UNDERSTANDING COST DIFFERENCES OF TRANSIT SERVICES BETWEEN URBAN AND RURAL CANADA

by

Siew Ten Ong

Thesis submitted in partial fulfillment of the

requirements for the Degree of

Bachelor of Science with

Honours in Economics

Acadia University

May, 2021

This thesis by Siew Ten Ong

is accepted in its present form by the

Department of Economics

as satisfying the thesis requirements for the degree of

Bachelor of Science with Honours

Approved by the Thesis Supervisor

Dr. Justin Beaudoin	Date

Approved by the Head or Director of the Department or School

Dr. Burc Kayahan	Date

Approved by the Chair, Senate Honours Committee

Dr. Joseph Hayes	Date

I, Siew Ten Ong, grant permission to the University Librarian at Acadia University to reproduce, loan or distribute copies of my thesis in microform, paper or electronic formats on a non-profit basis. I, however, retain the copyright in my thesis.

____

Signature of Author

____

Date

Acknowledgements

I would like to express my deepest appreciation to my supervisor Dr. Justin Beaudoin for providing invaluable guidance throughout this thesis.

I would like to thank Dr. Xiao Ting Wang for her helpful suggestions to further improve this thesis. Thank you Dr. Kayahan for his comprehensive assistance and encouragement when I was pursuing my Economics degree at Acadia.

Lastly, I would like to thank the entire Economics Department at Acadia University.

TABLE OF CONTENTS

List of Tables

Abstract

In 2021, the Canadian government announced a new policy regarding public transit funding that has put rural transit projects in a disadvantageous position as the funding guidelines ignore the differences of transit demand and supply between urban and rural Canada. In other words, urban areas that have higher population density and more job opportunities, will also have more complete and accessible public transit service that would attract more transit users, in contrast to rural areas. Hence, this thesis focuses on identifying the potential differences of ridership demand and supply, and the required subsidy, between urban and rural Canadian transit operations by examining the empirical relationship on: (1) the factors affecting transit demand, (2) the factors affecting transit supply costs, and (3) the factors affecting the required subsidy per trip from the public transit users' approach. A pooled cross-sectional dataset comprised exclusively of transit demand factors from 1996 to 2016 was used for the one-stage OLS regression for all three models. It should be noted that none of the existing literature has involved a cross-sectional comparison of public transit across Canadian cities, and none of these studies have focused on the rural context.

The results show that there are differences in ridership demand and supply, and the required subsidy, between urban and rural transit operations in Canada. Accounting for their respective difference in socioeconomic, built environment, and ridership levels this thesis provides evidence for policymakers that rural and urban areas may need different transit funding policies.

1. Introduction

In February 2021, Prime Minster Justin Trudeau announced a $14.9 billion permanent transit fund, in which $5.9 billion will be disbursed on a project-by-project basis. This $5.9 billion will not be divided up between provinces but will instead be put into a pot that can be accessed whenever a project is ready (Jones, 2021). While the remaining $9 billion is considered as a permanent transit fund of $3 billion per year starting in the year 2026, money from that fund will be earmarked following consultations with provinces, territories, municipalities and Indigenous communities. One of the key issues about this funding is that it is unclear how much money is being distributed for different transit projects across Canada. Therefore, a clear understanding of the factors affecting transit demand and transit supply is helpful to predict the feasibility of prospective transit projects across regions of Canada, and to decide which transit projects should be supported.

Most previous studies that have been conducted in the Canadian context did not involve a cross-sectional comparison of public transit across Canadian cities, and none have focused on the rural context (Boisjoly et al., 2018; Börjesson et al., 2020; Diab et al., 2020). This thesis uses empirical analysis to determine: (1) factors affecting public transit ridership, (2) factors affecting public transit operating costs, and (3) factors affecting per-trip subsidy for a pooled cross-sectional dataset consisting of a total of 103 urban and rural areas in Canada from 1996 to 2016.

The results shows that the ridership per capita in rural Canada is about 50% lower than in urban Canada. Rural Canada has a higher cost per trip, about 17% more than urban Canada. However, this number is reduced to 1% when controlling for economies of scale. Lastly, the required per-trip subsidy in rural Canada is 21% more than in urban Canada, though this difference is reduced to 1.6% when ridership per capita is controlled for. From the results, the impact of operating in a rural region drops significantly for both cost per trip and subsidy per trip after the model has controlled for economies of scale. This suggests that more transit users are needed to provide a sustainable public transit service in rural Canada.

This paper starts with a literature review of transit demand, transit supply, and transit subsidies. This is followed by a methodological section describing the data used and modelling approach. Finally, model results and policy implications are discussed.

2. Literature Review

2.1 Ridership Demand

Many studies have investigated the factors affecting transit ridership from different perspectives and at different scales (Diab et al., 2020). While the existing literature focus more on estimating changes in ridership within a given city, Diab et al. (2020) point out that only a small number of studies uses a city-wide transit system as the unit of analysis. In fact, most of the recent literature is at the multi-city level and is conducted in the US context (Boisjoly et al., 2018; Haire, 2009; Lee and Lee, 2013; Taylor et al., 2009; Thompson and Brown, 2006). In the Canadian context, both Boisjoly et al. (2018) and Diab et al. (2020) investigate the determinants of transit ridership across Canadian cities in 2018 and 2020, respectively; there is a 20-year gap from the last multi-city level study which was done by Kohn in 2000. Interestingly, none of the existing studies involve a comparison of demand for transit ridership between urban and rural Canada.

The literature has examined the factors affecting transit ridership at the system, station, and individual levels within a city or across different cities (Guerra and Cervero, 2011). Although system-level studies provide both across-system and over-time transit data, Guerra and Cervero (2011) argue that cross-sectional system-level studies may produce biased and inconsistent coefficient estimates due to the correlation between omitted relevant variables with both dependent variables and other independent variables in the model. Their argument is supported by the opposite direction of the estimated income variable in Taylor et al.'s (2009) and Gomez-Ibanez's (1996) studies; Taylor et al. (2009) find that a 1% increase in median household income is associated with a 0.65% increase in ridership in 265 US urbanized areas, while Gomez-Ibanez (1996) observes a 0.75% decline in ridership as real income increased by 1%. In addition, Taylor et al. (2009) admit under-specification bias is commonly found in cross-sectional models; however, Taylor et al. propose that this bias can be reduced through testing a large number of relevant independent variables that are helpful in explaining variation in dependent variables. Both Boisjoly et al. (2018) and Diab et al. (2020) use a stepwise process to check if their models remained stable after adding and removing independent variables.

According to Thompson and Brown (2006), the literature divides factors affecting transit ridership into two categories: external factors and internal factors. External factors are factors which are beyond the control of transit operator, such as the built environment and socioeconomic factors, for example. Some of the most used external factors in explaining transit ridership includes total population, population density, income, household rent, housing price, unemployment rate and labour force participation rate. Internal factors, on the other hand, are transit service factors that are under the transit operator's control (Thompson and Brown, 2006). These include transit fares, service hours, number of routes and others. In the literature, most studies have incorporated both the external and internal factors as the independent variables in their models to study transit ridership. It is important for researchers to account for the causal and two-directional relationship between transit supply and transit demand as Taylor and Fink (2013) highlight that "analyses concluding that transit service levels largely explain transit ridership levels tell us little about the underlying causality of transit use and can produce biased results" (p. 18). This argument is supported by the biased and inconsistent coefficient estimates of Lee and Lee's (2013) & Taylor and Fink's (2013) simple one stage ordinary least squares (OLS) regression model when this two-way causal relationship is not accounted in the model.

Different regression techniques are used by the researchers to study the factors affecting transit use. To account for the endogeneity of transit demand and supply that was mentioned earlier, one of the commonly used regression techniques is the two-staged least square (2LS) simultaneous equation models (Diab et al., 2020; Taylor et al., 2009).Diab et al. (2020) use both external factors and internal factors from 103 transit agencies in Canada between 2002 and 2016 as the explanatory variables to identify the variables that affect ridership changes at the transit agency level using longitudinal analysis. Among the external factors and the internal factors in their model, Diab et al. (2020) find that the predicted vehicle revenue hours (which measures the transit supply) is more strongly associated with transit ridership than other variables, and their results are supported by previous findings (Boisjoly et al., 2018; Taylor at al., 2009). These three studies observe a positive significant impact of the predicted vehicle revenue hours on transit ridership.

Boisjoly et al. (2018) conduct a 14-year longitudinal analysis of the determinants of public transport ridership in 25 North American cities. They find that internal factors like vehicle revenue kilometers and car ownership are the key determinants of transit ridership (Boisjoly et al., 2018). In their study, Boisjoly et al. (2018) state that "external factors such as unemployment rate are not significant as was found by Guerra and Cervero (2011) & Taylor et al. (2009)" (p. 439). Diab et al. (2020) also find it difficult to identify a relationship between ridership and unemployment rate.

Taylor et al. (2009) conduct a cross-sectional analysis of transit use in 265 urban areas in the US; they construct a two-stage simultaneous equations model to account for the endogeneity of transit supply on ridership. A wide number of internal and external factors are tested in their study, and their results show that most of the variation in transit ridership among urbanized areas can be explained by external factors (Taylor et al., 2009). Still, their results show that service frequency and fare levels contribute to about a quarter of the observed variance in per capita transit patronage across US urbanized areas when controlling for the fact that public transit use is highly correlated with urbanized area size (Taylor et al., 2009).

Nutley (1996) points out that “the study of rural areas by transport researchers appears a very much of a minority interest in relation to the overall body of transport literature” (p. 93). Studies focusing on the rural context are conducted to understand the factors affecting transit ridership in rural areas of a given country. In his study, Nutley highlights that the rural environments having a lower total population is the main cause of reduced transit service, which in turns reducing transit ridership in rural regions. Due to inadequacy of transit supply, the car ownership rates in rural areas are relatively high (Nutley, 1996). Stringham (1982) and Cervero (1994) observe similar trends, as their studies show that access trips are mainly done by private vehicles for commuters living beyond 1 mile of a suburban rail station in Toronto and the San Francisco Bay Area (Santoso et al., 2012). In Japan, the decline in bus ridership in rural areas is dominated by the prevalence of private cars (Sakai et al., 2010). It should be noted that these studies are almost exclusively outside of the North American context and none of the existing studies have involved a cross-sectional comparison of public transit across Canadian cities, and none have focused on the rural context.

Using only the external factors and a rural dummy variable, the objective of this thesis is, therefore, to investigate the partial effect of each external factor on the demand for transit ridership, while holding other variables constant at their mean values. By inserting a rural dummy variable, this thesis aims to investigate the difference in ridership demand between urban and rural areas across a total of 103 urban and rural areas in Canada from the years 1996 to 2016.

2.2 Transit Supply

Public transit supply, usually measured in terms of vehicle revenue hours or vehicle revenue miles, is commonly found in transit demand studies since it has a significant impact on public transit ridership. In the first stage of two-stage least square (2SLS) ridership demand models, Diab et al. (2020) and Taylor et al. (2009) construct different regression models to identify the variables that affect public transit supply. Some of the key predictors of vehicle revenue hours include total population, service area population, and total direct operating expenses (Diab et al., 2020; Mattson, 2017; Taylor et al., 2009).

After testing different variables, both Diab et al. (2020) and Taylor et al. (2009) arrive at a simple two-variable model for predicting total vehicle revenue hours. Although the variables used in their models slightly differ from one another, both models explain more than 80% of the variation in total vehicle revenue hours. Diab et al. (2020) uses total population and total direct operating expenses, and their study shows a 10% increase in total population and total direct operating expenses is associated with a 5.5% and a 4.7% increase in predicted revenue vehicle hours, respectively. In contrast, Taylor et al. (2009) use US urbanized area population and the percentage of the population vote for the Democrat in the 2000 presidential election to predict vehicle revenue hours. According to Taylor et al. (2020), "[d]emocratic-leaning areas are more likely to support public expenditures in transit subsidies," which in turn increases transit service supply (p. 69). Compared to Diab et al (2020), Taylor et al. (2009) observe a larger impact of urbanize area population (11.5%) on vehicle revenue hours.

Although public transit supply is strongly correlated to transit operating costs, however, there is very little evidence shown in the existing studies as to the extent in which one affects the other. Skinner (1981) points out that "transit supply parameters, such as vehicle miles, vehicle hours, and employees by category, are the major determinants of operating cost estimates that, in turn, are a principal factor in addressing operating feasibility implicitly or explicitly" (p. 24). Sale and Green (1979) analyze the US public transit operators' data from 1967 to 1977 to identify the main cause of rapid-rising operating costs in the US public transit service during that period. Their findings reveal that labor accounts for over 80% of total operating cost when fringe benefits and pensions are included. The combination of rapidly growing labor compensation and low average labor productivity is the main reason for the escalated operating costs during that period. Interestingly, service expansions account for only 1% of the total operating costs in their study. Kain and Liew (1999) provide estimates of operating costs per boarding and per passenger mile for Houston's bus operator and San Diego's bus and light rail operators 1968 to 1996; their results imply that operating costs can be reduced through eliminating low productivity routes and better labor arrangement. The implications from these analyses are limited because the models did not control for many other explanatory variables.

A number of studies have stressed the importance of economies of scale in public transit service production. Berechman (1983) identifies "the principal differences between studies are the specification of cost function, the set of independent variables used, and the specific output measures" (p. 8). The most used cost function forms in regression analysis include quadratic, linear, and logarithm functions. The linear cost model is used most frequently in the earlier studies (Koshal, 1970; Wabe and Coles, 1975). However, Berechman (1983) argue that the conclusions drawn from these linear cost model studies are limited since the cost structures of different-sized transit agencies are treated the same through a single cost function. To address this issue, Iseki (2018) uses a contracting variable to prove that cost per vehicle hours is better represented by different cost functions for different agency size groups. Both Izeki (2208) and Giuliano (1980) uses a quadratic function for the regression equation: Izeki (2008) finds diseconomies of scale for any transit agency size with any level of contracting while Giuliano (1980) find diseconomies of scale for medium size agencies and economies of scale for small and large agencies. Iseki's results suggest that the contracting variable might be an important variable, implying that future studies should construct different transit cost structures for different size transit agencies.

Berechman and Giuliano (1985) present that different output measures used in the cost model can affect the existence of economies of scales. Using a translog cost model to examine the existence of economies of scale in Israel's bus industry, Berechman and Giuliano (1985) found economies of scale with respect to revenue passengers, while they found diseconomies of scale with respect to vehicle miles provided as a measure of scale. According to Berechman and Giuliano (1985), "if measured on the basis of passenger-trips one would expect increasing returns if the number of trip possibilities increases more than proportionately with service increases. Under these conditions, ridership should increase more than proportionately as well" (p. 320). In fact, increasing returns to scale are frequently observed in studies using demand-related measures, such as passenger-trips and passenger-miles (Williams and Hall, 1981; Berechman, 1983). In contrast, most studies based on technical measures such as vehicle-miles and vehicle-hours have reported constant returns to scale (Berechman and Giuliano, 1985). Thus, mixed results produced from different approaches left the existence of economies of scale in transit service inconclusive.

In summary, most of the previous studies of economies of scale in public transit service focus more on the functional form of the cost models and the output measures, while a majority of them ignore the role of demand-settings variables. Using a large pooled cross-sectional dataset with a wide range of built environment and socioeconomic factors, plus the use of number of linked trips as the output measures, this thesis focus examines how demand factors affect cost per trip.

2.3 Public Transit Subsidy

In 2006, the Government of Canada introduced a public transit subsidy in the form of an income tax credit. This 15%, non-refundable income tax credit was promoted as a direct means of improving ridership on public transit systems and was discontinued in 2017. Three studies are conducted to examine the effectiveness of the public transit tax credit (Finance Canada, 2012; Chandler, 2014; Rivers and Plumptre, 2018). Using aggregate transit ridership data from seven major Canadian cities and controlling for city-specific events that may have affected transit use, Chandler's (2014) study shows no evidence that this targeted tax credit had significant effects in promoting transit ridership. In contrast, Finance Canada (2012) uses CUTA transit annual ridership data from 2001 to 2010 to compare the before (2001-2005) and after (2006-2010) public transit ridership rate; the study finds that the national transit ridership increased at an annual average rate of 1.9% from 2001 to 2005, compared to 2.9% from 2006 to 2010 (Rivers and Plumptre, 2018). Although the introduction of public transit tax credit seemed to have a positive effect on transit ridership, Chandler (2014) questions its effectiveness. In fact, the public transit ridership increased at a decreasing rate, and the resulting trend did not meet the expected trend that the public transit use should increase proportionally as the tax credit claim increases. On the other hand, the growth in transit ridership was at its lowest when the increase in credit claiming was at its highest (Chandler, 2014). According to Rivers and Plumptre (2018), these studies use a small number of observations and did not control for a wide range of variables that could also influence public transit use. Thus, these studies fail to identify the causal evidence of the effect of the public transit subsidy.

Controlling for detailed demographic characteristics of individuals as well as how the characteristics of the urban environment evolve over time, Rivers and Plumptre's (2018) quasi-experimental empirical analysis reveals that the public transit tax credit (PTTC) increased the mode share of public transit by a quarter to one percentage points. When excluding rural areas, the tax credit coefficient increases to 0.6%, suggesting that most of the increase in transit ridership due to the PTTC is contributed by urban commuters. Their study implies inequalities of public transit tax credit policy between urban and rural commuters in Canada as this policy ignored the fact that public transportation is not equally accessible across Canada. In other words, urban commuters who have greater access to public transportation are more likely to take advantage of this policy, in contrast to rural commuters who suffer from inadequate and incomplete transit supply services (Rivers and Plumptre (2018). Furthermore, Chandler (2014) highlights that the tax credit policy explicitly excludes low-income individuals who do not owe any income tax. Thus, subsidy in the form of a tax credit could incur a social cost and policymakers should research carefully before implementing such policies.

In their paper, Alfa and Clayton (1986) state that "there are two distinct areas to which transit subsidies are applied, one being the costs of operation of the transit system and the second being capital costs of initial purchases, replacement and improvement of the infrastructure of the transit system. These subsidies are generally referred to public transit operating subsidies (PTOS) and public transit capital subsidies (PTCS), respectively. The complementarity between PTOS and PTCS is recognized, as high investments in one area will subsequently reduce the costs incurred in the other area" (p. 224). Restricting to PTOS analysis, many papers argue PTOS may increase operational cost and cause transit operators to rely more on financial subsidy, leading to operational inefficiency while putting more burden on government (Yang et al. 2020). However, Yang et al. (2020) emphasize the difference between cost-based subsidy and ridership-based subsidy, which they find public transit operators had no monetary motivation to improve the quality and quantity of transit service, as ridership was not the source of profit. According to Yang et al. (2020), previous studies of PTOS are based on the cost-based subsidy, which explains the lack of operational efficiency in previous findings. Furthermore, Yang et al. (2020) demonstrate that cost-based subsidy is more economically sustainable to cover the upfront capital and operating costs of transit investment or expansion projects. Apart from this, they highlight the role of economies of scale to public transit investment. Coulombel & Monchambert (2019) study the effect of increasing levels of demand on the provision of service quality (frequency, vehicle size/capacity) and on economies of scale. As seen in their results, urban public transportation operations are characterized by economies of scale only up to a certain threshold demand level. If passing the critical demand level, the severity of crowding causes the marginal social cost of an additional passenger to exceed the average social cost, implying diseconomies of scale (Coulombel and Monchambert, 2019).

The existing literature of rural subsidization focus more on improving the existing model specification for better computation of optimal transit subsidies. Focusing specifically on whether urban transit subsidies should be reduced, Parry and Small (2009) construct an aggregate-level model that contains the supply and demand features most essential to measuring the factors that motivate transit subsidies. Compared to supply features, their model employs only a few demand factors, such as wait costs and transit modes. However, Parry and Small (2009) do not find strong evidence to support reduction in urban transit subsidies based on their results. Cooke and Behrens

(2017) examine the relationship between population density and public transport subsidization for developing cities and find that their result is not as clear as other studies. While other studies find a strong causal link, Cooke and Behrens (2017) states that "the connection to the subsidization levels of public transport services in developing cities is both less negative and almost uncorrelated" (p. 3007).

It should be noted that none of the existing studies uses a cross-sectional approach to compare public transit subsidization across rural and urban areas. While the transit supply factors are mostly examined and used in most public transit subsidization, this thesis evaluate the partial impact of various transit demand side factors on per-trip subsidy. On top of that, this thesis controls for the economies of scale on subsidy per trip using a transit demand factor approach.

3. Dataset Information

3.1 Data

This thesis uses two sources of data: (1) Canadian Urban Transit Association (CUTA) and (2) Census Canada. CUTA provides data for 103 transit agencies across Canada from 1996 to 2016. Three of the transit service variables from CUTA were used, which are total regular service passenger trips, total operating revenues, and total direct operating expenses. The dependent variables for the models were obtained after performing simple calculations using these three variables, and the methodology section covers more detailed information on the computation of the dependent variables.

Census Canada provides built environment, socioeconomic and demographic data at five-year intervals (1996, 2001, 2011 and 2016). Census variables for the regression analysis are selected by applying theoretical consideration: census variables that can explain the variation in transit ridership are used. For example, census variables such as marital status, type of mother tongue, and major of study have extremely weak or almost no impact on transit ridership in theory, and therefore these variables are excluded. A complete list of the census variables used in the regression can be found in Table 1. Using the list of transit systems from the CUTA Transit Fact Book as a guide, Census subdivisions (CSDs) falling within the service area of each transit agency were identified. Then, Census data corresponded to these CSDs were extracted and merged at the transit agency level from 1996 to 2016.

Table 1: Summary of the independent variables used in the regression

Variable	Obs.	Mean	Median	Standard Deviation	Min	Max
Total Population	317	425,716	76,407	122,9764	1774	854,8919
% Working population (age 15 to 64 years)	316	69.09	69.10	4.09	54.73	86.55
% Senior (age 65 and above)	316	12.72	12.34	5.52	1.06	33.37
% Postsecondary students	315	49.10	50.69	10.43	15.62	74.26
Employment rate	315	62.13	61.40	7.70	44.20	86.90
Average household income	316	72,729.63	69,319.50	26,176.36	34,356	210,417
Average owner's major payment	315	1,986.23	1020	16,328.23	520	290,800
% of tenants spending >=30% of its income on rent	315	15,426.06	135	71,577.91	18.50	803,135
Average monthly tenant rent	315	898.32	770	834.82	441	11,552
% of employed population who use private vehicle to work	315	76.03	78.65	10.22	32.07	91.18
% of employed population who rideshare to work	315	7.26	7.09	1.95	2.94	13.30
% of employed population who walk to work	315	6.98	5.51	6.08	1.64	55.17
Year(dummy), 1996 = 1	507	0.25	0	0.43	0	1
Year(dummy), 2001 = 1	507	0.25	0	0.43	0	1
Year(dummy), 2011 = 1	507	0.25	0	0.43	0	1
Rural = 1 (dummy)	317	0.28	0	0.45	0	1

3.2 Data Limitations

Each data source used in this thesis has its own limitations. For Census Canada, the census year of 2006 was excluded from the dataset since most of the 2006 census data are not available at the CSD level, and it is difficult to identify which transit agency these census data fall in. For example, the Alberni-Clayoquot is a regional district that consists of data from different communities (e.g. cities, district municipalities, and Indian Reserves), but only Port Alberni was found on the CUTA transit system. According to Statistics Canada (2012), there is a total population of 17,548 in Port Alberni; in contrast, the Alberni-Clayoquot has a total population of 30,664. The counts were almost doubled. Therefore, Alberni-Clayoquot has larger counts for the census variables than Port Alberni itself, and this will overestimate the dependent variables. To maintain consistency in the measurement of variables over time, observations from 2006 were excluded from the analysis.

For the CUTA data, the transit agencies and the service areas for British Columbia are aggregated into British Columbia Municipal Systems, making it impossible to extract the transit data for individual service areas. The aggregation of the service area data is not useful in observing the difference between rural and urban Canada or the effects of differences in transit operator size. Hence, British Columbia transit operators were excluded from the analysis.

Apart from this, there are many missing values and inconsistent information for the chosen CUTA variables. Diab et al. (2020) who use CUTA transit data in their study also identify this issue. After reviewing the merged census and CUTA data, variables that have below average number of observations (<150 observations) were removed as the statistical software will omit the row with missing values. Hence, removable of variables with below average observations helps to maximize the total number of observations of the regression, thereby producing more robust results. Table 1 summarizes the variables used in this thesis and their corresponding summary statistics.

4. Methodology

4.1 Model Development

The goal of this thesis is to produce three models to conduct empirical analysis on: (1) the factors affecting transit demand, (2) the factors affecting transit supply costs, and (3) the factors affecting the required subsidy per trip. Most importantly, this thesis focus on identifying the potential differences of ridership demand and supply, and the required subsidy, between urban and rural Canadian transit operations.

The independent variables and the dataset used in this thesis are the same for all three models, except that the dependent variable of model I (ridership) is incorporated as an independent variable in both models II (operating cost per trip) and III (subsidy per trip). All non-dummy variables and variables that are measured in terms of counts are transformed into the natural logarithm form since the log-log transformation allows for the interpretation of regression coefficients in terms of elasticities. Previous studies that investigate transit ridership have also applied the log-log transformation (Boisjoly et al., 2018; Diab et al., 2020, Guerra and Cervero, 2011; Taylor et al., 2009).

This thesis uses a one-step OLS regression method to estimate the partial impact of each independent variable on the dependent variable while holding other independent variables at their mean value. The next sections will look into each model specification.

4.2 Model I Specification

By regressing ridership per capita on a range of socioeconomic and demographic variables from 1996 to 2016 using a one-step OLS regression method, the goal of Model I is to estimate the demand for public transit in Canada. Ridership per capita is not directly observed from the CUTA transit data, and it is calculated by dividing annual total regular service passenger trips with annual total population of each region.

CUTA (2016) defines a passenger trip as “a linked trip, riding one way from origin to final destination”; CUTA (2016) defines total regular service passenger trips as “all passenger trips for which the fare system applied” (p. 253). Thus, the ridership per capita in this thesis is measured in terms of linked trips. Taylor et al. (2009) and Diab et al. (2020) highlight the importance of using linked trips to predict ridership since linked trips provide a more robust measure of transit ridership in contrast to unlinked trips.

This thesis uses only external factors to estimate ridership per capita, and hence this eliminates the issue arising from the endogeneity of transit supply and demand as mentioned in previous studies (Taylor and Fink, 2013; Lee and Lee, 2013). Model I is shown in equation (1):

$$\log(ridership\ per\ capita)_i = b_0 + b_j X_i + b_{j+1} D_{rural} + b_{j+2}(D_{rural} * X_i) + e_i \qquad (1)$$

where j and i denote regression coefficient and independent variable, respectively. The regression coefficients, as represented by j in equation 1, estimates the partial effect of a particular independent variable from a vector of independent variables (X_i) on ridership per capita, while holding other independent variables at their mean values. For instance, the first independent variable in Table 1 takes i = 1, and so on. The rural dummy variable indicates whether the region is rural or urban. According to Statistics Canada (2016). “an 'urban area' was defined as having a population of at least 1,000 and a density of 400 or more people per square kilometre. All territory outside an urban area was defined as rural area”. The definition of rural area in this thesis is different than Statistics Canada’s, where regions that satisfy both the following conditions are considered as rural area: 1) total population more than or equal to 30,000 and 2) population density of more than 400 per square kilometre. This difference in definition of rural areas is required to adjust for the fact that this study using Census data at the CSD level, which is a larger geographical area than used in the definition by Statistics Canada. Accounting for these two restrictions, 88 of the observations in the dataset are considered as rural areas, while the other 229 observations are classified as urban areas (see Table 2).

Table 2: Rural and urban areas observations

Rural (Yes = 1)	Frequency	Percent
0	229	72.24
1	88	27.76
Total	317	100.00

The fourth term in equation (1) is the interaction term between the rural dummy and the independent variable (X_i), and this interaction term helps to examine whether the impact of the rural dummy on ridership per capita depends on the corresponding independent variable (X_i) while holding other independent variables constant. The error term (e_i), represents the effect of the variables that were omitted from the regression (variables that are not in Table 1). It should be noted that the independent variables in equation (1) must be uncorrelated with the error term, as otherwise the regression coefficient may become biased and inconsistent.

4.3 Model II Specification

The goal of model II is to determine (1) the factors affecting costs of supplying public transit in Canada, (2) the difference in public transit cost between rural and urban Canada, and (3) the impact of economies of scale on public transit supply cost. In order to estimate the cost of supplying public transit, cost per trip is calculated by dividing total direct operating expenses by total regular service passenger trips. As in model I there are two individual equations that are developed to separate the effect of economies of scale. Equation (2) is similar to equation (1), as cost per trip is regressed on a range of socioeconomic and demographic variables, but this time with a different dependent variable. Ridership in logarithm form is incorporated into equation (3). Also, an interaction term is included to account for the dependency of other independent variables with the rural dummy on cost per trip.

$$\log(cost\ per\ trip)_i = b_0 + b_j X_i + b_{j+1} D_{rural} + b_{j+2}(D_{rural} * X_i) + e_i \tag{2}$$

$$\log(cost\ per\ trip)_i = b_0 + b_j X_i + b_{j+1} D_{rural} + b_{j+2}(D_{rural} X_i) + b_{j+3} \ln{(ridership)_i} + e_i \tag{3}$$

As in model I, the log-log transformation also applies to non-dummy variables and variables measured in counts.

4.4 Model III Specification

The goal of model III is to determine (1) the factors affecting public transit operating subsidies in Canada, (2) the difference between public transit operating subsidies in rural and urban Canada, and (3) the impact of economies of scale on operating subsidies. In order to estimate the cost of supplying public transit, cost per trip is calculated by dividing total direct operating expenses with total regular service passenger trips. Similar to model I, subsidy per trip is regressed on a range of socioeconomic and demographic variables but this time, two individual equations are developed to separate the effect of economies of scale. The baseline model is shown in equation (4), and the level of ridership is incorporated into equation (5). Also, an interaction term is included to account for the dependency of other independent variables with the rural dummy on cost per trip.

$$\log(subsidy\ per\ trip)_i = b_0 + b_j X_i + b_{j+1} D_{rural} + b_{j+2}(D_{rural} * X_i) + e_i \quad (4)$$

$$\log(subsidy\ per\ trip)_i = b_0 + b_j X_i + b_{j+1} D_{rural} + b_{j+2}(D_{rural} * X_i) + b_{j+3} \ln(ridership)_i + e_i \quad (5)$$

As in model I and II, the log-log transformation also applies to non-dummy variables and variable measured in counts.

5. Results

5.1 Model I Results

This section will present the empirical results of Model 1 and equation (1), which explores the effects of the independent variables on ridership per capita. The dependent variable in Model I is the log of the ridership per capita, which is used to estimate the transit demand. Table 3 presents the one-stage OLS results of Model I; the regression coefficients predict the impact of a certain independent variable on the expected ridership per capita while holding other independent variables constant. We assume that other variables do not change in order to allow for an evaluation of the partial variation in a dependent variable due to variation in a particular independent variable, while other variables do not change.

As seen in Table 3, Model I is based on 251 observations and explains about 56% of the variation in the log of the ridership per capita. A 10% increase in the total population is associated with a 0.8% increase in ridership per capita. This finding shows a limited impact of population growth in promoting transit ridership. A 10% increase in working population is associated with a 0.7% increase in ridership per capita, while a 10% increase in senior citizens is associated with a 0.08% increase in ridership per capita. This implies that the working population is more likely to use public transit than the retired population. Similarly, a 10% increase in percentage of postsecondary students is associated with a 0.29% increase in ridership per capita. This magnitude is about four times smaller than Diab et al. (2020)'s findings, as they found "a 10% increase in percentage of postsecondary students is associated with less than 1.17% increase in ridership (total number of linked trips by transit agency per year)". (p.108)

Table 3: Model I regression results

Source	SS	df	MS			
				Number of obs	=	251
				F(16,234)	=	102.92
Model	209.5224	16	13.09515	Prob > F	=	0
Residual	149.4197	234	0.638546	R-squared	=	0.5837
				Adj R-squared	=	0.5553
Total	358.9421	250	1.435768	Root MSE	=	0.79909

Log(ridershippercapita)	**Coef.**	**Std. Err.**	**t**	**P> | t |**
Log (totalpop)	0.0805	0.0564	1.43	0.155
pct15_64years	0.0755	0.0294	2.56	0.011
pctover65years	0.0081	0.0235	0.34	0.731
pctpostcert	0.0292	0.0088	3.33	0.001
empl_rate	0.0147	0.0139	1.06	0.292
Log(aver_inc_hhold)	-1.3188	0.4341	-3.04	0.003
Log(aver_ownrent)	-0.4008	0.1393	-2.88	0.004
pct_ten_rentmore30per	0.0000	0.0000	-4.19	0
Log(aver_tenrent)	-0.7951	0.2140	-3.71	0
pct_driv_cartruck	-0.0815	0.0100	-8.16	0
pct_pass_cartruck	-0.0445	0.0328	-1.36	0.176
pct_walk	-0.0825	0.0212	-3.90	0
yeardummy1	-1.0056	0.2823	-3.56	0
yeardummy2	-1.0791	0.2717	-3.97	0
yeardummy3	-0.3929	0.1543	-2.55	0.012
rural	-0.4987	0.1364	-3.66	0
constant	25.0795	4.8950	5.12	0

Although the employment rate has a positive relationship with ridership per capita, its impact (as indicated by 0.15% increase) on ridership per capita is smaller than might be expected. Also, the employment rate is not statistically significant at any level of significance. On the other hand, a few studies find that there is a stronger association between employment rate and transit supply, though the effect of one on the other is ambiguous and there is very little evidence of the degree to which one affects the other (Hughes, 1991; Paul, 1990; Thomas, 1999). These studies have considered the relative impacts of employment accessibility that result from public transportation availability. In his study, Thomas (1999) analyzes the impact of public transportation on labor force participation for Portland, Oregon, and Atlanta, Georgia; his results suggest that public transit accessibility is a significant factor in determining the average labor participation rates within two cities. Hence, future research is needed to better understand the connections between transit needs and employment activities, which would be helpful in estimating transit ridership.

A 10% increase in average household income is associated with a 13.138% reduction in ridership per capita. This finding is reasonable as people can afford a car when their income increases, which in turn discourages public transit use. While both average monthly tenant rent and owner rent is negatively associated with ridership per capita, tenant rent shows a steeper decline in ridership per capita when monthly tenant rent increases by 10%. As seen in Table 2, a 10% increase in average owner's major payment is associated with a 4% decrease in ridership per capita, while a 10% increase in average monthly shelter cost for rented dwellings is associated with a 7.9% decrease in ridership per capita.

For the commuting mode, a 10% increase in employed population who drives to work is associated with a 0.8% decrease in ridership per capita. A 10% increase in employed population who rideshares and walks to work is associated with a 0.4% and 0.5% decrease in ridership per capita, respectively. The negative sign of the commuting modes coefficients suggest they are the substitutes for public transit rather than complements. Using 2016 as the base year, ridership per capita in 1996, 2001 and 2011 is relatively lower than 2016. This increase in ridership may be due to the increase in transit services which results from the expansion of transit projects and building of new stations or subways in the recent years.

Ridership per capita in rural areas is 49.9% lower than in urban areas. The incorporated interaction terms fail to capture the impacts of the rural dummy with the corresponding independent variables on ridership per capita. Therefore, the interaction terms are dropped from model I.

5.2 Model II Results

This section will demonstrate the empirical relationship between cost per trip and public transit demand side variables using a one-stage OLS linear regression model as well as incorporating a dummy variable to distinguish per-trip cost between urban and rural regions in Canada. As mentioned earlier, two specification types are used here: equation (2) which does not control for economies of scale, and equation (3) which includes ridership as an independent variable to control for the economies of scale. Results for equations (2) and (3) are presented in Tables 4 and 5, respectively. The only difference between the two specifications is indicated by the presence of the logarithm of ridership, and it is important to emphasise here that what was measured is economies of output related to passenger trips. The ways that regression coefficients are interpreted follow Model I, as we assume other independent variables remain constant when we examine the partial effect of a corresponding independent variable.

Table 4: Model II's specification 1 regression results

Source	SS	df	MS			
				Number of obs	=	249
				F(16, 232)	=	25.82
Model	54.85674	16	3.428546	Prob > F	=	0
Residual	30.81092	232	0.132806	R-squared	=	0.6403
				Adj R-squared	=	0.6155
Total	85.66766	248	0.345434	Root MSE	=	0.36443

Log(costpertrip)	**Coef.**	**Std. Err.**	**t**	**P > |t|**
Log(totalpop)	-0.129	0.0258	-5.01	0
pct15_64years	-0.010	0.0135	-0.71	0.476
pctover65years	-0.008	0.0107	-0.79	0.433
pctpostcert	-0.005	0.0040	-1.23	0.222
empl_rate	-0.017	0.0065	-2.66	0.008
Log(aver_inc_hhold)	0.466	0.1994	2.34	0.02
Log(aver_ownrent)	0.187	0.0635	2.94	0.004
pct_ten_rentmore30per	0.000	0.0000	2.38	0.018
Log(aver_tenrent)	0.488	0.0976	5.00	0
pct_driv_cartruck	0.007	0.0046	1.56	0.121
pct_pass_cartruck	-0.021	0.0150	-1.43	0.153
pct_walk	-0.016	0.0097	-1.61	0.108
yeardummy1	-0.195	0.1289	-1.52	0.131
yeardummy2	-0.160	0.1242	-1.29	0.199
yeardummy3	-0.030	0.0706	-0.43	0.667
rural	0.171	0.0625	2.74	0.007
constant	-5.077	2.2479	-2.26	0.025

Table 5: Model II's specification 2 regression results

Source	SS	df	MS			
				Number of obs	=	249
				F(17, 231)	=	59.91
Model	69.82938	17	4.107611	Prob > F	=	0
Residual	15.83828	231	0.068564	R-squared	=	0.8151
				Adj R-squared	=	0.8015
Total	85.66766	248	0.345434	Root MSE	=	0.26185

Log(costpertrip)	**Coef.**	**Std. Err.**	**t**	**P > |t|**
Log(totalpop)	0.214	0.0186	-5.72	0
pct15_64years	0.017	0.0099	1.73	0.085
pctover65years	-0.006	0.0077	-0.77	0.441
pctpostcert	0.005	0.0030	1.53	0.128
empl_rate	-0.015	0.0047	-3.24	0.001
Log(aver_inc_hhold)	0.078	0.1457	0.54	0.593
Log(aver_ownrent)	0.058	0.0465	1.25	0.211
pct_ten_rentmore30per	0.000	0.0000	-0.77	0.444
Log(aver_tenrent)	0.231	0.0723	3.19	0.002
pct_driv_cartruck	-0.019	0.0037	-5.04	0
pct_pass_cartruck	-0.035	0.0108	-3.24	0.001
pct_walk	-0.042	0.0072	-5.79	0
yeardummy1	-0.507	0.0950	-5.34	0
yeardummy2	-0.492	0.0920	-5.35	0
yeardummy3	-0.148	0.0514	-2.89	0.004
rural	0.010	0.0462	0.21	0.83
Log(ridership)	-0.321	0.0217	-14.78	0
constant	2.581	1.6962	1.52	0.129

Focusing specifically on Table 4 (the specification without ridership), approximately three quarters of independent variables are negatively associated with cost per trip with mixed levels of statistical significance. Among these variables, total population has relatively high impact on cost per-trip: A 10% increase in total population is associated with a 12% decrease in cost per trip, and this relationship is statistically significant. Since the results in Table 4 do not control for the transit supply factors and scale economies, both might play their roles here. In other words, increasing returns to scale with respect to public transit supply may affect the relationship between cost per trip and total population and thus producing misleading results. Similar trends are observed for other independent variables, including the percentage of working population and percentage of postsecondary students. These three independent variables have a negative and statistically insignificant relationship with cost per trip. Taking the Model 1 results into account, ridership increases as percentage of working population and postsecondary students increases, respectively. Therefore, we would expect these two variables to move in the same direction as cost per trip.

Employment rate is negatively linked with cost per trip, and this association is statistically significant at 5% level of significance. Although a positive relationship between the two variables is expected, however, using employment rate alone is insufficient to explain the underlying impact of labor force characteristics on cost per trip. Instead, some studies which examine transportation accessibility using distance between workplace and transit service station provide more information for estimating operating cost (Thomas, 1999; Cooke and Behrens, 2017). For the percentage of employed group who uses private vehicles as a mode of transport for work, driving alone shows a very weak positive association with cost per trip while ridesharing presents a negative and relatively stronger impact on cost per trip. On the other hand, a 10% increase in the employed population walking to work is associated with a 0.16% reduction in cost per trip. Again, the implication from these results are very limited. We can only deduce that ridesharing and walking is a substitution for public transit use, while employees who drive to work complements public transit based on the signs of the regression coefficients.

When average household income increases by 10%, the expected cost per trip drop by 4.7%. As mentioned earlier, people can afford to buy a car when their income increases. Hence, transit ridership would fall as car ownerships increases. Since this model do not control for transit service factors, there is no evidence to prove whether the reduced operating cost is caused by the transit service adjustments (i.e. reduce service hours/ frequencies) in order to deal with the falling transit use. Both average monthly gross rent and owner's major payment have a positive and significant impact on cost per trip, respectively. For the year dummy, each of them presents lower cost per trip compared to the base year, 2016. One may argue that despite the rising transit ridership from 1996 to 2016, multiple public transportation modes that contributes scale economies (especially rails) and technology advancements which makes production factors cheaper has reduced the operating cost per trip. In fact, the regression coefficients for each year dummy are larger and statistically significant when controlling for economies of scales (Table 4). Focusing on the rural dummy, cost per trip in rural Canada is 17.09% more than in urban Canada. Due to weak quality and quantity of transit supply in rural area, most people rely more on car to get to their destination. Thus, lower per-trip passenger volume while maintaining the same service hours/frequencies leads to higher per-trip operating cost in rural area.

Table 5 presents the regression results after including the log of ridership in the regression equation, and there is a dramatic increase in R-squared and adjusted R-squared (approximately 20%) compared to the previous specification. This may imply that economies of scale is a relevant variable for explaining variation in cost per trip. The signs of regression coefficients change for a few independent variables such as total population, % of working population, % of postsecondary students, and % of employed population who drives to work, while a number of them changes in terms of magnitudes: average household income, average major payment for owner, average tenant gross rent, and year dummies.

After controlling for economies of scale in terms of the log of ridership, the regression results shows that a 10% increase in total population is associated with a 2.1% increase in cost per trip. The results meet our expectation made previously. The partial effects of percentage of working population and postsecondary students on cost per trip becomes positive after controlling for economies of scale, implying our previous guess is correct. Interestingly, employment rate is the least affected independent variable, and its coefficient estimates is nearly the same as before. Now, the percentage of employed group who drives to work has a negative and statistically significant relationship with cost per trip.

Average household income, average gross rent, and average major payments for owners exert a much smaller positive effect than before: 0.8%, 2.3%, and 0.6% on cost per trip when these three variables increase by 10%, respectively. For the rural dummy, the regression coefficient has dropped significantly from 17.09% to 1%, indicating than estimated cost per trip in rural transit is 1% more than in urban transit. In other words, the predicted per-trip cost difference becomes smaller after controlling for the differences in ridership levels across urban and rural Canada. Moreover, the rural dummy becomes statistically insignificant after the model controls for economies of scale. On the other hand, a 10% increase in ridership (measure of economies of scale) is associated with a 3.2% decrease in cost per trip, indicating the economies of scale in this model. Tying these findings together, public transit in rural area would need more ridership to reduce the high per-trip operating cost. Unfortunately, the incorporated interaction terms fail to capture the impacts of the rural dummy with the corresponding independent variables on cost per trip. Therefore, the attempt to evaluate the critical threshold value of public transit output level for rural and urban area in Canada was unsuccessful, and hence the interaction terms are dropped from both specifications of model II.

5.3 Model III Results

This section will examine the empirical relationship between subsidy per trip and a number of public transit demand side factors using a one-stage OLS linear regression model as well as incorporating a dummy variable to distinguish per-trip subsidy between urban and rural regions in Canada. Two specification types are used here: equation (4) which does not control for economies of scale, and equation (5) which includes ridership as an independent variable to control for the economies of scale. Results for equations (4) and (5) are presented in Tables 6 and 7, respectively. The only difference between the two specifications is indicated by the presence of the logarithm of ridership, and it is important to emphasise here that what was measured is economies of output related to passenger trips. The ways that regression coefficients are interpreted follow Model I and II, as we assume other independent variables remain constant when we examine the partial effect of a corresponding independent variable.

Table 6: Model III's specification 1 regression results

Source	SS	df	MS			
				Number of obs	=	248
				F(16, 231)	=	21.9
Model	97.2473	16	6.077956	Prob > F	=	0
Residual	64.1155	231	0.277556	R-squared	=	0.6027
				Adj R-squared	=	0.5751
Total	161.3628	247	0.653291	Root MSE	=	0.52684

Log(subsidypertrip)	**Coef.**	**Std. Err.**	**t**	**P > |t|**
Log(totalpop)	-0.186	0.0375	-4.95	0
pct15_64years	-0.021	0.0195	-1.08	0.282
pctover65years	-0.018	0.0155	-1.14	0.257
pctpostcert	-0.007	0.0058	-1.2	0.232
empl_rate	-0.022	0.0095	-2.33	0.021
Log(aver_inc_hhold)	0.803	0.2883	2.79	0.006
Log(aver_ownrent)	0.150	0.0919	1.63	0.104
pct_ten_rentmore30per	0.000	0.0000	1.11	0.269
Log(aver_tenrent)	0.314	0.1411	2.22	0.027
pct_driv_cartruck	0.010	0.0066	1.55	0.122
pct_pass_cartruck	-0.028	0.0217	-1.29	0.199
pct_walk	-0.018	0.0140	-1.25	0.211
yeardummy1	-0.379	0.1863	-2.03	0.043
yeardummy2	-0.343	0.1796	-1.91	0.057
yeardummy3	-0.030	0.1024	-0.3	0.767
rural	0.216	0.0905	2.39	0.018
constant	-6.160	3.2498	-1.9	0.059

Table 7: Model III's specification 2 regression results

Source	SS	df	MS			
				Number of obs	=	248
				F(17, 230)	=	41.01
Model	121.33367	17	7.1372749	Prob > F	=	0
Residual	40.029124	230	0.1740397	R-squared	=	0.7519
				Adj R-squared	=	0.7336
Total	161.3628	247	0.6532907	Root MSE	=	0.41718

lnsubsidypertrip	**Coef.**	**Std. Err.**	**t**	**P > |t|**
Log(totalpop)	0.249	0.0298	-5.36	0
pct15_64years	0.014	0.0158	0.86	0.388
pctover65years	-0.014	0.0123	-1.16	0.249
pctpostcert	0.005	0.0047	1.14	0.254
empl_rate	-0.020	0.0075	-2.66	0.008
Log(aver_inc_hhold)	0.308	0.2321	1.33	0.186
Log(aver_ownrent)	-0.015	0.0741	-0.2	0.839
pct_ten_rentmore30per	0.000	0.0000	-1.77	0.079
Log(aver_tenrent)	-0.015	0.1152	-0.13	0.895
pct_driv_cartruck	-0.023	0.0059	-3.84	0
pct_pass_cartruck	-0.044	0.0172	-2.55	0.011
pct_walk	-0.051	0.0115	-4.47	0
yeardummy1	-0.774	0.1513	-5.12	0
yeardummy2	-0.766	0.1467	-5.23	0
yeardummy3	-0.174	0.0820	-2.12	0.035
rural	0.016	0.0736	0.21	0.832
lnridershippercap	-0.409	0.0347	-11.76	0
constant	3.624	2.7044	1.34	0.182

Looking at the specification excluding the log of ridership (Table 6), a 10% increase in total population is associated with a 1.8% reduction in subsidy per trip, and this impact is statistically significant. A 10% increase in the percentage of working population, percentage of retired population, and percentage of postgraduate students are linked with a 0.21%, 0.17%, and 0.06% reduction in subsidy per trip, respectively. Most importantly, these variables are also negatively associated with cost per trip in Model II. The cost per trip and the subsidy per trip move in the same direction for these variables, implying consistent findings. Employment rate shows a negative and statistically significant relationship with subsidy per trip. As employment rate increases by 10%, subsidy per trip is reduced by 0.22%.

Although we expect average household income to have a negative relationship with subsidy per trip, the results shows that a 10% increase in average household income is linked with an 8% increase in subsidy per trip. This unexpected positive relationship may reflect the issue of income distributional effects of transit subsidies and using different income groups may better explain this relationship (Frankena, 1973).

For the percentage of employed group who uses private vehicles as a mode of transport for work, driving alone shows a very weak positive association with subsidy per trip while ridesharing and walking also presents a very weak but negative relationship with subsidy per trip, respectively. None of these three variables are statistically significant at any level of significance. Again, these findings are consistent with the Model II results (without controlling for economies of scale). On the other hand, both average owner's major payments and average gross rent presents a positive relationship with subsidy per trip.

For the year dummy, the estimated subsidy per trip for each year (1996, 2001, and 2011) are less than the base year of 2016. Model 2's year dummies results show that the predicted cost per trip for each year dummy is smaller than the base year, and hence the estimated subsidy per trip for year dummies are consistent with Model II results. In the context of rural area, regression results report rural transit requires 21.6% more per-trip subsidy than urban transit.

After controlling for economies of scale in terms of the log of ridership, R-squared and adjusted R-squared increase from 60% and 58% to 75% and 73%, respectively. This imply the relevance of economies of scale in explaining variation in per-trip subsidy. The signs for a number of independent variables have changed, including total population, percentage of working population, percentage of postsecondary students, average owner's major payments, and percentage of employed population who drives to work. Total population is now positively linked to per-trip subsidy, where a 10% increase in total population is associated with a 2.5% increase in per-trip subsidy. Higher population may lead to more transit supply, increasing operating costs and therefore increasing subsidy. Similar explanations can apply for the increased working population and postsecondary students. Following the Model II results, the regression coefficient for employment rate does not change much and remains statistically significant after controlling for economies of scale.

Both average owner's major payments and average gross rent becomes negatively associated with per-trip subsidy after controlling for ridership levels. As the employed population who drives to work increases by 10%, the estimated per-trip subsidy falls by 0.2%. This may reflect that driving complements public transit. In the context of average household income, its regression coefficient drops drastically from 8% to 3% and becomes statistically insignificant at any level of significance after the model control for ridership levels.

For the rural dummy, the regression coefficient has dropped significantly from 21.6%% to 1.5%, indicating than estimated subsidy per trip in rural transit is 1.5% more than in urban transit after controlling for economies of scale. In other words, the predicted per-trip subsidy difference becomes smaller after controlling for the differences in ridership levels across urban and rural Canada. Moreover, the rural dummy becomes statistically insignificant after the model control for economies of scale. On the other hand, a 10% increase in ridership level (measure of economies of scale) is associated with a 4.08% decrease in subsidy per trip. While this is clearly not a one-to-one ratio, it is important to compare transit supply factors such as passenger trip per vehicle-kilometer to the estimated cost per trip to evaluate the operational efficiency, and this would help in deciding if the predicted per-trip subsidy is sustainable.

Unfortunately, the incorporated interaction terms fail to capture the impacts of the rural dummy with the corresponding independent variables on subsidy per trip. Therefore, the attempt to evaluate the critical threshold value of each independent variable for rural and urban area in Canada was unsuccessful, and hence the interaction terms are dropped from both specifications of model III.

6. Conclusion and Policy Implications

The goal of this thesis is to examine the empirical relationship on: (1) the factors affecting transit demand, (2) the factors affecting transit supply costs, and (3) the factors affecting the required subsidy per trip. Most importantly, this thesis focuses on identifying the potential differences of ridership demand and supply, and the required subsidy, between urban and rural Canadian transit operations. Thereby, a pooled cross-sectional dataset comprised exclusively of transit demand factors from 1996 to 2016 was used for the one-stage OLS regression for all three models. It should be noted that none of the existing literature has compared the differences of transit demand and supply, and the required subsidy between rural and urban transit operators in Canada.

For (1), the factors affecting transit demand, many studies have been conducted to study factors affecting transit ridership. Most of the independent variables in this thesis display a similar relationship with transit ridership as the previous studies, although the magnitudes of the regression estimates differ. From the empirical analysis, this thesis find ridership per capita in rural Canada is 50% less than in urban Canada. In addition, this impact is statistically significant. Due to the existing transportation accessibility and high passenger volume, urban transit projects may generate a higher predicted revenue-cost ratio. Based on this result, the recent announced Federal public transit funds which lump urban and rural public transit funds together has put rural transit operators into a disadvantageous position. Hence, transit authorities and policymakers should be aware of the ridership gap between rural and urban Canada when implementing transit-related policies.

For (2), the factors affecting transit supply costs: unlike previous studies that focus more on transit supply factors, this thesis uses a transit demand side approach to examine the effect of transit demand factors on operating cost per trip. Two model specifications are developed to evaluate the impact of economies of scale on transit supply cost. The predicted per-trip cost difference becomes significantly smaller after controlling for the differences in ridership levels across urban and rural Canada. This thesis finds economies of scale since the cost per trip decreases when ridership levels increase. The regression coefficient of this model is difficult to analyze completely, as the model does not control for transit supply variables.

For (3), the factors affecting the required subsidy per trip: unlike previous studies that focus more on transit supply factors, this thesis uses a transit demand side approach to investigate the effect of transit demand factors on subsidy per trip. Two model specifications are developed to evaluate the impact of economies of scale on transit subsidy. The predicted per-trip subsidy difference drops significantly after controlling for the differences in ridership levels across urban and rural Canada. Since the subsidy per trip represents a less than proportionate change in ridership levels, it is important to compare transit supply factors such as passenger trip per vehicle-kilometer to the estimated cost per trip to evaluate the operational efficiency. This would help in deciding if the predicted per-trip subsidy is sustainable. Otherwise, the unsustainable transit subsidy resulting from operational inefficiency would cause unnecessary fiscal burdens on government and taxpayers.

This study has some data limitations. Due to the missing values of some relevant variables (for example, median commuting duration), it was not possible to investigate the association of these factors with the dependent variables. Due to the small number of observations, interaction terms fail to fail to capture the impacts of the rural dummy with the corresponding independent variables on the dependent variables. Future study with a larger number of observations may help to solve this issue. With that said, the empirical results highlight the importance of increasing transit ridership levels in order to reduce operational subsidies. Future research should compare the costs of policies designed to increase transit ridership with the effects of operating subsidies (and external impacts on congestion and the environment), to help guide policymakers.

Reference List

Primary Sources

Alfa, Attahiru Sule, & Clayton, Diana. (1986). An improved procedure for allocating public transit operating subsidies in Canada. *Journal of Advanced Transportation, 20*(3), 223-238.

Berechman, Joseph. (1983). Costs, Economies of Scale and Factor Demand in Bus Transport: An Analysis. *Journal of Transport Economics and Policy, 17*(1), 7-24.

Berechman, J, & Giuliano, G. (1985). Economies of scale in bus transit: A review of concepts and evidence. *Transportation (Dordrecht), 12*(4), 313-332.

Boisjoly, Geneviève, Grisé, Emily, Maguire, Meadhbh, Veillette, Marie-Pier, Deboosere, Robbin, Berrebi, Emma, & El-Geneidy, Ahmed. (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, 116*, 434-445.

Börjesson, Maria, Fung, Chau Man, & Proost, Stef. (2020). How rural is too rural for transit? Optimal transit subsidies and supply in rural areas. *Journal of Transport Geography, 88*, 102859.

Chen, Cynthia, Don Varley, & Jason Chen. (2011). What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags and Asymmetric Behaviour. *Urban Studies (Edinburgh, Scotland), 48*(9), 1893-1908.

Cooke, S, & Behrens, R. (2017). Correlation or cause? The limitations of population density as an indicator for public transport viability in the context of a rapidly growing developing city. *Transportation Research Procedia. 25*, 3003-3016.

Coulombel, Nicolas, & Monchambert, Guillaume. (2019). Diseconomies of scale and subsidies in urban public transportation. Retrieved from ⟨hal-02373768v1⟩

CUTA. (2016). Canadian Transit Fact Book.

Diab, Ehab, Kasraian, Dena, Miller, Eric J, & Shalaby, Amer. (2020). The rise and fall of transit ridership across Canada: Understanding the determinants. *Transport Policy, 96*, 101-112.

Frankena, Mark. (1973). Income Distributional Effects of Urban Transit Subsidies. *Journal of Transport Economics and Policy, 7*(3), 215-230.

Guerra, Erick, & Cervero, Robert. (2011). Cost of a Ride. *Journal of the American Planning Association, 77*(3), 267-290.

Haire, Ashley Raye. (2009). *A Methodology for Incorporating Fuel Price Impacts into Short-term Transit Ridership Forecasts*.

Iseki, Hiroyuki. (2008). Economies of scale in bus transit service in the USA: How does cost efficiency vary by agency size and level of contracting? *Transportation Research. Part A, Policy and Practice, 42*(8), 1086-1097.

Jones, R. P. (2021, February 10). Trudeau pledges billions in permanent funding for public transit. CBC News. Retrieved from https://www.cbc.ca/news/politics/trudeau-transit-fund-1.5908346

Kain, John F, & Liu, Zvi. (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, 33*(7), 601-624.

Mattson, Jeremy. (2017). Estimating Ridership of Rural Demand–Response Transit Services for the General Public. *Transportation Research Record, 2647*(1), 127-133.

Nutley, Stephen D. (1996). Rural transport problems and non-car populations in the USA: A UK perspective. *Journal of Transport Geography, 4*(2), 93-106.

Parry, Ian W.H, & Small, Kenneth A. (2009). Should Urban Transit Subsidies Be Reduced? *The American Economic Review, 99*(3), 700-724.

Rivers, Nicholas, & Plumptre, Bora. (2018). The effectiveness of public transit tax credits on commuting behaviour and the environment: Evidence from Canada. *Case Studies on Transport Policy, 6*(4), 651-662.

Sakai, Hiroki, & Shoji, Kenichi. (2010). The effect of governmental subsidies and the contractual model on the publicly-owned bus sector in Japan. *Research in Transportation Economics, 29*(1), 60-71.

Sale, James E, & Green, Bryan. (1979). Operating Costs and Performance of American Public Transit Systems. *Journal of the American Planning Association, 45*(1), 22-27.

Sanchez, Thomas W. (1999). The Connection Between Public Transit and Employment. *Journal of the American Planning Association, 65*(3), 284-296

Santoso, Djoen San, Yajima, Masaru, Sakamoto, Kunihiro, & Kubota, Hisashi. (2012). Opportunities and strategies for increasing bus ridership in rural Japan: A case study of Hidaka City. *Transport Policy, 24*, 320-329.

Skinner, R. E. (1981) Estimating Transit Supply Requirements for Alternatives Analysis. *Transportation Research Record*, *835*, 24–30.

Statistics Canada. (2012). Focus on Geography Series, 2011 Census. Statistics Canada Catalogue no. 98-310-XWE2011004. Ottawa, Ontario. Analytical products, 2011 Census. Retrieved from https://www12.statcan.gc.ca/census-recensement/2011/as-sa/fogs-spg/Facts-cma-eng.cfm?LANG=Eng&GK=CMA&GC=940

Statistics Canada. (2016). Geographic Classifications. Retrieved from https://www.statcan.gc.ca/eng/subjects/standard/pcrac/2016/introduction

Taylor, Brian D, Miller, Douglas, Iseki, Hiroyuki, & Fink, Camille. (2009). Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. *Transportation Research. Part A, Policy and Practice, 43*(1), 60-77.

Taylor, B D, & Fink, C N Y. (2013). Explaining transit ridership: What has the evidence shown? *Transportation Letters, 5*(1), 15-26.

Thompson, Gregory, & Brown, Jeffrey. (2006). Explaining Variation in Transit Ridership in U.S. Metropolitan Areas Between 1990 and 2000: Multivariate Analysis. *Transportation Research Record, 1986*, 172-181.

Yang, Jiawen, Zhou, Huaqing, & Zhou, Meng. (2020). Bus transit subsidy under China's transit metropolis initiative: The case of Shenzhen. *International Journal of Sustainable Transportation, 14*(1), 56-63.

Secondary Sources

Chandler, Vincent. (2014). The effectiveness and distributional effects of the tax credit for public transit. Can. *Public Policy 40*(3), 259–269.

Department of Finance Canada. (2012). Tax expenditures and evaluations 2011. Government of Canada. Retrieved from https://www.canada.ca/content/dam/fin/migration/taxexp-depfisc/2011/taxexp11-eng.pdf

Gómez-Ibáñez JA. (1996). Big-city transit, ridership, deficits, and politics, *J. Am. Plan. Ass.*, *62*(1), 30–50.

Hughes, Mark Alan. (1991). Employment Decentralization and Accessibility: A Strategy for Stimulating Regional Mobility. *Journal of the American Planning Association, 57*(3), 288-298.

Kain, John F, & Liu, Zvi. (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, 33*(7), 601-624.

Giuliano, Genevieve. (1980). Transit performance: the effect of environmental factors. Dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Social Sciences, University of California, Irvine, Irvine.

Gomez-Ibanez, J. (1996). Big-City Transit Ridership, Deficits, and Politics: Avoiding Reality In Boston. *Journal of the American Planning Association, 62*(1), 30.

Lee, Bumsoo, & Lee, Yongsung. (2013). Complementary Pricing and Land Use Policies: Does It Lead to Higher Transit Use? *Journal of the American Planning Association, 79*(4), 314-328.

Li, Jianling, & Wachs, Martin. (2004). The effects of federal transit subsidy policy on investment decisions: The case of San Francisco's Geary Corridor. *Transportation (Dordrecht), 31*(1), 43-67.

Liu Z. (1993). Determinants of Public Transit Ridership: Analysis of Post World War II Trends and Evaluation of Alternative Networks. Harvard University, Cambridge, MA, USA.

McLeod MS, Flannelly KJ, Flannelly L, Behnke RW. (1991). Multivariate time-series model of transit ridership based on historical, aggregate data: The past, present and future of Honolulu, *Transport. Res. Record*, *1297*, 76–84.

Paul R. BLACKLEY. (1990). Spatial Mismatch in Urban Labor Markets: Evidence from Large U.S. Metropolitan Areas. *Social Science Quarterly, 71*(1), 39-52

Rajindar K. Koshal. (1970). Economies of Scale in Bus Transport: II: Some Indian Experience. *Journal of Transport Economics and Policy, 4*(1), 29-36.

Wabe, S. J., and O. B. Coles. (1975). The Short and Long Run Cost of Bus Transport in Urban Areas. *Journal of Transport Economics and Policy. 9*(2), 127-140.