit demand could be further stimulated if the disutility associated with automobile use is increased via integrated transport pricing methods. A summary of these strategies, and their applicability to the study area, is outlined below.

A review of transport studies found that the availability and price of parking is the most significant variable that influences the utility of private automobile use (Morrall & Bolger, 1996; Taylor & Fink, 2003). Various pricing strategies including cost, supply, and incentive-based programs are various options that increase the disutility of private automobile use and increase transit demand (M. Bianco et al., 1997). Specific examples include the implementation of a tax on parking space use, increased tax rates on revenues earned by parking providers, the expansion of parking meters and residential permit programs, and the monetization of parking benefits generated choice transit users. Of these, Bianco et al. (1997) found that mode shift occurred most frequently when a tax on parking space use was implemented. Further, a study of pricing strategies in California found that private automobile users were 2.25 times more likely to change their transport behaviour if parking fees were increased, rather than if an incentive of the same amount was provided (Shoup, 1997). Regardless, the authors recommend that the use of mixed methods are most effective as politicians are less likely to implement policies that disproportionally impact a specific segment of the customer base.

Furthermore, several studies have found that pricing mechanisms such as congestion charges, road pricing, and the implementation of tolls can effectively stimulate mode shift and transit demand (Stopher, 2004; U.S. Department of Transportation, 2008). Planners in London, England, found that vehicle use in Central London decreased by 15% once a congestion charge was implemented (U.S. Department of Transportation, 2008). Further, they found that the majority of vehicle users switched to transit modes, meaning that more service could be provided by the transit agency as a result of increased revenues. A study of five major European cities found that once road tolls were implemented, traffic volumes decreased by 1 to 4.5% depending on initial congestion conditions that were experienced (Spears et al., 2012). Comparable results were identified by Arentze et al. (2004) who found that private auto users were likely to switch to public transit if road tolls were implemented. The results suggest that similar pricing mechanisms could be implemented in the GGH to shift private auto users to GO Transit services.

Notably, the potential of road pricing strategies in the GGH has been previously investigated. The Big Move, the original regional transportation plan developed by Metrolinx in 2008, suggested that High Occupancy Vehicle (HOV) lanes, road pricing schemes, and High Occupancy Toll (HOT) lanes should be explored and implemented throughout the Greater Golden Horseshoe. Specifically, a strategy aimed at assessing and implementing an inter-connected regional expressway network with the potential for HOT lanes was outlined (Government of Ontario, 2008). Strategies 3.7 and 3.8 in the 2041 Regional

Transportation Plan reiterated these proposals, and stated that Metrolinx will "continue to explore how mobility pricing (e.g., parking, road pricing, HOT lanes, and off-peak fares) could be used to shift travel behaviour", and that the use of HOT lanes and toll roads could be expended to include some form of charging on all major roads throughout the study area (Government of Ontario, 2018a).

An investment strategy released by Metrolinx in 2013 further stated that various pricing strategies, including a \$0.05 per litre regional fuel tax, a \$0.25 per day commercial parking space tax, and the implementation of a HOV lane system that would charge users \$0.30 cents per kilometer would be implemented throughout the GGH with the purpose of encouraging mode shift, but to also fund future transit investments such as the GO Expansion Program (Government of Ontario, 2013). Interest in transport pricing strategies has also been expressed by municipal stakeholders, as the city of Toronto previously recommended that tolling options including flat and distance-based fares be implemented on major highways throughout the area, such as the Gardiner Expressway and the Don Valley Parkway. Toll prices ranging in value from \$1.25 to \$3.25 for flat fares and \$0.10 to \$0.35 per kilometer for distance-based options were proposed (Livey & Rossini, 2015).

However, the implementation of these policies has been difficult. Despite approval from internal stakeholders, transport pricing strategies have often been halted or overturn by provincial governments. This is likely a result of motorists' low acceptability of transport pricing strategies, resulting in policy decisions that are made in the interest of political vitality rather than in the best interest of the public (Schade & Schlag, 2003). A study of HOT lane implementation in the GGH further identified political willingness as the key barrier to implementation, and noted that the coordination of road, transit, and urban planning, as well as cooperation from different departments and levels of government would be required (Lindsey, 2007). Therefore, despite being identified as an effective solution, the implementation of transport pricing methods in the GGH could prove difficult if political willingness to implement such a proposal remains low.

Various studies have suggested that motorists are more receptive to transport pricing schemes if the costs, benefits, and allocation of funds generated from such programs are openly communicated to the public. Further, those that prioritize revenues towards transit system improvements are more successful compared to those that are directed towards generalized public budgets (Schuitema & Steg, 2008; U.S. Department of Transportation, 2008). Based on the policies reviewed, it is clear that the transport pricing schemes proposed within the GGH have a purpose of reducing automobile reliance and stimulating transit demand, but this information might not be readily communicated to constitutes. It is recommended that Metrolinx continues to propose similar transport pricing strategies such as those previously proposed by the agency, and could further emphasize the connection between revenue allocation and transit improvements so that the benefits of such policies are positively communicated to the public. Furthermore, collaborative efforts could be explored with municipal partners to reinvestigate how the implementation of toll roads and congestion pricing could be further used to disincentivize the use of private automobile for regional transport and increase the demand for regional transit services.

7.3.3. Decrease Fare Price Within the City of Toronto

Finally, our results suggest that location had a significant impact on station-level demand in all time periods examined. Notably, demand during the a.m. peak and midday off-peak periods was concentrated at stations located outside of the city of Toronto. Previous studies have theorized that the price, abundance, and accessibility of local transit service available within major urban areas can explain these trends. A study of transit ridership in Chicago found that public transit users switched from bus to rail modes when bus fares were increased (Nowak & Savage, 2013). A review of New Jersey Transit's commuter rail system further stated that the extent and availability of local public transit in the area is extensive compared to other urban areas throughout North America (C. Chen et al., 2011). Therefore, they found that users were more likely to switch to alternative public transit modes when a fare price increase occurred. Hensher (1997) further explored these relationships and measured how mode-specific ridership is expected to vary when the price of another mode changes. He found that a 10% increase in rail fare prices should increase bus ridership by 0.57%. Balcombe (2004) also suggested that metro ridership is sensitive to the price of alternatives, as a 0.18 demand elasticity with respect to rail fares was found. Therefore, these results illustrate that demand could be negatively impacted if less expensive or more convenient transit options are available within the study area.

While cross-elasticities were not directly measured as part of this study, a review of alternative transit options and associated fare price schemes renders this theory plausible. For example, the Toronto Transit Commission operates extensive bus, light rail transit, and metro service within Toronto. Further, a flat fare price of $3.25 is charged for all trips regardless of distance travelled and number of transfers completed. In comparison, the price of a trip between GO Transit rail stations in the city of Toronto ranges between $3.70 - $11.06, with the average cost being $6.04 (GO Transit, 2021). Therefore, inner-city travellers may prefer to use Toronto Transit Commission services, as the cost of the GO Transit rail system is approximately 86% higher compared to the competitor.

Notably, Metrolinx has noticed that this discrepancy could impact the effectiveness of the GO Expansion Program. A draft memo released by the agency in the spring of 2018 found that current fare prices were uncompetitive with other transit agencies, thereby discouraging the use of GO Transit rail for short distance trips (Woo, 2018). As a result, the memo proposed that all trips between stations within Toronto be charged a flat fare of $3 per trip, and that trips less than 10km in length will be charged a similar price. The report projects that station-level ridership should increase by approximately 10-15% as a result of this change alone.

However, a media scan and use of Metrolinx's interactive fare price calculator revealed that these changes have not been implemented. Notably, several fare price changes have been implemented to reduce the cost of short-distance trips, as trips 10 kilometers or less in distance were reduced to a flat fare rate of $3.70 in 2019. This resulted in a 21% reduction in fare price compared to the previous price of $4.71 (GO Transit, 2021). However, this price is still substantially larger than the flat fare currently charged by the Toronto Transit Commission and other local service operators within the study area. Further, it is clear that the flat fare scheme originally proposed for trips between stations within Toronto has not been implemented, as distance-based fares are still being assigned to GO Transit rail users travelling within Toronto. It is recommended that Metrolinx continue to pursue competitive fare pricing options, including a flat fare policy for travel within the city of Toronto that is similar to the $3.25 flat fare currently charged by the Toronto Transit Commission. The implementation of such policies should increase ridership at stations located within the city of Toronto, therefore balancing network utilization and increasing overall ridership figures.

7.4. Limitations

7.4.1. Processing Capacity

As evidenced in Section 5, extensive work was needed to compile, collect, and estimate various independent variable datasets included in this study. Improved model fit and additional demand elasticities could have been realized if extra time and resources were available to gather and process supplementary datasets. To ensure that this thesis was completed with the time constraints of a two-year master's program, the inclusion of additional factors such as land use mix, education, service quality, and the availability of active transit options was not possible. However, since variables capturing both regional geography, economy, population characteristics, transit system aspects, and those of the surrounding automobile / highway system were included in the analysis, these are arguably the best models that could be specified within the time constraints of the study.

7.4.2. Extrapolation of Socioeconomic and Demographic Datasets

A time-series of January 2016 to December 2019 was selected as accurate and specific ridership data was available for analysis. As mentioned in Section 5.1.3, the majority of socioeconomic and demographic variable datasets were obtained from extrapolated estimates, as the time-series was located within an intercensal period. While extrapolation is commonly used when monthly values need to be estimated using previous observations, error can be introduced to the model as ridership is therefore explained by predicted values rather than real observations. Further, linear extrapolation could have resulted in the inaccurate specification of some external variable datasets, such as unemployment rate, as socioeconomic factors fluctuate heavily based on localized conditions and economic events. These limitations could be negated in future studies if more specific and disaggregated data products are used to formulate socioeconomic and demographic variable datasets.

7.4.3. Discrepancy in Dissemination Area Boundaries

The majority of socioeconomic and demographic variable datasets were extracted at the dissemination area scale using overlay analysis. As noted in Section 5.1.3, the size and extent of some dissemination areas changed between observation periods. This was the case for approximately 20% of the dissemination areas identified within station catchment boundaries, meaning that only a single data point obtained from 2016 Census products were associated with these entities.

As a result, extrapolated values could not be estimated for these entities. This inevitably introduces error into the model, as changing socioeconomic and demographic trends were not captured. However, the effect on model performance was negligible, as the proportion of those included is relatively small compared to those where linear extrapolation methods were used. Similar to the recommendation outlined in the previous section, this effect could be minimized if socioeconomic and demographic datasets were obtained from more specific sources. If monthly figures are estimated from census products in future studies, it is recommended to interpolate monthly estimates between two known census periods, rather than extrapolated using existing observations, to minimize the spatial variability of dissemination areas.

7.5. Directions for Future Research

7.5.1. Further Investigation of Ridership Determinants During A.M. Peak Period

Section 6.6 highlights that 85% of variance in ridership during the a.m. peak period was explained by factors not included in the analysis. Further research could be undertaken to identify additional factors associated with ridership during the a.m. peak period to better understand the behaviour of those engaged in home-based work trips.

7.5.2. Inclusion of Additional Factors

This research analyzed the sensitivity of GO Transit rail demand in relation to a variety of factors. As mentioned above, a subset of independent variables was chosen due to resource constraints and processing capacity. The inclusion of additional variables, specifically those related land use mix and business type, could be included to further understand how transit demand responds as changes to the built environment occur. Additional analysis could also be conducted using updated census products, including those expected to be released in 2021, to increase the reliability of model outputs.

7.5.3. Investigation of Multi-City Analysis

A study conducted by the Canadian Urban Transit Association found that within-city studies, including those that analyze ridership of a single network, are appropriate for answering research questions and formulating policy recommendations specific to a single transit agency or study area (E. J. Miller et al., 2018). However, results are typically non-transferable to other agencies, as variables specific to the regional context often influence model outputs. While this research has identified factors that are heavily associated with commuter rail ridership in southern Ontario, additional research could be completed that utilizes a multi-city approach to further understand the main determinants of commuter rail ridership across varying geographies. Literature reviewed as part of this study indicates that a multi-city study of commuter rail demand has not been undertaken in the North American context, indicating that this gap could be filled by future work. Additional cities that could be considered include Vancouver, New York City, Philadelphia, Boston, and Los Angeles, as commuter rail systems are operational in these areas.

7.5.4. The Role of Station Accessibility Indicators

The results show that transit demand is not overly responsive to station accessibility indicators, including park and ride capacity and the quality of feeder bus connection services. However, a comprehensive subset of station accessibility factors, such as the provision of pedestrian connections, cycling infrastructure, or the availability of pull-up-drop-off areas was not included in the analysis. Therefore, absolute conclusions regarding the role of station access on transit demand should not be reached until more detailed and complete analysis is undertaken. It is recommended that monthly, station-level observations with respect to a wide range of station-accessibility indicators be undertaken so that future demand elasticity models can better understand how first-and-last-mile connections influence station-level transit demand.

8. Conclusion

This study sought to identify factors significantly associated with GO Transit rail ridership. 12 independent variable datasets, including those related to the study area's regional geography, economy, population characteristics, transit system aspects, and the surrounding automobile / highway system were collected across a 48-month time-series. Datasets were separated into a.m. peak, midday off-peak, p.m. peak, and evening off-peak periods to determine if relationships differed depending on the trip type analyzed. A random effect linear panel data estimator was then used to determine the explanatory power of each variable.

Despite challenges encountered in terms of data availability and processing capacity, the results revealed that ridership was sensitive to service quantity, population density, fuel price, and station location regardless of trip type examined. Notably, the sign and significance of fare price, park and ride capacity, feeder bus connection quality, and vehicle ownership differed between models. This study highlights the importance of disaggregating demand elasticity estimates by trip type, as policies targeted towards stimulating demand in all time periods should be prioritized to ensure effective ridership growth.

Findings from these models were used to inform a critical review of the GO Expansion Program and associated land-use and transportation planning policies currently in-effect throughout the Greater Golden Horseshoe. The results suggest that a variety of planning policies, in addition to service quantity improvements outlined in the GO Expansion Program, could be implemented to stimulate station-level demand. Since demand was associated with increased population and employment densities, Metrolinx could collaborate with municipal stakeholders to ensure that planning objectives conform to transit-supportive guidelines as mandated by the province of Ontario. Metrolinx could further lobby for the implementation of a consistent land use planning framework to ensure that sprawled and uniform development does not occur on lands surrounding GO Transit stations. Demand was also associated with fuel price, indicating that users are sensitive to monetary costs associated with automobile use. Toll roads or congestion pricing schemes could be implemented to increase the amount of disutility associated with inter-regional trips, therefore increasing demand. Increased transparency and more effective communication methods could also be explored to increase the level of political willingness needed to implement such methods. Finally, station-level demand within the city of Toronto could be increased if competitive fare price options are offered to inner-city customers.

In conclusion, this study finds that a variety of factors are associated with commuter rail ridership. Therefore, integrated planning policies are needed stimulate mode shift and increase demand for regional transit in the Greater Golden Horseshoe. In the absence of this, negative externalities currently being realized throughout the study area may not be alleviated to the greatest extent possible. This research could be considered by other jurisdictions looking to transition their commuter rail networks to regional rail networks, and should be used to encourage integrated and complementary planning policies when undergoing similar network transitions.

Bibliography

Allen, J. G. (1998). From commuter rail to regional rail: Operating practices for the 21st century. *Transportation Research Record*. https://doi.org/10.3141/1623-17

Allen, J. G., & Levinson, H. S. (2014). Accommodation of long-term growth on North America's commuter railroads. In *Transportation Research Record*. https://doi.org/10.3114/2419-05

Allen, R. Campsie, P. (2013). *Implementing the Growth Plan for the Greater Golden Horseshoe - Has the strategic regional vision been compromised?* https://doi.org/978-0-9739888-6-4

Alshalalfah, B., & Shalaby, A. (2012). Feasibility of Flex-Route as a Feeder Transit Service to Rail Stations in the Suburbs: Case Study in Toronto. *Journal of Urban Planning and Development*. https://doi.org/10.1061/(asce)up.1943-5444.0000096

Andersen, J. L. E., & Landex, A. (2008). Catchment areas for public transport. *WIT Transactions on the Built Environment*. https://doi.org/10.2495/UT080171

Arentze, T., Hofman, F., & Timmermans, H. (2004). Predicting multi-faceted activity-travel adjustment strategies in response to possible congestion pricing scenarios using an Internet-based stated adaptation experiment. *Transport Policy*. https://doi.org/10.1016/S0967-070X(03)00016-7

Ashby, B. (2018). *Transportation Tomorrow Survey 2016: 2016, 2011, 2006, 2001, 1996 and 1986 Travel Summaries for the TTS Area*. http://dmg.utoronto.ca/pdf/tts/2016/2016TTS_Summaries_TTSarea.pdf

Athira, I. C., Muneera, C. P., Krishnamurthy, K., & Anjaneyulu, M. V. L. R. (2016). Estimation of Value of Travel Time for Work Trips. *Transportation Research Procedia*. https://doi.org/10.1016/j.trpro.2016.11.067

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., & White, P. (2004). The demand for public transport: a practical guide. *Transport Policy*.

Ben-Akiva, M., & Morikawa, T. (2002). Comparing ridership attraction of rail and bus. *Transport Policy*. https://doi.org/10.1016/S0967-070X(02)00009-4

Bergman, Å., Gliebe, J., & Strathman, J. (2011). Modeling access mode choice for inter-suburban commuter rail. *Journal of Public Transportation*. https://doi.org/10.5038/2375-0901.14.4.2

Bernal, M., Welch, E. W., & Sriraj, P. S. (2016). The effect of slow zones on ridership: An analysis of the Chicago Transit Authority "El" Blue Line. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2016.02.007

Bianco, M. J., Dueker, K., & Strathman, J. . G. (1997). *Parking Strategies to Attract Auto Users to Public Transportation.* http://pdxscholar.library.pdx.edu/cus_pubs/33?utm_source=pdxscholar.library.pdx.edu%2Fcus_pubs%2F33&utm_medium=PDF&utm_campaign=PDFCoverPages

Bianco, M., Dueker, K., & Strathman, J. (1997). *Parking Strategies to Attract Auto Users to Public Transportation* (No. 33). http://pdxscholar.library.pdx.edu/cus_pubs/33?utm_source=pdxscholar.library.pdx.edu%2Fcus_pubs%2F33&utm_medium=PDF&utm_campaign=PDFCoverPages

Bissonauth, N. (2019). *Travelling back in time: The remarkable evolution of your GO train.* Blog. https://blog.metrolinx.com/2019/03/20/travelling-back-in-time-the-remarkable-evolution-of-your-go-train/

Boisjoly, G., Grisé, E., Maguire, M., Veillette, M. P., Deboosere, R., Berrebi, E., & El-Geneidy, A. (2018). Invest in the ride: A 14 year longitudinal analysis of the determinants of public transport ridership in 25 North American cities. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2018.07.005

Bomberg, M. S., & Kockelman, K. M. (2012). Traveler response to the 2005 gas price spike. In *Household Energy: Economics, Consumption and Efficiency*.

Brown, J., Thompson, G., Bhattacharya, T., & Jaroszynski, M. (2014). Understanding Transit Ridership Demand for the Multidestination, Multimodal Transit Network in Atlanta, Georgia: Lessons for Increasing Rail Transit Choice Ridership while Maintaining Transit Dependent Bus Ridership. *Urban Studies*. https://doi.org/10.1177/0042098013493021

Caroline, H., & Yves, A. (2012). ERRAC ROADMAP WPO3: Urban, Suburban and Regional Rail and Urban Mobility. *Procedia - Social and Behavioral Sciences*. https://doi.org/10.1016/j.sbspro.2012.06.1201

Casello, J., & Hellinga, B. (2008). Impacts of Express Bus Service on Passenger Demand. *Journal of Public Transportation*. https://doi.org/10.5038/2375-0901.11.4.1

Chen, C., & McKnight, C. E. (2007). Does the built environment make a difference? Additional evidence from the daily activity and travel behavior of homemakers living in New York City and suburbs. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2006.11.010

Chen, C., Varley, D., & Chen, J. (2011). What affects transit ridership? a dynamic analysis involving multiple factors, lags and asymmetric behaviour. *Urban Studies*. https://doi.org/10.1177/0042098010379280

Chen, S. H., & Zegras, C. (2016). Rail transit ridership: Station-area analysis of Boston's Massachusetts Bay Transportation Authority. *Transportation Research Record*. https://doi.org/10.3141/2544-13

Chiang, W. C., Russell, R. A., & Urban, T. L. (2011). Forecasting ridership for a metropolitan transit authority. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2011.04.016

Chisholm, G. (2002). Germany's Track-Sharing Experience: Mixed Use of Rail Corridors. In *TCRP Research Results Digest 47*. http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rrd_47.pdf

Chow, L. F., Zhao, F., Liu, X., Li, M. T., & Ubaka, I. (2006). Transit ridership model based on geographically weighted regression. *Transportation Research Record*. https://doi.org/10.3141/1972-15

City of Brampton. (2020). *City of Brampton Official Plan - Schedule A - General Land Use Designations*. https://ago-item-storage.s3.us-east-1.amazonaws.com/bb774054f55e4a3d93c8c277c0cc0c3b/Schedule_A_General_Land_Use_Designations.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEAAaCXVzLWVhc3QtMSJHMEUCIFUQH52JX2PYN4JS3BJp5gDcy4uyF543M3uLW7tmy6iFAiEA132w6bJm8B5w1k1a

City of Brampton. (2020). *City of Brampton Official Plan - Schedule G - Secondary Plan Areas*. https://ago-item-storage.s3.us-east-1.amazonaws.com/6499526334074246825dc6794e27a276/Schedule_G_Secondary_Plan_Areas.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEAAaCXVzLWVhc3QtMSJHMEUCIQD8N7q7OYj2Z7DsKhNKuHE2Fc1EP%2FbWeBbHG2j%2BMV8WZwIgP98%2Fs%2BVTHwXdMgVOp

City of Brampton. (2020). *City of Brampton Official Plan*. https://www.brampton.ca/EN/City-Hall/Official-Plan/Documents/Sept2020_Consolidated_OP_2006.pdf

City of Burlington. (2020). *Mobility Hubs*. https://www.burlington.ca/en/services-for-you/mobility-hubs.asp

City of Burlington. (2019). *City of Burlington Official Plan Part III – Land Use Policies – Urban Planning Area*. https://www.burlington.ca/en/services-for-you/resources/Planning_and_Development/Official_Plan/december-2019-updates-to-OP-docs/Part-III_2019.pdf

City of Burlington. (2019). *City of Burlington Official Plan Schedule B - Comprehensive Land Use Plan Urban Planning Area*. https://www.burlington.ca/en/services-for-you/resources/Planning_and_Development/Official_Plan/december-2019-updates-to-OP-docs/COB_Official-Plan_Schedule-B.pdf

City of Burlington. (2018). *Aldershot GO Mobility Hub Draft Precinct Plan – May 2018.* City of Burlington. https://www.burlington.ca/en/your-city/resources/Grow-Bold/Mobility-Hubs/Appendix-A---Aldershot-GO-Draft-Precincts.pdf

City of Toronto. (2016). *The City of Toronto's Immediate and Longer-term Revenue Strategy Direction*. http://app.toronto.ca/tmmis/viewAgendaItemHistory.do?item=2016.EX20.2

City of Toronto. (2006). *Port Union Village Secondary Plan*. https://www.toronto.ca/wp-content/uploads/2017/11/9049-cp-official-plan-SP-4-PortUnionVillage.pdf

City of Toronto. (2012). *City of Toronto Regional Municipal Boundary*. City of Toronto, Information & Technology. https://open.toronto.ca/dataset/regional-municipal-boundary/

City of Toronto. (2019). *Secondary Plan Key Map*. https://www.toronto.ca/city-government/planning-development/official-plan-guidelines/official-plan/official-plan-maps-copy/

City of Toronto. (2006). *Port Union Village Secondary Plan - Urban Structure Plan*. City of Toronto. https://www.toronto.ca/wp-content/uploads/2017/11/9049-cp-official-plan-SP-4-PortUnionVillage.pdf

Collenette, D. (2016). *High Speed Rail in Ontario – Final Report*. http://www.mto.gov.on.ca/english/publications/high-speed-rail-in-ontario-final-report/pdfs/high-speed-rail-in-ontario-final-report.pdf

Croissant, Y., & Miinno, G. (2008). *Panel Data Econometrics in R: The plm Package* (p. 43). Journal of Statistical Software. https://doi.org/10.18637/jss.v027.i02

Currie, G., & Delbosc, A. (2011). Exploring the trip chaining behaviour of public transport users in Melbourne. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2010.08.003

Currie, G., & Phung, J. (2007). Transit ridership, auto gas prices, and world events new drivers of change? *Transportation Research Record*. https://doi.org/10.3141/1992-01

Curtin, J. F. (1968). Effect of Fares on Transit Riding. *Highway Research Record*.

Diab, E., Kasraian, D., Miller, E. J., & Shalaby, A. (2020). The rise and fall of transit ridership across Canada: Understanding the determinants. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2020.07.002

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., Mcclean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Durning, M., & Townsend, C. (2015). Direct ridership model of rail rapid transit systems in Canada. In *Transportation Research Record*. https://doi.org/10.3141/2537-11

El-Geneidy, A., Grimsrud, M., Wasfi, R., Tétreault, P., & Surprenant-Legault, J. (2014). New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas. *Transportation*. https://doi.org/10.1007/s11116-013-9508-z

Engel-Yan, J., Rudra, M., Livett, C., & Nagorsky, R. (2014). Strategic station access planning for commuter rail balancing park-and-ride with other modes. In *Transportation Research Record*. https://doi.org/10.3114/2419-08

ESRI. (2020). *Feature To Point (Data Management)*. https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/feature-to-point.htm

ESRI. (2019). *How To: Create heat maps in ArcMap using the Density toolset* https://support.esri.com/en/technical-article/000021720

ESRI. (2020). *About the ArcGIS Network Analyst extension tutorial*. https://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/about-the-network-analyst-tutorial-exercises.htm

ESRI. (2019). *ArcGIS Network Analyst Tutorial Data*. ArcGIS Team - Network Analyst. https://www.arcgis.com/home/item.html?id=d6bd91b2fddc483b8ccbc66942db84cb

Faraway, J. (2016). *faraway: Functions and Datasets for Books by Julian Faraway* (1.0.7). https://cran.r-project.org/package=faraway

Farber, S., Bartholomew, K., Li, X., Páez, A., & Nurul Habib, K. M. (2014). Assessing social equity in distance based transit fares using a model of travel behavior. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2014.07.013

Fox, J., & Weisberg, S. (2019). *An {R} Companion to Applied Regression* (No. 3rd). https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Frondel, M., & Vance, C. (2011). Rarely enjoyed? A count data analysis of ridership in Germany's public transport. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2010.09.009

Gill-Carcia, JR. Puron-Cid, G. (2014). Using Panel Data Techniques for Social Science Research: an Illustrative Case and Some Guidelines. *CIENCIA Ergo-Sum, Revista Científica Multidisciplinaria de Prospectiva*.

GO Transit. (2021). *Your Fare*. https://www.gotransit.com/en/trip-planning/calculate-fare/your-fare

Government of Canada. (2014). *Canadian Census Analyser, 2011 Census Profiles Files / Profile of Census Disseminations Areas*. Statistics Canada via the University of Toronto's Computing in the Humanities and Social Sciences (CHASS) Data Centre. http://dc1.chass.utoronto.ca/cgi-bin/census/2011/displayCensus.cgi?year=2011&geo=da

Government of Canada. (2020). *Canadian Census Analyser, Profile of Dissemination Areas - 2006 Census*. Statistics Canada via the University of Toronto's Computing in the Humanities and Social Sciences (CHASS) Data Centre. http://dc1.chass.utoronto.ca/cgi-bin/census/2006/displayCensusDA.cgi

Government of Canada. (2017). *Postal Code Boundary Shapefile*. Statistics Canada. Obtained from the University of Waterloo's Geospatial Centre

Government of Canada. (2017). *Canadian Census Analyser, 2016 Census Profiles Files / Profile of Provinces*. Statistics Canada via the University of Toronto's Computing in the Humanities and Social Sciences (CHASS) Data Centre. http://dc1.chass.utoronto.ca/cgi-bin/census/2016/displayCensus.cgi?year=2016&geo=prov

Government of Canada. (2019). *2016 Census - Boundary files*. Statistics Canada. https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm

Government of Canada. (2017). *Canadian Census Analyser, 2016 Census Profiles Files / Profile of Census Disseminations Areas*. Statistics Canada via the University of Toronto's Computing in the Humanities and Social Sciences (CHASS) Data Centre. http://dc1.chass.utoronto.ca/cgi-bin/census/2016/displayCensus.cgi?year=2016&geo=da

Government of Canada. (2013). *Canadian Census Analyser, 2011 National Household Survey (NHS) Profiles Files / Profile of Census Disseminations Areas*. Statistics Canada via the University of Toronto's Computing in the Humanities and Social Sciences (CHASS) Data Centre. http://dc1.chass.utoronto.ca/cgi-bin/census/2011nhs/displayCensus.cgi?year=2011&geo=da

Government of Canada. (2020). *Consumer Price Index, monthly, not seasonally adjusted, Canada*. Statistics Canada. https://doi.org/https://doi.org/10.25318/1810000401-eng

Government of Ontario. (2013). *Metrolinx Investment Strategy*. http://www.metrolinx.com/en/regionalplanning/funding/IS_Full_Report_EN.pdf

Government of Ontario. (2020). *Future Regional Transit Network Shapefile*. Metrolinx. http://www.metrolinx.com/en/aboutus/opendata/default.aspx

Government of Ontario. (2020). *GO Transit Rail Daily Ridership Statistics via PRESTO, Jan. 2016 - Dec. 2019.* Metrolinx. Obtained from Metrolinx, Service Planning Department.

Government of Ontario. (2020). *GO Transit Rail Fare Price Statistics via PRESTO, Jan. 2016 - Dec. 2019*. Metrolinx. Obtained from Metrolinx, Service Planning Department

Government of Ontario. (2020). *GO Rail Parking & Utilization*. Metrolinx. http://www.metrolinx.com/en/aboutus/opendata/default.aspx

Government of Ontario. (2019). *Fuel price survey information - Data*. Department of Energy, Northern Development and Mines via the Ontario Data Catalogue. https://data.ontario.ca/dataset/fuels-price-survey-information/resource/15b54ed9-f711-45c8-bd87-30eb626f7c0a?inner_span=True

Government of Ontario. (2020). *Ontario Road Network Segment With Addresses*. Government of Ontario, Ministry of Natural Resources. https://geohub.lio.gov.on.ca/datasets/923cb3294384488e8a4ffbeb3b8f6cb2_32

Government of Ontario. (2016). *GO Rail Station Access Plan*. http://www.metrolinx.com/en/regionalplanning/projectevaluation/studies/GO_Rail_Station_Access_Plan_EN.pdf

Greater Toronto Transportation Authority Act, (2006) (testimony of Government of Ontario). https://www.ontario.ca/laws/statute/06g16/v2

Government of Ontario. (2020). *Provincial Policy Statement*. https://doi.org/978-1-4868-4231-5

Government of Ontario. (2020). *A Place to Grow – Growth Plan for the Greater Golden Horseshoe, Office Consolidation.* https://doi.org/978-1-4868-4654-2

Government of Ontario. (2017). *Now arriving 1960s – GO Transit*. Government of Ontario, Metrolinx. http://goingstrong.gotransit.com/en/1960.html

Government of Ontario. (2017). *Greenbelt Plan*. https://doi.org/978-1-4606-9531-9

Government of Ontario. (2018). *GO Expansion Full Business Case*. http://www.metrolinx.com/en/docs/pdf/board_agenda/20181206/20181206_BoardMtg_GO_Expansion_Full_Business_Case.PDF

Government of Ontario. (2018). *2041 Regional Transportation Plan – for the Greater Toronto and Hamilton Area*. https://doi.org/(978-1-7751379-2-4

Planning Act, (1990) (testimony of Government of Ontario). https://www.ontario.ca/laws/statute/90p13

Government of Ontario. (2008). *The Big Move: Transforming transportation in the Greater Toronto and Hamilton Area*. https://doi.org/(978-1-4249-8602-6

Greater Toronto and Hamilton Area Transit Implementation Act, (2009) (testimony of Government of Ontario). https://www.ontario.ca/laws/statute/s09014

Grimsrud, M., & El-Geneidy, A. (2013). Driving transit retention to renaissance: Trends in Montreal commute public transport mode share and factors by age group and birth cohort. *Public Transport*. https://doi.org/10.1007/s12469-013-0075-7

Guelph Transit. (2016). *Guelph Transit General Transit Feed Specification File November 24, 2016*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20161124

Guerra, E., & Cervero, R. (2011). Cost of a ride. *Journal of the American Planning Association*. https://doi.org/10.1080/01944363.2011.589767

Guerra, E., Cervero, R., & Tischler, D. (2012). Half-Mile Circle Does It Best Represent Transit Station Catchments? *Transportation Research Record*.

Gutiérrez, J., Cardozo, O. D., & García-Palomares, J. C. (2011). Transit ridership forecasting at station level: An approach based on distance-decay weighted regression. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2011.05.004

Hanson, M. E. (1992). Automobile subsidies and land use estimates and policy responses. *Journal of the American Planning Association*. https://doi.org/10.1080/01944369208975535

Holmgren, J. (2007). Meta-analysis of public transport demand. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2007.06.003

Hsu, H. P., Boarnet, M. G., & Houston, D. (2019). Gender and Rail Transit Use: Influence of Environmental Beliefs and Safety Concerns. *Transportation Research Record*. https://doi.org/10.1177/0361198119837193

Hsu, H. (2008). How Does Fear of Sexual Harassment on Transit Affect Women's Use of Transit? *Women's Issues in Transportation*.

Ingvardson, J. B., & Nielsen, O. A. (2018). How urban density, network topology and socio-economy influence public transport ridership: Empirical evidence from 48 European metropolitan areas. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2018.07.002

Ionescu, E. (2020). *General Transit Feed Specification Processing Tool*.

Iseki, H., & Ali, R. (2014). Net effects of gasoline price changes on transit ridership in U.S. urban areas. In *Mineta Transportation Institute*. http://transweb.sjsu.edu/PDFs/research/1106-gas-price-changes-and-transit-ridership.pdf.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning - with Applications in R*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kain, J. F., & Liu, Z. (1999). Secrets of success: Assessing the large increases in transit ridership achieved by Houston and San Diego transit providers. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/S0965-8564(99)00009-9

Keijer, M. J. N., & Rietveld, P. (2000). How do people get to the railway station? The Dutch experience. *Transportation Planning and Technology*. https://doi.org/10.1080/03081060008717650

Kohn, H. (2000). Factors Affecting Urban Transit Ridership. *Statistics Canada, Archived Journals and Periodicals*, *53F0003-XI*. https://www150.statcan.gc.ca/n1/en/catalogue/53F0003X

Kuby, M., Barranda, A., & Upchurch, C. (2004). Factors influencing light-rail station boardings in the United States. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2003.10.006

Lane, B. W. (2012). A time-series analysis of gasoline prices and public transportation in US metropolitan areas. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2011.10.006

Lane, B. W. (2010). The relationship between recent gasoline price fluctuations and transit ridership in major US cities. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2009.04.002

Lascano Kežić, M. E., & Durango-Cohen, P. L. (2018). New ridership for old rail: An analysis of changes in the utilization of Chicago's urban rail system, 1990–2008. *Research in Transportation Economics*. https://doi.org/10.1016/j.retrec.2018.06.012

Lee, B., & Lee, Y. (2013). Complementary pricing and land use policies: Does it lead to higher transit use? *Journal of the American Planning Association*. https://doi.org/10.1080/01944363.2014.915629

Levinson, H. S., Ryus, P., Schofer, J. L., Semler, C., Parks, J., Coffel, K., Sampson, D., & Kachadoorian, C. (2012). Guidelines for Providing Access to Public Transportation Stations. In *Guidelines for Providing Access to Public Transportation Stations*. https://doi.org/10.17226/14614

Lewis, J. (2018). *Lecture 4 - Population Projections. Methods of Planning Analysis.* University of Waterloo.

Li, L., Kasraian, D., & Shalaby, A. (2020). Empirical Analysis of Long-Run Elasticities and Asymmetric Effects of Transit Demand Determinants. *Transportation Research Record*. https://doi.org/10.1177/0361198120925465

Lindsey, R. (2007). Congestion Relief: Assessing the Case for Road Tolls in Canada. *Commentary - C.D. Howe Institute*.

Lintern, G. (2019). *Toronto Official Plan*. https://www.toronto.ca/wp-content/uploads/2019/06/8f06-OfficialPlanAODA_Compiled-3.0.pdf

Liu, C., Erdogan, S., Ma, T., & Ducca, F. W. (2016). How to Increase Rail Ridership in Maryland: Direct Ridership Models for Policy Guidance. *Journal of Urban Planning and Development*. https://doi.org/10.1061/(asce)up.1943-5444.0000340

Liu, R. (2018). *Improving Demand Modeling in California Rail Transit System*. https://doi.org/WP 18-01

Livey, J., & Rossini, R. (2015). *Tolling Options for the Gardiner Expressway and Don Valley Parkway*. https://www.toronto.ca/legdocs/mmis/2015/ex/bgrd/backgroundfile-83671.pdf

Lysyk, B. (2016). *2016 Annual Report of the Office of the Auditor General of Ontario*. https://doi.org/978-1-4606-8937-0

Mahmoud, M. S., Habib, K. N., & Shalaby, A. (2014). Park-and-ride access station choice model for cross-regional commuting: Case study of greater Toronto and Hamilton area, Canada. In *Transportation Research Record*. https://doi.org/10.3114/2419-09

Maley, D. W., & Weinberger, R. (2009). Rising gas price and transit ridership: Case study of Philadelphia, Pennsylvania. *Transportation Research Record*. https://doi.org/10.3141/2139-21

Mankiw, N. G. (2013). *Principles of Microeconomics – Sixth Canadian Edition*. Nelson Thomson Learning.

Mayworm, P. D., Lago, A. M., & McEnroe, J. M. (1980). *Patronage Impacts of Changes in Transit Fares and Services.*

McCarter, J. (2012). *2012 Annual Report of the Office of the Auditor General of Ontario*. https://doi.org/978-1-4606-0349-9

McDonnel, G. (2015). *Locomotives: The Modern Diesel and Electric Reference.* (2nd Editio). Boston Mills Press.

Miller, C., & Savage, I. (2017). Does the demand response to transit fare increases vary by income? *Transport Policy*. https://doi.org/10.1016/j.tranpol.2017.01.006

Miller, E. J., Shalaby, A., Ehab Diab, Pe., & Kasraian, D. (2018). *Canadian Transit Ridership Trends Study*. https://cutaactu.ca/sites/default/files/cuta_ridership_report_final_october_2018_en.pdf

Moore, T., Thorsnes, P., & Appleyard, B. (2007). The transportation land use connection. *APA Planning Advisory Service Reports*.

Morrall, J., & Bolger, D. (1996). Relationship between downtown parking supply and transit use. *ITE Journal (Institute of Transportation Engineers)*.

Newman, P. W. G., & Kenworthy, J. R. (1996). The land use-transport connection: An overview. *Land Use Policy*. https://doi.org/10.1016/0264-8377(95)00027-5

Nowak, W. P., & Savage, I. (2013). The cross elasticity between gasoline prices and transit use: Evidence from Chicago. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2013.03.002

O'Fallon, C., Sullivan, C., & Hensher, D. A. (2004). Constraints affecting mode choices by morning car commuters. *Transport Policy*. https://doi.org/10.1016/S0967-070X(03)00015-5

Pasha, M., Rifaat, S. M., Tay, R., & De Barros, A. (2016). Effects of street pattern, traffic, road infrastructure, socioeconomic and demographic characteristics on public transit ridership. *KSCE Journal of Civil Engineering*. https://doi.org/10.1007/s12205-016-0693-6

Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., & White, P. (2006). The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2005.12.004

Rahman, M., Yasmin, S., & Eluru, N. (2019). Examining determinants of rail ridership: a case study of the Orlando SunRail system. *Transportation Planning and Technology*. https://doi.org/10.1080/03081060.2019.1622252

Schade, J., & Schlag, B. (2003). Acceptability of urban transport pricing strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/S1369-8478(02)00046-3

Schimek, P. (2015). Dynamic estimates of fare elasticity for U.S. Public Transit. *Transportation Research Record*. https://doi.org/10.3141/2538-11

Schuitema, G., & Steg, L. (2008). The role of revenue use in the acceptability of transport pricing policies. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/j.trf.2007.11.003

Schumann, J. W., & Phraner, S. D. (1994). Regional rail for U.S. metropolitan areas: concept and applications. In *Transportation Research Record*.

Shoup, D. C. (1997). Evaluating the effects of cashing out employer-paid parking: Eight case studies. *Transport Policy*. https://doi.org/10.1016/S0967-070X(97)00019-X

Singhal, A., Kamga, C., & Yazici, A. (2014). Impact of weather on urban transit ridership. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2014.09.008

Smith, A. (2020). *Personal communication*.

Spears, S., Boarnet, M. G., & Handy, S. (2012). *Brief on the Impacts of Voluntary Trav el Behavior Change Programs Based on a Review of the Empirical Literature*. https://www.semanticscholar.org/paper/Brief-on-the-Impacts-of-Voluntary-Trav-el-Behavior-Spears-Boarnet/089172aa1f17e512664c2885bb443aeedba54d5d?p2df

Statistics Canada. (2017). *Toronto, C [Census subdivision], Ontario and Canada [Country]. (Table). Census Profile*. Government of Canada, Statistics Canada. https://doi.org/Catalogue no. 98-316-X2016001

Stopher, P. R. (2004). Reducing road congestion: A reality check. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2003.09.002

Stover, V. W., & Christine Bae, C. H. (2011). Impact of gasoline prices on transit ridership in Washington State. *Transportation Research Record*. https://doi.org/10.3141/2217-02

Taylor, B. D., & Fink, C. (2003). *The Factors Influencing Transit Ridership: A Review and Analysis of the Ridership Literature*. https://escholarship.org/uc/item/3xk9j8m2

Taylor, B. D., Miller, D., Iseki, H., & Fink, C. (2009). Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2008.06.007

TomTom International BV. (2020). *Toronto Traffic Report – TomTom Traffic Index*. https://www.tomtom.com/en_gb/traffic-index/toronto-traffic

TomTom International BV. (2020). *Traffic Index 2019; North America*. https://www.tomtom.com/en_gb/traffic-index/ranking/?country=CA,MX,US

Torres-Reyna, O. (2010). *Getting started in fixed/random effects models using R*. Princeton University, Data and Statistical Services. http://www.princeton.edu/~otorres/Panel101.pdf

Torres-Reyna, O. (2007). *Panel data analysis fixed and random effects using Stata (v. 4.2)*. Princeton University, Data and Statistical Services. http://www.princeton.edu/~otorres/Panel101.pdf

Town of Whitchurch-Stouffville. (2017). *Town of Whitchurch-Stouffville Official Plan*. https://www.townofws.ca/en/town-hall/resources/Documents/Current-Official-Plan---July-2017.pdf

U.S. Department of Transportation. (2008). *Congestion Pricing — A Primer: Overview*. https://ops.fhwa.dot.gov/publications/fhwahop08039/fhwahop08039.pdf

University of Toronto. (2016). *Transportation Tomorrow Survey Cross Tabulation Query Form - Household - 2006, 2011, 2016*. University of Toronto's Data Management Group. https://dmg.utoronto.ca/idrs/ttsForm/Cros/hhld/2006,2011,2016

University of Toronto. (2014). *2011 Transportation Tomorrow Survey Area Summary: 2011, 2006, 1996 & 1986 Travel Survey Summaries*. http://dmg.utoronto.ca/pdf/tts/2011/regional_travel_summaries/TTS_report4_full.pdf

University of Toronto. (2003). *2001 Transportation Tomorrow Survey: 2001, & 1996 Travel Survey Summaries*. http://dmg.utoronto.ca/pdf/tts/2001/2001regional-travel-summaries.pdf

University of Toronto. (2009). *2006 Transportation Tomorrow Survey: 2006, 2001, & 1996 Travel Survey Summaries*. http://dmg.utoronto.ca/pdf/tts/2006/regional_travel_summaries/TTS_report4_full.pdf

Vijayakumar, N., El-Geneidy, A. M., & Patterson, Z. (2011). Driving to suburban rail stations: Understanding variables that affect driving distance and station demand. *Transportation Research Record*. https://doi.org/10.3141/2219-12

Voith, R. (1997). Fares, Service Levels, and Demographics: What Determines Commuter Rail Ridership in the Long Run? *Journal of Urban Economics*. https://doi.org/10.1006/juec.1996.1092

Vuchic, V. R. (2007). Urban Transit Systems and Technology. In *Urban Transit Systems and Technology*. https://doi.org/10.1002/9780470168066

Wang, Z., Chen, F., & Xu, T. (2016). Interchange between Metro and Other Modes: Access Distance and Catchment Area. *Journal of Urban Planning and Development*. https://doi.org/10.1061/(asce)up.1943-5444.0000330

Wasserman, J. (2019). *A Time and a Place for Every Rider?: Geographic and Temporal Changes in Bay Area Transit Ridership*. https://doi.org/http://dx.doi.org/10.17610/T6KW22

Webster, F. V., & Bly, P. H. (1981). The demand for public transport. *Transport Reviews*. https://doi.org/10.1080/01441648108716470

Weden, M. M., Peterson, C. E., Miles, J. N., & Shih, R. A. (2015). Evaluating Linearly Interpolated Intercensal Estimates of Demographic and Socioeconomic Characteristics of U.S. Counties and Census Tracts 2001–2009. *Population Research and Policy Review*. https://doi.org/10.1007/s11113-015-9359-8

Wei, T., & Simko, V. (2017). *R package "corrplot": Visualization of a Correlation Matrix* (0.84). https://github.com/taiyun/corrplotfaraway

Wiginton, L., Heyman, E., & Allen, J. (2018). *Fare Pricing - Exploring how road pricing on the DVP and Gardiner would impact income groups in the GTHA*. https://www.pembina.org/reports/2018-08-finalreport-farepricing.pdf

Woo, L. (2019). *Transit Oriented Development Implementation*. Government of Ontario, Metrolinx. http://www.metrolinx.com/en/docs/pdf/board_agenda/20181206/20181206_BoardMtg_TOD_Strategy.pdf

Woo, L. (2018). *GTHA Fare Integration Update*. http://www.metrolinx.com/en/docs/pdf/board_agenda/20180426/20180426_BoardMtg_Fare_Integration_EN.pdf

Woo, L., & Childs, M. (2019). *GO Fare Changes*. http://www.metrolinx.com/en/docs/pdf/board_agenda/20190410/20190410_BoardMtg_Bylaw_Changes_EN.pdf

Wooldridge, J. M. (2012). *Introductory Econometrics - A Modern Approach* (J. Sabatino (ed.)). South-Western, Cengage Learning. https://economics.ut.ac.ir/documents/3030266/14100645/Jeffrey_M._Wooldridge_Introductory_Econometrics_A_Modern_Approach__2012.pdf

World Bank Group. (2020). *Introduction to the General Transit Feed Specification (GTFS) and Informal Transit System Mapping (Self-Paced)*. World Bank Group. https://olc.worldbank.org/content/introduction-general-transit-feed-specification-gtfs-and-informal-transit-system-mapping#comments

Yang, Y., Wang, C., Liu, W., & Zhou, P. (2018). Understanding the determinants of travel mode choice of residents and its carbon mitigation potential. *Energy Policy*. https://doi.org/10.1016/j.enpol.2018.01.033

Yanmaz-Tuzel, O., & Ozbay, K. (2010). Impacts of gasoline prices on New Jersey transit ridership. *Transportation Research Record*. https://doi.org/10.3141/2144-07

Zeileis, A. (2002). *Diagnostic Checking in Regression Relationships.* (p. 4). R News. https://cran.r-project.org/doc/Rnews/

Zhang, D., & Wang, X. C. (2014). Transit ridership estimation with network Kriging: A case study of Second Avenue Subway, NYC. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2014.08.021

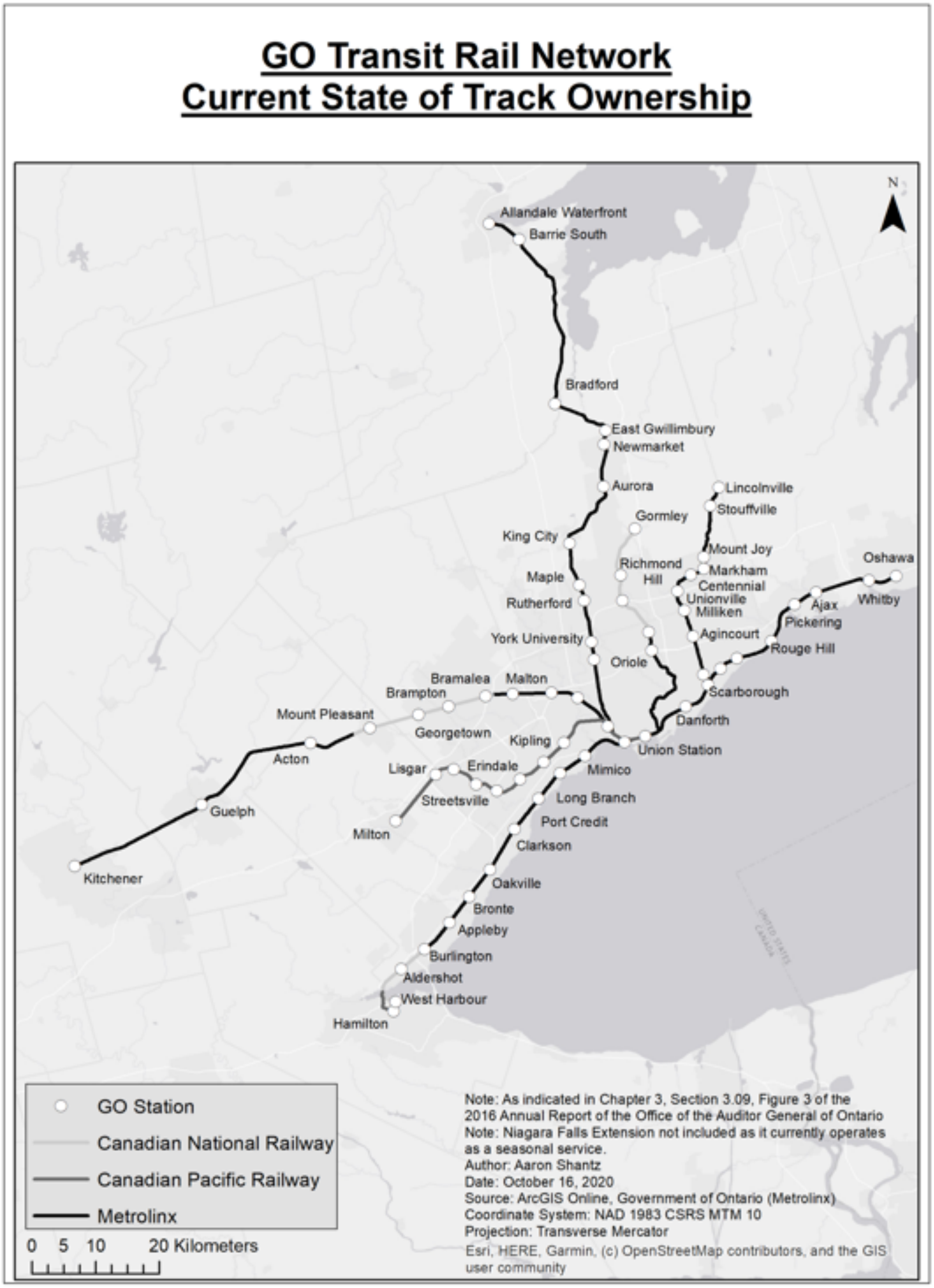
Zhao, F., Chow, L. F., Li, M. T., Ubaka, I., & Gan, A. (2003). Forecasting transit walk accessibility: Regression model alternative to buffer method. *Transportation Research Record*. https://doi.org/10.3141/1835-05

Appendix A – GO Transit Rail Network Additional Information

Table 18 - *GO Transit Rail Network - Additional Corridor Information*

Corridor Name	Length (km)	Terminus Station	Stations Serviced	Notes
Barrie	101	Allandale Waterfront GO Station	11	
Kitchener	101	Kitchener GO Station	11	
Lakeshore East	51	Oshawa GO Station	9	
Lakeshore West	67	West Harbour GO Station / Hamilton GO Centre.	11	The corridor branches into two separate lines after Aldershot GO Station. Therefore, both West Harbour GO Station and Hamilton GO Centre act as terminus stations for the corridor.
Milton	50	Milton GO Station	8	
Niagara Falls	70	Niagara Falls GO Station	3	A branch of the Lakeshore West corridor, which stretches between West Harbour GO Station and Niagara Falls GO Station. Service only provided on select weekends to satisfy seasonal travel demands.
Richmond Hill	43	Gormley GO Station	5	
Stouffville	50	Lincolnville GO Station	11	

Figure 17 - *GO Transit Rail Network - Current State of Track Ownership*



Appendix B – Descriptive Review of Transit Demand Elasticity Literature

Table 19 - *Descriptive Review of Transit Demand Elasticity Literature*

Authors	Geographical Scale	Geographical Context	Mode Examined	Time Period	Trip Type	Analysis Methods	Dependent Variable	Investigated Factors	Significant Factors
Balcombe et al. (2004)	Multi-city	United Kingdom, North America, Australia	All transit, bus, metro, suburban rail	Short-run	Varies	Metadata analysis	Ridership	Variety of internal and external factors	Results differed depending on mode examined / regional context of study
Bernal et al. (2016)	Within-city	Chicago, Illinois, United States	Heavy rail	Did not distinguish	Weekday A.M. Peak	Two stage ordinary least squares	Ridership	Slow zone delay, Slow zone delay squared, Number of trains, Reliability, Gas price, Holiday dummy, Friday dummy, Monday dummy, Seasonal dummies	**Total Northbound Ridership Model:** Reliability (+), Gas Price (+), Holidays (-), Fall (+) **Total Southbound Ridership Model:** Reliability (+), Gas price (+), Holidays (-), Friday (-), Fall (+)
Boisjoly et al. (2018)	Multi-city	North America (Canada and United States)	Heavy rail, bus, light rail, streetcar	Did not distinguish	Aggregated (total trips)	Longitudinal multilevel mixed-effect regression	Ridership	Total vehicle revenue kilometers, Rail vehicle revenue kilometers, Bus vehicle revenue kilometers, Fare price, Population, Area, Percent of households without a car, Unemployment rate, GDP per capita, Gas price, Highway milage, Presence of private bus operator, Presence of Uber, Presence of bicycle sharing system	**Final Model With Disaggregated Service Quantity Statistics:** Rail Vehicle Revenue Kilometers (+), Bus Vehicle Revenue Kilometers (+), Fare Price (-), Presence of private bus operator (+), Presence of bicycle sharing system (+), Percent of households without a car (+), Gas Price (+)
Bomberg et al. (2012)	Within-city	Austin, Texas, United States	All transit	Did not distinguish	Aggregated (total trips)	Qualitative (survey analysis, ordered probit model)	Probability of reducing overall driving distance	Current commute mode, Work at home 2+ times per week, Commute to work using different modes 2+ times per week, Take children to school, Travel time, Number of non-work related driving trips per week, Gas expenditure, Vehicle miles traveled per week, Average fuel economy of all household vehicles, Fuel economy less than 20 mpg dummy, Fuel economy greater	Bus (+), Work at home (+), Multiple modes (-), Number of home based not work related trips per week (-), Vehicle miles travelled per week (-), Employed dummy (-), Age dummy (-), College educated dummy (+), Income (+), Vehicles per driver (+), Number pf basic jobs (+), Number of retail jobs (+),

								than 30 mpg dummy, Age, Male gender dummy, Household income before taxes, Full time student dummy, Employed dummy, College educated dummy, Household size, Vehicles per driver, Local population, residential area, Commercial area, Number of basic jobs, Number of retail jobs, Number of service industry jobs, Total employment, Distance to Central Business District, Bus stop density, Mised use density	Number of service jobs (-), Total number of jobs (-)
Brown et al. (2014)	Within-city	Atlanta, Georgia, United States	Regional bus, regional rail	Did not distinguish	Aggregated (total trips)	Negative binomial regression	Transit use	Percent of population white, Population, Population Density, Median household income, Percentage of households without children, Number of vehicles per person, Residential vacancy rate, Unemployment rate, Employment, Employment density, Presence of transit oriented development, Presence of regional centre, Presence of central business district, Out of vehicle travel time, In-vehicle travel time, Transfer time	**Regional Rail Model:** Percent of population white (-), Population (+), Population density (+), Vehicles per person (+), Residential vacancy rate (-), Employment (+), Employment density (-), Transit oriented development (+), Presence of central business district (+), Out of vehicle travel time (-), In vehicle travel time (-), Transfer time (-)
Chen & McKnight (2007)	Within-city	New York City, New York, United States	Non-motorized, automobile, train, bus, cab	Not applicable	Not applicable	Qualitative (survey analysis)	Mode share	Spatial Location - Manhattan, Bronx Brooklyn and Queens Area, Suburbs	Significant difference in mode split by area
Chen & Zegras (2016)	Within-city	Boston, Massachusetts, United States	Commuter rail	Did not distinguish	Aggregated (total trips), A.M. Peak, P.M. Peak, Off-Peak, Weekend	Ordinary least squares	Ridership	Average household auto ownership, Median household income, Population, Population density, Employment, Employment density, Development mix, Retail mix, Entropy-land use mix, Intersection density, Four-way intersection density, Average sidewalk width, Sidewalk density, Average road width, Walk score, Walk index, Level of service, Interstation spacing, Transfer station, Terminal station, Number	**A.M. Peak Model:** Household income (+), Population density (+), Employment density (-), Percentage of four-way intersections (+), Sidewalk density (+), Interstation spacing (+), Number of feeder bus connections (+) **P.M. Peak Model:** Population (+), Employment density (+), Walk index (+), Average road width (+),

								of feeder bus connections, Parking availability, Accessibility, Distance to central business district	Interstation spacing (+), Number of feeder bus connections (+), Accessibility (+) **Off-Peak Model:** Population (+), Retail employment density (+), Percentage of four-way intersections (+) , Interstation spacing (+), Number of feeder bus connections (+), Distance to central business district (-)
Chen et al. (2011)	Within-city	New Jersey, New Jersey State, United States	Commuter rail	Short-run, long-run	Aggregated (total trips)	Autoregressive fractionally integrated moving average model	Ridership	Gas price, Fare price, Vehicle revenue miles, Size of labour force, seasonal dummy variables	**Short-run Model:** Gas price (+), Vehicle revenue hours (+), Fare price (-), Labour force (+) **Long-run Model:** Gas price (+), Vehicle revenue hours (+), Fare price (-), Labour force (+)
Chow et al. (2006)	Within-city	Brownard County, Florida, United States	All transit	Did not distinguish	Weekday A.M. Peak	Geographically weighted regression	Percentage of workers taking transit	Average walking distance from residence, Shortest walking distance from TAZ centroid to nearest bus stop, Highway accessibility, Average number of cars in households without children, Percent of population black, Median worker earnings, Median household income, Percent of households with no automobiles, Percent of roads with sidewalk within 025 mile buffer of bus stop, Average peak hour headway, Service level, Service coverage, Employment density, Average number of cars in households with children, Percent of households below poverty	Employment density (+/-), Regional accessibility (+), Percentage of zero car households (+), Average cars owned by households (+/-), Percent of population black (+/-)
Currie & Phung (2007)	Multi-city	United States	All transit, heavy rail, light rail, commuter rail, bus	Long run	Aggregated (total trips)	Ordinary least squares	Ridership	Gas price	Gas price (+/-)

Diab et al. (2020)	Multi-city	Canada	Did not distinguish	Did not distinguish	Aggregated (total trips)	Two stage ordinary least squares	Ridership	Total population, Number of businesses, Percent dwellings apartments, Percent dwellings row houses, Percent dwellings single family, Percent population recent immigrants, Percent population working from home, Percent population postsecondary students, Median household income, Personal expenditure on public transit, Percent of people who work outside CSD of residence, Total operating expenses, Vehicle revenue hours, Gas price, Presence of Uber, Presence of bike-sharing systems, Presence of automated fare collection system	**General Model:** Predicted vehicle revenue hours (+), Percent of dwellings apartments (+), Percent of dwellings row house (+), Percent of dwellings single family (-), Number of businesses (+), Percent of population postgraduate students (+), Percent of people who work outside CSD of residence (-), Personal expenditure on public transit (-), Gas price (+), Presence of automated fare collection system (-), Presence of bike sharing system (-) **Larger Agency Model:** Predicted vehicle revenue hours (+), Percent of dwellings apartments (-), Percent of dwellings row house (+), Percent of dwellings single family (-), Number of businesses (+), Percent of dwellings rented (+), Percent of population working from home (-), Percent of population senior (+), Percent of people who work outside CSD of residence (-), Personal expenditure on public transit (-), Gas price (+), Presence of automated fare collection system (-), Presence of Uber (+), Presence of bike sharing system (-) **Smaller Agency Model:** Predicted vehicle revenue hours (+), Percent of dwellings apartments (+), Percent of dwellings row house (+), Percent of

									dwellings single family (-), Percent of population recent immigrant (+), Percent of people work outside CSD of residence (-), Person expenditure on public transit (-), Presence of Uber (-)
During & Townsend (2015)	Multi-city	Canada	All rail (did not disaggregate)	Did not distinguish	Aggregated (total trips)	Ordinary least squares	Ridership	Unemployment rate, Median household income, Percentage of renter households, Age, Number of bus connections, Number of park and ride spaces, Terminal / transfer station, Distance to terminus, Relative distance to terminus, Station spacing, Presence of bike parking, Presence of car share, Population density, Population + Employment density, Number of nodes, long-node ratio, Total links, Total road length, Street density, Average block length, Intersection density, Percentage of area open, Percentage of area park, Percentage of area residential, Job density, Dwelling density, Percentage of area resource-industrial, Percentage of area government-institutional, Percentage of area commercial, Residential-nonresidential, Presence of university, presence of CBD, Land use mix, Land use entropy, Walkability index, Commercial site density, Presence of peak only service, Transit pass cost, Regular fare price	Population density (+), Intersection density (+), Street density (-), Number of bus connections (+), Number of park and ride spaces (+), Station is transfer station (+), Station only offers peak only service (-), Commercial site density (+), Residential ratio (+)
Engel-Yan et al. (2014)	Within-city	Toronto, Ontario, Canada	Commuter rail	Not applicable	Aggregated (total trips)	Ordinary least squares	Station access distance	Distance from boarding to alighting station, Number of parking spots, Vehicle headway, Presence of terminal station, Station located in City of Toronto, Presence of all-day two-way service	Distance from boarding to alighting station (+), Number of parking spots (+), Presence of terminal station (+)

Farber et al. (2014)	Within-city	Wasatch Front, Utah, United States	All transit	Not applicable	Not applicable	Ordinal / continuous regression model	Trip generation	Household income, Ethnicity, Race, Age, Employment, Education, Drivers license dummy, Limited mobility dummy, Number of vehicles owned, Home ownership status, Residency status, Self-reported place type, Residence type	Age less than 17 years (+), Age 18-24 (+), Age over 65 (+), Mobility limitation (-), Households with retirees (-), Self-employed (+), Student, employed 25+ hours per week (-), Unemployed/retired (+/-), Grad or post-grad degree (-), Female (+), Hispanic (+/-), No driver's license (+/-), Zero vehicle household (-), 2 vehicle household (+), 3+ vehicle household (+), Household rents + Distance to central business district (+), Household rents (-), Household tenure refusal (+), 3+ workers (-), 3+ children bikes (+), 6+ people (-), Suburban mixed neighbourhood (-)
Frondel & Vance (2011)	Multi-city	Germany	All transit	Not applicable	Not applicable	Qualitative (survey analysis, zero-inflated negative binomial model)	Mode share	Fuel price, Fare price, Public transit density, Age, Income, Number of children younger than 18, Walking time to nearest transit stop, High school diploma, Drivers license, Employed, Gender, City size, Parking space at home, Parking space at work, Workplace transit connection, Presence of rail transit, Number of vehicles owned	**Zero-Inflated Negative Binomial Model (marginal effects):** Female gender (+), Age (-), Employed (+), High school diploma (+), License (-), Employed w/ parking space at work (-), Parking space at home (-), Number of cars (-), Walking distance to transit (-), Transit connection at work (+), City size (+), Presence of rail transit (+), Number of children less than 18 (-), Income (-), Fuel price (+), Transit density (+)

Grimsrud & El-Geneidy (2013)	Within-city	Montreal, Quebec, Canada	Bus, metro, commuter rail	Not applicable	A.M. Peak	Probability analysis (logit model)	Mode share	Age, Female, School trip, Number of licensed drivers, Cars per license, Number of children less than 5, Number of commuters, Income, Distance to downtown from origin, Bus route count at origin, Average bus wait time at origin, Metro stop count at origin, Commuter rail stop count at origin, Distance to downtown from destination, Bus route count at destination, Average bus wait time at destination, Metro stop count at destination, Commuter rail stop count at destination	Transit Mode Share Model: Female (+), School trip (+), Number of licensed drivers (+), Cars per license (-), Number of children less than 5 (-), Number of commuters (-), Income (-), Distance to downtown from origin (-), Bus route count at origin (+), Average bus wait time at origin (-), Metro stop count at origin (+), Commuter rail stop count at origin (+), Distance to downtown from destination (-), Bus route count at destination (+), Average bus wait time at destination (-), Metro stop count at destination (+), Commuter rail stop count at destination (+), Various age groups differing effects
Guerra & Cervero (2011)	Multi-city	United States	All rail (did not disaggregate)	Did not distinguish	Aggregated (total trips)	Linear panel data estimators (ordinary least squares, fixed effect, random effect, between effect)	Ridership	Number of jobs within 05 mile of station, Population within 05 mile of station, Number of jobs within 5 miles of station, Population within 5 miles of station, Number of park and ride spots, Number of bus route connections, Average fare price, Average speed, Average frequency, Metropolitan economic growth, Average retail gas price, New corridor, Average distance to central station	**Random Effect Model:** Number of jobs within 05 miles (+), Population within 05 miles (+), Average fare price (-), Average frequency (+), Average speed (+), Fuel price (+), Average distance of stations to central business district (-), Metropolitan economic growth (-), New corridor (-)
Holmgren (2007)	Multi-city	United Kingdom, North America, Australia	All transit	Short-run, long-run	Aggregated (total trips)	Metadata analysis	Demand elasticities	Fare price, Vehicle kilometers, Income, Price of gas, Car ownership	Fare price (-), Vehicle kilometers (+), Income (-), Price of gas (+), Car ownership (-)
Hse (2008)	Within-city	Irvine, California, United States	All transit	Not applicable	Not applicable	Qualitative (interview analysis)	Transit use	Cultural differences, sexual harassment, experiences regarding sexual harassment, presence of policies / interventions aimed at	Responses varied

								addressing sexual harassment on public transit	
Hsu et al. (2019)	Within-city	Los Angeles, California, United States	Light rail transit	Not applicable	Not applicable	First differenced model using ordinary least squares	Transit use	Gender, Age, Ethnicity, Education level, Household income, Household size, Home ownership status, residence within 05 mile of station, Questions regarding attitudes and intentions regarding safety, security, and environmental concerns related to transit use	**First-Difference Panel Data Model (Model 4):** Environmental concerns (+), Residence within 05 mile of station (+), Female + Residence within 05 miles from station + Safety and security concerns (-)
Iseki & Ali (2014)	Multi-city	United States	All transit, bus, commuter rail, light rail, heavy rail	Did not distinguish	Aggregated (total trips)	Linear panel data estimators (fixed effect)	Ridership	Gas price, Fare price, Vehicle revenue hours, Service frequency, Total population, Number of federal highway miles, Mean household income, Unemployment rate, Percentage of households without vehicle	**Non-Constant Elasticity Model:** Gas price (+), Fare price (-), Vehicle revenue hours (+), Total population (+), Federal highway miles (-), Mean household income (+),
Kain & Liu (1999)	Multi-city	Houston, Texas & San Diego, California, United States	Did not distinguish	Long run	Aggregated (total trips)	Ordinary least squares	Ridership	Metropolitan area employment, Central city population, Bus and rail miles, Fare price	**Houston Ridership Model:** Metropolitan area employment (+), Central city population (+), Bus and rail miles (+), Fare price (-) **San Diego Ridership Model:** Metropolitan area employment (+), Central city population (+), Bus and rail miles (+), Fare price (-)
Kežić et al. (2018)	Within-city	Chicago, Illinois, United States	Heavy rail	Did not distinguish	Aggregated (total trips)	Qualitative review (review of change in variables over time, regression analysis not completed)	Ridership	Distance from central business district, Ethnic groups, Service quantity	Distance from central business district (-)

Kohn (2000)	Multi-city	Canada	All transit	Did not distinguish	Not applicable	Multiple regression	Ridership	Average fare price, Number of passengers, City size, Ridership rate per capita, Vehicle revenue hours, Vehicle revenue kilometers, Population	Average fare (-), Vehicle revenue hours (+)
Kuby et al. (2004)	Multi-city	United States	Light rail transit	Did not distinguish	Weekdays	Ordinary least squares	Ridership	Employment, Population, Presence of airport, Presence of international border, Number of college enrollments, Presence of central business district, Number of park and ride spaces, Number of bus connections, Presence of other rail lines, Temperature, Metropolitan area population, Presence of terminal station, Interstation spacing, Presence of transfer station, Accessibility, Employment coverage, Percent renters	Employment (+), Population (+), Presence of airport (+), Number of park and ride spaces (+), Number of bus connections (+), Temperature (-), Presence of terminal station (+), Presence of transfer station (+), Accessibility (-), Employment coverage (+), Percent renters (+)
Lane (2010)	Multi-city	United States	Bus, trolley bus, commuter rail, heavy rail, light rail	Long run	Aggregated (total trips)	Ordinary least squares	Ridership	Gas price, Standard deviation gas price, Vehicle revenue miles, Vehicles operated in maximum service, Time dummy, Seasonal dummy	Gas price (+/-)
Lane (2012)	Multi-city	United States	Bus (motor bus, trolley bus), Rail (commuter rail, heavy rail, light rail)	Long run	Aggregated (total trips)	Ordinary least squares	Ridership	Inflated gas price, Deflated gas price, Gas price range, Vehicle revenue miles, Linear trend, Log trend, Service coverage dummy variables	**Rail Ridership Model:** Gas price (+/-), Vehicle revenue miles (+/-)
Lee & Lee (2013)	Multi-city	United States	All transit	Long run	Aggregated (total trips)	Two stage ordinary least squares, random effect linear panel data estimator	Ridership	Vehicle revenue miles, Public operating subsidy per capita, Total urbanized area population, Gas price, Population density, Compactness index, Containment policy dummy, Average fare price, Freeway lane miles per capita, Share of residents who are postsecondary students, Unemployment rate, Trend (time-	R**andom Effect Model (Model 3):** Predicted vehicle revenue miles (+), Gas price (+), Population density (+), Containment policy (+), Fare price (-), Highway lane miles per capita (+), Proportion of population in postsecondary school (+), Unemployment rate (-), Time-series trend (-),

								series) dummy, Post gasoline price peak dummy, Seasonal dummies	Post gas price peak dummy (+), Seasonal effects (+/-)
Li et al. (2020)	Multi-city	Canada	Did not distinguish	Short-run, long-run	Aggregated (total trips)	Linear panel data estimators (fixed effect and ordinary least squares, dynamic panel data estimator (generalized methods of moments)	Ridership	Vehicle revenue kilometres, Fare price, Gas price, Median personal income	**Short-run Model:** Vehicle revenue kilometers (+), Fare price (-), Price of gas (+) **Long-run Model:** Vehicle revenue kilometers (+), Fare price (-)
Liu (2018)	Multi-city	Los Angeles, San Francisco, San Jose, Sacramento, California, Unites States	All rail (did not disaggregate)	Did not distinguish	Aggregated (total trips)	Autoregressive fractionally integrated moving average model, Linear panel data estimators (pooled ols, fixed effect, random effect)	Ridership	Vehicle revenue miles, Number of stations, Number of employees, Gas price, Fare price	**Fixed Effect Model:** Fare price (-), Vehicle revenue miles (+), Number of stations (+)
Liu et al. (2016)	Within-city	State of Maryland, United States	Commuter rail, light rail transit, light rail transit and metro, all transit excluding commuter rail	Did not distinguish	Aggregated (total trips)	Ordinary least squares	Ridership	Number of trips a.m. peak, Presence of park and ride, Presence of feeder bus connections, Terminal station, Connectivity index, Population density, Employment density, Land use index, Street connectivity, Automobile accessibility, Transit accessibility, Distance to central business district, Walk score, Vehicle ownership, Household income, Ethnicity, Median age, Percent of housing owned	**Commuter Rail Model:** Feeder bus connections (+)

Mahmoud et al. (2014)	Within-city	Toronto, Ontario, Canada	Commuter rail, metro	Not applicable	Not applicable	Probability analysis (multinominal logit model)	Station choice	Distance from household location to park and ride station location, Station direction relative to work-home location, Parking lot capacity, Parking cost in a.m. peak period, Presence of refreshment kiosk, Presence of washrooms, Presence of reserved parking, Presence of reserved carpool parking, Presence of regional transit, Presence of local transit, Presence of metro, Presence of metro pass, Presence of regional transit pass	**Overall Demand Model:** Distance from household location to park and ride station location (-), Station direction relative to work-home location (-), Parking lot capacity (+), Presence of regional transit (-), Distance from household location to park and ride station location with Presence of metro pass (+), Distance from household location to park and ride station location with Presence of regional transit pass (-) **Regional Rail Model:** Distance from household location to park and ride station location (-), Station direction relative to work-home location (-), Parking lot capacity (+), Presence of reserved parking and Presence of reserved carpool parking (+), Presence of refreshment kiosk and Presence of washrooms (+), Presence of local transit (+) **Metro Model:** Distance from household location to park and ride station location (-), Station direction relative to work-home location (-), Parking lot capacity (+), Presence of regional transit (+)
Maley & Weinberger (2009)	Within-city	Philadelphia, Pennsylvania, United States	Regional transit (rail), City transit (bus, trolley, metro)	Long run	Aggregated (total trips)	Ordinary least squares	Ridership	Gas price, seasonal dummy variables	Gas price (+), Summer seasonal effects (-)

Miller & Savage (2017)	Within-city	Chicago, Illinois, United States	Heavy rail, bus	Did not distinguish	Weekdays, Saturdays, Sunday / Holidays	Linear panel data estimators (Pooled ordinary least squares, fixed effect)	Ridership	Per capita income, Distance from downtown, Proportion of males, Proportion of elderly, Proportion of children, Year over year change in revenue per rider, Year over year change in total employment, Year over year change in price of gas, Year over year change in revenue vehicle miles	**Overall Ridership Model:** Population density (+), Distance from downtown (-), Proportion of males (+), Proportion of elderly (+), Year over year change in employment (+) **Weekday Ridership Model:** Distance from downtown (-) Saturday Ridership Model: Distance from downtown (-), Proportion of males (+) **Sundays / Holiday Ridership Model:** Distance from downtown (-), Proportion of males (+)
Nowak & Savage (2013)	Within-city	Chicago, Illinois, United States	City rail, city bus, commuter rail, suburban bus	Medium run	Aggregated (total trips)	Ordinary least squares (12-month difference in dependent variable)	Ridership	Gas price, Gas price per litre dummy variables, Average daily bus revenue miles, Fare price, Unemployment Rate, Proportion of weekdays in a month, Leap year dummy variable	**Commuter Rail Model:** Gas price >$3 (+), Gas price >$4 (+), Average daily bus revenue miles (+), Fare price (-), Unemployment rate (-), Proportions of weekdays in month (+)
Pasha et al. (2016)	Within-city	Calgary, Alberta, Canada	Light rail transit	Did not distinguish	Weekday A.M. Peak	Ordinary least squares	Transit share	Street pattern, Geographical size, Primary land use, Structural type, Heavy vehicle volume, Highway km, Train stations, Commercial area, Total occupied dwellings, Median income, Family size, share of couple families with / without children, Share of lone-parent families with / without children, Share of not living with spouse who are single / separated / divorced / widowed, Share of 65+ year olds not living in a census family who are living with relative / living with non-relative / living alone, Language spoken at home, Unemployment rate, Percent male, Number of children <14 years old	Curvilinear street pattern (+), Mixed street pattern (+), Geographical size <1 (-), Primary land use park (+), Type of dwelling, row house (+), Heavy vehicle volume (-), Highway km (+), Train station (+), Total commercial area (+), Total number of occupied dwellings (+), Median income <$40k (+), Median income >$125k (-), Share of couple families without children (+), Share of lone-parent families with female lone parent (+), Family size 3 (+), Share of not living with spouse who are divorced (-), share of 65+ not living in a census family and living with a relative (-), Multiple

									languages spoken at home (+), Unemployment rate (+), Percent male (+), Number of children (-)
Paulley et al. (2006)	Multi-city	United Kingdom, North America, Australia	Commuter rail	Short-run	Peak period	Metadata analysis	Ridership	Fare price,	Fare price (-),
Rahman et al. (2019)	Within-city	Orlando, Florida, United States	Commuter rail	Did not distinguish	Aggregated (total trips)	Linear panel data estimators (random effect)	Ridership	Day of week, month of year, Total roadway length, Number of bus stops, Presence of free parking facility, Number of commercial centers, Number of educational centers, Number of financial centers, Land use mix, Vehicle ownership, Average temperature, Average windspeed, Presence of rainfall	**Number of Boardings Model:** Monday (-), Friday (+), Jan-Aug (+), Total roadway length (-), Number of bus stops (+), Presence of free parking facility (+), Number of commercial centers (+), Number of educational centers (-), Number of financial centers (+), Land use mix (+), Vehicle ownership - no vehicle (-), Average temperature (+), Average wind speed (-), Rainfall (-)
Schimek (2015)	Multi-city	United States	Did not distinguish	Short-run, long-run	Aggregated (total trips)	Dynamic panel data estimator	Ridership	Average fare price, Vehicle revenue miles, Gas price, Metropolitan area employment, Personal income per capita	**Total Ridership Model:** Fare (-)(-), Vehicle revenue miles (+)(+), Gas price (+)(+), Employment (+)(+), Income (-)(-) **Large Urban Area Model:** Fare (-)(-), Vehicle revenue miles (+)(+), Gas price (+)(+), Employment (+)(+), Income (-)(-) **Small Urban Area Model:**

									Fare (-)(-), Vehicle revenue miles (+)(+), Gas price (+)(+), Income (-)(-)
Singhal et al. (2014)	Within-city	New York City, New York, United States	Heavy rail	Did not distinguish	Weekday and weekend (A.M. Peak, Midday Off-Peak, and P.M. Peak), holidays excluded	Ordinary least squares	Ridership	Temperature deviation, Hot temperature dummy, Cold temperature dummy, Wind speed, Strong breeze dummy, Rain, Heavy rainfall dummy, Snow, Heavy snow dummy, Snow in last 24h dummy, Fog dummy, seasonal dummies	**Weekday A.M. Peak Model:** Rain (-), Hot day (-), Cold day (-), Snow last 24h (-), Fall (+), Winter (+), Spring (+) **Weekday Midday Off-Peak Model:** Rain (-), Snow (-), Temperature deviation (+), Snow last 24h (-), Fall (+), Winter (+), Spring (+) **Weekday P.M. Peak Model:** Snow (-), Temperature deviation (-), Snow last 24h (-), Fall (+), Winter (+), Spring (+) **Underground Station Model:** Rain (-), Snow (-), Heavy rain (-), Heavy snow (+), Wind speed (-), Strong breeze (-), Temperature deviation (+), Hot day (-), Cold day (+), Fall (+), Winter (+), Spring (+) **Elevated Station Model**: Rain (-), Snow (-), Heavy rain (-), Wind speed (-), Temperature deviation (+), Hot day (-), Cold day (+), Fall (+), Spring (+)
Stover & Bae (2011)	Within-city	Washington State, United States	Did not distinguish	Did not distinguish	Aggregated (total trips)	Ordinary least squares, Linear panel data estimators (fixed effect)	Ridership	Gas price, Fare price, Vehicle revenue hours, Unemployment Rate, Size of labour force, Seasonal dummy variables	**Total Ridership Model:** Gas price (+), Vehicle revenue hours (+), Unemployment rate (+), Size of labour force (+) Winter (-), Summer (-)

Taylor et al. (2009)	Multi-city	United States	Did not distinguish	Did not distinguish	Aggregated (total trips)	Two stage ordinary least squares	Ridership	Area, Population, Population density, Regional location (ex UZA in the South), Median household income, Number of people unemployed, percept of population in collage, Percent of population in poverty, percent of population recent immigrants, political party affiliation, ethnic composition, freeway lane miles, fuel prices, number of non-transit trips, percent carless households, total lane miles of roads, vehicle miles per capita, Total revenue vehicle hours, Dominance of primary transit operator, Fare price, Service frequency, Predicted transit service levels, Route density	**Total Transit Demand Model:** Predicted vehicle revenue hours (+), Population density (+), UZA in the South (-), Percent of population in college (+), Percent of population recent immigrants (+), Percent carless households (+), Fare price (-), Service frequency (+) **Per Capita Transit Demand Model:** Predicted vehicle revenue hours (+), Geographic land area (+), Median household income (+), Number of non-transit trips (+), Fare price (-), Service frequency (+)
Vijayakumar et al. (2011)	Within-city	Montreal, Quebec, Canada	Commuter rail	Did not distinguish	Weekday A.M. Peak	Ordinary least squares	Driving station access distance	Male dummy, Age, Distance to destination, Number of parking spots, Number of inbound trains, Distance to downtown terminus	Male (+), Age (-), Distance to destination (+), Number of parking spots (+), Number of inbound trains (+)
Wasserman (2019)	Within-city	San Francisco, California, United States	Heavy rail	Did not distinguish	Weekday A.M. Peak, Weekday P.M. Peak, Off-Peak	Ordinary least squares	Ridership	Jobs at destination, BART travel time, Population at destination, Presence of transfer, Destination at a terminus, BART parking at origin, Population at origin, Origin at a terminus, Household income at origin, Jobs, at origin, Lines at destination, Household income at destination, Drive time to BART time ratio, Lines at origin	**A.M. Peak Model:** Jobs at destination (+), Presence of transfer (+), BART travel time (-), Population at destination (+), Destination at a terminus (+), BART parking at origin (+), Jobs at origin (+), Household income at origin (-), Population at origin (+), Origin at a terminus (+), Lines at destination (+), Drive time to BART time ratio (+), Household income at destination (-) **P.M. Peak Model:** Jobs at origin (+), Presence of transfer (-), BART travel time (-), Population at origin (+), Jobs at destination (+), Origin at a terminus (+), Population at destination (+), BART

									parking at destination (+), Household income at destination (-), Destination at a terminus (+), Lines at origin (+), Drive time to BART time ratio (-), Household income at origin (-)
Yanmaz-Tuzel et al. (2010)	Within-city	New Jersey, New Jersey State, United States	Did not distinguish	Short-run, medium-run	Aggregated (total trips)	Time-lag linear regression model	Ridership	Gas price, Average fare price, Employment total, Total vehicle hours	**2005 Short-run Model:** Gas price 3-month lag (+), Gas price 4-month lag (+), Employment rate (+), Average fare price (-), Vehicle revenue hours (+) **2005 Long-run Model:** Gas price 4-month lag (+), Employment rate (+), Average fare price (-), Vehicle revenue hours (+) **2008 Short-run Model:** Gas price 2-month lag (+), Employment rate (+) **2008 Long-run Model: Gas** price 3-month lag (+), Employment rate (+), Average fare price (-)
Zhang & Wang (2014)	Within-city	New York City, New York, United States	Metro	Did not distinguish	Weekdays	Network Kriging model	Ridership	Total population, Income, Total employment, Retail area, Storage area, Top attraction dummy, Number of subway lines, Multimodal connection dummy	**Network Kriging with Network Distance Model:** Total population (+), Income (+), Total employment (+), Retail area (+), Storage area (-), Top attraction dummy (+), Number of subway lines (+), Multimodal connection dummy (+)

* title of model and subsequent results summarized highlighted in bold if multiple models presented by author.

Appendix C – Characteristics as Downloaded from Statistics Canada

Table 20 – *Description of Characteristics as Downloaded from Statistics Canada*

Description of Data Sources Obtained from Statistics Canada			
Variable	**Indicator**	**Data Source**	**Characteristic**
Population Density	Population density	**2016, 2011, 2006** - Census of Population	**2016, 2011, 2006** - Total population
Employment Density	Density of incoming commuters	**2016** - Census of Population	**2016** - Place of work, total, mode of transportation
Gender - Female	Percentage of population female	**2016, 2011, 2006** - Census of Population	**2016, 2011, 2006** - Total population, female
Income	Median household income	**2016, 2006** - Census of Population **2011** - National Household Survey	**2016** - Median household income, **2011** - Median household total income, **2006** - Median earnings for economic families with earnings
Unemployment Rate	Unemployment rate	**2016, 2006** - Census of Population **2011** - National Household Survey	**2016, 2011, 2006** - Unemployment rate
Age	Median age	**2016, 2011** - Census of Population	**2016** - Average age, **2011** - Median age of population
Households With Children	Density of households with children	**2016, 2011, 2006** - Census of Population	**2016** - Total, couple census families in private households; Total, couples with children in private households; Total, lone-parent census families in private households **2011** - Total, number of census families in private households (Married couples, with children at home; Common-law couples, with children at home; Total lone-parent families by sex of parent and number of children) **2006** - Married couples with children; Common-law couples, with children at home

Appendix D – Calculations Used to Adjust Median Age Values

Table 21 - *Calculations Used to Adjust 2016 Average Age Values*

	Province of Ontario
Median age of the population	41.3
Average age of the population	41.0
Difference between values	0.3
Percentage difference between values	0.73
2016 Average Age values adjusted by	1.0073

Statistics Canada. 2017. *Ontario [Province] and Ontario [Province]* (table). *Census Profile*. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa. Released November 29, 2017. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E (accessed February 9, 2021).

When external variable datasets were obtained from Statistics Canada for inclusion in the demand models, it was found that differing measures of central tendency were used to report age statistics at the Dissemination Area scale between census periods. For example, Median Age was reported in 2011, while Average Age was reported in 2016. Therefore, age could not be reported using raw values as differences in measurement methodology could induce error into the model.

Notably, both Average Age and Median Age were reported at the provincial scale in the 2016 Census of Population. As shown in the above table, these two figures were compared and a difference of 0.73% was realized. Therefore, Median Age estimates at the Dissemination Area scale were obtained by adjusting all Average Age values by a factor of 1.0073%.

Appendix E – Spatial Parameters Used to Assign Station-Level Household Vehicle Ownership and Fuel Price Values

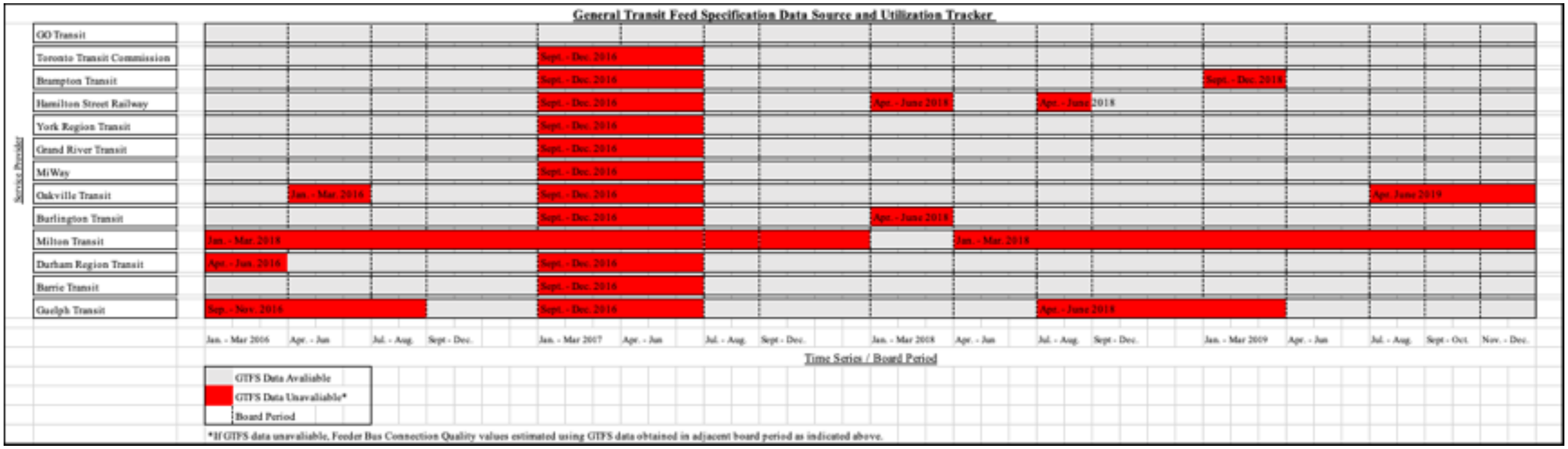
Table 22 - *Spatial Parameters Used to Assign Household Vehicle Ownership Values*

Delineation of Household Vehicle Ownership Values		
Upper / Single-tier Municipality	**Station Name**	
City of Barrie	• Allandale Waterfront GO Station	• Barrie South GO Station
County of Durham	• Ajax GO Station • Bronte GO Station • Burlington GO Station	• Oshawa GO Station • Pickering GO Station • Whitby GO Station
City of Guelph	• Guelph GO Station,	
Regional Municipality of Halton	• Acton GO Station • Aldershot GO Station • Appleby GO Station	• Georgetown GO Station • Milton GO Station • Oakville GO Station
Regional Municipality of Peel	• Bramalea GO Station • Brampton GO Station • Clarkson GO Station • Cooksville GO Station • Dixie GO Station • Erindale GO Station • Lisgar GO Station	• Long Branch GO Station • Malton GO Station • Meadowvale GO Station • Mount Pleasant GO Station, • Port Credit GO Station • Streetsville GO Station
Simcoe County	• Bradford GO Station	
City of Toronto	• Agincourt GO Station • Bloor GO Station • Danforth GO Station • Eglinton GO Station • Etobicoke North GO Station • Exhibition GO Station • Guildwood GO Station • Kennedy GO Station • Kipling GO Station	• Milliken GO Station • Mimico GO Station • Old Cummer GO Station • Oriole GO Station • Rouge Hill GO Station • Scarborough GO Station • Weston GO Station • York University GO Station
Regional Municipality of Waterloo	• Kitchener GO Station,	
Regional Municipality of York	• Aurora GO Station • Centennial GO Station • East Gwillimbury GO Station • King City GO Station • Langstaff GO Station • Lincolnville GO Station • Maple GO Station	• Markham GO Station • Mount Joy GO Station • Newmarket GO Station • Richmond Hill GO Station • Rutherford GO Station • Stouffville GO Station • Unionville GO Station

Table 23 - *Spatial Parameters Used to Assign Fuel Price Values*

Delineation of Fuel Price Values	
Station Location	**Geographical Region (per Ontario Fuel Price Survey)**
City of Toronto	Average of Toronto West + Toronto East
Lakeshore East corridor AND not in the City of Toronto	Toronto East
Kitchener corridor AND not in the City of Toronto	Southern Ontario
Milton Corridor AND not in the City of Toronto	Toronto West
Barrie corridor AND not in the City of Toronto	Southern Ontario
Stouffville corridor AND not in the City of Toronto	Southern Ontario
Richmond Hill corridor AND not in the City of Toronto	Southern Ontario

Appendix F – General Transit Feed Specification Files Used to Extract Service Quantity and Feeder Bus Connection Quality Data



GTFS Files Used to Extract Service Quantity and Feeder Bus Connection Quality Data – Separated by Board Period

Jan. – Mar. 2016

Barrie Transit. (2016). *Barrie Transit General Transit Feed Specification File January 14, 2016*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20160114

Brampton Transit. (2016). *Brampton Transit General Transit Feed Specification File February 3, 2016*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20160203

Burlington Transit. (2016). *Burlington Transit General Transit Feed Specification File March 15, 2016*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20160315

Durham Region Transit. (2016). *Durham Region Transit General Transit Feed Specification File May 3, 2016*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20160503

Grand River Transit. (2015). *Grand River Transit General Transit Feed Specification File December 31, 2015*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20151231

Guelph Transit. (2016). *Guelph Transit General Transit Feed Specification File November 24, 2016*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20161124

Hamilton Street Railway. (2016). *Hamilton Street Railway General Transit Feed Specification File January 8, 2016*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20160108

Metrolinx. (2016). *GO Transit General Transit Feed Specification File January 11, 2016*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20160111

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2016). *MiWay General Transit Feed Specification File June 8, 2016*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20160608

Oakville Transit. (2016). *Oakville Transit General Transit Feed Specification File March 8, 2016*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20160308

Toronto Transit Commission. (2016). *Toronto Transit Commission General Transit Feed Specification File January 12, 2016*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20160112

York Region Transit. (2016). *York Region Transit General Transit Feed Specification File January 12, 2016*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20160112

Apr. – Jun 2016

Barrie Transit. (2016). *Barrie Transit General Transit Feed Specification File May 4, 2016*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20160504

Brampton Transit. (2016). *Brampton Transit General Transit Feed Specification File April 20, 2016*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20160420

Burlington Transit. (2016). *Burlington Transit General Transit Feed Specification File March 15, 2016*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20160315

Durham Region Transit. (2016). *Durham Region Transit General Transit Feed Specification File May 3, 2016*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20160503

Grand River Transit. (2016). *Grand River Transit General Transit Feed Specification File April 25, 2016*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20160425

Guelph Transit. (2016). *Guelph Transit General Transit Feed Specification File November 24, 2016*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20161124

Hamilton Street Railway. (2016). *Hamilton Street Railway General Transit Feed Specification File April 22, 2016*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20160422

Metrolinx. (2016). *GO Transit General Transit Feed Specification File Apr. 26, 2016*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20160426

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2016). *MiWay General Transit Feed Specification File June 8, 2016*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20160608

Oakville Transit. (2016). *Oakville Transit General Transit Feed Specification File March 8, 2016*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20160308

Toronto Transit Commission. (2016). *Toronto Transit Commission General Transit Feed Specification File April 3, 2016*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20160403

York Region Transit. (2016). *York Region Transit General Transit Feed Specification File April 22, 2016*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20160422

Jul. – Aug. 2016

Barrie Transit. (2016). *Barrie Transit General Transit Feed Specification File June 23, 2016*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20160623

Brampton Transit. (2016). *Brampton Transit General Transit Feed Specification File August 18, 2016*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20160818

Burlington Transit. (2016). *Burlington Transit General Transit Feed Specification File June 27, 2016*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20160627

Durham Region Transit. (2016). *Durham Region Transit General Transit Feed Specification File August 16, 2016*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20160816

Grand River Transit. (2016). *Grand River Transit General Transit Feed Specification File July 12, 2016*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20160712

Guelph Transit. (2016). *Guelph Transit General Transit Feed Specification File November 24, 2016*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20161124

Hamilton Street Railway. (2016). *Hamilton Street Railway General Transit Feed Specification File, July 5 2016*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20160705

Metrolinx. (2016). *GO Transit General Transit Feed Specification File July 4, 2016*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20160704

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2016). *MiWay General Transit Feed Specification File June 30, 2016*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20160630

Oakville Transit. (2016). *Oakville Transit General Transit Feed Specification File June 27, 2016*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20160627

Toronto Transit Commission. (2016). *Toronto Transit Commission General Transit Feed Specification File July 23, 2016*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20160723

York Region Transit. (2016). *York Region Transit General Transit Feed Specification File June 30, 2016*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20160630

Sept. – Dec. 2016

Barrie Transit. (2016). *Barrie Transit General Transit Feed Specification File June 23, 2016*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20160623

Brampton Transit. (2016). *Brampton Transit General Transit Feed Specification File October 8, 2016*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20161008

Burlington Transit. (2016). *Burlington Transit General Transit Feed Specification File November 22, 2016*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20161122

Durham Region Transit. (2016). *Durham Region Transit General Transit Feed Specification File September 26, 2016*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20160926

Grand River Transit. (2016). *Grand River Transit General Transit Feed Specification File October 6, 2016*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20161006

Guelph Transit. (2016). *Guelph Transit General Transit Feed Specification File November 24, 2016*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20161124

Hamilton Street Railway. (2016). *Hamilton Street Railway General Transit Feed Specification File November 25, 2016*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20161125

Metrolinx. (2016). *GO Transit General Transit Feed Specification File September 26, 2016*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20160906

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2016). *MiWay General Transit Feed Specification File September 23, 2016*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20160923

Oakville Transit. (2016). *Oakville Transit General Transit Feed Specification File September 1, 2016*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20160901

Toronto Transit Commission. (2016). *Toronto Transit Commission General Transit Feed Specification File October 3, 2016*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20161003

York Region Transit. (2016). *York Region Transit General Transit Feed Specification File November 10, 2016*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20161110

Jan. – Mar. 2017

Barrie Transit. (2016). *Barrie Transit General Transit Feed Specification File June 23, 2016*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20160623

Brampton Transit. (2016). *Brampton Transit General Transit Feed Specification File December 29, 2016*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20161229

Burlington Transit. (2016). *Burlington Transit General Transit Feed Specification File November 22, 2016*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20161122

Durham Region Transit. (2017). *Durham Region Transit General Transit Feed Specification File January 17, 2017*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20170117

Grand River Transit. (2017). *Grand River Transit General Transit Feed Specification File January 6, 2017*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20170106

Guelph Transit. (2017). *Guelph Transit General Transit Feed Specification File January 10, 2017*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20170110

Hamilton Street Railway. (2017). *Hamilton Street Railway General Transit Feed Specification File March 7, 2017*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20170307

Metrolinx. (2017). *GO Transit General Transit Feed Specification File January 14, 2017*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20170114

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2017). *MiWay General Transit Feed Specification File January 30, 2017*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20170130

Oakville Transit. (2017). *Oakville Transit General Transit Feed Specification File February 18, 2017*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20170218

Toronto Transit Commission. (2017). *Toronto Transit Commission General Transit Feed Specification File January 5, 2017*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20170105

York Region Transit. (2017). *York Region Transit General Transit Feed Specification File January 18, 2017*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20170118

Apr. – Jun 2017

Barrie Transit. (2017). *Barrie Transit General Transit Feed Specification File May 22, 2017*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20170522

Brampton Transit. (2017). *Brampton Transit General Transit Feed Specification File April 10, 2017*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20170410

Burlington Transit. (2017). *Burlington Transit General Transit Feed Specification File June 16, 2017*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20170616

Durham Region Transit. (2017). *Durham Region Transit General Transit Feed Specification File April 225, 2017*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20170425

Grand River Transit. (2017). *Grand River Transit General Transit Feed Specification File April 5, 2017*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20170405

Guelph Transit. (2017). *Guelph Transit General Transit Feed Specification File May 8, 2017*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20170508

Hamilton Street Railway. (2017). *Hamilton Street Railway General Transit Feed Specification File March 7, 2017*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20170307

Metrolinx. (2017). *GO Transit General Transit Feed Specification File March 29, 2017*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20170329

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2017). *MiWay General Transit Feed Specification File April 5, 2017*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20170405

Oakville Transit. (2017). *Oakville Transit General Transit Feed Specification File April 25, 2017*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20170425

Toronto Transit Commission. (2017). *Toronto Transit Commission General Transit Feed Specification File March 16, 2017*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20170316

York Region Transit. (2017). *York Region Transit General Transit Feed Specification File April 20, 2017*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20170420

Jul. – Aug. 2017

Barrie Transit. (2017). *Barrie Transit General Transit Feed Specification File July 8, 2017*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20170708

Brampton Transit. (2017). *Brampton Transit General Transit Specification Feed June 28, 2017*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20170628

Burlington Transit. (2017). *Burlington Transit General Transit Feed Specification File June 16, 2017*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20170616

Durham Region Transit. (2017). *Durham Region Transit General Transit Feed Specification File July 22, 2017*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20170722

Grand River Transit. (2017). *Grand River Transit General Transit Feed Specification File July 11, 2017*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20170711

Guelph Transit. (2017). *Guelph Transit General Transit Feed Specification File May 8, 2017*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20170508

Hamilton Street Railway. (2017). *Hamilton Street Railway General Transit Feed Specification File July 31, 2017*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20170731

Metrolinx. (2017). *GO Transit General Transit Feed Specification File June 27, 2017*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20170627

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2017). *MiWay General Transit Feed Specification File June 30, 2017*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20170630

Oakville Transit. (2017). *Oakville Transit General Transit Feed Specification File June 27, 2017*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20170627

Toronto Transit Commission. (2017). *Toronto Transit Commission General Transit Feed Specification File March 16, 2017*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20170316

York Region Transit. (2017). *York Region Transit General Transit Feed Specification File July 4, 2017*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20170704

Sept. – Dec. 2017

Barrie Transit. (2017). *Barrie Transit General Transit Feed Specification File September 6, 2017*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20170906

Brampton Transit. (2017). *Brampton Transit General Transit Feed Specification File October 17, 2017*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20171017

Burlington Transit. (2017). *Burlington Transit General Transit Feed Specification File August 24, 2017*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20170824

Durham Region Transit. (2017). *Durham Region Transit General Transit Feed Specification File September 7, 2017*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20170907

Grand River Transit. (2017). *Grand River Transit General Transit Feed Specification File September 8, 2017*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20170908

Guelph Transit. (2017). *Guelph Transit General Transit Feed Specification File September 2, 2017*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20170902

Hamilton Street Railway. (2017). *Hamilton Street Railway General Transit Feed Specification File September 24, 2017*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20170924

Metrolinx. (2017). *GO Transit General Transit Feed Specification File September 8, 2017*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20170908

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2017). *MiWay General Transit Feed Specification File September 29, 2017*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20170929

Oakville Transit. (2017). *Oakville Transit General Transit Feed Specification File November 24, 2017*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20171124

Oakville Transit. (2017). *York Region Transit General Transit Feed Specification File October 10, 2017*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20171010

Toronto Transit Commission. (2017). *Toronto Transit Commission General Transit Feed Specification Data October 5, 2017*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20171005

Jan. – Mar. 2018

Barrie Transit. (2018). *Barrie Transit General Transit Feed Specification File January 5, 2018*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20180105

Brampton Transit. (2018). *Brampton Transit General Transit Feed Specification File January 31, 2018*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20180131

Burlington Transit. (2018). *Burlington Transit General Transit Feed Specification File April 25, 2018*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20180425

Durham Region Transit. (2017). *Durham Region Transit General Transit Feed Specification File December 16, 2017*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20171216

Grand River Transit. (2018). *Grand River Transit General Transit Feed Specification File January 4, 2018*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20180104

Guelph Transit. (2018). *Guelph Transit General Transit Feed Specification File January 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20180120

Hamilton Street Railway. (2018). *Hamilton Street Railway General Transit Feed Specification File April 17, 2018*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20180417

Metrolinx. (2018). *GO Transit General Transit Feed Specification File January 22, 2018*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20180122

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2018). *MiWay General Transit Feed Specification File February 1, 2018*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20180201

Oakville Transit. (2018). *Oakville Transit General Transit Feed Specification File February 28, 2018*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20180228

Toronto Transit Commission. (2018). *Toronto Transit Commission General Transit Feed Specification File February 10, 2018*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20180210

York Region Transit. (2018). *York Region Transit General Transit Feed Specification File January 8, 2018*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20180108

Apr. – Jun 2018

Barrie Transit. (2018). *Barrie Transit General Transit Feed Specification File April 7, 2018*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20180407

Brampton Transit. (2018). *Brampton Transit General Transit Feed Specification File April 14, 2018*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20180414

Burlington Transit. (2018). *Burlington Transit General Transit Feed Specification File April 25, 2018*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20180425

Durham Region Transit. (2018). *Durham Region Transit General Transit Feed Specification File April 30, 2018*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20180430

Grand River Transit. (2018). *Grand River Transit General Transit Feed Specification File April 7, 2018*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20180407

Guelph Transit. (2018). *Guelph Transit General Transit Feed Specification File April 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20180420

Hamilton Street Railway. (2018). *Hamilton Street Railway General Transit Feed Specification File April 17, 2018*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20180417

Metrolinx. (2018). *GO Transit General Transit Feed Specification File March 24, 2018*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20180324

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2018). *MiWay General Transit Feed Specification File April 3, 2018*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20180403

Oakville Transit. (2018). *Oakville Transit General Transit Feed Specifcation File April 19, 2018*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20180419

Toronto Transit Commission. (2018). *Toronto Transit Commission General Transit Feed Specification File May 5, 2018*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20180505

York Region Transit. (2018). *York Region General Transit Feed Specification File April 26, 2018*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20180426

Jul. – Aug. 2018

Barrie Transit. (2018). *Barrie Transit General Transit Feed Specification File July 8, 2018*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20180708

Brampton Transit. (2018). *Brampton Transit General Transit Feed Specification File July 12, 2018*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20180712

Burlington Transit. (2018). *Burlington Transit General Transit Feed Specification File July 8, 2018*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20180708

Durham Region Transit. (2018). *Durham Region Transit General Feed Specification File July 17, 2018*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20180717

Grand River Transit. (2018). *Grand River Transit General Transit Feed Specification File July 6, 2018*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20180706

Guelph Transit. (2018). *Guelph Transit General Transit Feed Specification File April 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20180420

Hamilton Street Railway. (2018). *Hamilton Street Railway General Transit Feed Specification File June 18, 2018*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20180618

Metrolinx. (2018). *GO Transit General Transit Feed Specification File July 8, 2018*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20180708

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2018). *MiWay General Transit Feed Specification File July 10, 2018*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20180710

Oakville Transit. (2018). *Oakville Transit General Transit Feed Specification File June 21, 2018*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20180621

Toronto Transit Commission. (2018). *Toronto Transit Commission General Transit Feed Specification File June 18, 2018*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20180618

York Region Transit. (2018). *York Region Transit General Transit Feed Specification File July 4, 2018*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20180704

Sept. – Dec. 2018

Barrie Transit. (2018). *Barrie Transit General Transit Feed Specification File September 4, 2018*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20180904

Brampton Transit. (2018). *Brampton Transit General Transit Feed Specification File October 10, 2018*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20181010

Burlington Transit. (2018). *Burlington Transit General Transit Feed Specification File August 13, 2018*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20180813

Durham Region Transit. (2018). *Durham Region Transit General Transit Feed Specification File September 12, 2018*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20180912

Grand River Transit. (2018). *Grand River Transit General Transit Feed Specification File September 3, 2018*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20180903

Guelph Transit. (2018). *Guelph Transit General Transit Feed Specification File August 22, 2018*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20180822

Hamilton Street Railway. (2018). *Hamilton Street Railway General Transit Feed Specification File June 18, 2018*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20180618

Metrolinx. (n.d.). *GO Transit General Transit Feed Specification File September 6, 2018*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20180906

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2018). *MiWay General Transit Feed Specification File October 5, 2018*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20181005

Oakville Transit. (2018). *Oakville Transit General Transit Feed Specification File November 29, 2018*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20181129

Toronto Transit Commission. (2018). *Toronto Transit Commission General Transit Feed Specification File November 9, 2018*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20181109

York Region Transit. (2018). *York Region Transit General Transit Feed Specification File September 2, 2018*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20180902

Jan. – Mar. 2019

Barrie Transit. (2019). *Barrie Transit General Transit Feed Specification File January 9, 2019*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20190109

Brampton Transit. (2018). *Brampton Transit General Transit Feed Specification File October 10, 2018*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20181010

Burlington Transit. (2019). *Burlington Transit General Transit Feed Specification File January 3, 2019*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20190103

Durham Region Transit. (2019). *Durham Region Transit General Transit Feed Specification File January 12, 2019*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20190112

Grand River Transit. (2019). *Grand River Transit General Transit Feed Specification File January 7, 2019*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20190107

Guelph Transit. (2019). *Guelph Transit General Transit Feed Specification File January 2, 2019*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20190102

Hamilton Street Railway. (2019). *Hamilton Street Railway General Transit Feed Specification File January 2, 2019*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20190102

Metrolinx. (2018). *GO Transit General Transit Feed Specification File December 25, 2018*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20181225

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2019). *MiWay General Transit Feed Specification File January 29, 2019*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20190129

Oakville Transit. (2019). *Oakville Transit General Transit Feed Specification File January 18, 2019*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20190118

Toronto Transit Commission. (2019). *Toronto Transit Commission General Transit Feed Specification File February 7, 2019*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20190207

York Region Transit. (2019). *York Region Transit General Transit Feed Specification File January 11, 2019*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20190111

Apr. – Jun 2019

Barrie Transit. (2019). *Barrie Transit General Transit Feed Specification File April 20, 2019*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20190420

Brampton Transit. (2019). *Brampton Transit General Transit Feed Specification File April 20, 2019*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20190420

Burlington Transit. (2019). *Burlington Transit General Transit Feed Specification File May 4, 2019*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20190504

Durham Region Transit. (2019). *Durham Region Transit General Transit Feed Specification File April 15, 2019*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20190415

Grand River Transit. (2019). *Grand River Transit General Transit Feed Specification File April 16, 2019*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20190416

Guelph Transit. (2019). *Guelph Transit General Transit Feed Specification File May 6, 2019*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20190506

Hamilton Street Railway. (2019). *Hamilton Street Railway General Transit Feed Specification File April 20, 2019*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20190420

Metrolinx. (2019). *GO Transit General Transit Feed Specification File April 15, 2019*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20190415

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2019). *MiWay General Transit Feed Specification File April 20, 2019*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20190420

Oakville Transit. (2019). *Oakville Transit General Transit Feed Specification File April 27, 2019*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20190427

Toronto Transit Commission. (2019). *Toronto Transit Commission General Transit Feed Specification File March 27, 2019*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20190327

York Region Transit. (2019). *York Region Transit General Transit Feed Specification File April 29, 2019*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20190429

Jul. – Aug. 2019

Barrie Transit. (2019). *Barrie Transit General Transit Feed Specification File July 5, 2019*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20190705

Brampton Transit. (2019). *Brampton Transit General Transit Feed Specification File July 16, 2019*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20190716

Burlington Transit. (2019). *Burlington Transit General Transit Feed Specification File July 5, 2019*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20190705

Durham Region Transit. (2019). *Durham Region Transit General Transit Feed Specification File June 24, 2019*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20190624

Grand River Transit. (2019). *Grand River Transit General Transit Feed Specification File July 5, 2019*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20190705

Guelph Transit. (2019). *Guelph Transit General Transit Feed Specification File May 6, 2019*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20190506

Hamilton Street Railway. (2019). *Hamilton Street Railway General Transit Feed Specification File June 19, 2019*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20190619

Metrolinx. (2019). *GO Transit General Transit Feed Specification File July 5, 2019*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20190705

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2019). *MiWay General Transit Feed Specification File July 12, 2019*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20190712

Oakville Transit. (2019). *Oakville Transit General Transit Feed Specification File April 27, 2019*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20190427

Toronto Transit Commission. (2019). *Toronto Transit Commission General Transit Feed Specification File July 5, 2019*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20190705

York Region Transit. (2019). *York Region Transit General Transit Feed Specification File July 5, 2019*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20190705

Sept. – Oct. 2019

Barrie Transit. (2019). *Barrie Transit General Transit Feed Specification File September 3, 2019*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20190903

Brampton Transit. (2019). *Brampton Transit General Transit Feed Specification File August 14, 2019*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20190814

Burlington Transit. (2019). *Burlington Transit General Transit Feed Specification File August 30, 2019*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20190830

Durham Region Transit. (2019). *Durham Region Transit General Transit Feed Specification File November 20, 2019*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20191120

Grand River Transit. (2019). *Grand River Transit General Transit Feed Specification File September 9, 2019*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20190909

Guelph Transit. (2019). *Guelph Transit General Transit Feed Specification File September 9, 2019*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20190909

Hamilton Street Railway. (2019). *Hamilton Street Railway General Transit Feed Specification File December 24, 2019*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20191224

Metrolinx. (2019). *GO Transit General Transit Feed Specification File September 6, 2019*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20190906

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2019). *MiWay General Transit Feed Specification File September 30, 2019*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20190930

Oakville Transit. (2019). *Oakville Transit General Transit Feed Specification File April 27, 2019*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20190427

Toronto Transit Commission. (2019). *Toronto Transit Commission General Transit Feed Specification File August 21, 2019*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20190821

York Region Transit. (2019). *York Region Transit General Transit Feed Specification File September 13, 2019*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20190913

Nov. – Dec. 2019

Barrie Transit. (2019). *Barrie Transit General Transit Feed Specification File October 24, 2019*. transitfeeds.com. http://transitfeeds.com/p/barrie-transit/522/20191024

Brampton Transit. (2019). *Brampton Transit General Transit Feed Specification File December 25, 2019*. transitfeeds.com. http://transitfeeds.com/p/brampton-transit/35/20191225

Burlington Transit. (2019). *Burlington Transit General Transit Feed Specification File October 24, 2019*. transitfeeds.com. http://transitfeeds.com/p/burlington-transit/294/20191024

Durham Region Transit. (2019). *Durham Region Transit General Transit Feed Specification File November 20, 2019*. transitfeeds.com. http://transitfeeds.com/p/durham-region-transit/642/20191120

Grand River Transit. (2019). *Grand River Transit General Transit Feed Specification File November 15, 2019*. transitfeeds.com. http://transitfeeds.com/p/grand-river-transit/203/20191115

Guelph Transit. (2019). *Guelph Transit General Transit Feed Specification File September 9, 2019*. transitfeeds.com. http://transitfeeds.com/p/guelph-transit/53/20190909

Hamilton Street Railway. (2019). *Hamilton Street Railway General Transit Feed Specification File December 24, 2019*. transitfeeds.com. http://transitfeeds.com/p/hamilton-street-railway/31/20191224

Metrolinx. (2019). *GO Transit General Transit Feed Specification File November 3, 2019*. transitfeeds.com. http://transitfeeds.com/p/go-transit/32/20191104

Milton Transit. (2018). *Milton Transit General Transit Feed Specification File March 20, 2018*. transitfeeds.com. http://transitfeeds.com/p/milton-transit/929/20180320

MiWay. (2019). *MiWay General Transit Feed Specification File November 9, 2019*. transitfeeds.com. http://transitfeeds.com/p/miway/641/20191109

Oakville Transit. (2019). *Oakville Transit General Transit Feed Specification File April 27, 2019*. transitfeeds.com. http://transitfeeds.com/p/oakville-transit/615/20190427

Toronto Transit Commission. (2019). *Toronto Transit Commission General Transit Feed Specification File November 14, 2019*. transitfeeds.com. http://transitfeeds.com/p/ttc/33/20191114

York Region Transit. (2019). *York Region Transit General Transit Feed Specification File December 14, 2019*. transitfeeds.com. http://transitfeeds.com/p/york-regional-transit/34/20191214

Appendix G – Steps Taken to Calculate Customer Origin Density and Conduct Feeder Bus Connection Quality Analysis in ArcGIS

As mentioned in Section 5.1.2, station catchment areas were estimated using customer origin data as downloaded from Metrolinx's PRESTO smartcard system. The system provides the postal code address of PRESTO users that register their card online, the total number of boardings, and the access station associated with these boardings. The following text outlines the process followed in ArcMap to create station catchment areas using this dataset A database containing all relevant datasets and station catchment area outputs can obtained by contacting the author.

Section 5.1.9.5 further stations that station-level feeder bus connection quality was measured by calculating the transit access time from intense customer origin locations to the access station. As part of this process, customer origin data as indicated by PRESTO was analyzed a second time to identify common points of customer origin within a 0.8-10km radius of each GO Transit Rail station. After points were digitized over these areas, the Network Analyst tool in ArcGIS was used to determine the transit travel time between each of these points and the associated access station. The average of these values was calculated, and used to estimate the level of feeder bus connection quality at the station-level. A database containing all relevant datasets and outputs relevant to the feeder bus connection quality analysis can obtained by contacting the author.

The process outlined below describes the steps taken in ArcMap to obtain customer origin density estimates and associated station catchment boundary outputs. Following this, the steps taken to obtain feeder bus connection quality estimates in ArcGIS are also outlined.

Process Used to Delineate Station Catchment Boundaries Using PRESTO Data

Step 1 – Obtain Customer Origin Data from Metrolinx

- Obtain customer origin dataset from PRESTO. A Microsoft Excel spreadsheet should be provided, which indicates the postal code of the user and the number of boardings attributed to each postal code. Separate spreadsheets should be provided for each station analyzed. This data can be obtained from the Customer Analytics team at Metrolinx.

Step 2 – Georeference Customer Origin Data in ArcMap

- Data provided at the postal code scale cannot be georeferenced unless additional spatial data is provided, such as latitude and longitude coordinates. Unfortunately, the customer origin dataset provided by Metrolinx only specifies the postal code address of the user. The dataset needs to be joined with a postal code baselayer, which delineates all postal code locations and boundaries in the study area, so that it can be georeferenced to the appropriate location. This dataset can be obtained by contacting Canada Post, or in this case was obtained from the University of Waterloo's Geospatial Centre.

- Load the postal code boundary shapefile into ArcMap. It is important to note that postal code boundaries are not mutually exclusive. For example, several polygons within the shapefile could represent a single postal code address, as natural and physical breaks in the environment can prevent postal code boundaries from being continuous. Postal code boundaries containing the same postal code address first need to be aggregated into a single polygon to ensure that customer origin counts are only attributed to each postal code address once.
  - Open the "Dissolve" tool in ArcMap.
  - Dissolve postal code boundaries by Postal Code ID – this will dissolve multiple polygons that share the same Postal Code ID into a single polygon and will ensure that only a single polygon represents each postal code address within the study area.
- Load the customer origin dataset into ArcMap. Join the customer origin dataset with the cleaned postal code boundary shapefile. When completing the join, ensure to exclude records where a match does not exist, so that postal code addresses that are not associated with any riders are excluded from the analysis. Once completed, all customer origin locations and the number of boardings associated with each location will be properly georeferenced.

Step 3 – Clean Customer Origin Data to Only Include Home-Based Trips

- Next, ensure that only home-based trips are included in the analysis. Notably, customer origin locations are distributed throughout the country, a result of visitors using the system. To eliminate these outliers and reduce processing capacity, only customer origin data within a 10km radius of the station will be included in the analysis. Create a 10km buffer around the station using the "Buffer" tool. Once created, select all postal code polygons that intersect this feature. Use the "Export" tool to extract this dataset, and include it as a separate file in the analysis. This shapefile should therefore contain customer origin data, georeferenced at the postal code scale, only within 10km of the station being analyzed.
  - If customer origin density is being visualized for the purpose of feeder bus connection quality analysis, also exclude observations located within an 800m buffer of the station. This is done as customer origin density was often concentrated in these areas, therefore limiting the level of analysis that could be completed with respect to feeder bus

connections. Follow the same steps as outlined in the above paragraph to eliminate these observations.

Step 4 – Create Heatmap Illustrating Customer Origin Density, Weighted by Number of Boardings

- Using the dataset as exported above, create a heatmap, weighted by number of boardings, to illustrate where customer origin is most intense.
    - Heatmaps can only be created using point features. Since postal code addresses are currently stored as a polygon feature, convert the polygon feature class to a point feature class using the "Feature to Point" tool. The points created are referenced to the centroid of each postal code polygon included in the analysis and will still indicate the postal code address of the user and number of boardings associated with this location.
    - Create a heatmap using the newly created point feature class. Complete this using the "Point Density" tool. Select the "Number of Boardings" field in the "Population" drop down menu to ensure that the heatmap is weighted by the number of boardings associated with each postal code address. This step is important, as otherwise the software will simply create a heatmap illustrating where postal code addresses are most concentrated.

Step 5 – Digitize Station Catchment Area

- Once the heatmap has been created, digitize a polygon around areas where customer origin data is shown to be most intense. This could be a single polygon or could be several if the concentration of riders is located in several fragmented areas. This is possible, as natural or physical breaks in the environment (ex. the presence of rivers, streams, highways, or uniform employment / industrial land uses) could result in this occurring.
- Save and export the digitized polygon. The area within the polygon therefore represents the station catchment area of the analyzed station.

Process Used to Obtain Feeder Bus Connection Quality Estimates

Step 1 – Create Heatmap Illustrating Customer Origin Density, Weighted by Number of Boardings

- Follow steps 1-4 in previous section to estimate customer origin density surrounding the station being analyzed

Step 2 – Digitize 5 Dense Customer Origin Locations

- Next, digitize five locations where customer origin is most concentrated. This is accomplished because travel time estimates can only be calculated between a pair of point features. In this analysis, the customer origin location acts as the origin point, while the access station acts as the destination point. Notably, heat map outputs as computed by ArcGIS are produced as "raster" outputs, meaning that the output illustrates the data being analyzed but has no quantitative standing.

Therefore, the user has to manually digitize customer origin points over areas where customer origin is observed to be most dense.

- Conduct a qualitative scan of the heatmap produced in the previous step. Using the "Create Point" feature, digitize five locations where rider origin is most concentrated within the station catchment area. Early in the analysis, it was decided that five locations were to be selected at each station so that an "average" transit access time value could be estimated. Ensure that the point shapefile is referenced to the "GCS_North_American_1983_CSRS" projection to ensure data consistency between outputs.

Step 3 – Construct Virtual Transit Network in ArcGIS

- A virtual transit network will then need to be constructed in ArcGIS to obtain transit travel time estimates between customer origin locations and the associated access station. Instructions outlining this process were obtained from the ArcGIS guidebook (ESRI, 2020). These steps are summarized below.
- Download a shapefile of the road network within the study area. For the purpose of this study, a shapefile of Ontario's road network was downloaded from the Province of Ontario's geospatial portal (insert citation).
- Convert the road network shapefile to a geodatabase. In ArcGIS:
    - Open the Catalog pane,
    - Right-click on the folder where you want to store all data / files relevant to the analysis,
    - Click New > File Geodatabase,
    - Name the new geodatabase appropriately (ex. Ontario Road Network).
    - Right-click the Ontario Road Segment shapefile in the table of contents pane,
    - Select Data -> Export Data,
    - Under save type, select File and Personal Geodatabase feature class,
    - Navigate to the geodatabase that was previous created – export the road network shapefile into the new geodatabase
    - Save the new road network geodatabase appropriately
- Add appropriate columns in roads geodatabase.
    - Open the roads geodatabase.
    - Add "RestrictPedestrians" field (text field) – leave values "null"
    - Add "ROAD_CLASS" field (short integer field) – leave values "null"
    - These columns will allow pedestrians to walk on all streets, and will allow the software to configure directions using the road network once the Network Analyst tool is used.

- Download and unzip General Transit Feed Specification files
  - Download General Transit Feed Specification files for the time period in which you want to obtain feeder bus connection quality estimates. For example, to obtain feeder bus connection quality estimates for the board period spanning January – March 2016, download the GTFS files as indicated in the Jan-Mar 2016 reference paragraph above. The metadata of these files has been checked and confirmed that these files and the transit schedules associated with these transit providers were in-effect between January and March, 2016.
  - Once downloaded, unzip the GTFS files.
- Create a file geodatabase and feature database that the GTFS files can reference. To create a file geodatabase:
  - Open the catalog pane,
  - Right-click the databases folder – click New File Geodatabase,
  - browse to location where you want to store the database, enter an appropriate name (ex. GTFS Analysis Jan-Mar2016) and save,
- To create a Feature Database:
  - Right-click on the newly created File Geodatabase,
  - Click New > Feature Dataset
  - Select a coordinate system identical to the one that the road network geodatabase is currently referenced too
- Place the road network geodatabase into the newly created Feature Database
  - Load road network geodatabase into ArcGIS
  - Copy and paste into the Feature Database
  - Ensure that it is named "Streets" – the GTFS files will not reference the roads baselayer if it is not named "Streets"
- Obtain schedule, stop, and route information from previously downloaded GTFS files
  - Upload GTFS files using GTFS to Network Dataset Transit Sources tool
  - Select unzipped GTFS feeds that you will use to construct the virtual transit network
  - Under Target Feature Dataset dropdown menu, select the Feature Database previously created
  - Run the tool – a virtual Network Dataset will be created
- Georeference schedule, stop, and route information from GTFS files to previously uploaded street network geodatabase
  - Use the Connect Network Dataset Transit Sources To Streets tool

    - Select the previously created Network Dataset
    - Reference this file to the Streets file geodatabase previously created
- Create virtual transit network
    - Download the Transit Network Template provided in the ESRI Network Analyst Tutorial
    - Select the Create Network Dataset From Template tool
    - Under the input dropdown menu, select the downloaded template
    - Under the output dropdown menu, select the Feature Dataset previously created
    - Run the tool
- Build the network
    - Open the Build Network Tool
    - Select the Feature Dataset previously created
- A fully functioning virtual transit network should now be constructed

Step 4 – Obtain Travel Time Estimates

- Now that a virtual transit network has been constructed, transit travel times between origin and destination pairs can be estimated using the schedule, stop, and route information that has been extracted from the previously downloaded GTFS files
- Uploaded the identified points of customer origin as estimated in Step 2
- Uploaded a point shapefile of the station location – station locations were obtained from the Metrolinx Open Data Portal
- Select the OD-Matrix tool under the Network Analyst toolbox
- For origins, select the shapefile containing the identified points of customer origin. For destination, select the shapefile illustrating the access station
- Select departure time / data parameters – for the purpose of this analysis, a generic weekday was selected, while time parameters as indicated below were selected depending on the trip type examined
- Specify that each origin should only map to the closest destination – this prevents the software from calculating transit travel time estimates between origin points
- Run the tool 4 times, specifying a different time parameter to obtain separate estimates for each trip type:
    - A.M. Peak – 7:00am
    - Midday Off-Peak – 12:30pm
    - P.M. Peak – 4:00pm
    - Evening Off-Peak – 7:00pm

- Export file containing travel time estimates for further analysis

Step 5 – Calculate Average Transit Travel Time (this value will be used as the indicator for feeder bus connection quality)

- Once the transit travel time estimates were calculated between each customer origin location and the access station, data was cleaned to account for any discrepancies caused by the software
- For example, sometimes the Network Analyst tool would calculate the transit travel time between a customer origin location and a GO Transit rail station other than the associated access station. This would occur if the software determined that it was faster to access an alternate station that the associated access station. In these situations, a value of 60 minutes was assigned to these origin-destination pairs.
- Further, extremely large transit travel times were found for some stations, specifically those located in rural locations where no local bus service is provided. In these situations, the software would assume that a user would walk between the customer origin location and the associated access station, resulting in extremely inflated transit access time values. In these situations, an upper value of 60 minutes were also assigned to these origin-destination pairs.
- Once cleaned, the average transit access time for each station in a given time period was calculated. These estimated were obtained using the "Average" function in Microsoft Excel. These results were exported as the final feeder bus connection quality values for use in the demand model.

Appendix H – Summary Statistics

Table 24 – *A.M. Peak Model All Variables Summary Statistics*

A.M. Peak Model All Variables Summary Statistics					
	Mean	Median	Minimum	Maximum	Range
Ridership A.M. Peak	1498.08	1013.26	38.90	5272.05	5233.15
Service Quantity	10.93	9.00	2.00	24.00	22.00
Fare Price	6.28	6.24	2.54	12.99	10.45
Feeder Bus Connection Quality	35.59	33.15	13.14	60.00	46.86
Population Density	3538.26	3206.26	458.98	12875.39	12416.41
Gender - Female	52.02	51.35	49.95	1885.12	1835.17
Households With Children	659.73	660.14	92.06	1632.46	1540.40
Unemployment Rate	7.65	7.64	4.10	10.73	6.63
Income	94066.35	95700.89	58667.12	130295.92	71628.80
Age	40.66	40.93	34.11	44.34	10.23
Employment Density	474.88	414.15	69.99	1846.81	1776.83
Fuel Price	109.22	108.24	86.99	130.63	43.63
Vehicle Ownership	755171.65	743433.43	83906.00	1224959.58	1141053.58
Number of Parking Spaces	992.41	657.00	1.00	4540.00	4539.00
Distance to CBD - Near	0.30	0.00	0.00	1.00	1.00
Distance to CBD - Far	0.71	1.00	0.00	1.00	1.00
Winter	0.25	0.00	0.00	1.00	1.00
Spring	0.25	0.00	0.00	1.00	1.00
Summer	0.25	0.00	0.00	1.00	1.00
Fall	0.25	0.00	0.00	1.00	1.00
n = 2928					

Table 25 - *Midday Off-Peak Model All Variables Summary Statistics*

Midday Off-Peak Model All Variables Summary Statistics					
	Mean	Median	Minimum	Maximum	Range
Ridership Midday Off-Peak	371.72	315.62	4.14	1438.86	1434.72
Service Quantity	16.26	12.00	1.00	45.00	44.00
Fare Price	5.39	5.28	2.44	11.09	8.66
Feeder Bus Connection Quality	35.00	34.25	13.60	60.00	46.40
Population Density	3927.66	3352.03	708.01	12875.39	12167.38
Gender - Female	52.47	51.42	49.95	1885.12	1835.17
Households With Children	711.67	680.85	140.75	1632.46	1491.71
Unemployment Rate	8.01	8.05	4.14	10.73	6.59
Income	91463.04	86399.56	58667.12	130295.92	71628.80
Age	40.78	41.12	34.11	44.28	10.17
Employment Density	427.45	365.70	69.99	1070.26	1000.27
Fuel Price	109.97	108.55	86.99	130.63	43.63
Vehicle Ownership	823375.96	760936.99	92014.71	1224959.58	1132944.87
Number of Parking Spaces	1183.55	783.00	1.00	4540.00	4539.00
Distance to CBD - Near	0.39	0.00	0.00	1.00	1.00
Distance to CBD - Far	0.61	1.00	0.00	1.00	1.00
Winter	0.25	0.00	0.00	1.00	1.00
Spring	0.24	0.00	0.00	1.00	1.00
Summer	0.25	0.00	0.00	1.00	1.00
Fall	0.26	0.00	0.00	1.00	1.00
n = 1735					

Table 26 – *P.M. Peak Model All Variables Summary Statistics*

P.M. Peak Model All Variables Summary Statistics					
	Mean	Median	Minimum	Maximum	Range
Ridership P.M. Peak	247.13	89.12	1.40	2390.05	2388.65
Service Quantity	11.11	8.00	1.00	30.00	29.00
Fare Price	3.99	3.87	1.01	9.66	8.65
Feeder Bus Connection Quality	35.20	34.48	13.40	60.00	46.60
Population Density	3644.44	3341.03	458.98	12875.39	12416.41
Gender - Female	52.08	51.36	49.95	1885.12	1835.17
Households With Children	678.43	677.38	92.06	1632.46	1540.40
Unemployment Rate	7.70	7.69	4.10	10.73	6.63
Income	94266.56	95972.54	58667.12	130295.92	71628.80
Age	40.69	41.03	34.11	44.34	10.23
Employment Density	477.74	401.74	69.99	1846.81	1776.83
Fuel Price	109.30	108.26	86.99	130.63	43.63
Vehicle Ownership	775615.34	748360.48	83906.00	1224959.58	1141053.58
Number of Parking Spaces	1032.67	701.50	1.00	4540.00	4539.00
Distance to CBD - Near	0.31	0.00	0.00	1.00	1.00
Distance to CBD - Far	0.69	1.00	0.00	1.00	1.00
Winter	0.25	0.00	0.00	1.00	1.00
Spring	0.25	0.00	0.00	1.00	1.00
Summer	0.25	0.00	0.00	1.00	1.00
Fall	0.25	0.00	0.00	1.00	1.00
n = 2690					

Table 27 - *Evening Off-Peak Model All Variables Summary Statistics*

Evening Off-Peak Model All Variables Summary Statistics					
	Mean	Median	Minimum	Maximum	Range
Ridership Evening Off-Peak	101.85	15.86	1.00	2489.43	2488.43
Service Quantity	9.35	4.00	1.00	25.00	24.00
Fare Price	4.75	4.83	1.01	8.86	7.86
Feeder Bus Connection Quality	35.55	34.14	12.00	60.00	48.00
Population Density	3564.98	3291.73	458.98	12875.39	12416.41
Gender - Female	52.12	51.39	49.95	1885.12	1835.17
Households With Children	669.14	669.02	92.06	1632.46	1540.40
Unemployment Rate	7.75	7.73	4.10	10.73	6.63
Income	94622.24	96985.89	58667.12	130295.92	71628.80
Age	40.70	41.04	34.11	43.37	9.25
Employment Density	487.80	404.48	69.99	1846.81	1776.83
Fuel Price	109.42	108.29	86.99	130.63	43.63
Vehicle Ownership	761971.34	746715.37	83906.00	1224959.58	1141053.58
Number of Parking Spaces	1078.44	774.00	1.00	4540.00	4539.00
Distance to CBD - Near	0.29	0.00	0.00	1.00	1.00
Distance to CBD - Far	0.72	1.00	0.00	1.00	1.00
Winter	0.25	0.00	0.00	1.00	1.00
Spring	0.25	0.00	0.00	1.00	1.00
Summer	0.26	0.00	0.00	1.00	1.00
Fall	0.25	0.00	0.00	1.00	1.00
n = 2515					

Appendix I – Calculations Used to Adjust For Inflation

Fare price, income, and fuel price values were adjusted for inflation before inclusion in the demand model. This was done to ensure that change in the purchasing power of Canadian currency did not influence model results.

Statistics Canada maintains a Consumer Price Index (CPI) which indicates changes in consumer prices experienced by the Canadian public. Over time, the cost of a fixed basket of goods and services, referenced to the price paid in 2002, is tabulated. As a result, price movement of the goods and services represented in this basket are representative of inflation costs being realized by the public.

Archived CPI values can be obtained online via the Statistics Canada website. Using the Statistics Canada Consumer Price Index (CPI) deflator tool, CPI values in a given month can be compared to a CPI value in a previous month to determine the level of inflation that has occurred between these periods. A deflator value is then estimated, which allows for consistent dollar estimated between values obtained in each observation period.

CPI values were downloaded at monthly intervals for the duration of the time series. All values were referenced to the CPI value shown in January 2016, as this is the first month of the time-series analyzed. The CPI value in January 2016 was then divided by the CPI value in a given month to obtain an associated deflator value. All fare price, income, and fuel price values were then multiplied by the deflator value that was estimated in a given month to obtain inflation adjusted estimated. This process is outlined in the table below.

Table 28 - *Adjusting For Inflation - Consumer Price Index and Estimated Deflator Values*

Adjusting For Inflation - Consumer Price Index and Estimated Deflator Values												
	Jan-16	Feb-16	Mar-16	Apr-16	May-16	Jun-16	Jul-16	Aug-16	Sep-16	Oct-16	Nov-16	Dec-16
CPI Value	126.8	127.1	127.9	128.3	128.8	129.1	128.9	128.7	128.8	129.1	128.6	128.4
Deflator Value	1.000	0.998	0.991	0.988	0.984	0.982	0.984	0.985	0.984	0.982	0.986	0.988
	Jan-17	Feb-17	Mar-17	Apr-17	May-17	Jun-17	Jul-17	Aug-17	Sep-17	Oct-17	Nov-17	Dec-17
CPI Value	129.5	129.7	129.9	130.4	130.5	130.4	130.4	130.5	130.8	130.9	131.3	130.8
Deflator Value	0.979	0.978	0.976	0.972	0.972	0.972	0.972	0.972	0.969	0.969	0.966	0.969
	Jan-18	Feb-18	Mar-18	Apr-18	May-18	Jun-18	Jul-18	Aug-18	Sep-18	Oct-18	Nov-18	Dec-18
CPI Value	131.7	132.5	132.9	133.3	133.4	133.6	134.3	134.2	133.7	134.1	133.5	133.4
Deflator Value	0.963	0.957	0.954	0.951	0.951	0.949	0.944	0.945	0.948	0.946	0.950	0.951
	Jan-19	Feb-19	Mar-19	Apr-19	May-19	Jun-19	Jul-19	Aug-19	Sep-19	Oct-19	Nov-19	Dec-19
CPI Value	133.6	134.5	135.4	136.0	136.6	136.3	137.0	136.8	136.2	136.6	136.4	136.4
Deflator Value	0.949	0.943	0.936	0.932	0.928	0.930	0.926	0.927	0.931	0.928	0.930	0.930

Appendix J – Correlation Plots and Variance Inflation Factor Scores

Table 29 *– A.M. Peak Model Correlation Plots and VIF Scores*

AM Peak - Initial Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Gender - Female	Households With Children	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																			
Fare Price	-0.16																		
Feeder Bus Connection Qual	0.14	0.30																	
Population Density	0.23	-0.66	-0.22																
Gender - Female	0.01	-0.08	-0.02	0.06															
Households With Children	0.21	-0.64	-0.23	0.98	0.07														
Unemployment Rate	0.26	-0.34	0.06	0.49	0.07	0.53													
Income	0.05	0.39	0.07	-0.43	-0.02	-0.33	-0.44												
Age	0.00	-0.36	-0.12	0.11	0.11	0.06	0.20	-0.24											
Employment Density	0.09	-0.11	0.12	0.36	0.04	0.31	0.23	-0.13	0.23										
Fuel Price	0.12	-0.07	0.07	0.04	0.00	0.04	0.06	-0.02	0.02	0.03									
Vehicle Ownership	0.15	-0.75	-0.28	0.53	0.05	0.56	0.33	-0.18	0.26	0.01	0.04								
Park and Ride Capacity	0.27	0.41	0.09	-0.30	0.00	-0.20	0.11	0.42	-0.09	-0.04	0.00	-0.15							
Distance to CBD - Near	0.11	-0.76	-0.35	0.54	0.07	0.49	0.36	-0.56	0.40	-0.11	0.03	0.62	-0.49						
Distance to CBD - Far	-0.11	0.76	0.35	-0.54	-0.07	-0.49	-0.36	0.56	-0.40	0.11	-0.03	-0.62	0.49	-1*					
Winter	-0.02	0.03	-0.03	0.00	-0.01	0.00	-0.02	0.00	0.00	0.00	-0.32	-0.01	0.00	0.00	0.00				
Spring	-0.02	0.00	0.05	0.00	-0.01	0.00	-0.01	-0.01	0.00	0.00	0.25	0.00	0.00	0.00	0.00	-0.33			
Summer	0.01	-0.01	-0.01	0.00	-0.01	0.00	0.01	0.00	0.00	0.00	0.12	0.00	-0.01	0.00	0.00	-0.33	-0.33		
Fall	0.03	-0.01	-0.01	0.00	0.03	0.00	0.02	0.01	0.00	0.00	-0.06	0.01	0.01	0.00	0.00	-0.33	-0.33	-0.33	
VIF Score	1.70	4.42	1.39	56.77	1.02	51.84	2.72	2.92	1.67	2.04	1.19	2.86	2.51	na	na	na	na	na	na

AM Peak - Final Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Gender - Female	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																		
Fare Price	-0.16																	
Feeder Bus Connection Qual	0.14	0.30																
Population Density	0.23	-0.66	-0.22															
Gender - Female	0.01	-0.08	-0.02	0.06														
Unemployment Rate	0.26	-0.34	0.06	0.49	0.07													
Income	0.05	0.39	0.07	-0.43	-0.02	-0.44												
Age	0.00	-0.36	-0.12	0.11	0.11	0.20	-0.24											
Employment Density	0.09	-0.11	0.12	0.36	0.04	0.23	-0.13	0.23										
Fuel Price	0.12	-0.07	0.07	0.04	0.00	0.06	-0.02	0.02	0.03									
Vehicle Ownership	0.15	-0.75	-0.28	0.53	0.05	0.33	-0.18	0.26	0.01	0.04								
Park and Ride Capacity	0.27	0.41	0.09	-0.30	0.00	0.11	0.42	-0.09	-0.04	0.00	-0.15							
Distance to CBD - Near	0.11	-0.76	-0.35	0.54	0.07	0.36	-0.56	0.40	-0.11	0.03	0.62	-0.49						
Distance to CBD - Far	-0.11	0.76	0.35	-0.54	-0.07	-0.36	0.56	-0.40	0.11	-0.03	-0.62	0.49	-1*					
Winter	-0.02	0.03	-0.03	0.00	-0.01	-0.02	0.00	0.00	0.00	-0.32	-0.01	0.00	0.00	0.00				
Spring	-0.02	0.00	0.05	0.00	-0.01	-0.01	-0.01	0.00	0.00	0.25	0.00	0.00	0.00	0.00	-0.33			
Summer	0.01	-0.01	-0.01	0.00	-0.01	-0.01	0.00	0.00	0.00	0.12	0.00	-0.01	0.00	0.00	-0.33	-0.33		
Fall	0.03	-0.01	-0.01	0.00	0.03	0.03	0.01	0.00	0.00	-0.06	0.01	0.01	0.00	0.00	-0.33	-0.33	-0.33	
VIF Score	1.44	4.39	1.39	3.07	1.02	2.05	2.12	1.64	1.84	1.19	2.83	2.34	na	na	na	na	na	na

Table 30 - *Midday Off-Peak Model Correlation Plots and VIF Scores*

Midday Off Peak - Initial Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Gender - Female	Households With Children	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																			
Fare Price	-0.27																		
Feeder Bus Connection Quality	0.07	0.27																	
Population Density	0.19	-0.60	-0.23																
Gender - Female	0.01	-0.01	0.00	0.01															
Households With Children	0.13	-0.61	-0.26	0.95	0.03														
Unemployment Rate	0.21	-0.22	-0.01	0.24	0.03	0.35													
Income	-0.06	0.44	0.21	-0.60	0.00	-0.51	-0.48												
Age	-0.01	-0.11	-0.09	-0.01	0.06	-0.06	0.01	-0.19											
Employment Density	0.09	0.09	0.15	0.21	0.01	0.08	-0.05	-0.09	0.18										
Fuel Price	0.02	-0.04	0.04	-0.05	-0.01	-0.05	-0.05	0.05	0.02	-0.01									
Vehicle Ownership	-0.05	-0.70	-0.43	0.61	0.02	0.65	0.31	-0.52	0.13	-0.37	0.00								
Park and Ride Capacity	0.11	0.51	0.06	-0.67	0.01	-0.57	0.12	0.53	0.01	-0.02	0.00	-0.52							
Distance to CBD - Near	0.19	-0.67	-0.43	0.62	0.02	0.58	0.29	-0.66	0.35	-0.26	-0.05	0.80	-0.57						
Distance to CBD - Far	-0.19	0.67	0.43	-0.62	-0.02	-0.58	-0.29	0.66	-0.35	0.26	0.05	-0.80	0.57	-1*					
Winter	-0.01	0.00	0.00	0.02	-0.01	0.02	0.01	0.00	0.00	0.01	-0.31	0.01	0.00	0.02	-0.02				
Spring	0.02	0.02	0.04	0.02	-0.01	0.02	0.02	-0.02	0.00	0.01	0.27	0.01	0.01	0.01	-0.01	-0.32			
Summer	0.01	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00	-0.01	0.14	0.00	-0.01	-0.01	0.01	-0.33	-0.33		
Fall	-0.02	0.00	-0.02	-0.03	0.04	-0.02	-0.02	0.02	0.00	-0.01	-0.09	-0.02	-0.02	-0.02	0.02	-0.34	-0.33	-0.34	
VIF Score	2.37	3.08	1.85	####	1.01	####	3.65	4.31	2.22	3.54	1.21	6.12	4.64	na	na	na	na	na	na

Midday Off Peak - Final Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Female	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																		
Fare Price	-0.27																	
Feeder Bus Connection Quality	0.07	0.27																
Population Density	0.19	-0.60	-0.23															
Female	0.01	-0.01	0.00	0.01														
Unemployment Rate	0.21	-0.22	-0.01	0.24	0.03													
Income	-0.06	0.44	0.21	-0.60	0.00	-0.48												
Age	-0.01	-0.11	-0.09	-0.01	0.06	0.01	-0.19											
Employment Density	0.09	0.09	0.15	0.21	0.01	-0.05	-0.09	0.18										
Fuel Price	0.02	-0.04	0.04	-0.05	-0.01	-0.05	0.05	0.02	-0.01									
Vehicle Ownership	-0.05	-0.70	-0.43	0.61	0.02	0.31	-0.52	0.13	-0.37	0.00								
Park and Ride Capacity	0.11	0.51	0.06	-0.67	0.01	0.12	0.53	0.01	-0.02	0.00	-0.52							
Distance to CBD - Near	0.19	-0.67	-0.43	0.62	0.02	0.29	-0.66	0.35	-0.26	-0.05	0.80	-0.57						
Distance to CBD - Far	-0.19	0.67	0.43	-0.62	-0.02	-0.29	0.66	-0.35	0.26	0.05	-0.80	0.57	-1*					
Winter	-0.01	0.00	0.00	0.02	-0.01	0.01	0.00	0.00	0.01	-0.31	0.01	0.00	0.02	-0.02				
Spring	0.02	0.02	0.04	0.02	-0.01	0.02	-0.02	0.00	0.01	0.27	0.01	0.01	0.01	-0.01	-0.32			
Summer	0.01	-0.02	-0.01	-0.01	-0.01	-0.01	0.00	0.00	-0.01	0.14	0.00	-0.01	-0.01	0.01	-0.33	-0.33		
Fall	-0.02	0.00	-0.02	-0.03	-0.02	-0.02	0.02	0.00	-0.01	-0.09	-0.02	0.01	0.02	0.02	-0.34	-0.33	-0.34	
VIF Score	1.97	2.79	1.79	4.55	1.01	2.37	3.12	2.21	2.91	1.21	5.95	4.60	na	na	na	na	na	na

Table 31 *– P.M. Peak Model Correlation Plots and VIF Scores*

PM Peak - Initial Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Gender - Female	Households With Children	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																			
Fare Price	0.11																		
Feeder Bus Connection Quality	0.10	0.05																	
Population Density	0.31	-0.25	-0.17																
Gender - Female	0.03	-0.03	-0.02	0.06															
Households With Children	0.28	-0.33	-0.17	0.97	0.06														
Unemployment Rate	0.37	-0.29	0.09	0.50	0.06	0.56													
Income	-0.05	0.08	0.12	-0.46	-0.02	-0.36	-0.43												
Age	0.08	-0.12	-0.02	0.12	0.10	0.08	0.11	-0.25											
Employment Density	0.18	0.04	0.16	0.36	0.04	0.31	0.22	-0.10	0.28										
Fuel Price	0.09	-0.07	0.05	0.04	0.00	0.03	0.07	-0.02	0.03	0.03									
Vehicle Ownership	0.23	-0.49	-0.24	0.55	0.05	0.57	0.38	-0.26	0.23	0.05	0.04								
Park and Ride Capacity	0.19	-0.15	0.17	-0.38	-0.02	-0.29	0.13	0.41	-0.10	-0.02	0.00	-0.24							
Distance to CBD - Near	0.21	-0.21	-0.33	0.53	0.06	0.48	0.34	-0.59	0.39	-0.12	0.02	0.63	-0.56						
Distance to CBD - Far	-0.21	0.21	0.33	-0.53	-0.06	-0.48	-0.34	0.59	-0.39	0.12	-0.02	-0.63	0.56	-1*					
Winter	-0.01	-0.03	-0.02	0.00	-0.01	0.00	-0.02	0.00	-0.01	0.00	-0.31	-0.01	0.00	0.00	0.00				
Spring	-0.02	0.02	0.04	0.00	-0.01	0.00	0.00	-0.01	0.00	0.00	0.25	0.00	0.00	0.00	0.00	-0.33			
Summer	0.00	0.03	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.12	0.00	-0.01	0.00	0.00	-0.33	-0.34		
Fall	0.03	-0.02	-0.01	0.00	0.03	0.00	0.02	0.01	0.01	0.00	-0.06	0.01	0.01	0.00	0.00	-0.33	-0.33	-0.33	
VIF Score	2.19	2.04	1.39	####	1.02	####	3.50	3.01	1.75	2.23	1.20	2.62	3.51	na	na	na	na	na	na

PM Peak - Final Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Gender - Female	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																		
Fare Price	0.11																	
Feeder Bus Connection Quality	0.10	0.05																
Population Density	0.31	-0.25	-0.17															
Gender - Female	0.03	-0.03	-0.02	0.06														
Unemployment Rate	0.37	-0.29	0.09	0.50	0.06													
Income	-0.05	0.08	0.12	-0.46	-0.02	-0.43												
Age	0.08	-0.12	-0.02	0.12	0.10	0.11	-0.25											
Employment Density	0.18	0.04	0.16	0.36	0.04	0.22	-0.10	0.28										
Fuel Price	0.09	-0.07	0.05	0.04	0.00	0.07	-0.02	0.03	0.03									
Vehicle Ownership	0.23	-0.49	-0.24	0.55	0.05	0.38	-0.26	0.23	0.05	0.04								
Park and Ride Capacity	0.19	-0.15	0.17	-0.38	-0.02	0.31	0.41	-0.10	-0.02	0.00	-0.24							
Distance to CBD - Near	0.21	-0.21	-0.33	0.53	0.06	0.34	-0.59	0.39	-0.12	0.02	0.63	-0.56						
Distance to CBD - Far	-0.21	0.21	0.33	-0.53	-0.06	-0.34	0.59	-0.39	0.12	-0.02	-0.63	0.56	-1*					
Winter	-0.01	-0.03	-0.02	0.00	-0.01	-0.02	0.00	-0.01	0.00	-0.31	-0.01	0.00	0.00	0.00				
Spring	-0.02	0.02	0.04	0.00	-0.01	0.00	-0.01	0.00	0.00	0.25	0.00	0.00	0.00	0.00	-0.33			
Summer	0.00	0.03	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	0.12	0.00	-0.01	0.00	0.00	-0.33	-0.34		
Fall	0.03	-0.02	-0.01	0.00	0.03	0.02	0.01	0.01	0.00	-0.06	0.01	0.01	0.00	0.00	-0.33	-0.33	-0.33	
VIF Score	1.98	1.98	1.37	2.93	1.02	2.39	2.05	1.75	2.01	1.20	2.60	3.50	na	na	na	na	na	na

Table 32 - *Evening Off-Peak Model Correlation Plots and VIF Scores*

Evening Off Peak - Initial Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Gender - Female	Households With Children	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																			
Fare Price	0.11																		
Feeder Bus Connection Quality	0.04	0.25																	
Population Density	0.04	-0.28	-0.13																
Gender - Female	0.03	-0.06	-0.02	0.05															
Households With Children	-0.03	-0.30	-0.14	0.97	0.06														
Unemployment Rate	0.06	-0.17	0.08	0.50	0.05	0.56													
Income	-0.02	0.22	0.06	-0.44	-0.02	-0.34	-0.46												
Age	0.18	-0.29	-0.12	0.12	0.09	0.06	0.06	-0.25											
Employment Density	-0.08	0.10	0.20	0.33	0.03	0.27	0.15	-0.10	0.25										
Fuel Price	0.10	-0.05	0.06	0.00	0.00	0.00	0.05	-0.01	0.01	-0.01									
Vehicle Ownership	-0.01	-0.42	-0.23	0.53	0.05	0.55	0.38	-0.24	0.22	0.03	0.03								
Park and Ride Capacity	0.14	0.33	0.09	-0.36	-0.02	-0.26	0.08	0.40	-0.13	-0.05	0.01	-0.22							
Distance to CBD - Near	0.20	-0.57	-0.33	0.50	0.05	0.45	0.35	-0.57	0.39	-0.16	0.01	0.61	-0.55						
Distance to CBD - Far	-0.20	0.57	0.33	-0.50	-0.05	-0.45	-0.35	0.57	-0.39	0.16	-0.01	-0.61	0.55	-1*					
Winter	-0.02	0.00	-0.02	0.01	-0.01	0.01	-0.01	0.00	0.00	0.01	-0.31	-0.01	0.00	0.00	0.00				
Spring	-0.02	0.03	0.05	0.01	-0.01	0.01	0.00	0.01	0.00	0.01	0.25	0.00	0.01	0.00	0.00	-0.33			
Summer	0.00	0.00	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	-0.01	0.12	0.00	-0.01	0.00	0.00	-0.33	-0.34		
Fall	0.04	-0.33	-0.01	0.03	0.03	-0.01	0.01	0.01	0.01	0.00	-0.07	0.01	0.00	0.00	0.00	-0.33	-0.33	-0.34	
VIF Score	1.77	1.80	1.34	####	1.02	####	3.04	2.98	1.66	2.21	1.19	2.24	2.59	na	na	na	na	na	na

Evening Off Peak - Final Correlation Plot	Service Quantity	Fare Price	Feeder Bus Connection Quality	Population Density	Gender - Female	Unemployment Rate	Income	Age	Employment Density	Fuel Price	Vehicle Ownership	Park and Ride Capacity	Distance to CBD - Near	Distance to CBD - Far	Winter	Spring	Summer	Fall
Service Quantity																		
Fare Price	0.11																	
Feeder Bus Connection Quality	0.04	0.25																
Population Density	0.04	-0.28	-0.13															
Gender - Female	0.03	-0.06	-0.02	0.05														
Unemployment Rate	0.06	-0.17	0.08	0.50	0.05													
Income	-0.02	0.22	0.06	-0.44	-0.02	-0.46												
Age	0.18	-0.29	-0.12	0.12	0.09	0.06	-0.25											
Employment Density	-0.08	0.10	0.20	0.33	0.03	0.15	-0.10	0.25										
Fuel Price	0.10	-0.05	0.06	0.00	0.00	0.05	-0.01	0.01	-0.01									
Vehicle Ownership	-0.01	-0.42	-0.23	0.53	0.05	0.38	-0.24	0.22	0.03	0.03								
Park and Ride Capacity	0.14	0.33	0.09	-0.36	-0.02	0.08	0.40	-0.13	-0.05	0.01	-0.22							
Distance to CBD - Near	0.20	-0.57	-0.33	0.50	0.05	0.35	-0.57	0.39	-0.16	0.01	0.61	-0.55						
Distance to CBD - Far	-0.20	0.57	0.33	-0.50	-0.05	-0.35	0.57	-0.39	0.16	-0.01	-0.61	0.55	-1*					
Winter	-0.02	0.00	-0.02	0.01	-0.01	-0.01	0.00	0.00	0.01	-0.31	-0.11	0.00	0.00	0.00				
Spring	-0.02	0.03	0.05	0.01	-0.01	0.00	-0.01	0.00	0.01	0.25	0.00	0.01	0.00	0.00	-0.33			
Summer	0.00	0.00	-0.01	-0.01	-0.01	0.00	0.00	0.00	-0.01	0.12	0.00	-0.01	0.00	0.00	-0.33	-0.34		
Fall	0.04	-0.03	-0.01	0.00	0.03	0.01	0.01	0.01	0.00	-0.07	0.01	0.00	0.00	0.00	-0.33	-0.33	-0.34	
VIF Score	1.51	1.76	1.33	2.45	1.01	2.25	2.14	1.66	1.86	1.19	2.22	2.55	na	na	na	na	na	na

Appendix K – Outputs Used to Inform Backwards Stepwise Regression

Table 33 - *Outputs Used to Inform Backwards Stepwise Regression*

A.M. Peak Unrestricted Model Estimated Using Robust Standard Errors				
	Coefficient	SE	t-value	p-value
(Intercept)	-6.763	5.893	-1.148	0.251
Service Quantity	0.208	0.033	6.248	< 0.001
Fare Price	-0.283	0.112	-2.522	0.012
Feeder Bus Connection Quality	-0.008	0.015	-0.546	0.585
Population Density	0.397	0.139	2.861	0.004
Gender - Female	-0.057	0.005	-11.696	< 0.001
Unemployment Rate	0.568	0.273	2.080	0.038
Income	0.151	0.203	0.743	0.458
Age	1.201	1.217	0.987	0.324
Employment Density	-0.134	0.089	-1.510	0.131
Fuel Price	0.199	0.036	5.524	< 0.001
Vehicle Ownership	0.265	0.072	3.651	< 0.001
Park and Ride Capacity	0.043	0.010	4.282	< 0.001
Distance to CBD - Near	-1.626	0.134	-12.140	< 0.001
Winter	0.022	0.009	2.511	0.012
Spring	0.022	0.009	2.435	0.015
Summer	-0.007	0.009	-0.780	0.435
Action: Eliminate Age. **Rationale:** P-value greater than predetermined cut-off, large coefficient that does not align with previous estimates as seen in literature.				

A.M. Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Age				
	Coefficient	SE	t-value	p-value
(Intercept)	-2.112	2.539	-0.832	0.406
Service Quantity	0.209	0.034	6.230	< 0.001
Fare Price	-0.291	0.114	-2.555	0.011
Feeder Bus Connection Quality	-0.008	0.015	-0.523	0.601
Population Density	0.353	0.129	2.731	0.006
Gender - Female	-0.055	0.005	-11.828	< 0.001
Unemployment Rate	0.592	0.261	2.267	0.023
Income	0.128	0.196	0.653	0.514
Employment Density	-0.097	0.078	-1.243	0.214
Fuel Price	0.201	0.037	5.479	< 0.001
Vehicle Ownership	0.274	0.073	3.734	< 0.001
Park and Ride Capacity	0.043	0.010	4.268	< 0.001
Distance to CBD - Near	-1.558	0.128	-12.208	< 0.001
Winter	0.021	0.009	2.474	0.013

Spring	0.022	0.009	2.383	0.017
Summer	-0.008	0.009	-0.808	0.419
Action: Eliminate Feeder Bus Connection Quality. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

A.M. Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Feeder Bus Connection Quality				
	Coefficient	SE	t-value	p-value
(Intercept)	-1.985	2.550	-0.779	0.436
Service Quantity	0.208	0.033	6.225	< 0.001
Fare Price	-0.290	0.113	-2.558	0.011
Population Density	0.352	0.130	2.709	0.007
Gender - Female	-0.056	0.005	-12.126	< 0.001
Unemployment Rate	0.585	0.263	2.229	0.026
Income	0.114	0.196	0.579	0.563
Employment Density	-0.098	0.079	-1.237	0.216
Fuel Price	0.200	0.037	5.479	< 0.001
Vehicle Ownership	0.277	0.074	3.749	< 0.001
Park and Ride Capacity	0.043	0.010	4.256	< 0.001
Distance to CBD - Near	-1.560	0.128	-12.143	< 0.001
Winter	0.021	0.009	2.465	0.014
Spring	0.021	0.009	2.339	0.019
Summer	-0.008	0.009	-0.817	0.414
Action: Eliminate Unemployment Rate. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

A.M. Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Income				
	Coefficient	SE	t-value	p-value
(Intercept)	-0.719	2.550	-0.282	0.778
Service Quantity	0.205	0.033	6.143	< 0.001
Fare Price	-0.297	0.113	-2.621	0.009
Population Density	0.347	0.130	2.670	0.008
Gender - Female	-0.056	0.005	-12.197	< 0.001
Unemployment Rate	0.537	0.263	2.046	0.041
Employment Density	-0.100	0.079	-1.271	0.204
Fuel Price	0.203	0.037	5.547	< 0.001
Vehicle Ownership	0.293	0.074	3.975	< 0.001
Park and Ride Capacity	0.040	0.010	4.002	< 0.001
Distance to CBD - Near	-1.598	0.128	-12.440	< 0.001
Winter	0.021	0.009	2.413	0.016
Spring	0.020	0.009	2.224	0.026
Summer	-0.009	0.009	-0.899	0.369

Action: Summer demonstrates p-value greater than predetermined cut-off and is largest p-value in current model output. Investigate if seasonal dummy variables are jointly significant in explaining ridership.

Joint Test of Significance for Seasonal Dummy Variables				
Hypothesis				
Winter = 0				
Spring = 0				
Summer = 0				
Model 1	restricted model			
Model 2	Ridership.AM.Peak ~ Service Quantity + Fare Price + Population Density + Gender - Female + Unemployment Rate + Employment Density + Fuel Price + Vehicle Ownership + Park and Ride Capacity + Station Location - Near + Winter + Spring + Summer			
	Res.Df	Df	Chisq	Pr(>Chisq)
Model 1	2917			
Model 2	2914	3	53.053	< 0.001
Action: Eliminate Employment Density. **Rationale:** Seasonal dummy variables shown to jointly significant in explaining ridership, therefore Summer should not be removed from the regression model. Employment Density therefore eliminated as it has a p-value greater than predetermined cut-off, and second largest p-value (after summer) in current model output.				

A.M. Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Employment Density				
	Coefficient	SE	t-value	p-value
(Intercept)	-0.983	1.154	-0.852	0.394
Service Quantity	0.207	0.033	6.261	< 0.001
Fare Price	-0.296	0.114	-2.591	0.010
Population Density	0.284	0.111	2.550	0.011
Gender - Female	-0.056	0.004	-12.625	< 0.001
Unemployment Rate	0.532	0.268	1.990	0.047
Fuel Price	0.203	0.036	5.642	< 0.001
Vehicle Ownership	0.305	0.080	3.816	< 0.001
Park and Ride Capacity	0.041	0.010	4.194	< 0.001
Distance to CBD - Near	-1.543	0.145	-10.611	< 0.001
Winter	0.021	0.009	2.396	0.017
Spring	0.020	0.009	2.213	0.027
Summer	-0.009	0.009	-0.898	0.369
Action: Summer demonstrates p-value greater than predetermined cut-off and is largest p-value in current model output. Investigate if seasonal dummy variables are jointly significant in explaining ridership.				

Joint Test of Significance for Seasonal Dummy Variables				
Hypothesis				
Winter = 0				
Spring = 0				
Summer = 0				
Model 1	restricted model			
Model 2	Ridership.AM.Peak ~ Service Quantity + Fare Price + Population Density + Gender - Female + Unemployment Rate + fuel Price + Vehicle Ownership + Park and Ride Capacity + Station Location - Near + Winter + Spring + Summer			
	Res.Df	Df	Chisq	Pr(>Chisq)
Model 1	2918			
Model 2	2915	3	52.465	< 0.001
Seasonal dummy variables shown to be statistically significant, therefore Summer should not be removed from regression model. All remaining variables demonstrate p-values less than the predetermined significance cut-off, therefore stepwise regression is completed. The restricted AM Peak model therefore takes the final form as outlined in Equation 27. Complete results are further shown in Table 16.				

Midday Off-Peak Unrestricted Model Estimated Using Robust Standard Errors				
	Coefficient	SE	t-value	p-value
(Intercept)	-57.082	9.683	-5.895	< 0.001
Service Quantity	0.565	0.029	19.402	< 0.001
Fare Price	-0.232	0.095	-2.446	0.015
Feeder Bus Connection Quality	-0.091	0.034	-2.657	0.008
Population Density	1.442	0.234	6.166	< 0.001
Gender - Female	0.010	0.011	0.869	0.385
Unemployment Rate	0.427	0.285	1.499	0.134
Income	1.593	0.406	3.927	< 0.001
Age	11.824	1.754	6.740	< 0.001
Employment Density	-0.504	0.160	-3.150	0.002
Fuel Price	0.495	0.066	7.452	< 0.001
Vehicle Ownership	-0.859	0.199	-4.309	< 0.001
Park and Ride Capacity	0.002	0.013	0.162	0.871
Distance to CBD - Near	-1.689	0.272	-6.212	< 0.001
Winter	-0.136	0.014	-9.837	< 0.001
Spring	-0.121	0.015	-8.229	< 0.001
Summer	0.094	0.015	6.460	< 0.001
Action: Eliminate Age. **Rationale:** Large coefficient that does not align with previous estimates as seen in literature.				

Midday Off-Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Age				
	Coefficient	SE	t-value	p-value
(Intercept)	-13.073	5.421	-2.412	0.016
Service Quantity	0.570	0.030	19.270	< 0.001
Fare Price	-0.247	0.095	-2.591	0.010
Feeder Bus Connection Quality	-0.085	0.035	-2.447	0.015
Population Density	0.773	0.193	4.002	< 0.001
Gender - Female	0.023	0.008	2.874	0.004
Unemployment Rate	0.576	0.279	2.063	0.039
Income	1.497	0.397	3.776	< 0.001
Employment Density	0.015	0.136	0.111	0.911
Fuel Price	0.533	0.068	7.857	< 0.001
Vehicle Ownership	-0.683	0.183	-3.734	< 0.001
Park and Ride Capacity	0.010	0.014	0.685	0.494
Distance to CBD - Near	-0.868	0.238	-3.642	< 0.001
Winter	-0.139	0.014	-9.855	< 0.001
Spring	-0.126	0.015	-8.538	< 0.001
Summer	0.091	0.015	6.231	< 0.001
Action: Eliminate Employment Density. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

Midday Off-Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Employment Density				
	Coefficient	SE	t-value	p-value
(Intercept)	-12.821	5.212	-2.460	0.014
Service Quantity	0.571	0.030	19.306	< 0.001
Fare Price	-0.246	0.095	-2.579	0.010
Feeder Bus Connection Quality	-0.084	0.035	-2.441	0.015
Population Density	0.784	0.149	5.249	< 0.001
Gender - Female	0.023	0.008	2.863	0.004
Unemployment Rate	0.585	0.278	2.104	0.036
Income	1.485	0.389	3.816	< 0.001
Fuel Price	0.533	0.068	7.849	< 0.001
Vehicle Ownership	-0.692	0.174	-3.980	< 0.001
Park and Ride Capacity	0.010	0.014	0.703	0.482
Distance to CBD - Near	-0.875	0.219	-4.002	< 0.001
Winter	-0.139	0.014	-9.846	< 0.001
Spring	-0.126	0.015	-8.530	< 0.001
Summer	0.091	0.015	6.225	< 0.001
Action: Eliminate Park and Ride Capacity. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

Midday Off-Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Park and Ride Capacity				
	Coefficient	SE	t-value	p-value
(Intercept)	-13.069	5.143	-2.541	0.011
Service Quantity	0.571	0.030	19.334	< 0.001
Fare Price	-0.244	0.095	-2.566	0.010
Feeder Bus Connection Quality	-0.085	0.035	-2.449	0.014
Population Density	0.770	0.147	5.240	< 0.001
Gender - Female	0.022	0.008	2.811	0.005
Unemployment Rate	0.625	0.274	2.278	0.023
Income	1.515	0.383	3.953	< 0.001
Fuel Price	0.532	0.068	7.833	< 0.001
Vehicle Ownership	-0.692	0.173	-3.992	< 0.001
Distance to CBD - Near	-0.882	0.215	-4.100	< 0.001
Winter	-0.139	0.014	-9.847	< 0.001
Spring	-0.126	0.015	-8.515	< 0.001
Summer	0.091	0.015	6.210	< 0.001
All remaining variables demonstrate p-values less than the predetermined significance cut-off, therefore stepwise regression is completed. The restricted Midday Off Peak model therefore takes the final form as outlined in Equation 28. Complete results are further shown in Table 16.				

P.M. Peak Unrestricted Model Estimated Using Robust Standard Errors				
	Coefficient	SE	t-value	p-value
(Intercept)	-32.576	7.182	-4.536	< 0.001
Service Quantity	0.553	0.055	10.154	< 0.001
Fare Price	0.132	0.045	2.910	0.004
Feeder Bus Connection Quality	0.091	0.024	3.763	< 0.001
Population Density	0.390	0.187	2.084	0.037
Gender - Female	0.013	0.006	2.263	0.024
Unemployment Rate	2.662	0.246	10.806	< 0.001
Income	1.591	0.341	4.670	< 0.001
Age	2.935	1.508	1.947	0.052
Employment Density	0.259	0.144	1.801	0.072
Fuel Price	0.169	0.054	3.139	0.002
Vehicle Ownership	-0.352	0.185	-1.902	0.057
Park and Ride Capacity	-0.002	0.013	-0.119	0.906
Distance to CBD - Near	0.333	0.271	1.227	0.220
Winter	-0.099	0.011	-9.203	< 0.001
Spring	0.026	0.011	2.298	0.022
Summer	0.205	0.012	17.347	< 0.001
Action: Eliminate Age. **Rationale:** Large coefficient that does not align with previous estimates as seen in literature.				

P.M. Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Age				
	Coefficient	SE	t-value	p-value
(Intercept)	-21.083	4.502	-4.683	< 0.001
Service Quantity	0.559	0.054	10.377	< 0.001
Fare Price	0.134	0.045	2.974	0.003
Feeder Bus Connection Quality	0.094	0.024	3.910	< 0.001
Population Density	0.302	0.173	1.749	0.080
Gender - Female	0.016	0.006	2.617	0.009
Unemployment Rate	2.710	0.240	11.299	< 0.001
Income	1.512	0.338	4.471	< 0.001
Employment Density	0.343	0.138	2.487	0.013
Fuel Price	0.176	0.054	3.272	0.001
Vehicle Ownership	-0.331	0.181	-1.830	0.067
Park and Ride Capacity	-0.001	0.013	-0.097	0.923
Distance to CBD - Near	0.481	0.264	1.821	0.069
Winter	-0.100	0.011	-9.239	< 0.001
Spring	0.024	0.011	2.188	0.029
Summer	0.204	0.012	17.265	< 0.001
Action: Eliminate Park and Ride Capacity. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

P.M. Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Park and Ride Capacity				
	Coefficient	SE	t-value	p-value
(Intercept)	-20.945	4.479	-4.676	< 0.001
Service Quantity	0.562	0.054	10.448	< 0.001
Fare Price	0.135	0.045	2.993	0.003
Feeder Bus Connection Quality	0.094	0.024	3.902	< 0.001
Population Density	0.307	0.171	1.790	0.074
Gender - Female	0.016	0.006	2.650	0.008
Unemployment Rate	2.703	0.238	11.344	< 0.001
Income	1.506	0.337	4.467	< 0.001
Employment Density	0.342	0.136	2.503	0.012
Fuel Price	0.176	0.054	3.292	0.001
Vehicle Ownership	-0.339	0.180	-1.887	0.059
Distance to CBD - Near	0.485	0.258	1.878	0.060
Winter	-0.100	0.011	-9.235	< 0.001
Spring	0.024	0.011	2.178	0.030
Summer	0.204	0.012	17.326	< 0.001
All remaining variables demonstrate p-values less than the predetermined significance cut-off, therefore stepwise regression is completed. The restricted PM Peak Model therefore takes the final form as outlined in Equation 29. Complete results are further shown in Table 17.				

Evening Off-Peak Unrestricted Model Estimated Using Robust Standard Errors				
	Coefficient	SE	t-value	p-value
(Intercept)	-45.478	13.414	-3.390	0.001
Service Quantity	0.488	0.027	17.884	< 0.001
Fare Price	0.118	0.046	2.568	0.010
Feeder Bus Connection Quality	-0.060	0.045	-1.338	0.181
Population Density	0.567	0.243	2.329	0.020
Gender - Female	-0.006	0.019	-0.344	0.731
Unemployment Rate	2.346	0.360	6.513	< 0.001
Income	3.130	0.552	5.672	< 0.001
Age	5.397	2.579	2.092	0.036
Employment Density	0.264	0.206	1.284	0.199
Fuel Price	0.204	0.086	2.366	0.018
Vehicle Ownership	-1.493	0.219	-6.809	< 0.001
Park and Ride Capacity	0.000	0.017	0.023	0.982
Distance to CBD - Near	1.065	0.386	2.758	0.006
Winter	-0.131	0.017	-7.679	< 0.001
Spring	0.103	0.018	5.789	< 0.001
Summer	0.307	0.019	16.296	< 0.001
Action: Eliminate Age. **Rationale:** Large coefficient that does not align with previous estimates as seen in literature.				

Evening Off-Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Age				
	Coefficient	SE	t-value	p-value
(Intercept)	-23.786	6.788	-3.504	< 0.001
Service Quantity	0.493	0.027	18.093	< 0.001
Fare Price	0.123	0.046	2.671	0.008
Feeder Bus Connection Quality	-0.056	0.045	-1.251	0.211
Population Density	0.404	0.216	1.876	0.061
Gender - Female	-0.003	0.017	-0.169	0.866
Unemployment Rate	2.430	0.357	6.807	< 0.001
Income	2.917	0.530	5.507	< 0.001
Employment Density	0.413	0.179	2.307	0.021
Fuel Price	0.210	0.086	2.440	0.015
Vehicle Ownership	-1.430	0.217	-6.583	< 0.001
Park and Ride Capacity	0.001	0.017	0.089	0.929
Distance to CBD - Near	1.294	0.374	3.457	0.001
Winter	-0.132	0.017	-7.728	< 0.001
Spring	0.101	0.018	5.674	< 0.001
Summer	0.306	0.019	16.225	< 0.001
Action: Eliminate Park and Ride Capacity. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

Evening Off-Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Park and Ride Capacity				
	Coefficient	SE	t-value	p-value
(Intercept)	-23.659	6.772	-3.494	< 0.001
Service Quantity	0.493	0.027	18.118	< 0.001
Fare Price	0.123	0.046	2.680	0.007
Feeder Bus Connection Quality	-0.057	0.045	-1.254	0.210
Population Density	0.407	0.214	1.907	0.057
Gender - Female	-0.003	0.017	-0.185	0.854
Unemployment Rate	2.423	0.354	6.848	< 0.001
Income	2.900	0.529	5.481	< 0.001
Employment Density	0.411	0.179	2.300	0.022
Fuel Price	0.209	0.086	2.433	0.015
Vehicle Ownership	-1.423	0.216	-6.581	< 0.001
Distance to CBD - Near	1.279	0.369	3.464	0.001
Winter	-0.132	0.017	-7.730	< 0.001
Spring	0.101	0.018	5.669	< 0.001
Summer	0.306	0.019	16.252	< 0.001
Action: Eliminate Gender - Female. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

Evening Off-Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Gender - Female				
	Coefficient	SE	t-value	p-value
(Intercept)	-23.503	6.704	-3.506	< 0.001
Service Quantity	0.494	0.027	18.283	< 0.001
Fare Price	0.124	0.046	2.693	0.007
Feeder Bus Connection Quality	-0.057	0.045	-1.253	0.210
Population Density	0.412	0.212	1.946	0.052
Unemployment Rate	2.412	0.352	6.846	< 0.001
Income	2.876	0.526	5.472	< 0.001
Employment Density	0.408	0.177	2.308	0.021
Fuel Price	0.208	0.086	2.427	0.015
Vehicle Ownership	-1.415	0.214	-6.600	< 0.001
Distance to CBD - Near	1.266	0.366	3.459	0.001
Winter	-0.132	0.017	-7.730	< 0.001
Spring	0.101	0.018	5.666	< 0.001
Summer	0.306	0.019	16.242	< 0.001
Action: Eliminate Feeder Bus Connection Quality. **Rationale:** P-value greater than predetermined cut-off, largest p-value in current model output.				

Evening Off-Peak Unrestricted Model Estimated Using Robust Standard Errors After Elimination of Feeder Bus Connection Quality				
	Coefficient	SE	t-value	p-value
(Intercept)	-23.063	6.666	-3.460	0.001
Service Quantity	0.493	0.027	18.115	< 0.001
Fare Price	0.121	0.046	2.639	0.008
Population Density	0.410	0.210	1.950	0.051
Unemployment Rate	2.388	0.350	6.828	< 0.001
Income	2.824	0.522	5.414	< 0.001
Employment Density	0.403	0.175	2.300	0.022
Fuel Price	0.197	0.086	2.306	0.021
Vehicle Ownership	-1.407	0.213	-6.593	< 0.001
Distance to CBD - Near	1.267	0.362	3.496	< 0.001
Winter	-0.132	0.017	-7.761	< 0.001
Spring	0.099	0.018	5.621	< 0.001
Summer	0.306	0.019	16.198	< 0.001
All remaining variables demonstrate p-values less than the predetermined significance cut-off, therefore stepwise regression is completed. The restricted Evening Off Peak model therefore takes the final form as outlined in Equation 30. Complete results are further shown in Table 17.				